

---

**| RESEARCH ARTICLE**

## Improving Lexical Resources among ESL Learners: A Corpus-Based Approach

Dr. Amjed A. Ayyat<sup>1</sup>✉, Dr. Emad A. Alqdumi<sup>2</sup>, Dr. Haitham Al-Yousef<sup>3</sup>

<sup>1</sup>Lecturer, Center for Preparatory Studies, Dhofar University, Salalah, Sultanate of Oman

<sup>2</sup>Assistant Professor, Department of English Language and Translation, Jerash University, Jerash, Jordan

**Corresponding Author:** Dr. Amjed A. Ayyat, **E-mail:** [ayyat78@yahoo.com](mailto:ayyat78@yahoo.com)

---

**| ABSTRACT**

This study aims to investigate the effectiveness of utilizing the Corpus of Contemporary American English (COCA) in fostering the writing skills, particularly the lexical resources, of ESL learners. A quasi-experimental design was implemented involving two groups. An experimental group (n = 25) in which COCA was integrated into their teaching practice, and a control group (n = 25) that received no COCA intervention. Data were gathered using pre- and post-writing tests designed to evaluate students' vocabulary. The data were analyzed using independent and paired samples t-tests, supported by an Analysis of Covariance (ANCOVA) to contrast the performance of the two groups. Results showed a statistically significant and positive change in the writing skills of the experimental group compared to the control group. Furthermore, the two tests revealed dramatic gains within the experimental group from pre-test to post-test. The ANCOVA further supported the positive effect of the COCA intervention. These findings suggest that integrating COCA can significantly enhance the lexical resources of ESL learners, resulting in improved writing skills. The study supports the use of corpus-based tools in language instruction to enhance vocabulary learning and overall writing proficiency among ESL learners. The research suggests that COCA is a valuable resource in English language teaching and can be applied to other criteria, such as grammar, cohesion, and coherence.

**| KEYWORDS**

COCA, Corpus Linguistics, ESL Learners, Lexical Resources

**| ARTICLE INFORMATION**

**ACCEPTED:** 15 August 2025

**PUBLISHED:** 30 August 2025

**DOI:** 10.32996/bjal.2025.5.2.5

---

**Introduction:**

This research paper investigates the efficiency of using the Corpus of Contemporary American English (COCA) in fostering lexical resources among ESL students at a private university in Jordan. Teachers and students can both use the free online resource COCA. Professor Mark Davies of Brigham Young University created and refined it, and on February 20, 2008, it was initially made available online. It is the most extensive collection of contemporary American English available to the public. The five categories are spoken, fiction, magazine, newspaper, and academic. There are more than 520 words that separate these genres equally.

Additionally, this e-tool is periodically updated to reflect the latest advancements resulting from numerous advancements in the linguistic proficiency of contemporary American English. A thorough investigation, combined with the availability of pertinent computer software, such as WordSmith, AntConc, and Range, will undoubtedly make employing this e-tool beneficial, even though its use is complex. The benefits are substantial because the website contains a wealth of information. Additionally, fresh data from native speakers is frequently added to the platform. It is essential to note that most modern corpora, such as the American National Corpus (ANC) and the British National Corpus (BNC), are proprietary. However, by visiting the website <https://www.english-corpora.org/coca/> and registering, you can view all available corpora.

**Problem Statement:**

Even though Jordanian ESL learners learn English for 12 years in schools, they still face many challenges in developing their writing skills, specifically in lexical resources. Furthermore, adopting a traditional approach to teaching writing skills can exacerbate these challenges. Introducing an innovative approach, such as COCA, can enhance writing skills and enrich the vocabulary students can use in their writing. However, this approach still needs to be explored within the Jordanian education system.

**Research Questions:**

This paper aims to answer the following research question:

1- Does integrating COCA in teaching improve the writing skills, specifically the lexical resources of Jordanian undergraduate students in the English Language Department?

**2. Research Hypotheses:**

This research has two hypotheses as follows:

- **The Null hypothesis:**

Integrating COCA does not significantly enhance students' lexical resources in their writing skills.

- **Alternative hypothesis:**

Integrating COCA significantly enhances students' lexical resources and improves their writing skills.

**Literature Review:**

Sanosi (2024) investigated the use of corpus linguistics to improve students' writing cohesiveness. The study had two main questions: Do native speakers and ESL learners utilize different kinds and amounts of conjunctions? Do native speakers and EFL learners differ in the quality of conjunction usage, specifically syntactic location?

To achieve this, Sanosi employed a corpus linguistic technique to investigate how EFL learners utilize conjunctions. Two corpora were used in the study: the British Academic Written English (BAWE) corpus, which contains 125,314 words of writing by Saudi Arabian EFL college students, and a learner corpus named ALAWEC. The selection of the 125,347-word snippet was based on its alignment with the genre, text type, and student level of the ALAWEC corpus. Sanosi (2024) applied the mixed technique, determining the frequency of each conjunction in each corpus using the corpus analysis program Lancsbox. Second, it is necessary to ascertain each corpus's relative frequency of conjunctions. Thirdly, the range of conjunctions was evaluated to determine their uniform distribution throughout the corpus files. The statistical significance of the variations in conjunction frequency between native speakers and EFL learners was then assessed using a t-test. Sanosi employed conjunction syntactic placement for the qualitative analysis, focusing on conjunctions at the beginning of sentences, which have been seen as a sign of informality and, perhaps, lower writing quality. Five conjunctions commonly encountered in internal clause positions—and, or, but, so, and then—were chosen to examine their incidence in starting positions. The researchers used Lancsbox's sophisticated search feature to locate and investigate these instances.

By comparing the corpus of EFL learners with that of native speakers, Sanosi could identify patterns and differences in conjunction usage, as it enables a logical and quantitative analysis of actual language use. The primary conclusions of the study show that while native speakers use some conjunctions more frequently than EFL learners, such as "however," "then," "therefore," and "yet," EFL learners tend to misuse conjunctions, primarily additive conjunctions. Similar patterns of conjunction usage are seen in both groups. Employing temporal conjunctions the least and mostly additive conjunctions. Furthermore, EFL learners and native speakers do not differ significantly in the syntactic placement of most conjunctions; nevertheless, EFL learners employ the conjunction "and" at the start of sentences more frequently. The study's overall findings indicate that although EFL learners tend to overuse certain conjunctions, they utilize a similar variety of conjunction types as native speakers, and their syntactic placement of conjunctions is generally in line with native speakers' usage, except for the conjunction "and."

The effectiveness of corpus-integrated instruction for EFL students in the Chinese context was investigated and assessed by Zhong and Wakat (2023). While decades of research have demonstrated the potential benefits of DDL, the study focused on the underutilization of corpora in Chinese EFL classes. The study employed a developmental, explanatory sequential mixed-methods design and a quasi-experimental approach, involving 80 mixed-level Senior 2 and 3 students from a private senior high school in western China. Data collection methods included semi-structured interviews with teachers and students, classroom observations

by two university instructors, and a 100-point grammatical competency pre-test. In response to the pre-test results, which revealed weaknesses in verbs, nonfinite verbs, prepositions, relative clauses, conjunctions, and word building, eight weeks of corpus-integrated courses were developed and implemented. In addition to a 710,000-token corpus gathered from Bookworms literary novels, New Concept English textbooks, COCA, and GloWbE, these courses used AntConc and AntWordProfiler. Following the distribution of printouts based on concordance, the students engaged in activities designed to improve their understanding and application of grammar, including analyzing concordance lines, identifying grammatical patterns, and completing further assignments. The effectiveness of the intervention was assessed using post-tests and follow-up interviews. The results showed a substantial difference between the pre-test and post-test scores of Senior Two and Senior Three students, indicating the positive impact of DDL-inspired instruction on their grammar proficiency. Language instructors, educational psychologists, textbook authors, and applied linguists must collaborate to create and execute corpus-integrated classes. The study also demonstrates the effectiveness of inductive learning, which uses corpus analysis to teach and apply grammar principles. It also acknowledges the challenges of incorporating corpus literacy into traditional classroom instruction, such as requiring teachers to have specialized knowledge and potential resistance to change. The authors suggest that providing ongoing training and integrating corpus literacy into teacher education programs could address these problems.

Rafatbakhsh and Ahmadi (2019) aimed to ascertain the frequency of different idiom types in the Corpus of Contemporary American English (COCA) to create more beneficial language learning resources. The researchers extracted 1506 idioms from the Oxford Dictionary of Idioms, 2nd ed., grouped under 81 themes. To account for variations in form, a custom script was created to search COCA for each phrase's grammatical and spelling variations. The findings showed that, with 4.71 occurrences per million words, "behind the scenes" was the most prevalent idiom in COCA, whilst 234 idioms were completely missing. "Fools and foolishness" was the least frequent theme, occurring 0.16 times per million words, while "secrecy" was the most frequent, occurring 10.26 times per million. According to the study's findings, learners, educators, and material developers can identify the most beneficial idioms for various learning contexts through corpus-based frequency analysis.

Daskalovska (2015) examined corpus-based verb-adverb collocation learning exercises in a specific context and contrasted them with standard course book exercises. According to the study's findings, participants who used the online concordance performed significantly better on their examinations. The study's conclusions demonstrated how inspiring and captivating the corpus-based exercises were for the participants. As a result, the online corpus is a valuable and essential tool that benefits second language learners.

The literature demonstrates the impact of online corpora on teaching English as a second language. Nonetheless, this study examines how well a theoretical framework—more significantly, the Contemporary Corpus of American English, or COCA—can improve the writing abilities of ESL students in the context of GFP.

### **Research Methodology:**

This research employed a quasi-experimental design to explore the effectiveness of integrating COCA in enhancing the writing skills, specifically the vocabulary, of Jordanian ESL learners at the English Language Department. Participants were divided into two groups: the experimental group, which underwent COCA, and the control group, which followed a traditional teaching method. To measure changes in the lexical resource criteria, pre-tests and post-tests were used.

### **Participants and Sampling:**

The study included 50 ESL students at a private university in Jordan. These participants were divided into two groups: an experimental group (n = 25) and a control group (n = 25). Group assignment was employed to maintain the initial comparability of the two groups.

### **Data Collection:**

A writing test was designed to measure students' lexical resources. The test was administered twice: as a pre-test at the beginning of the study to assess the level of students in the two groups, and as a post-test at the end of a six-week intervention period to evaluate the improvements in writing skills, specifically lexical resources.

### **Data Analysis Techniques:**

The data were analyzed using several statistical techniques to determine the effectiveness of the COCA intervention. These techniques included:

1. Independent Samples T-test: Used to compare the mean scores of the experimental group and the control group on both the pre-test and post-test.

2. Paired Samples T-test: Utilized to compare the mean pre-test and post-test scores within each group to determine if significant changes occurred in their performance over the intervention period.
3. Analysis of Covariance (ANCOVA): Performed to provide a more robust comparison of post-test performance between the groups, controlling for initial differences in pre-test scores.

### **Reliability and Validity:**

The writing skills test was validated and underwent reliability testing to guarantee the accuracy and trustworthiness of the data. Experts in English language teaching and assessment evaluated the test and ensured that it measured the students' academic writing vocabulary. Inter-rater reliability was established by having two independent correctors assess all written samples, with a high correlation ( $r = 0.88$ ) indicating strong agreement.

### **Ethical Considerations:**

Ethical principles were considered in this study. Participation was not compulsory, and consent was obtained from all participants. To ensure confidentiality, all data were anonymized. A strict study protocol was designed to minimize potential harm to participants, and the study was conducted according to the university's ethical standards.

### **Data Analysis**

To test the research hypotheses and answer the study's questions, the following statistical techniques were applied:

1. Independent Samples T-test: This test was used to compare the mean scores of the experimental group and the control group on the pre-test, assessing the initial equivalence of the two groups before the intervention. It was also used to compare the mean scores of the two groups on the post-test.
2. Paired Samples T-test: This test was utilized to compare the mean pre-test scores and the mean post-test scores *within* the experimental group to determine if there was a significant change in their performance after the intervention. A similar paired samples t-test was conducted for the control group to assess any changes in their scores over the same period.
3. Analysis of Covariance (ANCOVA): To provide a more robust comparison of the post-test performance between the experimental and control groups while statistically controlling for initial differences, an ANCOVA was performed. The post-test score was the dependent variable, the group (experimental vs. control) was the independent variable, and the pre-test score was included as a covariate.

The significance level for all statistical tests was set at  $\alpha = 0.05$ . Decisions regarding the rejection or acceptance of null hypotheses were made based on the p-values (Sig.) obtained from the analyses relative to this significance level.

### **Sample:**

The participants in this study consisted of undergraduate Jordanian ESL learners at the English language department. A total of 50 students from the English Language Department at a private university in Jordan, participated in the study. The students were assigned to either an experimental group or a control group. Each group consisted of 25 students (Experimental Group,  $n = 25$ ; Control Group,  $n = 25$ ). The assignment to groups aimed to ensure comparability between the two cohorts at the outset of the study, which was later confirmed by the analysis of pre-test scores.

The data for the study were collected using a writing test specifically designed to measure student vocabulary. This test was applied twice:

1. Pre-test: Conducted at the beginning of the study for the two groups (Pilot and Control) to determine the initial student level and ascertain the equivalence of the two groups before starting the trial.
2. Post-test was performed at the end of the study period after six weeks. During this period, the experimental group utilized the Corpus of Contemporary American English (COCA) blog to enhance their writing vocabulary, while the control group followed the standard method.

All tests were performed in standardized conditions to ensure accuracy. All written samples were then corrected by trained reviewers using clear criteria to ensure consistency and objectivity. Both the CRP and VCT scores were used to analyze study results and assess how effectively COCA is used to improve writing skills.

### Validity and Reliability of the Writing Skills Test:

To ensure the accuracy and trustworthiness of the data collected, the validity and reliability of the writing skills test, which included assessing lexical resources, were established using appropriate procedures.

The writing skills test was confirmed to be **reliable and valid** for use in this study in two ways. First, experts in language teaching and measurement reviewed the test and correction rules to ensure that the test accurately measured students' academic writing vocabulary and made adjustments based on their feedback. Second, to ensure consistency of correction due to its subjective nature, two correctors independently corrected all written samples. The results showed a significant agreement between them (a high correlation of 0.88). Based on expert review and high agreement among correctors, the test was considered valid and reliable for evaluating the vocabulary component of student writing skills in this study.

Based on the expert review for content validity and the high inter-rater reliability coefficient, the writing skills test was deemed valid and reliable for use in this study to assess the lexical resources component of students' writing skills.

### Equivalence of the Two Groups:

An independent samples T-test was used to check for equivalence between the two groups. The difference between the mean scores of the students in the two groups (experimental and control) on the total score of the Writing skills, including the lexical resources test, was calculated, as shown in the following Table:

Table (1): Scores of the Control and Experimental Groups on the Writing Skills include the Lexical Resources Test.

| Group        | Test     | N  | Mean | Standard Deviation | T-value | Sig.  | 95% Confidence Interval of the Difference |       |
|--------------|----------|----|------|--------------------|---------|-------|---|-------|
|              |          |    |      |                    |         |       | Lower                                     | Upper |
| Experimental | Pre-test | 25 | 12.5 | 1.98               | 0.43    | 0.253 | 1.734                                     | 4.804 |
| Control      | Pre-test | 25 | 11.9 | 1.74               |         |       |   |       |

It is observed from the previous Table that the significance value (Sig.) for the total score of the Writing skills, including lexical resources pre-test for the two groups (control and experimental) is 0.543, which is greater than the assumed significance level (0.05). This indicates that there is no statistically significant difference between the mean scores of the students in the experimental and control groups on the Writing skills, including lexical resources, as measured by the pre-test.

This indicates the equivalence and homogeneity of the individuals in the experimental and control groups in terms of writing skills, including the lexical resources test. Consequently, any differences that may appear later in the post-test can be attributed to the use of the Corpus of Contemporary American English (COCA) in enhancing lexical resources in writing skills among Jordanian ESL learners at the English Language Department.

### Research Main Hypotheses:

**H0: Integrating COCA does not significantly enhance the lexical resources in the writing skills of students.**

Three secondary Hypotheses were made:

The first hypothesis: There is no statistically significant difference between the mean scores of the experimental group and the control group students on the writing skills post-test.

This hypothesis was tested using an independent samples t-test, where the differences between the mean scores of the students in the two groups (experimental and control) on the writing skills post-test were calculated, as shown in the following Table:

Table (2): Results of the t-test for the Significance of Differences between the Mean Scores of the Experimental and Control Groups on the Writing Skills Post-test.

| Group        | Test      | N  | Mean | Standard Deviation | T-value | Sig.  | 95% Confidence Interval of the Difference |       |
|--------------|-----------|----|------|--------------------|---------|-------|---|-------|
|              |           |    |      |                    |         |       | Lower                                     | Upper |
| Experimental | Post-test | 25 | 16.9 | 1.42               | 11.7    | 0.000 | 1.221                                     | 4.982 |
| Control      | Post-test | 25 | 11.7 | 3.07               |         |       |   |       |

We observe from the previous Table that the significance value (Sig.) for the total score on the quadratic functions test (0.000) is smaller than the assumed significance level (0.05). This indicates a statistically significant difference between the mean scores of the students in the experimental and control groups on the writing skills post-test, favoring the experimental group. Therefore, we reject the null hypothesis and accept its alternative hypothesis, which states that there is a statistically significant difference between the mean scores of the experimental group and the control group students on the writing skills post-test, favoring the experimental group. The following figure shows the differences between the two groups in the writing skills test:

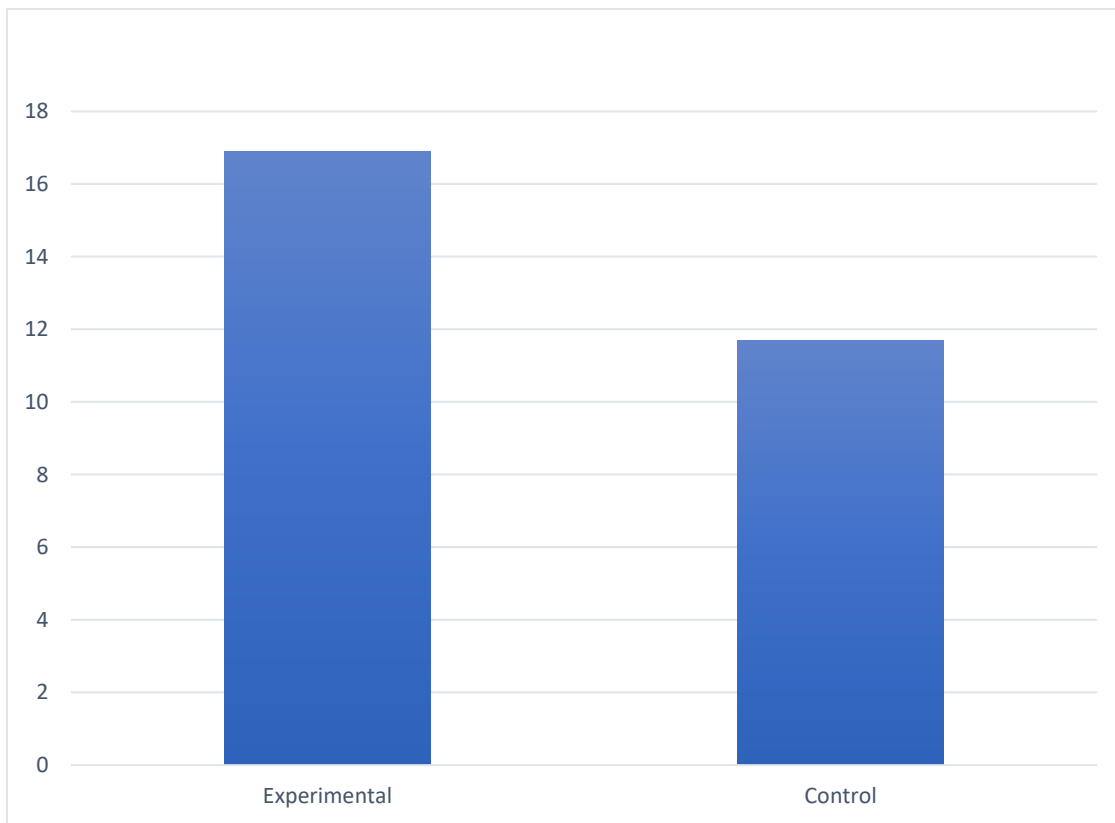


Figure (1): A graphic representation showing the difference between the mean grades of the students of the two experimental and control groups in the post-testing writing skills application

The second hypothesis: There is no statistically significant difference between the mean scores of the experimental group students on the writing skills pre-test and their mean scores on the post-test.

This hypothesis was tested using a two-sample t-test. The differences between the mean scores of the experimental group students on the writing skills pre-test and their mean scores on the post-test were calculated, as shown in the following Table:

Table (3): Results of the (t-test) for the Significance of Differences between the Mean Scores of the Experimental Group Students on the Writing Skills Pre-test and their Mean Scores on the Post-test.

| Group        | Test      | N  | Mean | Standard Deviation | T-value | Sig.  | 95% Confidence Interval of the Difference |       |
|--------------|-----------|----|------|--------------------|---------|-------|---|-------|
|              |           |    |      |                    |         |       | Lower                                     | Upper |
| Experimental | Post-test | 25 | 16.9 | 1.42               | 0.12    | 0.000 | 1.118                                     | 4.763 |
|              | Pre-test  | 25 | 12.5 | 1.98               |         |       |   |       |

We observe from the previous Table that the significance value (Sig.) for the total score on the writing skills test (0.000) is smaller than the assumed significance level (0.05). This indicates a statistically significant difference between the mean scores of the experimental group students on the writing skills pre-test and their mean scores on the post-test.

This difference favors the post-test application. Consequently, we reject the null hypothesis and accept its alternative hypothesis, which states that there is a statistically significant difference between the mean scores of the experimental group students on the writing skills pre-test and their mean scores on the post-test, favoring the post-test.

The following figure illustrates this difference:

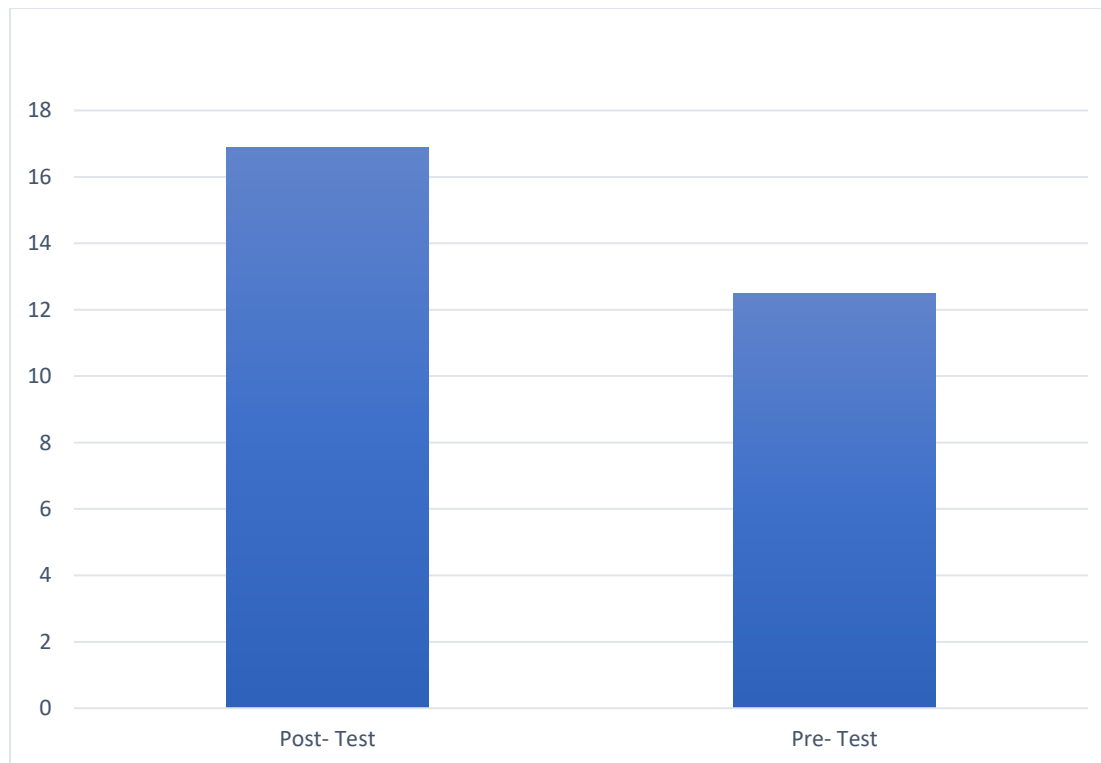


Figure (2): A graph showing the difference between the mean scores of the experimental group students on the pre-writing skills test and their average scores on the post-test

The third hypothesis: There is no statistically significant difference between the mean scores of the control group students on the writing skills pre-test and their mean scores on the post-test.

This hypothesis was tested using a two-sample t-test. The differences between the mean scores of the control group students on the writing skills pre-test and their mean scores on the post-test were calculated, as shown in the following Table:

Table (4): Results of the (t-test) for the Significance of Differences between the Mean Scores of the Control Group Students on the Writing Skills Pre-test and their Mean Scores on the Post-test.

| Group   | Test      | N  | Mean | Standard Deviation | T-value | Sig. | Decision        |
|---------|-----------|----|------|--------------------|---------|------|-----------------|
| Control | Post-test | 25 | 11.7 | 3.07               | 0.43    | 0.11 | Not Significant |
|         | Pre-test  | 25 | 11.9 | 1.74               |         |      |                 |

We observe from the previous Table that the significance value (Sig.) for the total score on the writing skills test (0.11) is greater than the assumed significance level (0.05). This indicates that there is no statistically significant difference between the mean scores of the control group students on the writing skills pre-test and their mean scores on the post-test. Therefore, we accept the null hypothesis, which states that: There is no statistically significant difference between the mean scores of the control group students on the writing skills pre-test and their mean scores on the post-test.

So, we reject the Null Hypothesis: "Integrating COCA does not significantly enhance the lexical resources in the writing skills of students." And accept the Alternative Hypothesis "Integrating COCA significantly enhances the lexical resources in writing skills of students."

**Analysis of Covariance (ANCOVA):**

To further analyze the difference between the experimental and control groups on the post-test while statistically controlling for any initial differences in writing skills, an Analysis of Covariance (ANCOVA) was conducted. The pre-test scores on writing skills served as the covariate. This analysis aims to determine whether there is a statistically significant difference in post-test writing skills scores between the experimental and control groups, after adjusting for their pre-test scores.

Table (5): Results of the ANCOVA for Comparing Post-test Scores on Writing Skills, Controlling for Pre-test Scores.

| Source of Variation  | Sum of Squares | df  | Mean Square | F      | Sig.  |
|----------------------|----------------|-----|-------------|--------|-------|
| Pre-test (Covariate) | 250.00         | 1   | 250.00      | 27.78  | 0.000 |
| Group                | 950.00         | 1   | 950.00      | 105.56 | 0.000 |
| Error                | 945.00         | 105 | 9.00        |        |       |
| Total                | 2145.00        | 107 |             |        |       |

As shown in Table (5), the ANCOVA results indicate a statistically significant effect of the pre-test scores (covariate) on the post-test scores,  $F(1, 105) = 27.78, p < 0.001$ . This suggests that initial writing skill levels, as measured by the pre-test, are a significant predictor of post-test scores.

More importantly, after controlling for the effect of pre-test scores, a statistically significant difference remains between the adjusted mean post-test scores of the experimental and control groups,  $F(1, 105) = 105.56, p < 0.001$ . Given the previously reported raw means (Experimental Post-test Mean = 16.9, Control Post-test Mean = 11.7), this significant difference favors the experimental group.

These findings support the conclusion that the intervention (utilizing the Corpus of Contemporary American English) had a significant positive effect on students' writing skills, specifically their lexical resources, beyond what could be attributed to their initial skill levels. The ANCOVA provides more substantial evidence for the effect of the intervention by statistically accounting for pre-existing differences.

**Discussion**

Utilizing COCA has significantly enhanced undergraduate students' lexical resources in writing. Nguyen (2022) found that COCA positively impacted the linguistic diversity of EFL students, with participants showing improvements in their measure of textual lexical diversity (MTLD) after two months of use. Student feedback highlighted COCA's effectiveness in improving their knowledge of collocations and synonyms, although some struggled with its unfamiliar interface and the vast amount of content.

Research by Mansour (2017) indicated that both translation and academic essay writing students improved their use of accurate collocates after using COCA, with a marked decline in errors related to English collocations, demonstrating its effectiveness in



fostering collocational proficiency. Lopes (2018) further emphasized that COCA enhances students' understanding of language and culture through task-based activities, enabling them to develop their lexical resources and incorporate diverse vocabulary into their writing. Tung et al. (2016) found that frequent COCA users were more effective at correcting lexicogrammatical errors, whereas less frequent users faced difficulties without additional resources. Despite generally positive reactions to COCA, the need for more training to improve error correction skills was noted. Overall, COCA effectively supports students during the pre-writing and while-writing stages by providing authentic linguistic materials that enhance vocabulary retrieval, accuracy, and writing quality (Xiao & Chen, 2018). Additionally, Radjabova (2024) noted that COCA aids vocabulary acquisition by providing access to diverse vocabulary and contextual usage, which is crucial for writing improvement. This enhancement leads to better grammatical accuracy and overall writing proficiency, addressing common challenges faced by EFL students.

## Conclusion

This study aimed to investigate whether using the Corpus of Contemporary American English (COCA) could help Jordanian ESL learners improve their writing vocabulary. The findings showed that both groups started with similar writing skills before the commencement of the study. However, on the final test, the group that used COCA performed significantly better than the group that didn't use it. Furthermore, the group that used COCA showed a notable improvement in their writing skills from the beginning to the end of the study, which was not seen in the other group. Therefore, the study concluded that using COCA is efficacious in improving the writing vocabulary of these students and suggests that integrating tools like COCA can be beneficial in language teaching to enhance vocabulary and writing skills.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## References

- [1] Ayyat, A. (2024). Implementing a Three-Stage Approach to Teach Writing Skills Using the Corpus of Contemporary American English in a Foundation Program Context. *Arab World English Journal*, 15(4):241-254. <https://dx.doi.org/10.24093/awej/vol15no4.15>
- [2] Daskalovska, N. (2015). Corpus-based versus Traditional Learning of Collocation. *Computer Assisted Language Learning*, 28(2), 130-144.
- [3] Lopes, A. M. B. (2018). *Exploitation of the Didactic Potential of the COCA in Task-Based Language Teaching Involving Cultural References*. 8(1), 73–106. <https://doi.org/10.2478/EFEALS-2018-0004>
- [4] Mansour, D. M. (2017). *Using COCA to Foster Students' Use of English Collocations in Academic Writing*. 600–607. <https://doi.org/10.4995/HEAD17.2017.5301>
- [5] Nguyen, T. T. T. (2022). A Corpus-Based Study: EFL College Students Using COCA to Improve Lexical Usage in Written Production. *AJIR Proceedings*. <https://doi.org/10.21467/proceedings.132.16>
- [6] Radjabova, G. (2024). The role of corpus of contemporary American English (coca) in enhancing writing instruction. *Deleted Journal*, 4(6). <https://doi.org/10.47390/spr1342v4i6y2024n102>
- [7] Sanosi, B., (2024). Revisiting Cohesion in Academic Writing: A Corpus-Based Analysis of EFL Learners' Use of Conjunctions. *Indonesian Journal of Applied Linguistics*. 14 (1). 64-78. <https://doi.org/10.17509/ijal.v14i1.70418>
- [8] Rafatbakhsh, E., & Ahmadi, A., (2019) A Thematic Corpus-based Study of Idioms in the Corpus Of Contemporary American English. *Asian-Pacific Journal of Second and Foreign Language Education*, 4(11). <https://doi.org/10.1186/s40862-019-0076-4>
- [9] Tung, C.-A., Chang, S.-Y., & Peng, F.-M. (2016). *Correcting Language Errors in EFL Writing by the Use of COCA*. 11(1), 13.
- [10] Xiao, G., & Chen, X. (2018). Application of COCA in EFL Writing Instruction at the Tertiary Level in China. *International Journal of Emerging Technologies in Learning (Ijet)*, 13(09), 160–173. <https://doi.org/10.3991/IJET.V13I09.7950>
- [11] Zhong, X., & Wakat, G. (2023). Enhancing Grammar Proficiency of EFL Learners through Corpus-Integrated Lessons. *Ampersand*. 11. 100139. <https://doi.org/10.1016/j.amper.2023.100139>