

21. Comparing Digital Flashcards and Contextual Instruction for L2 Vocabulary Retention¹

İhsan ÜNALDI² & Bilal ÜSTÜN³ & Aydan IRGATOĞLU⁴ & Lokman TANRIKULU⁵ & Ercan DEMİRCİ⁶

APA: Ünalı, İ. & Üstün, B. & Irgatoğlu, A. & Tanrikulu, L. & Demirci, E. (2025). Comparing Digital Flashcards and Contextual Instruction for L2 Vocabulary Retention. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (49), 389-412. **DOI:** <https://doi.org/10.5281/zenodo.17973309>

Abstract

This quasi-experimental study compared the effects of digital flashcards and traditional contextual instruction on the retention of L2 vocabulary. The setting was a preparatory program at a Turkish university where English is taught as a foreign language. At the beginning of the 2023-2024 academic year, the students in this institution were grouped according to their proficiency levels, and two of these groups at the A2 proficiency level were assigned as the experimental and control groups for the current study (total n=30). During the eight-week intervention, the experimental group studied 110 target words using digital flashcards on Quizlet, a web-based learning platform, and the control group studied the same target words in contextual settings through reading activities. ANCOVA calculations for the posttest and delayed posttest (4 weeks later) indicated that both groups made notable progress during the intervention, and the short-term effects of digital flashcard learning and contextual instruction were similar. However, it was clear that the long-term impacts of contextual instruction on retention were significantly better. These findings suggest that learning vocabulary through digital flashcards is as effective as contextual instruction in the short term, both in basic recognition and

¹ **Statement (Thesis / Paper):** It is declared that scientific and ethical principles were followed during the preparation process of this study and all the studies utilised are indicated in the bibliography. This article was produced as part of the "Project to Develop Foreign Language Teaching Materials Focused on the Region's Natural and Cultural Heritage for Use in Tourism Education Programs in the Cappadocia Region" (Specialization Project).

Conflict of Interest: No conflict of interest is declared.

Funding: No external funding was used to support this research.

Copyright & Licence: The authors own the copyright of their work published in the journal and their work is published under the CC BY-NC 4.0 licence.

Source: It is declared that scientific and ethical principles were followed during the preparation of this study and all the studies used are stated in the bibliography.

Similarity Report: Received – Turnitin, Rate: 18%

Ethics Complaint: editor@rumelide.com

Article Type: Research article, **Article Registration Date:** 09.11.2025-**Acceptance Date:** 18.12.2025-**Publication Date:** 19.12.2025; **DOI:** <https://doi.org/10.5281/zenodo.17973309>

Peer Review: Two External Referees / Double Blind

² Doç. Dr., Nevşehir Hacı Bektaş Veli Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü / Assoc. Prof., Nevşehir Hacı Bektaş Veli University, Faculty of Education, Department of Foreign Language Education (Nevşehir, Türkiye), **eposta:** ihsan@nevsehir.edu.tr **ORCID ID:** <https://orcid.org/0000-0003-0009-3537> **ROR ID:** <https://ror.org/019jds967> **ISNI:** 0000 0004 0386 1930 **Crossref Funder ID:** 100019891

³ Doç. Dr., Nevşehir Hacı Bektaş Veli Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü / Assoc. Prof., Nevşehir Hacı Bektaş Veli University, Faculty of Education, Department of Foreign Language Education (Nevşehir, Türkiye), **eposta:** bustun@nevsehir.edu.tr **ORCID ID:** <https://orcid.org/0000-0002-9490-3109> **ROR ID:** <https://ror.org/019jds967> **ISNI:** 0000 0004 0386 1930 **Crossref Funder ID:** 100019891

⁴ Doç. Dr., Ankara Hacı Bayram Veli Üniversitesi, Yabancı Diller Yüksekokulu / Assoc. Prof., Ankara Hacı Bayram Veli University, School of Foreign Languages (Ankara, Türkiye), **eposta:** aydan.irtatoglu@hbyv.edu.tr **ORCID ID:** <https://orcid.org/0000-0002-0773-4619> **ROR ID:** <https://ror.org/05mskc574> **ISNI:** 0000 0004 7221 6011

⁵ Doç. Dr., Nevşehir Hacı Bektaş Veli Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü / Assoc. Prof., Nevşehir Hacı Bektaş Veli University, Faculty of Education, Department of Foreign Language Education (Nevşehir, Türkiye), **eposta:** ltanrikulu@nevsehir.edu.tr **ORCID ID:** <https://orcid.org/0000-0002-1862-8176> **ROR ID:** <https://ror.org/019jds967> **ISNI:** 0000 0004 0386 1930 **Crossref Funder ID:** 100019891

⁶ Doç. Dr., Nevşehir Hacı Bektaş Veli Üniversitesi, Yabancı Diller Yüksekokulu / Assoc. Prof., Nevşehir Hacı Bektaş Veli University, School of Foreign Languages (Nevşehir, Türkiye), **eposta:** e.demirci@nevsehir.edu.tr **ORCID ID:** <https://orcid.org/0000-0002-6702-0247> **ROR ID:** <https://ror.org/019jds967> **ISNI:** 0000 0004 0386 1930 **Crossref Funder ID:** 100019891

contextual application of the target words; however, contextual instruction is significantly more effective in the long term in both aspects.

Keywords: L2 vocabulary, digital flashcards, contextual learning, Quizlet

Yabancı Sözcüklerin Kalıcılığı Açısından Dijital Kelime Kartları ile Bağlamsal Öğretimin Karşılaştırılması⁷

Öz

Bu yarı deneysel çalışma, dijital flaş kartların ve geleneksel bağlamsal öğretimin ikinci dil kelime dađarcığının kalıcılığı üzerindeki etkilerini karşılaştırmıştır. Çalışma, İngilizenin yabancı dil olarak öğretildiği bir Türk üniversitesindeki hazırlık programında gerçekleştirilmiştir. 2023-2024 akademik yılının başında, bu kurumdaki öğrenciler yeterlilik seviyelerine göre gruplara ayrılmış ve A2 yeterlilik seviyesindeki iki grup, mevcut çalışma için deney grubu ve kontrol grubu olarak belirlenmiştir (toplam n=30). Sekiz haftalık müdahale süresince, deney grubu 110 hedef kelimeyi web tabanlı bir öğrenme platformu olan Quizlet'te dijital flaş kartlar kullanarak çalıştı ve kontrol grubu da aynı hedef kelimeleri okuma etkinlikleri yoluyla bağlamsal ortamlarda çalıştı. Son test için yapılan ANCOVA hesaplamaları, her iki grubun da müdahale süresince kayda değer ilerlemeler kaydettiğini ve dijital öğrenme ile bağlamsal öğretimin kısa dönem etkilerinin benzer olduğunu gösterdi. Ancak sonuçlar bağlamsal öğretimin kalıcılık üzerindeki uzun vadeli etkilerinin önemli ölçüde daha iyi olduğunu da gösterdi. Bu bulgular, dijital flaş kartlarla kelime öğrenmenin hem temel tanıma hem de hedef kelimelerin bağlamsal uygulaması açısından kısa vadede bağlamsal kelime öğretimi kadar etkili olduğunu göstermektedir; ancak, bağlamsal öğretim her iki açıdan da uzun vadede önemli ölçüde daha etkilidir.

Anahtar kelimeler: Yabancı dil kelime bilgisi, dijital kelime kartları, bağlamsal öğrenme, Quizlet

⁷ **Beyan (Tez/ Bildiri):** Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur. Bu makale "Kapadokya Bölgesinde Turizm Eğitimi Veren Programlarda Kullanılabilecek Bölgenin Doğal ve Kültürel Mirasına Odaklı Yabancı Dil Öğretim Materyali Geliştirme Projesi" (İhtisarlaşıma Projesi) kapsamında üretilmiştir.

Çıkar Çatışması: Çıkar çatışması beyan edilmemiştir.

Finansman: Bu araştırmayı desteklemek için dış fon kullanılmamıştır.

Telif Hakkı & Lisans: Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları CC BY-NC 4.0 lisansı altında yayımlanmaktadır.

Kaynak: Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Benzerlik Raporu: Alındı – Turnitin, Oran: 18%

Etik Şikayeti: editor@rumelide.com

Makale Türü: Araştırma makalesi, **Makale Kayıt Tarihi:** 09.11.2025-**Kabul Tarihi:** 18.12.2025-**Yayın Tarihi:** 19.12.2025; **DOI:** <https://doi.org/10.5281/zenodo.17973309>

Hakem Deđerlendirmesi: İki Dış Hakem / Çift Taraflı Körleme

1. Introduction

Decades of research in the field have demonstrated that L2 lexical development is a complex and multidimensional process, and researchers continue to debate a comprehensive definition of lexical proficiency (M. Kim et al., 2018). Given this complexity, Schmitt and Schmitt (2020) suggest that our understanding of L2 vocabulary is still incomplete. Additionally, according to Nation (2018, p. 145) “Some of our most treasured ideas on vocabulary size, text coverage, fluency development and vocabulary learning are based on one or two rather shaky pieces of research.” These observations imply the need for more research related to different aspects of L2 vocabulary acquisition. For instance, retention of L2 vocabulary is one of these aspects, as the results of related studies directly impact instructional settings. Rote learning is one of the techniques that has been used for vocabulary retention, and researchers have repeatedly questioned its effectiveness. Some researchers believe that rote learning is beneficial at low proficiency levels (Nation, 2022). In contrast, others claim that, irrespective of proficiency level, rote learning is not an effective way of learning vocabulary because this approach fails to create sufficient lexical connections that are necessary in real communicative settings (Schmitt, 2008).

Bahari et al. (2022) suggest that technology use in vocabulary instruction faces challenges. The body of research concerning technology use in L2 vocabulary instruction offers varied perspectives. For example, in recent meta-analysis studies, Hao et al. (2021) and Yu and Trainin (2022) report better results in vocabulary acquisition when technology is integrated into the process, claiming that the use of technology can enhance long-term retention. However, this claim is not supported by longitudinal or experimental studies in different settings. Thus, the long-term effects of technology use on vocabulary development calls for further investigation, particularly L2 vocabulary instruction at higher education, as there is also a gap in the related literature regarding young adult L2 learners.

Despite these discussions, questions remain regarding the comparative effectiveness of digital rote learning and contextualized instruction on vocabulary retention. The current study is important as it aims to fill a part of this gap in the related literature. To our knowledge, this study is among the first quasi-experimental comparisons of digital rote learning of L2 vocabulary with traditional contextual vocabulary instruction in the Turkish English as a Foreign Language (EFL) instructional setting.

2. Literature Review

Over the years, issues related to rote vs. contextual approaches have been debated from various perspectives, yet practical applications of L2 vocabulary research remain underexplored. Some researchers have tried to put related research findings directly into classroom settings. Hilton (2022), for example, suggested a four-stage model for vocabulary instruction by taking into account the accumulation of research findings. According to this model, teaching vocabulary starts with a systematic planning of the target words. Next, these words are clearly presented in the classroom environment, embedded in various tasks, activities, games, and exercises. The last stage of instruction is less communicative and dwells more on metalinguistic activities. At this stage, the learners try to tackle morphological and/or grammatical problems concerning the target words. In the same vein, Schmitt and Schmitt (2020) advised practitioners not to forget that vocabulary instruction should be both intentional and incidental, planned, and include systematic recycling for long-term retention.

The related literature clearly demonstrates that vocabulary learning and instruction are

multidimensional and complex, implying the need for diverse methods (Wi and Boers 2025). Similarly, the importance of engagement and exposure in vocabulary learning, and group work activities in reading classes were highlighted as effective strategies for vocabulary learning (Alghamdi, 2019). Butler (2022) claimed that a multimodal approach in vocabulary will be more beneficial to learners and emphasize the importance of contextual practices. This implies that when various sensory modalities are involved and a context is provided in the learning process, learners learn better and remember more. In support of this perspective, the need for explicit instruction and repeated exposure to new words were also emphasized (Butler, 2022).

The role of context in L2 vocabulary instruction and learning is also an important topic of discussion. Recent studies have focused on the use of contextual support in vocabulary instruction together with word characteristics as the most important parameters, and the comparison of contextual vocabulary instruction with non-contextual methods like rote learning is also among these discussions (Mulder et al., 2018). The discussions concerning rote learning is also not new. Laufer and Hulstijn (2001) claimed that rote learning would yield better results in the short term compared to contextual learning. Their study is interesting at this point because the results indicated that language learners who practiced isolated words were significantly better at recall tests compared to those who practiced the same words in a contextualized setting through reading. Similarly, Ganesan et al. (2025) compared the effectiveness of rote memorization and contextual learning in an ESL setting and found that rote learning improved immediate vocabulary recall significantly but is limited for long-term retention and use in meaningful contexts. As a support to these claims, Nation (2022) suggested that language instructors could benefit from rote learning especially at the initial stages of vocabulary acquisition because of its simplicity. However, the counterclaim that rote learning is ineffective has more supporters, as rote learning is regarded as shallow and having a low long-term retention rate. Furthermore, rote learning is believed to fail in assisting learners in establishing deeper semantic connections, which is crucial to enhance language proficiency (Schmitt, 2008). Additionally, research concerning human learning indicates that deeper cognitive processes result in a stronger formation of new information (Goldstein, 2024). Rote learning is also criticized as it does not allow language learners the flexibility to use vocabulary in different contexts (Ellis, 2008).

Another recent discussion is the effectiveness of using digital technology in L2 instruction in general terms (Bećirović et al., 2021; Chang & Hung, 2019; Daly, 2022; Li & Lan, 2022). The application of digital technology to vocabulary instruction is a relatively new area, which has also been explored from multiple perspectives. Researchers have repeatedly reported significant vocabulary gains in digital platforms (Yudintseva, 2015; Zou et al., 2019). The literature further supports the effectiveness of mobile-assisted learning platforms in vocabulary retention, and language learners' increasing interest in these platforms is often emphasized (Zou et al., 2019). The efficacy of mobile-assisted learning platforms in vocabulary acquisition and retention is also supported in the related literature. For example, Ji and Abdul Aziz (2021) noted in their systematic literature review that "...the mobile-assisted learning platforms aid ESL/EFL learners in vocabulary acquisition, enhance vocabulary retention, increase motivation and provide rich and supplemental language learning materials and experience" (p. 1504). Learners further perceive them as efficient, easy, flexible, accessible, satisfying, entertaining and interesting.

Flashcards, commonly used for L2 vocabulary practice, are typically regarded as a rote-learning tool, and their effectiveness has been discussed from several perspectives. Khan (2022), for instance, discussed their advantages in online and offline learning environments and reported their positive

effects on the learning process. Ashcroft et al. (2018) also reported that digital flashcards were more effective compared to the traditional paper ones. Other studies supported the use of flashcards by emphasizing their value in collaborative learning tasks (Hung, 2015; Supriatin & Rizkilillah, 2018). Wilkinson (2020) focused on the use of flashcards as a deliberate study technique for L2 vocabulary gain and found that structured repetition-based techniques are effective but their effectiveness mostly depend on repetition and sustained engagement. In a review concerning adult L2 vocabulary training, Rice and Tokowicz (2020) found that when the focus is only on the form connections between L1 and L2 words through massed repetition alone (as in flashcards), the learning process becomes less effective. However, when such activities are integrated into the process with other activities that also focus on form-meaning connections, such as spaced repetition or semantic elaboration, they will facilitate learning, as they promote learner participation. In the same vein, some researchers acknowledge the potential of the use of flashcards in enhancing vocabulary mastery while claiming that the creation of flashcards is time-consuming (Hamer & Rohimajaya, 2018). These discussions have led some researchers to highlight a need for more empirical data in L2 vocabulary learning (Zhu & Irwin, 2024). We conducted this study with these discussions in mind, and the following research questions were our primary concerns:

1. What are the short-term effects of digital flashcards compared to contextual vocabulary instruction in terms of recognition of target vocabulary items?
2. What are the long-term effects of digital flashcards compared to contextual vocabulary instruction in terms of recognition of target vocabulary items?
3. What are the short-term effects of digital flashcards compared to contextual vocabulary instruction in terms of application of target vocabulary items?
4. What are the long-term effects of digital flashcards compared to contextual vocabulary instruction in terms of application of target vocabulary items?

3. Method

3.1. The Context of the Study and the Participants

This quasi-experimental study was conducted in a preparatory program at a Turkish university where English is taught as a foreign language for native speakers of Turkish. At the beginning of the 2023-2024 academic year, the students in this institution were grouped according to their proficiency levels, and two of these groups at the A2 proficiency level were assigned as the experimental ($n = 15$) and control groups ($n = 15$) for the current study. The participants' ages varied between 19 to 21, with approximately equal gender distribution. The same EFL teacher instructed the groups for the whole term with exactly the same syllabus.

3.2. Data Collection Tools and Procedures

Three tests were used in the data collection process. The first of these tests was the New Vocabulary Levels Test (NVLTL), a multiple-choice test developed by McLean and Kramer (2015), which is used to measure general English vocabulary. We used this test to see whether there were any initial differences between the experimental and control groups in terms of general English vocabulary. The second tool that we used was a 50-item vocabulary test, which was developed by the research team (See Appendix

1). It was designed to assess participants' ability to apply their vocabulary knowledge in brief contexts. The third test, also developed by the research team, consisted of productive items requiring the participants to provide English definitions and Turkish translations of the target words. A sample item from the test is provided below (See Appendix 2). Two EFL experts reviewed the two tests for content validity. Reliability was assessed via Cronbach's Alpha (α), and the results yielded ($\alpha = .75$), which meets criterion for acceptable reliability (Field 2017, p. 784).

3.3. The Intervention Process

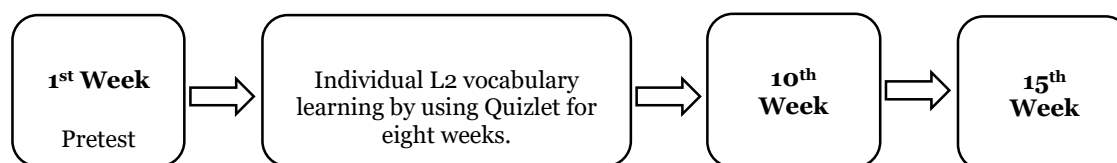
An eight-week intervention plan was prepared before the study, which is a sufficient duration for vocabulary retention (Nation, 2022). In the first week, the participants were informed about the content and purpose of the study, and their consents were taken. Additionally, the experimental group received brief training on Quizlet, a digital learning platform, while the control group was introduced to the reading set that they would use. The language instructor was also informed in detail about the intervention process. In the second week, the three tests mentioned previously were administered, and two of these tests were used as the pretests.

Every week during the intervention, each group dedicated two hours to two sets of new words, each consisting of five to ten items. The experimental group studied the new word through Quizlet (*see a screenshot in Appendix 3*), while the control group studied the same words using thematic reading materials. This reading material was an outcome of a specialization project that was produced to teach vocabulary at the university level (*see sample pages in Appendix 4*). In each group, the participants were informed that the study would focus on learning new words in English and that they were going to focus on 110 new words over eight weeks.

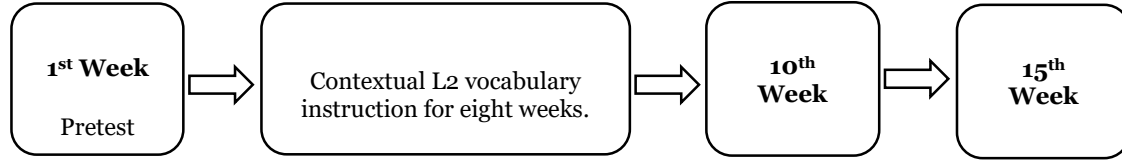
The experimental group used a variety of activities on Quizlet, which provides learners with digital flashcards. Equipped with a multimodal learning paradigm, Quizlet incorporates audiovisuals, and support group interaction. Although all these features of Quizlet were encouraged during the lessons in the experimental group, no extended narrative context was provided to the participants, though brief peer discussions were encouraged to elaborate on the target words. While the experimental group used Quizlet's interactive features without meaningful contexts, the control group engaged in traditional contextualized pre-, while-, and post-reading activities. The activities in the experimental group were mostly individual although group interactions were encouraged, while the activities in the control group were mostly traditional. After the last day of the intervention, the second and the third tests mentioned previously were administered again to both groups as the posttests. Additionally, five weeks after the administration of the posttest, the same tests were readministered to the groups to assess retention decay. The following figure provides an overview of the process for both groups.

Figure 1 *The Intervention Process*

The Experimental Group



The Control Group



4. Results

To compare the change in the mean scores of the experimental and control groups, analysis of covariance (ANCOVA) was used, as this statistical method enables us “to reduce within-group error variance” and helps eliminate confounds in experimental studies (Field, 2017, p. 747). Best and Kahn (2006) suggest that this is also one of the most common methods used in pretest/posttest designs and that “the use of pretest mean scores as covariates is considered preferable to the use of gain scores” (p. 193). Before performing the main analysis, the data were checked for the assumptions required to conduct ANCOVA. The Shapiro-Wilk test of normality was used as this test is ideal for sample sizes smaller than 50 (González-Estrada & Cosmes, 2019), and the results showed that the scores were normally distributed ($p > .05$). Other assumptions like homogeneity of regression slopes and linearity between the covariates and the dependent variable, were also checked; no violations were detected. At the next step, NVLT test results of the groups were compared through an independent samples *t*-test, and the results are given in the following table. The results imply that there is statistically no significant difference between the groups ($t = 1.70$, $df = 28$, $p = .10$).

Table 1 *t*-test Results for the New Vocabulary Levels Test

Groups	N	Mean	sd	SEM	<i>t</i> test		
					<i>t</i>	df	<i>p</i>
Experimental	15	60.27	20.30	5.24	1.70	28	.10
Control	15	44.60	29.31	7.57			

Research Question 1: What are the short-term effects of digital flashcards compared to contextual vocabulary instruction in terms of recognition of target vocabulary items?

To address the first research question, we analyzed the pretest and posttest mean scores of the two groups. Mean scores of the groups were adjusted to eliminate the possible confounds, as explained previously, and the results are presented in the following table.

Table 2 Adjusted Posttest Mean Scores after Taking the Recognition Pretest Mean Scores as the Covariate

Groups	Posttest scores		Adjusted posttest scores	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Experimental	32.27	8.21	36.10	4.91
Control	36.67	25.53	32.83	4.91

As is clear from Table 2, there is a difference between the mean scores of the groups ($M_E = 32.27$, $SD = 8.21$; $M_C = 36.67$, $SD = 25.53$). Since this difference is likely to cause unwanted confounds, the posttest mean scores were adjusted by controlling the pretest scores ($M_E = 36.10$, $SE = 4.91$; $M_C = 32.83$, $SE =$

4.91). After this adjustment, the posttest scores of the groups were compared through ANCOVA, and the results are given in the following table.

Table 3 ANCOVA Test Results Comparing Recognition of Target Words with Pretest as the Covariate

Source	SS	df	MS	F	p	η^2
Corrected Model	1676.49 ^a	2	838.24	2.65	.08	.16
Intercept	1197.53	1	1197.53	3.78	.06	.12
Recognition Pretest	1531.29	1	1531.29	4.84	.37	.15
Treatment	62.21	1	62.21	.19	.66	.00
Error	8538.97	27	316.25			
Total	45854	30				
Corrected Total	10215.47	29				

a. *R Squared* = .16 (*Adjusted R Squared* = .10)

It is clear from Table 3 that the difference between the experimental and control groups is not statistically significant when the recognition pretest scores are taken as the covariate, $F(1, 27) = .19, p = .66$.

Research Question 2: What are the long-term effects of digital flashcards compared to contextual vocabulary instruction in terms of recognition of target vocabulary items?

The second research question investigated the delayed effects of digital and contextual L2 vocabulary learning on recognition. Delayed posttest scores were compared using the same statistical procedure applied to the first research question. Table 4 presents the delayed posttest mean scores and adjusted mean scores, with pretest scores included as a covariate to ensure a valid comparison.

Table 4 Adjusted Delayed Posttest Mean Scores after Taking the Recognition Posttest Mean Scores as the Covariate

Groups	Delayed posttest scores		Adjusted delayed posttest scores	
	<i>M</i>	<i>SD</i>	<i>X</i>	<i>SE</i>
Experimental	21.00	10.06	21.15	2.30
Control	31.67	7.39	31.51	2.30

In Table 4, we can see the difference between delayed posttest mean scores of the participants ($M_E = 21.00, SD = 10.06; M_C = 31.67, SD = 7.39$). The obvious difference in the mean scores is adjusted again by taking the posttest mean scores as the covariate ($M_E = 21.15, SE = 2.30; M_C = 31.51, SE = 2.30$). After this adjustment, an ANCOVA was performed, and Table 5 displays the results comparing delayed posttest mean scores of the two groups.

Table 5 ANCOVA Test Results Comparing Delayed Recognition Mean Scores with Posttest as the Covariate

Source	SS	df	MS	F	p	η^2
Corrected Model	902.32 ^a	2	451.16	5.70	.00	.29
Intercept	3784.69	1	3784.69	47.88	.00	.64
Recognition Pretest	48.98	1	48.98	.62	.43	.02

Treatment	793.49	1	793.49	10.04	.01	.17
Error	2134.35	27	79.05			
Total	23840	30				
Corrected Total	3036.67	29				

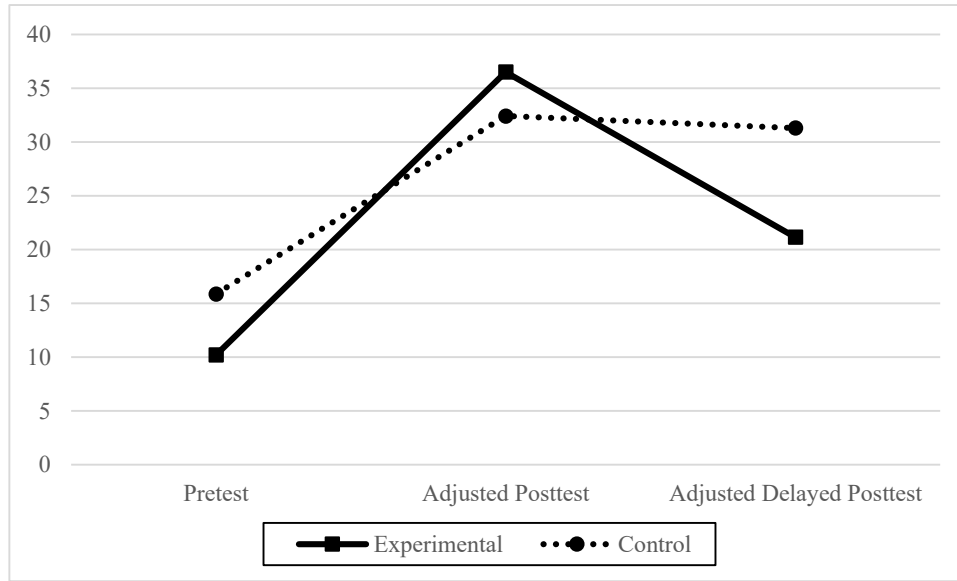
a. R Squared = .29 (Adjusted R Squared = .25)

The comparison of the delayed posttest mean scores reveals a statistically significant difference, $F(1,27) = 10.04$, $p = .01$; $F(1, 27) = 10.04$, $p = .01$; $F(1,27) = 10.04$, $p = .01$; the control group scored significantly higher than the experimental group in the delayed recognition test, with a large effect size ($\eta^2 = .17$). To illustrate these findings, mean scores from the pretest, posttest, and delayed posttest are presented in Table 6 and visualized in Figure 2.

Table 6 Means Scores of Recognition Pretest, Posttest and Delayed Posttest

Group	Pretest	Posttest	Delayed Posttest
Experimental	10.20	36.10	21.15
Control	15.87	32.83	31.51

Figure 2 Recognition Pretest, Posttest and Delayed Posttest Mean Scores



The line graph in Figure 2 illustrates the progress that the two groups made in learning target vocabulary, specifically in recognizing the target words. The experimental group studied the target vocabulary in a digital learning environment, and the control group received instruction through contextualized reading activities. It is clear from the figure that both groups made considerable progress. However, there is some difference in the initial mean scores of the two groups, and this difference was adjusted via ANCOVA using pretest scores as covariates, as detailed earlier (see Table 5). The delayed posttest results of the groups were also calculated by taking their pretest mean scores into account. Although the mean score of the experimental was higher in the posttest, the control group retained the target words significantly better than the experimental group in the delayed posttest.

Research Question 3: What are the short-term effects of digital flashcards compared to contextual vocabulary instruction in terms of application of target vocabulary items?

To address the third research question, which analyzes contextual application of the target vocabulary, initially, the pretest and posttest mean scores of the two groups were compared. Before the comparison, initial differences between the two groups were adjusted. Table 6 shows the results of this process.

Table 7 Adjusted Posttest Mean Scores after Taking the Contextual Application Pretest Mean Scores as the Covariate

Group	Posttest scores		Adjusted delayed posttest scores	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Experimental	14.53	5.13	11.41	.71
Control	17.33	10.58	17.85	.71

The figures in Table 7 indicate a difference between the contextual application posttest mean scores of the two groups ($M_E = 14.53$, $SD = 5.13$; $M_C = 17.33$, $SD = 10.58$). To prevent this difference from causing unwanted variance in the posttest mean scores, the posttest mean scores were adjusted by controlling the pretest scores ($M_E = 11.41$, $SE = .71$; $M_C = 17.85$, $SE = .71$). After this adjustment, the mean scores of the two groups were compared through ANCOVA, and the results are given in the following table.

Table 8 ANCOVA Results Comparing Contextual Application Means Scores with Pretest as the Covariate

Source	SS	df	MS	F	<i>p</i>	η^2
Corrected Model	86.31 ^a	2	43.15	.61	.55	.04
Intercept	571.17	1	571.17	8.01	.00	.23
Contextual Pretest	27.51	1	27.51	.39	.54	.01
Treatment	21.23	1	21.23	.30	.59	.01
Error	1907.52	27	70.65			
Total	9610	30				
Corrected Total	1993.87	29				

a. R Squared = .04 (Adjusted R Squared = .03)

The figures in Table 8 clearly show that the difference between the mean scores of the experimental and control groups is statistically insignificant, $F(1, 27) = .01$, $p = .59$.

Research Question 4: What are the long-term effects of digital flashcards compared to contextual vocabulary instruction in terms of application of target vocabulary items?

The fourth research question examines the delayed effects of digital learning and contextualized L2 vocabulary instruction on the contextual application of the target words. To ensure a valid comparison, the delayed posttest results of the groups were adjusted by using their pretest scores as the covariate. Table 9 shows the adjusted posttest mean scores for the groups.

Table 9

Adjusted Delayed Posttest Mean Scores after Taking the Contextual Application Pretest Mean Scores as the Covariate

Group	Posttest scores		Adjusted posttest scores	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Experimental	1.73	2.49	2.44	2.88
Control	23.07	14.69	22.36	2.88

The figures in Table 9 indicate a difference between the contextual application posttest mean scores of the two groups ($M_E = 1.73$, $SD = 2.49$; $M_C = 23.07$, $SD = 14.69$). To prevent this difference from causing unwanted variance in the posttest mean scores, the posttest mean scores were adjusted by controlling the pretest scores ($M_E = 2.44$, $SE = 2.88$; $M_C = 22.36$, $SE = 2.88$). After these calculations, the mean scores for the two groups were compared through ANCOVA, and the results are given in the following table.

Table 10 ANCOVA Results Comparing Contextual Application Means Scores with Pretest as the Covariate

Source	SS	df	MS	F	<i>p</i>	η^2
Corrected Model	333.14 ^a	2	1663.57	21.85	.00	.42
Intercept	1143.09	1	1143.09	149.95	.00	.84
Contextual Pretest	6.44	1	6.44	.85	.36	.03
Treatment	301.73	1	301.73	39.58	.01	.19
Error	205.82	27	7.62			
Total	6963	30				
Corrected Total	538.97	29				

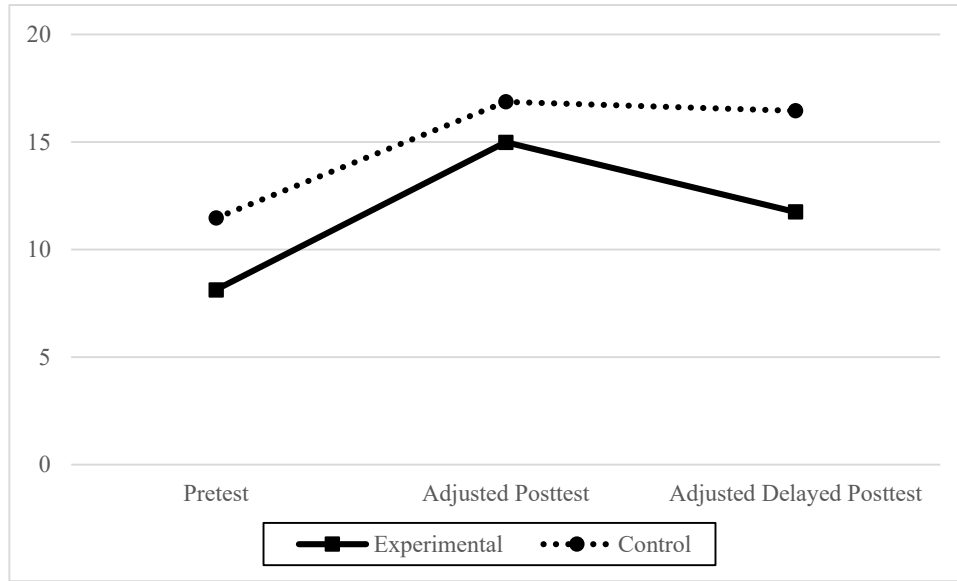
a. *R Squared* = .42 (*Adjusted R Squared* = .39)

The figures in Table 10 show that the difference between delayed posttest scores of the experimental and control groups is statistically significant $F(1, 27) = 39.58$, $p = .01$, with a large effect size ($\eta^2 = .19$). The delayed posttest scores of the participants are not affected by prior performance indicated in the pretest results as they were adjusted, and the overall model presented in Table 9 explains a large proportion of the variance (42%). The adjusted mean scores are presented in Table 11.

Table 11 Mean Scores of Contextual Application Pretest, Posttest and Delayed Posttest

Group	Pretest	Posttest	Delayed Posttest
Experimental	8.13	14.99	11.41
Control	11.47	16.87	17.85

The figures in Table 10 show an obvious difference between the pretest mean scores of the two groups ($M_E = 8,13$, $M_C = 11,47$). Although posttest mean scores of the two groups are similar ($M_E = 14,99$, $M_C = 16,87$), their adjusted posttest and delayed posttest scores show a significant difference ($M_E = 11,41$, $M_C = 17,85$). Both groups made some progress during the intervention, and Figure 3 below summarizes this progress and illustrates the delayed effects of the intervention.

Figure 3 *Contextual Application Pretest, Posttest and Delayed Posttest Mean Scores*

The statistical analysis in Table 10, which indicates a significant difference between the mean scores of the groups, is displayed in a graphic in Figure 3. In the figure, target vocabulary learning progress in terms of contextual application made by the experimental and control groups is displayed. Both groups seem to have made considerable progress. The line graph clearly shows that both groups made similar progress during the treatment process; the short-term effects of the intervention are also comparable in both groups, and their mean scores declined in the delayed posttest. However, the decline in the experimental group seems to be greater than in the control group, which indicates that delayed retention of the target words in the experimental group was weaker.

5. Discussions

In this study, we compared the effectiveness of vocabulary learning through digital flashcards with contextualized vocabulary instruction using a quasi-experimental design. Two groups with similar English proficiency levels were chosen for the study as the experimental and control groups. The experimental group studied the target words via an online digital flashcard platform focusing on rote learning, and the control group studied them in contextual settings using reading passages. Recognition and contextual application of the target words were assessed using two separate tests, administered as a pretest, a posttest, and a delayed posttest. During the eight-week intervention, both groups made considerable progress regarding recognition and contextual application of the target words, with some differences in their posttest scores. However, these differences were statistically insignificant when the pretest results were considered as the covariate. A delayed posttest was conducted for both groups, again by taking the pretest scores as the covariate, and this time the results revealed a significant difference between the groups both in terms of recognition and contextual application, with the control group performing significantly better.

The present findings align with several perspectives in the vocabulary acquisition literature. For example, our findings support Wilkinson's (2020) claim that structured and repetition-based vocabulary techniques can lead to vocabulary gains with enough repetition and sustained engagement.

They also support Nation's (2022) claim that rote learning could be an effective option for short-term retention at lower levels. Furthermore, as Ganesan et al. (2025) discussed, studying L2 vocabulary via rote learning activities support short-term gains, which is also in agreement with our findings. The present findings also align with the body of research related to contextual instruction of vocabulary. Although our results indicate that the use of digital flashcards is as effective as contextualized instruction for short-term retention, they are significantly less effective for long-term retention. This was particularly evident in the delayed posttest results as the control group outperformed the experimental group in both recognition and contextual application of the target words. This finding is also consistent with previous research indicating that contextual L2 vocabulary learning is more effective for long-term retention (Ganesan et al. 2025). Contextual vocabulary learning is justifiable both in theory and practice; however, when we regard the complexity of the acquisition process (Schmitt & Schmitt, 2020), it becomes clear that learning new vocabulary through contexts alone is also not enough, and contextualization of target vocabulary takes considerable amount of time. We believe that digital learning sources have significant potential for facilitating this process and making it faster.

The results of our study indicate that the difference in efficiency between contextualized and decontextualized L2 vocabulary instruction and learning process is not as straightforward as is generally assumed. The results suggest that the experimental group performed equally well in the posttest, which assessed recognition and contextual application of the target words. Digital flashcards helped the participants to achieve similar short-term vocabulary retention with the contextual group. This similarity is consistent with previous research emphasizing the multidimensional and complex nature of L2 vocabulary learning process (Wi & Boers, 2025). As Mulder et al. (2018) noted, even though contextual support and word characteristics are crucial in the process, isolation-based techniques can lead to rapid and effective vocabulary gains. The positive effects of digital technology repeatedly mentioned in the related literature (Bećirović et al., 2021; Chang & Hung, 2019; Li & Lan, 2022) also align with our findings; as digital platforms offer accessible and motivational tools that accelerate vocabulary acquisition (Yudintseva, 2015; Zou et al., 2019). As a result, the similarity in the short-term does not indicate that one approach is more effective than the other, but it implies that these two approaches can be used as complementary to each other depending on the proficiency levels of the learners. This insight can enable L2 instructors to strategically evaluate digital tools based on rote-learning both from theoretical and practical aspects.

Aside from the statistical results reported above, qualitative data provided indirect insights into learner motivation. Systematic observations during the intervention and semi-structured interviews with the implementing teacher and select participants indicated higher motivation in the digital group, evidenced by increased voluntary app usage and enthusiasm during sessions. This aligns with literature showing mobile-assisted tools enhance engagement over traditional methods (Ji & Abdul Aziz, 2021); however, as our findings suggest, motivation alone may not suffice for long-term retention of target L2 words.

Taking all the points that have been discussed so far together, we can highlight two findings: The first one is that compared to traditional methods, digital learning environments have great potential for facilitating and speeding up the vocabulary acquisition process. The second one is that the use of digital tools in the process is more engaging and motivating for learners compared to traditional instructional settings.

6. Limitations and Future Directions

This study provides some valuable insights into L2 vocabulary acquisition; however, it has some limitations. First of all, results of only one study with a relatively small sample of 30 participants may not be sufficient for generalizations. The same intervention procedures can be repeated with three or more larger groups. Additionally, there is another potentially problematic aspect of this study. One of the variables tested in the study was the application of the target words in new contexts, a skill that might not be as straightforward as it appears. Our insight is that some of the participants might already be competent in contextual application of words in their native language, and this might have affected the intervention outcomes and testing procedures. This is a small but important possibility that needs to be considered in further studies by checking them for this skill for higher reliability.

L2 vocabulary can be analyzed within the framework of taxonomy of knowledge, which asserts that input becomes knowledge after following a predictable path (Anderson & Krathwohl, 2001). This path involves lower-level operations like *remembering* and higher-level ones like *evaluation* and *creation*. The results of our study show that contextualized and non-contextualized L2 vocabulary training show no significant differences at lower levels of the taxonomy for the short term, but the difference is significant at the application level. This study falls short in answering the question as to how the effectiveness of both approaches would change when it comes to higher levels of the taxonomy, like analysis and evaluation, and this problem could be the concern of further studies. In addition, to test our insights suggesting that a combination of digital and contextual learning will make the learning process faster and lead to longer retention, a combination of digital flashcards and contextual instruction could be tested against digital learning and contextual instructional settings.

The results of this study also imply a need for more longitudinal studies concerning the relationship between digital learning and vocabulary retention. The rapid advancements in digital learning platforms need to be analyzed for a deeper understanding of how adaptive learning technologies impact vocabulary retention as these platforms offer a wide range of possibilities through gamification and spaced learning. We need to see the impacts of integration of strategies and methods like collaboration, spaced learning, semantic elaboration and flipped vocabulary instruction in the learning and retention process (see a detailed discussion in Ünalı, 2023).

7. Pedagogical Implications

This study provides practical insights for L2 practitioners by highlighting the complexity of the vocabulary acquisition process which should involve both rote and contextual learning depending on the instructional setting. We suggest that while digital tools like Quizlet flashcards can foster engagement and benefit at initial stages, they may not be effective in long-term retention of L2 words. Such tools are user-friendly, engaging and freely available to most language learners, which makes them valid options for L2 instructors to be used as supplementary resources rather than standalone instructional methods. By nature, learning via traditional flashcards is an individual activity, which might be criticized from a social-constructivist perspective. However, like most digital vocabulary learning tools, Quizlet is equipped with social features through which learners can share their experience and even play competitive vocabulary games with their peers through Quizlet's gamification feature. This would allow learners to benefit from motivating and easy to use digital learning tools, while also incorporating tasks that promote deeper cognitive processes.

The findings of our study suggesting that contextualized vocabulary instruction yields better retention also align with Hilton's (2022) model of vocabulary instruction, which highlights the systematic planning, presentation, application, and review of target words in meaningful contexts. With a similar approach, L2 instructors can ensure that vocabulary learning is both systematic and embedded within the language context, which will allow them to transfer their vocabulary knowledge to real-world communicative settings while not ignoring the systematic aspect of the learning process.

The main implication of our study is that if digital learning tools are not used properly, they will not help longer retention of vocabulary; on the other hand, if they are used properly, they will act as a strong supplementary practice in the retention of vocabulary and help both instructors and learners promote motivation, which is one of the most crucial aspects of L2 acquisition process (Dörnyei & Ryan, 2015).

8. Conclusions

This study offers strong empirical evidence that the integration of digital technologies into L2 vocabulary instruction must align with the established principles, and the effective use of these technologies calls for sound strategies. Although there is growing consensus that the future of L2 learning will mainly be shaped by digital technologies, currently there is no sound theoretical background on which L2 vocabulary acquisition and related practices will be based. Further empirical and longitudinal studies related to the integration of digital technologies into the L2 acquisition process in general, and L2 vocabulary in specific, are not merely a necessity but a crucial step forward.

Ethical considerations: Ethical approvals were obtained from Nevsehir Hacı Bektas Veli University, Scientific Research Ethical Committee with the approval number 2024.03.34.

Conflict of interest: We hereby declare that there are no relevant financial or non-financial competing interests to report.

References

- Alghamdi, H. H. (2019). Exploring second language vocabulary learning in ESL classes. *English Language Teaching*, 12(1), 78–84. <https://doi.org/10.5539/elt.v12n1p78>
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Ashcroft, R. J., Cvitkovic, R., & Praver, M. (2018). Digital flashcard L2 vocabulary learning outperforms traditional flashcards at lower proficiency levels: A mixed-methods study of 139 Japanese university students. *The EuroCALL Review*, 26(1), 14–28. <https://doi.org/10.4995/eurocall.2018.7881>
- Bahari, A., Eamer, A., & Hughes, J. (2022). An analysis of current research on computer-assisted L2 vocabulary learning. *CALICO Journal*, 39(3), 257–280.
- Bećirović, S., Brdarević-Čeljo, A., & Delić, H. (2021). The use of digital technology in foreign language learning. *SN Social Sciences*, 1(10), 246. <https://doi.org/10.1007/s43545-021-00254-y>
- Best, J. W., & Kahn, J. V. (2006). *Research in education* (10th ed.). Pearson/Allyn & Bacon.
- Butler, Y. G. (2022). Teaching vocabulary to young second- or foreign-language learners. *Language Teaching for Young Learners*, 4, 4–33.
- Chang, M., & Hung, H. (2019). Effects of technology-enhanced language learning on second language acquisition: A meta-analysis. *Journal of Educational Technology & Society*, 22(4), 1–17.
- Daly, N. (2022). Investigating learner autonomy and vocabulary learning efficiency with MALL. *Language Learning and Technology*, 26(1).
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Ghafar, Z. N., & Mohamedamin, A. A. (2022). Vocabulary acquisition techniques to learn second language English speakers: A literature review. *International Journal of Middle Eastern Research*, 1(1), 17–24.
- Goldstein, E. B., & Hale, R. G. (2024). *Cognitive psychology: Connecting mind, research, and everyday experience* (6th ed.). Cengage Learning.
- González-Estrada, E., & Cosmes, W. (2019). Shapiro–Wilk test for skew normal distributions based on data transformations. *Journal of Statistical Computation and Simulation*, 89(17), 3258–3272. <https://doi.org/10.1080/00949655.2019.1658763>
- Hamer, W., & Rohimajaya, N. A. (2018). Using flash card as instructional media to enrich the students' vocabulary mastery in learning English. *Journal of English Language Studies*, 3(2), 167–177.
- Hilton, H. (2022). *Enseigner les langues avec les apports des sciences cognitives*. Hachette Éducation.
- Hung, H.-T. (2015). Intentional vocabulary learning using digital flashcards. *English Language Teaching*, 8(10), 107–112.
- Ji, P., & Abdul Aziz, A. (2021). A systematic review of vocabulary learning with mobile-assisted learning platforms. *International Journal of Academic Research in Business and Social Sciences*, 11(11), 1503–1521.
- Khan, R. M. I. (2022). The use of flashcards in teaching EFL vocabulary in online learning. *Register Journal*, 15(1), 109–125. <https://doi.org/10.18326/rgt.v15i1.109-125>
- Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality. *The Modern Language Journal*, 102(1), 120–141. <https://doi.org/10.1111/modl.12447>
- Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92(1), 114–130. <https://doi.org/10.1111/j.1540->

4781.2008.00690.x

- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. <https://doi.org/10.1093/applin/22.1.1>
- Li, P., & Lan, Y.-J. (2022). Digital language learning (DLL): Insights from behavior, cognition, and the brain. *Bilingualism: Language and Cognition*, 25(3), 361–378. <https://doi.org/10.1017/S1366728921000353>
- McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken*, 19(2).
- Mulder, E., Van de Ven, M., Segers, E., & Verhoven, L. (2018). Context, word, and student predictors in second language vocabulary learning. *Applied Psycholinguistics*, 40(1), 137–166.
- Nation, I. S. P. (2018). Keeping it practical and keeping it simple. *Language Teaching*, 51(1), 138–146. <https://doi.org/10.1017/S0261444817000349>
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Pascual, L. C., Dionisio, G., & Ilustre, R. (2022). Vocabulary acquisition and learning strategies in second language learning: A review paper. *International Journal of English Language Studies*, 4(3), 58–62.
- Rice, C. A., & Tokowicz, N. (2020). A review of laboratory studies of adult second language vocabulary training. *Studies in Second Language Acquisition*, 42(2), 439–470. <https://doi.org/10.1017/S0272263119000500>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.
- Sousa, D. A. (2011). *How the brain learns* (4th ed.). Corwin Press.
- Supriatin, T., & Rizkilillah, V. P. (2018). Teaching vocabulary using flashcard. *Professional Journal of English Education*, 1(4), 479–484.
- Ünalı, İ. (2023). Flipped vocabulary instruction experience with adult language learners. *International Journal of Education, Technology and Science*, 3(3), 1014–1031.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232–245.
- Wi, I., & Boers, F. (2025). A synthesis of research on L2 vocabulary learning through audiovisual input and on-screen text. *Digital Applied Linguistics*, 3, 102774. <https://doi.org/10.29140/dal.v3.102774>
- Yudintseva, A. (2015). Game-enhanced second language vocabulary acquisition strategies: A systematic review. *Open Journal of Social Sciences*, 03(10), 101–109.
- Zhang, P. (2006). An overview of second language vocabulary acquisition research in China. *Foreign Language and Their Teaching*, 207, 21–26.
- Zhu, T., Zhang, Y., & Irwin, D. (2024). Second and foreign language vocabulary learning through digital reading: A meta-analysis. *Education and Information Technologies*, 29(2), 4531–4563. <https://doi.org/10.1007/s10639-023-11969-1>
- Zou, D., Huang, Y., & Xie, H. (2019). Digital game-based vocabulary learning: Where are we and where are we going? *Computer Assisted Language Learning*, 34(5–6), 751–777. <https://doi.org/10.1080/09588221.2019.1640745>

Appendix 1*The Multiple-Choice Test*

Student ID:

Group:

Date: / /

-Choose the best alternative for each blank.**-You have 60 minutes.**

1. The facts about the caseafter a long investigation.
a) emerged b) preserved c) evolved d) survived
2. The war lasted twenty years, but the buildings in this town remained
a) massive b) tolerant c) humble d) intact
3. My cousinfigures from wood and sells them in the market.
a) interprets b) carves c) hikes d) insulates
4. An aspirin will give you a/anrelief from the pain.
a) vast b) resistant c) temporary d) solid
5. There are manyabout sport among people.
a) layers b) misconceptions c) seclusions d) networks
6. There are alwaysto many rules.
a) exceptions b) ventilations c) civilizations d) invasions
7. The volcanoand many people in the city were killed.
a) survived b) collapsed c) descended d) erupted
8. We are living in a/an of fast communication.
a) pollution b) era c) ignorance d) nest
9. The climbers completed their to the mountain on a rainy morning in April.
a) ascent b) erosion c) shelter d) empire
10. The birds over the surface of the lake.
a) hiked b) abandoned c) glided d) restored
11. There are remains of baths and a of Roman date.
a) cistern b) valley c) wicker d) creek
12. The famous singer has got a large store of wine in his
a) community b) cellar c) envelope d) treasure
13. The on the side of the building was a beautiful work of art
a) layer b) monastery c) synthesis d) mural
14. The famous actor has been living insince he retired.
a) seclusion b) destination c) attraction d) crucifixion
15. The island has morein plant life than other islands nearby.
a) diversity b) phenomenon c) civilization d) tolerance

16. Despite all his great achievements, he has remainedall his life.
a) wise b) holy c) humble d) resistant
17. They decided tothe new office building near the airport.
a) preserve b) situate c) illustrate d) insulate
18. There is an estimated shortfall of about five millionacross the country.
a) borders b) figures c) provinces d) dwellings
19. When I retire, I want to buy aand raise a lot of animals.
a) throne b) pigeon c) barn d) crossroad
20. The room was stuffy, and the onlycame from a tiny window.
a) ventilation b) process c) trail d) climate
21. Most of the money was in storage in bank in different cities.
a) burners b) tunnels c) baskets d) vaults
22. The aluminum is melted in a at a temperature of 1400 degrees.
a) grave b) castle c) furnace d) fresco
23. The road to the village gently to the right.
a) evolves b) operates c) curves d) settles
24. The city is built on the flat of a rocky hill, which makes it resistant to earthquakes.
a) valley b) lava c) canyon d) summit
25. There are no on hot air balloons; they go where the winds take them.
a) steers b) monks c) structures d) storerooms

END OF THE TEST

Appendix 2*Productive test for the target items*

Student ID:

Group:

Date: / /

-Write the English and the Turkish meanings of the target words in the blanks provided.
 -Mark X in the blank if you do not know the answer.
 -You have 30 minutes.

Target Word	Meaning in English or synonym	Turkish meaning
1) abandon		
2) ancient		
3) apostle		
4) ascent		
5) barn		
6) carve		
7) cellar		
8) cistern		
9) collapse		
10) creek		
11) curve		
12) diversity		
13) dwelling		
14) emerge		
15) era		
16) erupt		
17) evolve		
18) exception		
19) furnace		
20) glide		
21) grave		
22) hike		
23) humble		
24) insulate		
25) intact		

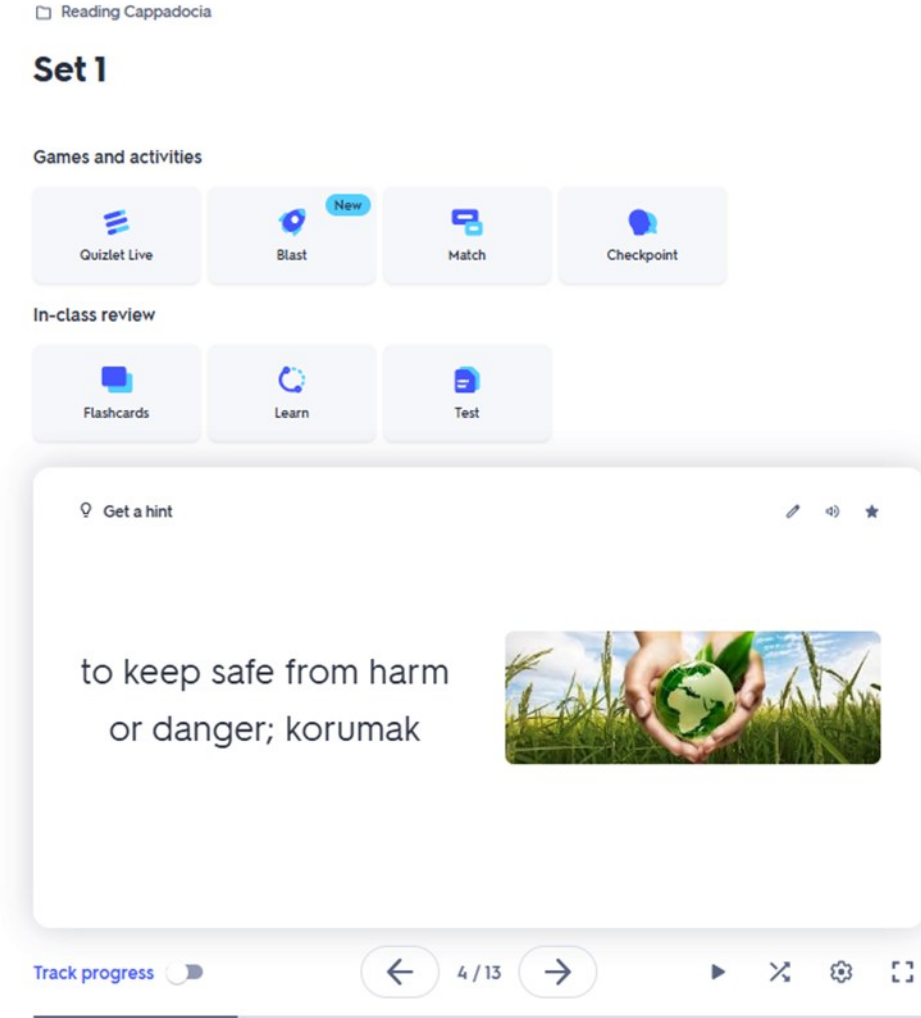
1

Target Word	Meaning in English or synonym	Meaning in Turkish
26) interpret		
27) invasion		
28) massive		
29) monk		
30) mural		
31) nest		
32) pigeon		
33) pilgrim		
34) preserve		
35) province		
36) seclusion		
37) settle		
38) shelter		
39) situate		
40) steer		
41) summit		
42) temporary		
43) trail		
44) treasure		
45) vast		
46) vault		
47) ventilation		
48) wicker		
49) wisdom		
50) worship		

END OF THE LIST

Appendix 3

A Screenshot of a Quizlet Digital Flashcard Used by the Experimental Group



Appendix 4*Sample Pages from the Reading Material Used by the Control Group*

6 READING CAPPADOCIA TOURISM VOCABULARY IN CONTEXT

► These are the **target words** for this unit. Look at them carefully and decide how familiar you are with them. Descriptions for each option in the scale are given below the table. Mark each word from ① to ⑥.

TARGET WORD/PHRASE	①	②	③	④	⑤	⑥
fairy chimney	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
destination	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tourist attraction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
melting pot	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
erosion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
carve	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
church	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
emerge	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
preserve	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
intact	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
vast	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
evolve	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ancient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- ① I have never heard or seen this word/phrase before.
 ② I have heard or seen this word/phrase before, but I don't know its meaning.
 ③ I know the meaning of this word/phrase, but I cannot use it.
 ④ I can use this word/phrase while speaking.
 ⑤ I can use this word/phrase while writing.
 ⑥ I can use this word/phrase both while speaking and writing.



A Dream Destination

A FAIRY TALE

Cappadocia... Colorful hot-air balloons, beautiful sunrises, magical sunsets, horse rides, living history, and **fairy chimneys**... For many people all around the world, it is a dream **destination**. Actually, the region is one of the most popular **tourist attractions** in the world. Every year, millions of people from different countries visit this wonderful place. But why is Cappadocia so special?

15

First, the Cappadocia region is in the center of Anatolia. The city in the east is called Malatya, and the city in the west is Aksaray. To the south you can see the Taurus Mountains, and in the north, the region reaches up to the shores of the Black Sea. Geographically speaking, most of its parts are away from the seas. Historically speaking, being away from the seas made Cappadocia a safe place for people during wars. This is why Cappadocia has been home to many civilizations throughout history. Hittites, Persians, Greeks, Romans, Byzantines, and early Christians are examples of these ancient civilizations. Because of these reasons, some even call Cappadocia a **melting pot** of civilizations.

20

Secondly, its unique nature makes Cappadocia a very special place. Scientists talk about millions of years when the topic is the geological history of Cappadocia. They say it formed in about sixty million years as a result of volcanic activities, **erosions**, and weather conditions. People also played a role in this process. They **carved** homes and **churches** into the soft rocks, and in time, entire cities **emerged**. Some of these cities were even completely under the ground.

25

A third point that makes Cappadocia special is that it is partly **preserved**. UNESCO added Cappadocia to the World Heritage list in 1985. Since those years, local people and official authorities have been trying to keep the region alive and **intact**. The government has very strict rules about buildings and farming. People in the area also care about the region because many of them make their living through tourism. As a result, the entire region has turned into a **vast** open-air museum.

30

People have different ideas about where the name Cappadocia came from. No one is sure about it. It is a Greek word for sure. Some people believe that this word **evolved** from the Persian word *Katpatuka* which means "the land of fine horses." Some others think that it came from other **ancient** civilizations like Assyrians or Hittites. However, when we think of the famous wild horses that have been living in the region for hundreds of years, it is easy to see why it is called *the land of fine horses*.