Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Quasi-Experimental Comparison of Paper-Based and Computer-Based Vocabulary Layering Techniques on EFL University Students' Vocabulary Development

Khadija Abdalla

nour.algoare@gmail.com

Cyprus International University

Mehmet Ali YAVUZ

Cyprus International University

Article

Keywords: Computer-Based Instruction, Data-Driven Learning, Paper-Based Instruction, University EFL, and Vocabulary Learning

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10884127/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Quasi-Experimental Comparison of Paper-Based and Computer-Based Vocabulary Layering Techniques on EFL University Students' Vocabulary Development

[1*] Khadija Abdalla [2] Mehmet Ali YAVUZ

[1*] Khadija Abdalla, Faculty of Education, Cyprus International University, Nicosia, Türkiye
***Correspondence*:** nour.algoare@gmail.com

[2] Prof. Dr. Mehmet Ali YAVUZ, Faculty of Education, Cyprus International University, Nicosia, Türkiye mayavuz@ciu.edu.tr

Abstract

This study has aimed to compare paper-based and computer-based vocabulary layering, and to examine vocabulary outcomes and learner attitudes under controlled instructional conditions. Vocabulary learning remains a major challenge for university EFL learners, particularly in developing collocational knowledge and achieving long-term retention. Although corpus-informed Data-Driven Learning (DDL) has been associated with improved vocabulary acquisition, previous studies have primarily compared DDL with traditional instruction rather than examining paper-based and computer-based implementations of the same approach. This study compared these two delivery modes and investigated their effects on vocabulary learning and learner attitudes. A quasi-experimental pre-test–post-test design was conducted with 90 EFL students (aged 19–24) from three intact classes in Northern Cyprus. Participants were assigned to a control group, a paper-based DDL group, and a computer-based DDL group (30 students each). Baseline equivalence was confirmed ($p = .590$). Over 14 weeks, both experimental groups received identical vocabulary instruction, differing only in the mode of delivery. A 40-item vocabulary test and a 12-item attitude questionnaire were administered, and the data were analysed using ANOVA, paired and independent t-tests, effect sizes, and reliability analysis. Both experimental groups demonstrated significant vocabulary gains from pre-test to post-test. The paper-based group improved from $M = 38.93$ to 58.60 ($p = .0001$, $d = -1.42$), while the computer-based group increased from $M = 38.86$ to 71.66 ($p = .0001$, $d = -2.55$). Both groups significantly outperformed the control group, with the computer-based group demonstrating the larger effect. The computer-based group also achieved significantly

higher scores than the paper-based group ($p = .0005$). Both DDL groups reported positive attitudes toward corpus-based DDL, with no significant difference between them ($p = .93$). These findings indicate that vocabulary layering effectively enhances lexical development in both instructional modes. At the same time, computer-based DDL produces greater vocabulary gains. The results support integrating technology-enhanced, corpus-informed vocabulary instruction in higher education to promote more effective, evidence-based language learning.

Keywords: Computer-Based Instruction, Data-Driven Learning, Paper-Based Instruction, University EFL, and Vocabulary Learning.

Background

Vocabulary knowledge is widely recognised as a core component of second language proficiency because lexical breadth and depth shape comprehension, academic writing, and spoken fluency (Webb & Nation, 2017). In practice, many university EFL learners continue to report difficulty with selecting appropriate words, using collocations accurately, and retaining new items over time, particularly when teaching relies heavily on decontextualized lists or short, test-focused exposure (Nation, 2022). As a result, vocabulary instruction has increasingly shifted towards approaches that promote richer contextual input, repeated encounters, and guided noticing of usage patterns, rather than treating word learning as memorisation alone (Schmitt & Schmitt, 2020).

Technology-enhanced language learning has expanded the range of tools available for vocabulary instruction. For example, Computer-Assisted Language Learning (CALL) environments can provide multimodal support, quick access to meanings, and repeated practice opportunities that may increase engagement and retention (Chapelle & Chapelle, 2001). Similarly, corpus-informed pedagogy has gained attention because corpora offer authentic examples of how words behave in real discourse, including frequency, collocation, and register-sensitive usage (McEnery & Hardie, 2012). In this context, Data-Driven Learning (DDL) has been positioned as an approach in which learners examine concordance lines to infer meaning and usage, shifting them from passive recipients to active analysts of language evidence (Johns, 1991). However, the label “DDL” has sometimes been used broadly to cover tasks ranging from direct corpus searching to teacher-curated worksheets, which can blur what is actually being tested when outcomes are reported.

Evidence suggests that corpus-informed approaches can support vocabulary development. For instance, DDL-style instruction has been associated with improvements in receptive vocabulary and collocational knowledge when compared with dictionary-based or traditional instruction, especially when learners receive structured support for interpreting concordance evidence (Boulton, 2010; Al-Mahbashi, Noor & Amir, 2015). Likewise, research in non-English foreign language contexts has indicated that concordance-based materials can promote more accurate word choice through exposure to multiple authentic contexts. However, retention may still decline without spaced practice or sustained use (Yao, 2019). At the same time, the feasibility of corpus-informed learning varies by setting; limited access to devices and internet connectivity has encouraged "soft" approaches in which concordance lines are printed, pre-selected, and scaffolded by the teacher (Boulton, 2012). Consequently, pedagogy in many university contexts involves a practical decision: whether to prioritise computer-based corpus interaction or to implement paper-based corpus materials that aim to achieve similar noticing outcomes with fewer technical demands.

A related strand of research has compared paper and computer delivery in vocabulary teaching more generally. For example, studies comparing paper-based and computer-based contextualised vocabulary presentation have reported gains in both conditions, with computer-based delivery sometimes showing higher post-test performance, potentially linked to multimodal cues and heightened attention (Ahmadian, Amerian & Goodarzi, 2015). Nevertheless, computer-based delivery in such studies is often operationalised as a multimedia presentation rather than a learner-led investigation, meaning that reported advantages may reflect presentation effects rather than bigger differences in how lexical knowledge is constructed. In contrast, corpus-based work emphasises inductive learning through multiple exemplars and pattern recognition, which aligns more closely with vocabulary development as a gradual process involving form–meaning mapping, collocational constraints, and contextual appropriacy (Schmitt & Schmitt, 2020; Nation, 2022). Therefore, comparisons between paper and computer formats need careful alignment at the level of instructional technique, not only at the level of delivery medium.

Vocabulary layering techniques, which structure learning through sequenced stages such as contextual exposure, guided noticing, retrieval practice, and consolidation tasks, fit this need for clearer alignment. In other words, the technique can remain constant while the medium changes: one format can deliver layered input and practice through printed materials, whereas

the other can deliver the same layering through computer-based tools such as interactive exercises, electronic dictionaries, or corpus interfaces. Despite this, existing evidence tends to compare DDL with traditional instruction, rather than comparing paper-based versus computer-based implementations of the same corpus-informed or layering-oriented pedagogy. Moreover, attitude and acceptance are not consistently measured, even though learner perceptions can determine sustained engagement and the likelihood that strategies are used beyond the intervention period (Teo, 2009). Although positive attitudes towards CALL are often assumed, research has also highlighted practical barriers, including cognitive overload when learners face dense concordance lines, limited analytical skills, and frustration with unfamiliar interfaces (Boulton, 2012; Yao, 2019). Accordingly, a modality comparison that includes both achievement and attitudes can clarify whether one format supports learning more efficiently and more acceptably for university learners.

What is currently established is that technology-mediated vocabulary instruction can be effective and that corpus-informed learning can support deeper lexical understanding (Chapelle & Chapelle, 2001; McEnery & Hardie, 2012). However, a clear gap remains regarding whether computer-based delivery provides measurable advantages over carefully designed paper-based delivery when the underlying instructional technique is held constant and when outcomes include both vocabulary development and learner attitudes. In addition, many studies focus on secondary school populations or single-site samples, limiting the transferability to university EFL contexts, where academic vocabulary demands, learner autonomy expectations, and digital practices differ (Ahmadian, Amerian & Goodarzi, 2015). Similarly, some research measures vocabulary primarily through short recognition tests, whereas university learning outcomes may require stronger evidence of durable gains and meaningful engagement with usage.

Furthermore, few studies have operationalised vocabulary layering as a structured, replicable pedagogical model while controlling for instructional medium, leaving uncertainty about whether the observed effects are attributable to the technique, the technology, or contextual variation. Equally, limited research has integrated controlled quasi-experimental comparisons with systematic attitudinal measurement within the same design at the university level, particularly in EFL settings where access to technology is uneven.

This gap supports a study that directly compares paper- and computer-based vocabulary layering techniques in a quasi-experimental design, with equivalent lexical targets,

instructional time, and assessment procedures across groups. On the one hand, paper-based layering may reduce technical barriers and enable tighter control over task sequencing and input complexity, benefiting learners who require structured guidance. On the other hand, computer-based layering may provide interactive features, faster feedback, and multimodal reinforcement that could enhance noticing and retention. However, without systematic comparison under controlled conditions, instructional decisions remain guided more by institutional preference than empirical clarity.

The novelty of the present study lies in isolating vocabulary layering as the constant pedagogical technique while experimentally manipulating only the delivery mode, thereby distinguishing medium effects from instructional design effects. In addition, the study introduces a parallel measurement of cognitive–affective response through structured attitude analysis, enabling examination of how modality influences both measurable vocabulary gains and learner engagement within the same cohort.

Accordingly, the major aim of this research is to conduct a quasi-experimental comparison of paper-based and computer-based vocabulary layering techniques in enhancing EFL university students' vocabulary development. The study addresses six research questions examining: (1) whether students in the paper-based group demonstrate significant pre–post improvement; (2) whether their post-test scores differ from those of a control group; (3) whether students in the computer-based group demonstrate significant pre–post improvement; (4) whether their post-test scores differ from those of a control group; (5) whether post-test differences emerge between the two experimental groups; and (6) whether learner attitudes vary according to instructional modality.

The significance of this study lies in strengthening vocabulary pedagogy in EFL higher education through a focused, methodologically aligned comparison of two delivery formats that implement the same instructional technique. Rather than relying on general assumptions about technology-enhanced learning, the research isolates instructional medium as a variable while maintaining pedagogical consistency, thereby producing clearer evidence for curriculum planning. In addition, integrating achievement outcomes with attitudinal analysis enables a more balanced understanding of both measurable learning gains and learner engagement. Consequently, the findings are expected to support informed decisions regarding resource allocation, instructional design, and sustainable vocabulary development practices in university EFL settings.

Related Studies

Research on corpus-informed and technology-mediated vocabulary instruction has developed across several strands, including comparisons between Data-Driven Learning (DDL) and traditional instruction, modality-based contrasts, and large-scale trend analyses. Al-Mahbashi, Noor, and Amir (2015) examined paper-based DDL using concordance worksheets in a Yemeni university context and reported statistically significant gains in receptive vocabulary and collocational knowledge compared to dictionary-based instruction. **Similarly**, Yao (2019) found that paper-based concordance materials outperformed dictionary use in a Spanish-as-a-foreign-language setting, particularly in immediate post-test performance. **However**, both studies contrasted DDL with traditional methods rather than comparing different delivery modes of the same pedagogical approach.

A comparable pattern appears in Soruç and Tekin (2017), who demonstrated that corpus-centred tasks led to higher vocabulary gains and more positive learner perceptions than traditional instruction in a secondary school context. **Likewise**, Boontam (2022) reported significant improvements in near-synonym use through paper-based DDL worksheets in a Thai university setting, alongside favourable attitudes towards corpus-informed learning. **Nevertheless**, Boontam employed a one-group pre–post design, which limits causal inference, while Soruç and Tekin focused on DDL versus traditional instruction rather than on modality comparisons. **Consequently**, although these studies confirm the pedagogical value of corpus-based techniques, they do not isolate the effect of instructional medium.

In contrast, Ahmadian, Amerian, and Goodarzi (2015) directly compared paper-based and computer-based contextualised vocabulary presentation. Both groups improved, yet the computer-based group achieved higher post-test and delayed post-test scores. **However**, the computer condition relied on multimedia presentation rather than learner-led corpus exploration, raising questions about whether the observed advantages were attributable to the presentation format rather than to instructional structure. **Compared to** DDL-oriented research, this study focused more on delivery technology than on inductive lexical analysis.

Broader syntheses also provide contextual insight. Huang and Ma (2025), through a meta-analysis of 45 DDL studies, reported large overall effect sizes across age groups, with stronger outcomes associated with hands-on corpus interaction. **However,** the review did not systematically differentiate between paper-based and computer-based implementations in tertiary education. **Similarly**, Chen et al. (2021), using bibliometric analysis of CALL research from 1995 to 2019, identified vocabulary learning and mobile-assisted language learning as prominent trends. **However**, their analysis mapped research growth rather than testing comparative instructional effectiveness.

Across these studies, consistent evidence supports corpus-informed vocabulary instruction over traditional approaches. **At the same time**, differences emerge in participant age, research design, intervention length, and the operationalisation of the instructional medium. University-level quasi-experimental comparisons between paper-based and computer-based implementations of the same structured vocabulary technique remain limited. **While** some research measures achievement outcomes, attitudinal analysis is often peripheral or based on small qualitative samples. **Therefore**, the literature reveals a methodological gap: the absence of controlled, university-based studies that simultaneously compare delivery modes of a consistent vocabulary layering technique and examine both achievement and learner attitudes under equivalent instructional conditions.

Materials and Methods

Study design and justification

A quasi-experimental, nonequivalent-groups pretest–posttest design was used because group allocation followed intact university classes rather than individual random assignment. Such a design was selected in an educational setting where timetabling, existing class membership, and administrative constraints made randomisation impractical, **while** still allowing estimation of instructional effects through baseline testing and a comparison group (Shadish, Cook, & Campbell, 2002). A three-group structure was implemented to align with the study aims: one control group received regular vocabulary instruction, and two experimental groups received the same vocabulary layering technique delivered via different media (paper-based versus computer-based). Pre-testing was used to examine baseline comparability and reduce selection threats. In contrast, post-testing was used to estimate learning gains attributable to the intervention, controlling for equivalent instructional time and lexical targets (Creswell & Creswell, 2017).

Setting and participant eligibility

The study was conducted in Northern Cyprus with EFL students enrolled in credit-bearing English courses. Eligibility criteria included: (a) age 19–24; (b) enrolment in the selected EFL course during the same semester; (c) regular attendance expectations (at least 80% of sessions); **and** (d) consent to participate. Exclusion criteria included: (a) prior formal training in corpus tools or concordance (to reduce prior-exposure bias); (b) documented learning accommodations that required substantially different assessment procedures unless equivalent testing could be provided; **and** (c) absence from either the pre-test or post-test administration.

Sampling technique and sample size

A cluster convenience sampling approach was used because intact classes formed the intervention units. Three comparable classes were selected in coordination with the programme timetable. One class was designated as the control group, one class received the paper-based vocabulary layering technique, **and** one class received the computer-based vocabulary layering technique. The intended sample included 90 students (30 per class).

Table 1: Demographic Characteristics of Participants by Group

Group	**Male (n)**	**Female (n)**	**Total (n)**
Control Group	11	19	30
Paper-Based DDL Group	12	18	30
Computer-Based DDL Group	17	13	30
Total	**40**	**50**	**90**

This sample size was considered adequate for detecting medium differences across three groups using standard parametric tests, provided that the normality assumption was reasonably met and attrition remained low (Field, 2024).

Group allocation

Group allocation followed existing class placement rather than individual randomisation. **However**, baseline equivalence was assessed using pretest scores and key demographics (gender, age, department, nationality). Where meaningful imbalances were present, relevant covariates (e.g., pre-test score) were included in supplementary analyses to strengthen inference.

Table 2: One-Way ANOVA Results for Pre-Test Scores Across Groups

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	188.156	2	94.078	.530	.590
Within Groups	15435.500	87	177.420		
Total	15623.656	89			

The analysis indicated no statistically significant differences among the three groups at pre-test level ($p = .590$), suggesting homogeneity prior to the intervention.

Table 3 Descriptive Statistics for Pre-Test Scores by Group

Group	N	Mean (M)	Standard Deviation (SD)
Paper-Based DDL Group	30	38.93	15.51
Computer-Based DDL Group	30	38.86	13.70
Control Group	30	35.83	12.

The mean scores were comparable across groups, further confirming baseline equivalence before treatment implementation.

1.1.Instruments and materials

Course textbook: *Inside Reading: The Academic Word List in Context* guided reading passages and target vocabulary selection across all groups, ensuring identical lexical input.

Corpus tool: Sketch Engine was used with the British Academic Written English (BAWE) corpus to generate concordance lines and collocation evidence for the target items. Corpus extracts were standardised in length and complexity across conditions to maintain equivalence between paper and computer delivery (Kilgarriff et al., 2004).

Vocabulary measures (pre-test and post-test): A 40-item vocabulary knowledge rating scale was administered to assess meaning, sentence use, and collocational production. Parallel forms were created to reduce test–retest effects, **while** retaining comparable difficulty and coverage (Nation, 2022). Internal consistency was assessed using Cronbach's alpha, **and** piloting with a small comparable cohort was used to check clarity and timing.

Attitude questionnaire: A 12-item Likert-scale questionnaire was administered to the two experimental groups to measure perceived usefulness, difficulty, engagement, and willingness

to continue using corpus-informed activities. The instrument was adapted from established CALL/DDL attitude measures and refined for the study context, with reliability estimated via Cronbach's alpha (Turan et al., 2017).

1.2.Procedures and recruitment

Recruitment occurred at the beginning of the semester through an in-class briefing and an information sheet. Participation was voluntary, **and** students were informed that course grades were not affected by participation decisions. Written informed consent was collected before data collection. Pre-testing was administered to all three groups in the same week under identical conditions (30 minutes, supervised classroom setting).

Following the pre-test, a 14-week intervention was delivered. Two introductory training sessions (one hour each) were provided to both experimental groups to develop consistent strategy awareness for interpreting concordance evidence and using contextual clues. The control group followed standard vocabulary practices used in the course (e.g., textbook exercises and teacher-led explanations) **without** corpus-based layering activities.

The paper-based vocabulary layering condition used printed worksheets containing (a) concordance lines, (b) guided noticing prompts, (c) collocation focus tasks, (d) retrieval practice, **and** (e) consolidation exercises. Concordance examples were extracted from the BAWE corpus and printed in a standardised format.

The computer-based vocabulary layering condition used the same layering stages, **but** delivery occurred through supervised corpus searches in Sketch Engine and aligned computer-based exercises. Task sequencing, target vocabulary, and instructional time were matched across both experimental groups to ensure that delivery mode was the primary manipulated variable.

At the end of the intervention, the post-test was administered to all groups under identical conditions. The attitude questionnaire was then completed by the two experimental groups in the same week, with clear guidance provided to minimise misunderstanding.

1.3.Data analysis

SPSS Version 27 was used to analyse the quantitative data. Preliminary screening included checking for missing values, outliers, and the normality of score distributions. Normality was assessed using skewness and kurtosis indices, histograms, and the Shapiro–Wilk test. Because the vocabulary and overall attitude test scores were normally distributed and satisfied the assumption of homogeneity of variance, parametric statistical procedures were selected. Descriptive statistics were first calculated, including means, standard deviations, and score

distributions for all groups at pre-test and post-test stages. Baseline equivalence among the three groups was examined using a one-way ANOVA on pre-test scores. Within-group development was analysed using paired-samples t-tests to compare pre-test and post-test scores for each experimental group.

Between-group comparisons at post-test were conducted using one-way ANOVA, followed by independent-samples t-tests for focused comparisons between the paper-based and control groups, the computer-based and control groups, and the two experimental groups. For each test, t-values, degrees of freedom, and exact p-values were reported. Effect sizes were calculated using Cohen's d for t-tests and partial eta squared for ANOVA to indicate the magnitude of differences. The attitude questionnaire data were analysed using descriptive statistics, and internal consistency reliability was examined using Cronbach's alpha. All statistical analyses were conducted at a significance level of $\alpha = .05$.

1.4.Ethical considerations

Ethical approval was obtained from Cyprus International University, Scientific Research and Publication Ethics Committee (Decision No: EKK2526/01/05, Date: 07/10/2025). The PhD study titled "Comparing the Impact of Paper-Based vs. Computer-Based Data-Driven Learning on University Students' Vocabulary Development," conducted by Khadija Abdalla Abubaker Abdalla (ID: 22115705), was reviewed and approved as ethically appropriate. The study was conducted in accordance with internationally recognized ethical principles for research involving human participants, including the principles outlined in the Declaration of Helsinki These principles included voluntary participation, informed consent, confidentiality, protection of participants' privacy, and the right to withdraw from the study at any time without penalty. All data were securely stored and used solely for research purposes.

Consent to Participate: Informed consent was obtained from all participants involved in the study prior to their participation. Participants were informed about the purpose and procedures of the study, the voluntary nature of their participation, the confidentiality of their information, and their right to withdraw from the study without penalty.

Results

This section reports the quantitative findings in direct alignment with the study aim and the six research questions. Descriptive statistics and inferential analyses are presented sequentially, beginning with within-group pre–post comparisons and followed by between-group post-test comparisons. Paired-samples and independent-samples t-tests were conducted, and effect sizes are reported where applicable. The results are presented objectively in tables that detail means, standard deviations, t-values, degrees of freedom, and exact p-values.

RQ 1

(1) Whether students in the paper-based group demonstrated significant pre–post improvement.

Table 1 displays the paired-samples t-test results comparing the pre-test and post-test scores of the paper-based vocabulary layering group. The pre-test mean was **38.93** (SD = **13.70**), and the post-test mean was 58.60 (SD = 13.97), with **N = 30** participants. The paired-samples t-test yielded **t(29) = 5.50, p = 0.0001**. The mean difference between post-test and pre-test scores was **19.67**. The calculated effect size was **Cohen's d =-1.42**.

Table 1. Paper-Based Group: Pre-Test and Post-Test Comparison (Paired-Samples t-Test)

Measure	**N**	**Mean**	**SD**	**t**	**p**	**Mean Difference (Post − Pre)**	**Cohen's d**
Pre-test	30	38.93	13.70	5.50	0.0001	19.67	-1.42
Post-test	30	58.60	13.97				

RQ2

(2) Whether the paper-based group's post-test scores differed significantly from those of the control group.

Table 2 presents the results of the independent-samples t-test comparing post-test scores between the paper-based group and the control group. The paper-based group (N = 30) obtained a mean score of **58.60** (SD = **13.97**), whereas the control group (N = 30) obtained a mean score of **46.87**(SD = **15.85**). The mean difference between the two groups was **-11.73**. The independent-samples t-test yielded **t(58) = 3.04, p = 0.003**. The effect size was **Cohen's d = -0.785**.

Table 2. Paper-Based Group Vs Control Group: Post-Test Comparison (Independent-Samples t-Test)

Group	N	Mean	SD	t	p	Mean Difference (Paper − Control)	Cohen's d
Paper-based	30	58.60	13.97	3.04	0.003	-11.73	-0.785
Control	30	46.87	15.85				

RQ 3

(3) Whether students in the computer-based group demonstrated significant pre–post improvement.

Table 3 reports the results of the paired-samples t-test comparing pre- and post-test scores for the computer-based group. The pre-test mean for the computer-based group (N = 30) was **38.86** (SD = **12.17**), while the post-test mean was **71.66** (SD = **13.50**). The mean difference (post−Pre) was -32.80. The paired-samples t-test yielded t(29) = 9.88, **p = 0.0001**. The calculated effect size was **Cohen's d = -2.55**.

Table 3 Computer-Based Group: Pre-Test and Post-Test Comparison (Paired-Samples t-Test)

Measure	N	Mean	SD	t	p	Mean Difference (Post − Pre)	Cohen's d
Pre-test	30	38.86	12.17	9.88	0.0001	-32.80	-2.55
Post-test	30	71.66	13.50				

RQ4

Whether the computer-based group's post-test scores differed significantly from the control group's.

Table 4 presents the results of the independent-samples t-test comparing post-test scores between the computer-based and control groups. The computer-based group (N = 30) obtained a mean score of **71.67** (SD = **13.50**), whereas the control group (N = 30) obtained a mean score of **46.87** (SD = **15.85**). The mean difference between the two groups was **-24.80**. The independent-samples t-test yielded **t(58) = 6.52, p = 0.0001**. The effect size was **Cohen's d = -1.68**.

Table 4. Computer-Based Group Vs Control Group: Post-Test Comparison (Independent-Samples t-Test)

Group	N	Mean	SD	t	p	Mean Difference (Computer − Control)	Cohen's d
Computer-based	30	71.67	13.50	6.52	0.0001	-24.80	-1.68
Control	30	46.87	15.85				

RQ5

Whether post-test differences emerged between the two experimental groups.

Table 5 reports the results of the independent-samples t-test comparing post-test scores between the paper-based and computer-based groups. The paper-based group (N = 30) obtained a mean score of **58.60** (SD = **13.97**), while the computer-based group (N = 30) obtained a mean score of **71.67** (SD = **13.50**). The mean difference between the two groups was **-13.07**. The independent-samples t-test yielded **t(58) = 3.68, p = 0.0005**. The effect size was **Cohen's d = -0.951**.

Table 5 Paper-Based Vs Computer-Based Groups: Post-Test Comparison (Independent-Samples t-Test)

Group	N	Mean	SD	t	p	Mean Difference (Computer − Paper)	Cohen's d
Paper-based	30	58.60	13.97	3.68	0.0005	-13.07	-0.951
Computer-based	30	71.67	13.50				

RQ6

Whether learner attitudes varied according to instructional modality.

Table 6 summarises the independent-samples t-test comparing overall attitude scores between the paper-based and computer-based groups. The paper-based group (N = 30) reported a mean attitude score of **3.72** (SD = **0.88**), whereas the computer-based group (N = 30) reported a mean of **3.70** (SD = **0.87**). The mean difference was **0.02**. The independent-samples t-test yielded **t(58) = 0.08**, with **p = 0.93**.

Table 6 Learner Attitudes: Paper-Based Vs Computer-Based (Independent-Samples t-Test)

Group	N	Mean	SD	t	p	Mean Difference (Paper − Computer)
Paper-based	30	3.72	0.88	0.08	0.93	0.02
Computer-based	30	3.70				

Discussion

The primary aim of this study was to investigate whether paper-based and computer-based Data-Driven Learning (DDL) differentially influence university EFL students' vocabulary development and attitudes in comparison with regular instruction. To address this aim, a quasi-experimental, nonequivalent-groups pretest–posttest design was implemented with three intact classes at a university in Northern Cyprus. Ninety EFL students aged between 19 and 24 years participated in the study. One group received regular vocabulary instruction and served as the control group. In contrast, the two experimental groups received the same vocabulary layering intervention delivered through either paper-based or computer-based DDL. Baseline equivalence among the groups was established before the intervention. Overall, the findings revealed significant vocabulary gains across all groups from pre-test to post-test. However, significant differences emerged in post-test performance, indicating that the mode of instruction influenced the extent of vocabulary development.

The findings demonstrated that learners in the paper-based DDL condition showed significant improvement in vocabulary knowledge following the intervention. This result suggests that a "soft DDL" approach, built around carefully selected concordance lines and structured noticing activities, can effectively support vocabulary learning even in the absence of digital technology. This finding is consistent with previous research indicating that paper-based corpus materials can facilitate vocabulary acquisition through repeated exposure to authentic language data, enhanced pattern salience, and guided discovery learning (Al-Mahbashi et al., 2015; Boontam, 2022; Yao, 2019). The effectiveness of the paper-based condition supports the argument that the pedagogical value of DDL lies primarily in learners' engagement with authentic linguistic evidence and their ability to identify recurring lexical patterns rather than in technology itself. Consequently, corpus-informed instruction may remain effective in educational contexts where technological resources are limited.

Similarly, the computer-based DDL group demonstrated significant gains in vocabulary development from pre-test to post-test. This finding supports previous research suggesting that

direct interaction with corpus tools can facilitate vocabulary learning by encouraging learners to explore authentic language use independently and engage in inductive learning processes (Huang & Ma, 2025; Soruç & Tekin, 2017). By searching, filtering, and examining concordance lines, learners can notice lexical patterns, test hypotheses, and refine their understanding of word meanings and usage. Such processes are central to the theoretical foundations of DDL and may contribute to stronger vocabulary retention and deeper lexical knowledge. The result also aligns with broader Computer-Assisted Language Learning (CALL) research demonstrating the positive contribution of technology-enhanced learning environments to vocabulary acquisition (Barani, 2013; Chen et al., 2021).

More importantly, the comparison between the post-test results of the two DDL conditions revealed that computer-based DDL produced significantly greater gains in vocabulary than paper-based DDL. The computer-based group achieved a significantly higher post-test mean score ($M = 71.67$, $SD = 13.50$) than the paper-based group ($M = 58.60$, $SD = 13.97$), with the difference statistically significant ($t(58) = 3.68$, $p = .0005$). Furthermore, the large effect size ($d = 0.95$) indicates that the advantage of computer-based DDL was not only statistically reliable but also educationally meaningful.

Several explanations may account for the superior performance of the computer-based DDL group. First, digital corpus environments provide learners with immediate access to large quantities of authentic language data. Unlike paper-based materials, which are limited to pre-selected examples, computer-based corpus tools allow learners to conduct independent searches, examine a wider range of concordance lines, and explore lexical items from multiple perspectives. This increased exposure to contextualised language use may facilitate deeper lexical processing and provide greater opportunities to notice linguistic patterns, both of which are considered essential mechanisms in vocabulary acquisition.

Second, the findings may be interpreted through the lens of learner autonomy. A fundamental principle of DDL is that learners function as language researchers who actively discover linguistic regularities from authentic data. Computer-based environments appear particularly well suited to fostering this role because they enable learners to manipulate corpus data directly, formulate hypotheses, and verify their observations through further searches. Such active engagement may enhance learners' sense of ownership over the learning process and promote deeper understanding of vocabulary meanings, collocations, and usage constraints.

The findings also support cognitive perspectives on language learning that emphasise repeated exposure, pattern recognition, and meaningful interaction with linguistic input. Computer-based DDL allows learners to access a larger quantity and greater variety of examples within a relatively short period of time, thereby increasing opportunities for inductive learning. The availability of multiple contextualised instances may help learners establish stronger form–meaning connections and develop a more sophisticated understanding of lexical usage. Consequently, while both paper-based and computer-based DDL are grounded in the same pedagogical principles, the technological affordances of digital corpus tools appear to enhance their effectiveness.

Despite the superiority of the computer-based condition, the findings should not be interpreted as evidence that paper-based DDL is ineffective. On the contrary, learners in the paper-based group also demonstrated substantial vocabulary gains, confirming the value of corpus-informed instruction regardless of delivery mode. Rather, the results suggest that computer-based DDL offers additional benefits that extend beyond those available through static concordance materials. These benefits may include faster retrieval of examples, greater input quantity, increased learner control, and more opportunities for independent exploration. Therefore, while paper-based DDL remains a valuable option in contexts with limited technological infrastructure, computer-based DDL may provide optimal conditions for vocabulary development when adequate technological resources and learner support are available.

Regarding learner attitudes, the findings showed no statistically significant differences between the paper-based and computer-based DDL groups. Students in both conditions reported generally positive perceptions of the instructional approach, suggesting that acceptance of DDL may be influenced more by the perceived usefulness of discovery learning and corpus evidence than by the delivery platform itself. This finding is consistent with previous research demonstrating that learners often appreciate DDL for promoting autonomy, engagement, and a deeper understanding of language patterns, while acknowledging the cognitive demands of corpus analysis (Boontam, 2022; Huang & Ma, 2025; Yao, 2019).

The absence of significant attitudinal differences suggests that technological delivery alone may not determine learners' perceptions of DDL. Instead, factors such as task design, instructional scaffolding, and the perceived relevance of corpus activities may play a more important role in shaping learner attitudes. Consequently, instructors may select either paper-

based or computer-based DDL based on contextual constraints while maintaining positive learner engagement through carefully structured, discovery-oriented tasks.

Overall, the findings provide strong empirical support for the effectiveness of DDL as a vocabulary learning approach. Although both delivery modes contributed significantly to vocabulary development, computer-based DDL produced substantially greater gains, highlighting the pedagogical value of direct corpus interaction. At the same time, the positive outcomes observed in the paper-based condition demonstrate that meaningful vocabulary learning can still occur when corpus evidence is presented through non-digital formats. These findings suggest that DDL represents a flexible and effective pedagogical framework that can be adapted to diverse educational contexts while promoting vocabulary growth, learner autonomy, and engagement with authentic language use.

Theoretical and Practical Implications

Theoretical Implications

The study will refine DDL theory by separating the core discovery mechanism from delivery medium, showing how "soft" and "hard" implementations will operate as equivalent routes to inductive noticing. **Similarly**, the work will extend debates on modality by treating paper-based corpus evidence as a technology in its own right, **rather than** a reduced substitute. The findings will challenge assumptions that digital affordances will be necessary for corpus-based learning effects, **while** recognising that interface control will shape learner agency. Conceptually, the study will position vocabulary layering as a scalable design principle linking input enhancement, retrieval practice, and pattern abstraction in higher education.

Practical Implications

The results will inform language policy by supporting corpus-informed vocabulary development as a standard option within credit-bearing EFL provision. Policymakers will be able to reference these outcomes when setting minimum national digital equity benchmarks and vocabulary standards for foundation programmes. **In addition**, university management will be able to align resourcing decisions with teaching goals by adopting paper-based concordance packs where computer access is limited, or by scheduling supervised lab sessions where platforms are available. Lecturers will be able to plan layered lessons that will combine

concordance evidence, guided noticing, collocation focus, and retrieval practice, **while** keeping lexical targets consistent across cohorts. Students will be able to use the same learning routine across formats, reducing uncertainty and supporting sustained engagement. **Moreover**, higher education institutions will be able to embed short corpus-literacy training into induction weeks to normalise DDL practices and reduce cognitive overload (Huang & Ma, 2025).

Strengths and Limitations

Strengths

The study will demonstrate originality by testing the same DDL vocabulary layering technique across paper and computer delivery, with matched lexical input and instructional time. **Moreover**, the quasi-experimental three-group structure will strengthen causal inference within timetabled classes, **while** baseline testing will reduce selection threats (Shadish et al., 2002). The sample size will support estimation of group differences and will permit effect-size reporting alongside significance testing (Field, 2024). **In addition**, instrument development will employ parallel test forms and reliability checks to enhance measurement consistency. Analytical rigour will be reinforced through assumption screening, reporting of exact p-values, and effect size metrics.

Limitations

The study will be limited by the use of intact classes and cluster convenience sampling, which will constrain generalisability beyond similar university EFL settings. **Although** baseline equivalence will be assessed, unmeasured differences in prior exposure, study habits, or teacher–student dynamics will still confound estimated effects. The work will be conducted within a single regional context and a single semester, so transfer to other curricula, proficiency ranges, or institutional cultures will remain uncertain. **In addition**, outcomes will rely mainly on a short vocabulary measure and a self-report attitude scale, which will not fully capture productive use in extended writing or speaking. Intervention fidelity will depend on consistent scaffolding and task pacing, and minor deviations across sessions will introduce noise. **Similarly**, attitudes will be sensitive to novelty and perceptions of workload, so stable

acceptance over time cannot be confirmed. **Finally**, the absence of a delayed post-test will limit claims about long-term retention.

Recommendations for Future Study

Future research will replicate the comparison across multiple universities and countries to test whether modality equivalence holds under different infrastructure and pedagogical cultures. **In addition**, studies will employ probability-based sampling where feasible and will include teacher-level variables, such as prior corpus experience, instructional style, and fidelity of scaffolding. Researchers will extend outcomes beyond receptive scores by assessing collocational accuracy in writing, spoken fluency with target items, and transfer to discipline-specific academic tasks. **Moreover**, longitudinal designs will include delayed posttests and spaced-review conditions to determine whether retention trajectories differ across delivery modes. Mixed-method work will integrate learning analytics from corpus logs with interviews to clarify how search behaviour, cognitive load, and autonomy will mediate gains (Huang & Ma, 2025). Comparative trials will examine “hands-on” querying against “hands-off” curated lines within the same computer condition to isolate interface effects. **Likewise**, research will test mobile-based corpus access and low-bandwidth solutions to address digital equity. Measurement models will refine attitude scales by separating perceived usefulness, perceived difficulty, and willingness to continue, and will explore links to achievement over semesters. **Finally**, experimental variants will test different corpus sizes, concordance line lengths, and prompt types to optimise the layering sequence for diverse proficiency levels.

Conclusion

The study examined whether paper-based and computer-based Data-Driven Learning (DDL) vocabulary layering produced different outcomes in university EFL students' vocabulary development and learner attitudes compared with regular vocabulary instruction. A quasi-experimental, nonequivalent-groups pretest–posttest design provided a realistic evaluation of the instructional intervention within intact university classes. Baseline testing confirmed that the groups were comparable before the intervention, thereby strengthening the interpretation of the post-test findings.

The results demonstrated that both DDL conditions produced significant improvements in vocabulary learning and outperformed regular instruction at the post-test. However, the

computer-based DDL group achieved significantly greater vocabulary gains than the paper-based DDL group, indicating that technology-enhanced DDL can further strengthen vocabulary development when combined with effective teacher guidance and structured autonomous learning opportunities. The interactive features of digital corpus tools enable learners to access a wider range of authentic language examples, conduct independent corpus searches, and engage more actively in noticing lexical patterns, thereby promoting deeper vocabulary learning.

Regarding learner attitudes, no statistically significant differences were found between the paper-based and computer-based DDL groups. This finding suggests that learners' positive perceptions of DDL are influenced more by the value of guided discovery learning and authentic corpus evidence than by the instructional platform itself.

The findings also demonstrate that paper-based concordance materials provide a practical and pedagogically sound alternative in contexts where access to digital resources is limited. In contrast, supervised computer-based DDL offers additional opportunities to develop corpus literacy and independent learning skills. Overall, the study contributes empirical evidence to an underexplored area of tertiary DDL research by comparing two modes of DDL delivery within the same instructional framework. These findings have important implications for curriculum designers, teacher educators, and policymakers by informing the selection of appropriate DDL implementation strategies and supporting the integration of corpus-informed instruction and corpus literacy into university language programs.

Author Contributions: Data curation, K.A.A.; Formal analysis, K.A.A.; Investigation, A.Y.; Methodology, K.A.A. and A.Y.; Resources, K.A.A.; Software, K.A.A.; Supervision, A.Y.; Validation, A.Y.; Writing—review and editing, K.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement:

Ethical approval was obtained from Cyprus International University, Scientific Research and Publication Ethics Committee (Decision No: EKK2526/01/05, Date: 07/10/2025). The PhD study titled “Comparing the Impact of Paper-Based vs. Computer-Based Data-Driven Learning on University Students' Vocabulary Development”, conducted by Khadija Abdalla Abubaker Abdalla (ID: 22115705), was reviewed and approved as ethically appropriate. The study adhered to internationally recognised ethical principles for social science research, including

voluntary participation, informed consent, confidentiality, and the right to withdraw without penalty (World Medical Association, 2013). All data were securely stored and used solely for research purposes.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: I would like to express my sincere and profound gratitude to **Professor Ali Yavuz** for his invaluable guidance, continuous support, and insightful academic supervision throughout this research journey. His expertise, constructive feedback, and unwavering encouragement have significantly strengthened this work and contributed greatly to its successful completion. I also extend my deepest appreciation to my family for their unconditional love, patience, and constant support. Their belief in me has been a continuous source of strength and motivation, enabling me to overcome challenges and achieve this important milestone.

Conflicts of Interest: All authors certify that they have no affiliations with or involvement in anyorganization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. Ahmadian, M., Amerian, M., & Goodarzi, A. (2015). A comparative study of paper-based and computer-based contextualization in vocabulary learning of EFL students. *Advances in Language and Literary Studies, 6*(2), 96–102.

2. Al-Mahbashi, A., Noor, N. M., & Amir, Z. (2015). The effect of data driven learning on receptive vocabulary knowledge of Yemeni University learners. *3L: Language, Linguistics, Literature: The Southeast Asian Journal of English Language Studies, 21*(3), 13–24.

3. Barani, G. (2013). The impact of computer-assisted language learning (CALL) on vocabulary achievement of Iranian university EFL learners. *International Journal of Basic Sciences & Applied Research, 2*(5), 531–537.

4. Boontam, P. (2022). The effect of teaching English synonyms through data-driven learning (DDL) on Thai EFL students' vocabulary learning. *Shanlax International Journal of Education, 10*(2), 80–91.

5. Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534–572.

6. Boulton, A. (2012). Corpus consultation for ESP: A review of empirical research. In A. Boulton, S. Carter-Thomas, & E. Rowley-Jolivet (Eds.), *Corpus-informed research and learning in ESP: Issues and applications* (Studies in Corpus Linguistics, Vol. 52, pp. 261–292). John Benjamins Publishing Company. https://doi.org/10.1075/scl.52.11bou

7. Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.

8. Chen, X., Zou, D., Xie, H. R., & Su, F. (2021). Twenty-five years of computer-assisted language learning: A topic modeling analysis. *Language Learning & Technology, 25*(3), 151–185.

9. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.

10. Field, A. (2024). *Discovering statistics using IBM SPSS Statistics*. Sage.

11. Huang, S., & Ma, Q. (2025). A systematic review of data-driven learning research on language learning and teaching for pre-tertiary learners: Balancing qualitative and quantitative research. *Learning and Individual Differences, 122*, 102752.

12. Johns, T. F. (1991). Should you be persuaded: Two samples of data-driven learning materials. *ELR Journal, 4*, 1–16.

13. Kilgarriff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. In G. Williams & S. Vessier (Eds.), *Proceedings of the 11th EURALEX International Congress* (pp. 105–116). Université de Bretagne-Sud.

14. Nation, I. S. P. (2022). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

15. Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.

16. Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

17. Soruç, A., & Tekin, B. (2017). Vocabulary learning through data-driven learning in an English as a second language setting. *Educational Sciences: Theory & Practice, 17*(6), 1811–1832. https://doi.org/10.12738/estp.2017.6.0305

18. Teo, T. (2006). Attitudes toward computers: A study of post-secondary students in Singapore. *Interactive Learning Environments, 14*(1), 17–24.

19. Paker, T., & Özcan, Y. E. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. *International Journal of Language Academy, 5*(1), 62–81. https://doi.org/10.18033/ijla.3494

20. Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.

21. Yao, G. (2019). Vocabulary learning through data-driven learning in the context of Spanish as a foreign language. *Research in Corpus Linguistics, 7*, 18–46.