



Research Article

Effects of paper-based concordance instruction on students' Lexical Bundle use and argumentative writing performance

Habtamu Gashaw¹, Yoseph Mezgebu¹, Gashaw Nigusie¹

¹Department of English Language and Literature, University of Gondar, P. O. Box 196, Gondar, Ethiopia

ARTICLE INFO

Received: 29 March 2026

Revised: 11 June 2026

Accepted: 25 August 2026

Available online: 15 September 2026

KEYWORDS

Argumentative essays;

Corpus;

Genre-based approach;

Lexical bundles;

Paper-based concordance

ABSTRACT

This study aimed to investigate the effects of paper-based concordance instruction on students' lexical bundle use and overall performance in argumentative writing. The study involved all 25 third-year English majors at Gondar College of Teachers Education, selected through purposive and comprehensive intact-class sampling. The intervention lasted five weeks (three hours per week). The study employed a genre-based approach. Participants were taught 120 lexical bundles using concordance lines taken from small corpora of IELTS (International English Language Testing System) argumentative essays and LOCNESS (Louvain Corpus of Native English Essays). The researchers used a single-group pretest-posttest design with repeated measures in which students were required to write argumentative essays for three pretests and three posttests. The data were analyzed using corpus software and an IELTS Task 2 rubric, and the findings were evaluated using descriptive statistics and the non-parametric Chi-square and Friedman tests. Findings indicated that the intervention was associated with an increase in both the frequency (from 28 to 244 tokens) and structural complexity of lexical bundles (with the emergence of four- and five-word bundles). The study also showed a statistically significant improvement in the participants' ability to write argumentatively ($\chi^2(5) = 112.74$, $p < .001$, Kendall's $W = 0.90$) across IELTS criteria. Finally, these findings imply that paper-based concordance instruction is a highly practical method for schools without computer access; therefore, it is recommended that teacher training colleges design training programs for pre-service and in-service teachers to improve learners' writing performance.

1. Introduction

1.1. Background of the study

Writing is a key component of academic discourse. In addition to being a means of communication, writing fosters intellectual development (Nunan, 2015). In higher educational institutions, the quality of writing in projects and daily assignments is a key factor in determining students' grades (Coffin et al., 2005; Reppen, 2010). This appears to be the case in Ethiopian higher educational institutions, according to the current researchers' experience. Building

* Corresponding author. habtamugashaw46@gmail.com (Gashaw H.)

on this concept, Yelay (2017) stated that strong writing is a hallmark of a high-quality graduate in Ethiopian higher education.

Regarding text genre, an essential form for undergraduate writers, particularly in the humanities, arts, and social sciences, is the argumentative essay (Mei, 2006; Wingate, 2012). This genre holds great importance for EFL teacher trainees because teaching argumentative writing is crucial for enhancing their own language proficiency as well as for enabling them to be better practitioners when teaching their future students (Graham & Perin, 2007).

Argumentative writing is thought to be demanding genre because it needs not only structural knowledge of writing, but also an understanding of the issue, the ability to develop thesis statements, editing skills, persuasiveness, and the capacity to put ideas into writing (Shin, 2018). To enhance this skill, the integration of technology, particularly corpus (plural *corpora*) tools, is assumed to bring about the compositions needed for it. Research by Weisser (2016) and Vyatkina and Boulton (2017) suggest that corpus linguistics can help writers analyze language patterns and improve their structural and rhetorical skills in argumentation, especially in the EFL classroom.

The fundamental concept of corpus linguistics is that meaning is frequently constructed through a series of words, or formulaic language (Sinclair, 1991). Among these forms of formulaic language, lexical bundles are frequent multi-word sequences that combine grammar and vocabulary. They are three or more word combinations, such as “on the other hand” or “it can be concluded that,” which writers use frequently in academic texts. They help students produce more fluent and natural argumentative writing (Biber et al., 1999; Wright, 2019) and their frequent use is characteristic of proficient writers (Amare, 2022). However, Chen and Baker (2010) found that second language writers primarily use lexical bundles that are structurally and functionally comparable to their first languages.

In addition to lexical bundle frequency, the proportional use of bundles of different lengths determines one’s writing level. Many studies examining lexical bundles focus on three- to five-word bundles (Chen & Baker, 2010; Cortes, 2004). However, according to these scholars, four-word bundles (e.g., some people think that, on the other hand) are preferred over shorter ones. Smaller bundles like three-word bundles (e.g. In my opinion, to sum up) are often part of four-word bundles. In addition, longer bundles, such as five-word bundles (e.g., in summary the evidence shows, it can be concluded that) are less common (Allan, 2016).

Therefore, concordance instruction, in TEFL context, refers to teaching students writing skills by showing them repeated patterns of words or phrases, including lexical bundles, taken from authentic texts. Concordance instruction and lexical bundles are, thus, assumed to support teaching argumentative writing. For effective instruction, numerous studies have suggested that learners should be taught lexical bundles explicitly (e.g., Cortes, 2004; Hirvela, 2004; Biber & Barbieri, 2007). According to these scholars, teaching bundles explicitly leads to their acquisition and improves learners’ writing ability.

Although global research highlights the benefits of lexical bundles in enhancing student writing, the extent to which these findings interpret into effective classroom practice depends on instructional resources. This connection is especially significant in the Ethiopian higher education context, where writing instruction often struggles to meet academic expectations.

1.2. Statement of the Problem

In Ethiopia, students face considerable difficulties with English writing. Although writing lessons are taught as an independent skill in textbooks at all levels of education, students who complete preparatory schools and enroll in university often perform below expectations, based on the current researchers’ experience supported by local empirical evidence. Teshome’s (2007) study, conducted at a teachers’ training college, showed that the writing instruction was not effective. The researcher attributes the problem to the teaching method. In addition, the teaching materials were inappropriate because they did not encourage learners to use the language. Supporting this view, Aklilu (2015) in his study affirmed that students in higher education have writing skills problem, and concluded that the problem is resulted from teaching materials that lack authentic academic writing vocabulary.

As part of a preliminary study, the present researchers also examined the writing activities in the recently developed EFL writing courses in an Ethiopian Teachers' Training College and concluded that they lacked authenticity. Thus, it is likely that the writing difficulties stemmed from the teaching material because "much of the language teaching in Ethiopia relies heavily on textbooks" (Ebabu, 2019, p. 5). These findings suggest that the writing performance of students at Ethiopian teacher training colleges might be affected by the absence of lexical bundles (kinds of authentic language) in writing instructional materials. Consequently, the current researchers deemed it essential to support L2 teacher trainees in acquiring frequent, genre-specific lexical bundles through a paper-based concordance approach. Corpus indexing software was used to discover these genre patterns, including lexical bundles. Concordancers, such as AntConc, are software tools that locate, identify, and count every occurrence of potential lexical bundles (Cortes, 2004; Johnston, 2017).

The identified lexical bundles can be presented to the students either directly, where learners access the chunks and their context from the corpus/computer, or indirectly (paper-based concordance instruction), where teachers prepare printed material from corpus data. However, the current researchers chose to implement paper-based concordance instruction for two reasons: First, paper-based materials are preferable when computer rooms are too small or non-existent (Boulton, 2009). In many cases, these infrastructure problems are common in Ethiopia. Second, especially for students with no prior experience using corpora, paper-based (indirect) tasks are more efficient (Boulton, 2009). Corpus-based instruction is unfamiliar to students at the Ethiopian teacher training college where the current study was carried out, based on both the available literature and the researchers' own knowledge.

Regarding the instructional method, teaching lexical bundles explicitly through a genre-based approach can enhance EFL learners' writing abilities. Solomon (2019) and Rashtchi (2019) defined a genre-based approach as the explicit instruction of genre conventions for EFL/ESL student writers. However, despite the benefits associated with using genres in language instruction, some researchers have pointed out potential drawbacks. For example, Freedman (1993) argued that explicitly teaching genres to students can be unnecessary or could even have a detrimental effect on their performance if the teachers are unfamiliar with the text type. Other scholars, such as Merris (2019), question the assumption that knowledge is transferred in genre-based training, arguing that automaticity in learning transfer might not happen quickly.

Moreover, acquiring frequent bundles alone is insufficient for students to fully generalize bundle use in their writing. Although L2 learners use common lexical bundles and other formulaic sequences, this quantitative gain is not matched by a qualitative counterpart (Paquot & Granger, 2012). This idea aligns with the findings of Adel and Erman (2012), who asserted that L2 students employ restricted and less varied bundles and may overuse or underuse certain types. Consequently, one may infer that using lexical bundles alone is not a guarantee of effective writing. Therefore, further exploration of the students' mastery of lexical bundles and their effective application in writing was needed.

Regarding previous research supporting the current study, many scholars have acknowledged the integration of corpora into language instruction. Nam (2010) investigated how corpus-based instruction on selected words fostered EFL learners' use of lexical bundles in writing. The findings revealed that there were variations in writing performance in favor of the experimental group. Ackerley (2017) conducted a comparative study on the genre of opinion survey reports by using both indirect (paper-based) and direct (hands-on computer exploration) methods with a reference corpus. The study indicated that students used more phrases from the indirect method than from the direct one.

According to the literature available, the only study that focused on corpus-based writing instruction in the EFL context in Ethiopia to date was conducted by Amare (2022). He investigated whether corpus-based instruction had an effect on EFL students' academic writing skills. He prepared a specialized corpus by compiling 120 articles taken from eight different mechanical engineering journals. The findings revealed that the intervention was an effective approach for enhancing students' academic writing skills within this disciplinary context.

However, regardless of differences in objectives, setting, course type delivered, population, and corpora sources, Amare's (2022) study, like other previous research, did not detect the distribution and frequency of lexical bundles. Nevertheless, researchers like Flowerdew (2012) recommend that studies in L2 or EFL should make use of learner corpora to deepen our understanding of how learners develop linguistically.

In addition, according to the current researchers' review of the literature, no corpus-focused research has been conducted in teacher training college settings. Studies show that teachers have a tendency to implement instructional strategies that have been modeled for them by trainers or teacher educators (Supovitz, Mayer, & Kahle, 2000; Faten, 2019).

Therefore, considering the scope and focus limitations of previous research, and recognizing the benefits of building a strong writing foundation for pre-service teachers, this study addressed the effect of paper-based concordance instruction on students' lexical bundle use and argumentative writing performance.

1.3. Research Questions

The following research questions guided this study:

1. To what extent does paper-based concordance instruction influence the distribution of lexical bundles in students' argumentative essays?
2. To what extent does paper-based concordance instruction influence students' argumentative writing performance?

1.4. Significance of the Study

The findings of the study are assumed to shed light on several points: First, they will aid teachers and prospective teachers to explore technology-based novel approaches for enhancing writing performance. Second, this work can provide new directions for EFL specialists and curriculum developers to integrate corpus-based activities into courses and programs. Finally, it can inspire other researchers interested in investigating the effects of corpus-based instruction on writing performance.

2. Methods

2.1. Research Design

This study employed a single-group pretest-posttest design with repeated measures, which was a type of quasi-experimental design. A single group design was chosen because the study population consisted of a single intact class. A time series design was used to strengthen internal validity by establishing a stable pre-intervention baseline. A quasi-experimental design is appropriate for examining how independent variables influence dependent variables over time (Cohen et al., 2011). The independent variable was paper-based concordance instruction, which was assumed to have a direct impact on the dependent variables: lexical bundle use and argumentative writing performance. The study employed a quantitative data analysis method, guided by the positivist paradigm.

2.2. Participants and Sampling Techniques

The participants of this study were all 25 third-year English major students enrolled in a degree program at Gondar College of Teachers' Education, Ethiopia. The study site was selected purposively because this teachers' training college is closely affiliated with the University of Gondar, where the researchers work, making it highly accessible for the study. To select the participants, the researchers used comprehensive intact-class sampling. This means the entire class was selected because they were the only group taking the 'Advanced Writing Skills' course during the first semester of the 2025 academic year. Using this entire existing class as a single group allowed the researchers to keep the classroom environment natural, which fits the requirements of a quasi-experimental design. The Advanced Writing course was offered for three hours per week.

The course and the third-year students were ideal because, according to the preliminary study of the current researchers, the participants had not learned the concept of lexical bundles and corpus linguistics explicitly. Moreover, the teacher would not need to pay extra effort to teach the basics of writing skills, as the students had a semester of paragraph writing experience in the Basic Writing course previously.

2.3. Data Collection Instruments

For the pretests and posttests, six different argumentative prompts (one for each phase) were administered. These prompts were selected based on students' familiarity with the topics, argumentative nature, linguistic complexity, and engaging quality, as reviewed by two EFL instructors.

Regarding the writing tests, they were primarily used to compile six learner corpora (three pretest corpora and three posttest corpora) and compare the students' use of lexical bundles in terms of frequency and distribution patterns (focusing on three- to five-word bundles). The software best suited to this purpose was AntConc 4.3.1 (Anthony, 2024). This freely available and comprehensive software for concordancing and text analysis helped the current researchers to answer research question one.

Pretest and posttest writing were also used to measure students' progress in writing ability after paper-based lessons. The writing progress was evaluated in terms of how fully the writer addressed their thoughts (task achievement), managed paragraphs (coherence and cohesion), kept the fluency of the writing (coherence and cohesion), and made error-free structures (grammatical range and accuracy). This methodological choice was preferred, based on previous research. Studies indicate that corpus-based instruction can increase language accuracy and fluency (Liu & Jiang, 2009), help in improving fluency (Boulton & Cobb, 2017), bring better proficiency ratings (Boers *et al.*, 2006), combine grammar and lexis (Ucar, 2017), and illustrate the lexico-grammatical patterns that the academic community recognizes (Hyland, 2012). To make the data gathering clear, every essay prompt was carefully checked to match the students' language level and background. The rubric that best contains criteria relevant to the above anticipated achievements in the participants' writing is IELTS Task 2 Writing criteria which include task response, coherence and cohesion, lexical resource, and grammatical range and accuracy. These four specific criteria were strictly followed step-by-step during the evaluation process to score the writing items accurately. The researchers analyzed the students' writing according to this rubric, and the total score of each student's writing was evaluated out of nine (9 = highest, 1 = lowest). All the writing samples were independently marked by two experienced TEFL teachers. These two markers were university EFL lecturers who received orientation on the rubric rules to grade all the essays blindly without knowing the students' names.

2.4. Procedure of the study

The study procedures were designed to deliver trustworthy and valid data that addressed the research questions. The following steps summarize these procedures:

Before the intervention, the researchers administered three writing tests to establish baseline writing skills. These tests were used to determine changes in lexical bundle distribution and writing performance after the intervention.

Following this, participants were taught an advanced writing skills course with materials and tasks designed through a paper-based concordance approach. The instruction was delivered by a departmental instructor assigned. The instruction relied on printed materials prepared indirectly from special corpora compiled from argumentative essays in the Louvain Corpus of Native English Essays (LOCNESS) and IELTS model essays. In these corpora, 190 essays (95 from each source) were selected by two experienced EFL teachers. This sample size of 190 was chosen because scholars (e.g., Flowerdew, 2015) recommend that the size of corpora for academic texts can range from as few as 10 to hundreds of texts, and because of the belief that 'the larger the better' (Sinclair 1991).

The researchers used the AntConc 4.3.1 (Anthony, 2024) corpus tool to generate the lexical bundles. This concordance software provides useful resources for examining word lists, collocations, concordances, and key words. Moreover, it offers free access to pre-existing corpora and allows users to compile their own. These features were needed to conduct the current study.

Therefore, the researchers then compiled and analyzed the selected essays using the specified corpus tool, identified lexical bundles, and designed the paper-based instruction as previous researchers suggested (e.g., Bennett, 2010) and scholars in the field recommended (Flowerdew, 2005).

This research included three-to-five word bundles as these are considered the most manageable size for manual categorization and analysis (Lee, 2020). The study also used Cortes' (2013) frequency threshold of at least 40 occurrences per million words for the bundles. This threshold was applied only to the reference corpus to select target bundles for instruction. Regarding text distribution, Huang (2015) emphasized that lexical bundles should

occur in no fewer than five different texts to avoid individual writer repetition. Accordingly, the current researchers adopted five as a dispersion threshold.

Based on this, Participants learned 120 lexical bundles in authentic contexts through concordance lines drawn from small corpora of IELTS argumentative essays and LOCNESS (Louvain Corpus of Native English Essays). They were practiced through the genre-based approach which followed the five stages proposed by Hyland (2007): Setting the context, Modeling, Joint construction, Independent construction, and Comparing. To clearly explain what happened during the five-week intervention period, these five stages were carried out step-by-step as follows:

First, in the introduction or setting the context stage, the teacher explained the characteristics of an argumentative essay, after completing brainstorming tasks. Second, in the modeling stage, the teacher's role was providing argumentative essays for students in order to deconstruct them for rhetorical patterns (particularly lexical bundles) and the moves (parts of the essays) using the authentic concordance lines. Third, during the collaborative (joint construction) stage, learners discussed the common lexical bundles used in different parts of argumentative essay; the teacher introduced the concordance lines at different stages in the form of gap filling activities, and helped students use lexical bundles with examples. Fourth, at the independent writing stage, students wrote argumentative essays using appropriate lexical bundles keeping the model moves. Finally, fifth, at the linking related tasks or comparing stage, students compared their writing with their friends' to evaluate their text choices.

Each part of an argumentative essay (introduction, body, and conclusion) was thoroughly discussed and supported with expressions, lexical bundles, and examples taken from the concordance lines. The intervention lasted five weeks (three hours per week).

Regarding the specific testing activities, they were administered one-by-one as follows: Before the intervention, the researchers administered three writing tests (Pretest 1, Pretest 2, and Pretest 3) to establish baseline writing skills. These tests were used to determine changes in lexical bundle distribution and writing performance after the intervention. After the intervention, three posttests (Posttest 1, Posttest 2, and Posttest 3) were administered one-by-one, analyzed and compared with the pretest scores.

2.5. Methods of Data Analysis

The research questions were addressed through analysis of argumentative writing test scores. This study initially adopted a corpus-based analytical approach to answer the first research question. Using the corpus tool AntConc, the researchers compiled six learner corpora consisting of essays written by the research participants. Lexical bundles were then identified based on the criteria established for them (as seen in the procedure section). Following this identification, variations in the proportional use of different bundle lengths before and after the treatment were computed quantitatively using descriptive statistics, specifically frequency counts and percentages. Next, the researchers used a Pearson Chi-Square Test to determine if the distribution of lexical bundle types changed significantly across the six testing phases (three pretests and three posttests).

After analyzing the lexical bundles in the students' essays quantitatively, the researchers examined the students' argumentative essays quantitatively in a way that answered the second research question. To do this, four marking criteria: Task Achievement, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy with nine descriptors were adopted from the [British Council \(2020a\)](#). The overall score for each essay was calculated as the mean of these four criteria.

Since the data did not meet the normality assumptions required for a repeated measure ANOVA, the non-parametric Friedman test was employed to analyze the test results. Data analysis was performed using SPSS, version 27. Prior to the main implementation, a pilot study was conducted with a different batch of third-year English majors from the previous academic year; this ensured that the instruments were tested on a population with the exact same language proficiency level without contaminating the main study sample. Two experienced EFL instructors verified the tests for face and content validity. The researchers also tested the Scoring Reliability (Inter-rater Reliability) to ensure consistent scoring. The resulting coefficient (Pearson's $r = 0.81$) indicated that the raters were consistent.

3. Results and Discussion

3.1.RQ1: Analysis of the distribution of lexical bundles in students' Argumentative writing

The distribution of lexical bundles was examined along two dimensions: frequency of usage and structural type (i.e., 3-word, 4-word, and 5-word), comparing results of pre- and post- intervention. Lexical bundle tokens were aggregated for all participants. The following table presents the resulting frequencies.

Table 1. The Frequency of Lexical Bundle Distribution in Pretests and Posttests.

Phase	3-Word LBs		4-Word LBs		5-Word LBs		Row total
	Frequency	%	Frequency	%	Frequency	%	
Pretest 1	24	85.71	4	14.29	0	0	28
Pretest 2	23	76.67	7	23.33	0	0	30
Pretest 3	29	78.38	8	21.62	0	0	37
Posttest 1	112	52.83	64	30.19	36	16.98	212
Posttest 2	121	51.49	71	30.21	43	18.30	235
Posttest 3	124	50.82	75	30.74	45	17.67	244
Column total	433	55.09	229	29.13	124	15.78	786

As shown in Table1, before the intervention, students used very few lexical bundles in their writing. In the first pretest, there were only 28 bundle tokens, and almost all (85.71%) of these were three-word structures (e.g., I think that, in my opinion), while the remaining 14.29% were four-word bundles (e.g., on the other hand). These essays contained no 5-word bundles. The result of the second pretest also demonstrated that there were a similar number of bundles (30 tokens), with 3-word bundles dominating (23 tokens, 76.67%) and seven 4-word bundles (7 tokens, 23.33%). The result of the third pretest also showed a comparable number of bundles (37 tokens), with 3-word bundles again dominating (29 tokens, 78.38%) and eight 4-word bundles (8 tokens, 21.62%).

The consistency of using a similar number of bundles and similar patterns (i.e., length) across pretests indicated that there was a constant pattern of lexical bundles before the intervention. This pattern aligns with previous research (e.g., Schmitt, 2010), which notes that using a limited number of lexical bundles and relying on short bundles is common, especially among learners who have not been taught about these patterns. The pretest results thus confirmed that the participants lacked such exposure.

Following the paper-based concordance instruction, a great number of lexical bundles was observed in the posttests. In posttest 1, students used 212 bundles, which was an increase compared to Pretest 1. The improvement included not only frequency but also the length of bundles used. In addition to more three-word bundles (112 tokens, 52.83%), a great number of four-word (64 tokens, 30.19%) and five-word (36 tokens, 16.98 %) bundles also emerged.

The improvement continued into posttest 2, where the frequency of bundles rose to 235 tokens. Regarding the distribution, the three-word bundles remained dominant (121 tokens, 51.49%), while four-word bundles showed proportional improvement (71 tokens, 30.21%). Likewise, the five-word bundles (43 tokens, 18.30%) showed a slight increase from posttest 1. The improvement was sustained in posttest 3, where the frequency of bundles reached 244 tokens. Three-word bundles accounted for 124 tokens (50.82%), four-word bundles 75 tokens (30.74%), and five-word bundles 45 tokens (17.67%), maintaining the pattern of increased bundle length and frequency. The post-intervention data showed a clear increase in the number of lexical bundles used. This progress was evident not only in frequency but also in the structural complexity of the bundles.

To determine the statistical significance of this change, a chi-square test of independence was performed. Because expected frequencies for 5-word bundles in pretests were below 5, a Pearson chi square test of independence was performed on a 2 (testing phase: pretest vs. posttest) $\times$ 3 (bundle type: 3-word, 4 word, 5- word) contingency table. Table 2 presents the test results.

The results in Table 2 showed a strong link between the types of word bundles students used and the testing phases, $\chi^2(2, N = 786) = 32.32, p < .001$, with a moderate effect size (Cramer's $V = .203$). This means there was a clear shift from the pretests to the posttests. Before the intervention, learners relied almost entirely on short, 3-word bundles. After the paper-based concordance lessons, however, they began using longer and more complex four- and five-word bundles. The noticeable rise in four and five-word bundles shows that explicit concordance instruction helped

students acquire more complex academic writing patterns. While shorter bundles were still dominantly used, this makes sense according to Wray's (2002) principle of cognitive effort, which is interpreted as learners exert higher effort for longer bundles.

Table 2. Cross-tabulation of lexical bundle length by Test phase with Chi-Square test results.

Phase	3-word bundles	%	4-word bundles	%	5-word bundles	%	Total
Pretests	76	80.0	19	20.0	0	0.0	95
Posttests	357	51.7	210	30.4	124	17.9	691
Total	433	55.1	229	29.1	124	15.8	786
Chi-Square Tests							
Pearson χ^2 (2, N = 786) = 32.32, $p < .001$							
Cramer's V = .203							

However, in addition to studying the frequency and distribution, examining the proper use of these bundles was needed to comprehensively understand the effect of the intervention. Therefore, the second research question was computed.

3.2.RQ2: Analysis of the effect of paper-based concordance instruction on students' writing performance

The students' essays were assessed using the IELTS Task 2 rubric to investigate the effect of paper-based concordance instruction on argumentative writing. As the data violated the assumptions of repeated measure ANOVA (e.g., normality and sphericity), the non-parametric Friedman test was employed to compare pre- and post-intervention writing performance. The mean scores represent averages across the four IELTS criteria and reflect the current proficiency level in this teacher training context. The following table shows the results of descriptive statistics and the Friedman Test.

Table 3. Results of descriptive statistics and Friedman Test for writing performance.

Phase	N	Mean	SD	Min	Max	Mean Rank
Pretest 1	25	2.36	0.59	1.5	3.5	1.52
Pretest 2	25	2.5	0.6	1.5	3.5	1.94
Pretest 3	25	2.76	0.77	2	4.5	2.74
Posttest 1	25	3.32	0.78	2.5	4.5	4.4
Posttest 2	25	3.48	0.77	2.5	4.5	4.84
Posttest 3	25	3.72	0.75	2.5	5	5.56
Friedman Test: $\chi^2(5) = 112.74$, $p < .001$, Kendall's W = 0.902						

As shown in Table 3, the descriptive statistics revealed that the first, second, and third pretests showed mean scores of 2.36 (SD = 0.59), 2.5 (SD = 0.6), and 2.76 (SD = 0.77), respectively. This gradual increase across the three pretests shows that students' writing ability remained relatively stable before the intervention, with the third pretest indicating a slight upward trend. On the other hand, the means and standard deviations of posttest 1 (M = 3.32, SD = 0.78), posttest 2 (M = 3.48, SD = 0.77), and posttest 3 (M = 3.72, SD = 0.75) were higher.

The mean ranks from the Friedman test also indicated an improvement in students' writing performance on the posttests. As shown in Table 3, the mean rank for Pretest 1 was 1.52, for Pretest 2 it was 1.94, and for Pretest 3 it was 2.74, which is progressively higher. This confirms that students' performance before the intervention was lower compared to the posttest phases. Additionally, the Friedman test results indicated a statistically significant difference in students' argumentative writing performance across the testing phases, $\chi^2(5) = 112.74$, $p < .001$, with a very large effect size (Kendall's W = 0.90). The increase in lexical bundle frequency and length after the intervention shows that paper based concordance instruction had a direct effect on students' writing performance. Specifically, the rise

in bundle frequency and complexity, and the students' writing performance according to IELTS rubric scores, confirm that the intervention directly enhanced students' argumentative writing ability.

These findings are consistent with previous research suggesting that well-designed paper-based instructional materials incorporating lexical bundles contribute to the development of effective academic writing skills by supporting creativity and maintaining fluency (Nation, 2006; Coxhead & Byrd, 2007; Khazaali, 2019). Similarly, Huang (2014) and Ungwaraphong (2021) found that paper-based DDL activities helped students notice lexical collocations, which directly supported the argument, thereby improving grammatical accuracy in writing. Therefore, the changes observed in the current study most likely occurred because students employed the lexical bundles they were taught in the structured paper-based materials.

Following this significant result, post-hoc analysis using Wilcoxon signed-rank tests with a Bonferroni correction (adjusted $\alpha = 0.0083$) was conducted to identify which specific testing phases differed significantly. The table below shows this result.

Table 4. Results of Post-Hoc Pairwise Comparisons using Wilcoxon Signed-Rank Test with Bonferroni Corrections.

Comparison	+Ranks (N)	Ties (N)	Z	p-value	Adj. p-value	Effect Size (r)
Pre2 – Pre1	7	18	-2.65	0.008	0.12	0.53
Pre3 – Pre1	17	8	-3.88	< .001	< .001	0.78
Post1 – Pre1	25	0	-4.44	< .001	< .001	0.89
Post2 – Pre1	25	0	-4.45	< .001	< .001	0.89
Post3 – Pre1	25	0	-4.53	< .001	< .001	0.91
Pre3 – Pre2	10	15	-2.92	0.004	0.06	0.58
Post1 – Pre2	25	0	-4.47	< .001	< .001	0.89
Post2 – Pre2	25	0	-4.50	< .001	< .001	0.9
Post3 – Pre2	25	0	-4.48	< .001	< .001	0.9
Post1 – Pre3	20	5	-4.06	< .001	< .001	0.81
Post2 – Pre3	21	4	-4.13	< .001	< .001	0.83
Post3 – Pre3	24	1	-4.48	< .001	< .001	0.9
Post2 – Post1	8	17	-2.83	0.005	0.075	0.57
Post3 – Post1	17	8	-3.88	< .001	< .001	0.78
Post3 – Post2	12	13	-3.46	< .001	.015	0.69

Bonferroni- corrected significance level: adj. $p = p \times 15$. Significance is indicated at $\alpha = .05$: $p < .05$, $p < .01$, $p < .001$.

As shown in Table 4, the Pairwise comparisons did not indicate any significant difference between Pretest 2 and Pretest 1 (Adj. $p = 0.12$) or between Pretest 3 and Pretest 2 (Adj. $p = 0.06$), but a significant difference was found between Pretest 3 and Pretest 1 (Adj. $p < .001$). Among the posttests, no significant difference was found between Posttest 2 and Posttest 1 (Adj. $p = 0.075$) or between Posttest 3 and Posttest 2 (Adj. $p = 0.015$), whereas Posttest 3 was significantly higher than Posttest 1 (Adj. $p < .001$). However, the scores of all posttests were significantly higher compared with all pretests (Adj. $p < .001$ for each cross-phase comparison). The highest score was particularly observed in posttest 3 against pretest 1 ($Z = -4.532$, $p < .001$). This demonstrates that paper-based concordance instruction had a direct and sustained positive effect on students' argumentative writing performance.

Although the findings indicated that paper-based concordance instruction brought about significant changes in the participants' lexical bundle use and argumentative writing performance, the study has limitations. First, the intervention was conducted with only 25 EFL teacher trainees from a single college. This limits the generalizability of the findings to other populations and institutional contexts. Second, because paper-based concordance instruction is a form of indirect data-driven learning, some studies have identified this as a limitation. Additionally, the paper-based material was designed using IELTS and LOCNESS essays, and the assessment used IELTS rubric. This may help the good results, but it is not generalizable to other writing tasks.

Therefore, future research should address the gaps identified by including more samples and institutions, and comparing paper-based (indirect data-driven learning) with the direct corpus approach over extended time.

4. Conclusion

This study investigated the impact of a corpus-informed, paper-based instruction on students' lexical bundle use and argumentative writing performance.

The findings indicated that paper-based concordance was an effective instructional approach, associated with enhancements in both the frequency and structural complexity of lexical bundles and the quality of students' writing. Specifically, the intervention was associated with a substantial increase in lexical bundle frequency from the pretest to the posttest phases, and a chi-square test confirmed that this change in the distribution of bundle lengths was statistically significant. More importantly, students moved from reliance on three-word bundles to more diversified and complex bundles, such as 4-and 5-word bundles.

Furthermore, the study provided evidence that paper-based concordance materials can serve as an effective pedagogical tool that contributes to improving students' argumentative writing performance. Analysis using the IELTS Task 2 rubric revealed that a statistically significant improvement in writing scores from the pretest to the posttest phases, with post-hoc comparisons further showing that all posttest scores were significantly higher than all pretest scores. In conclusion, despite its limitations (a single-group intact class pretest-posttest design without a control group, small sample size, and brief intervention period), paper-based concordance instruction was associated with significant improvements in students' argumentative writing performance.

5. Recommendation

Given the positive findings associated with paper-based concordance instruction, it is recommended that EFL teachers, particularly those with limited computer or internet access, consider incorporating this approach. Teacher training programs should introduce both in-service and pre-service teachers to paper-based concordance methods. This approach helps students learn lexis and syntax together while understanding the rhetorical structure of the target genre. As a result, learners gain exposure to authentic language, which can enhance their writing performance. Additionally, teacher training curriculum developers are encouraged to support technology-based language teaching by revising the writing curricula to integrate authentic academic inputs through corpus-based activities. This would equip trainees with the professional skills necessary for more effective writing instruction.

Finally, to address the limitations of this study, future research should employ larger sample sizes and longer intervention periods. Future studies should also compare the relative effectiveness of paper-based and direct corpus approaches.

Acknowledgements

The researchers would like to thank the University of Gondar for funding this research. They would also like to thank Gondar College of Teachers Education for permission to conduct this study at the institution.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

The first author acknowledges the University of Gondar for providing financial support.

Ethical consent

To keep the required ethical standards of this study, the current researchers first obtained ethical clearance from responsible authorities at Gondar College of Teachers Education. They also established a clear written agreement (using a consent form) with all participants before they participated in the study. The researchers respected the participants' right to withdraw from the study. The researchers also guaranteed the confidentiality and privacy of participants by removing their names and identifying details from all data collection forms. Instead, numerical codes were used to refer to each participant in the study.

References

- Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology*, 21(3), 195-216. <http://hdl.handle.net/10125/44624>

- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31(2), 81–92. <https://doi.org/10.1016/j.esp.2011.08.004>
- Aklilu, G. (2015). An investigation into the problems of teaching and learning writing skills. *Ethiopian International Journal of Multidisciplinary Research*, 2(3), 33–46.
- Al-Khazaali, H. (2019). The impact of lexical chunks instruction on developing Iraqi students' writing fluency. *Advances in Language and Literary Studies*, 10(5), 70–78. <https://doi.org/10.7575/aiac.all.v.10n.5p.70>
- Allan, R. (2016). Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning? *System*, 59, 61–72. <https://doi.org/10.1016/j.system.2016.04.005>
- Amare, T. (2022). *Effects of corpus-based instruction on EFL students' academic writing skills, critical thinking skills, engagement, and perception* [Unpublished doctoral dissertation]. Bahir Dar University.
- Anthony, L. (2024). *AntConc* (Version 4.3.1) [Computer software]. Waseda University. <http://www.laurenceanthony.net/software>
- Bennett, G. R. (2010). *Using corpora in the language learning classroom*. University of Michigan Press. <https://doi.org/10.3998/mpub.371534>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman <https://doi.org/10.1162/089120101300346831>
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26 (3), 263–286. <https://doi.org/10.1016/j.esp.2006.08.003>
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10(3), 245–261. <https://doi.org/10.1191/1362168806lr195oa>
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37–54. <https://doi.org/10.1017/S0958344009000068>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- British Council. (2020a). *IELTS Task 2 writing band descriptors*. <https://www.ielts.org/for-teachers/descriptives>
- Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30–49. <https://doi.org/10.125/44213>
- Coffin, C., Curry, M. J., Goodman, S., Hewings, A., Lillis, T., & Swann, J. (2005). *Teaching academic writing: A toolkit for higher education*. Routledge. <https://doi.org/10.4324/9780203994890>
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Routledge. <https://doi.org/10.4324/9780203720967>
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397–423. <https://doi.org/10.1016/j.esp.2003.12.001>
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, 12(1), 33–43. <https://doi.org/10.1016/j.jeap.2012.11.002>
- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16(3), 129–147. <https://doi.org/10.1016/j.jslw.2007.07.002>
- Ebabu, T. (2019). Theories vis-à-vis practices in English as a foreign language writing classes in the Ethiopian context. *PASAA*, 58, 183–212. <https://www.culi.chula.ac.th/en/pasaa/2>
- Faten, A. (2019). The impact of corpus on EFL pre-service teachers' self-directed learning and oral proficiency. *Language Teaching Research Quarterly*, 13, 85–105. <https://www.eurokd.com/journal/jd/1>

- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24 (3), 321–332. <https://doi.org/10.1016/j.esp.2004.09.002>
- Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan. <https://doi.org/10.1057/9780230355569>
- Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, 48(1), 99–116. <https://doi.org/10.1017/S0261444813000037>
- Freedman, A. (1993). Show and tell? The role of explicit teaching in the learning of new genres. *Research in the Teaching of English*, 27(3), 222–251. <https://doi.org/10.58680/rte199315402>
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools—A report to Carnegie Corporation of New York*. Alliance for Excellent Education. https://www.carnegie.org/media/filer_public/9d/e2/9de20604-a055-42da-bc00-77da949b29d7/ccny_report_2010_writing.pdf
- Hirvela, A. (2004). *Connecting reading and writing*. University of Michigan Press. <https://doi.org/10.3998/mpub.23736>
- Huang, L. (2015). Corpus-aided language learning. *ELT Journal*, 69(4), 481–484. <https://doi.org/10.1093/elt/ccv031>
- Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, 26(2), 163–183. <https://doi.org/10.1017/S0958344014000020>
- Hyland, K. (2007). Genre pedagogy: Language, literacy, and L2 writing instruction. *Journal of Second Language Writing*, 16(3), 148–164. <https://doi.org/10.1016/j.jslw.2007.07.005>
- Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics*, 32, 150–169. <https://doi.org/10.1017/S0267190512000037>
- Johnston, K. M. (2017). *Lexical bundles in applied linguistics and literature writing: A comparison of intermediate English learners and professionals* [Master's thesis, Portland State University]. ProQuest Dissertations and Theses Global. <https://doi.org/10.15760/etd.3353>
- Lee, H.-K. (2020). Lexical bundles in linguistics textbooks. *Linguistic Research*, 37(1), 121–145. <https://doi.org/10.17250/khisli.37.1.202003.005>
- Liu, D., & Jiang, P. (2009). Using a corpus-based lexico-grammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93(1), 61–78. <https://doi.org/10.1111/j.1540-4781.2009.00828.x>
- Mei, W. S. (2006). Creating a contrastive rhetorical stance: Investigating the strategy of problematization in students' argumentation. *RELC Journal*, 37(3), 329–353. <https://doi.org/10.1177/0033688206071316>
- Merris, S. (2019). Questioning genre-based pedagogy: From text to task in the EAP classroom. *ELT Journal*, 73(4), 437–446. <https://doi.org/10.1093/elt/ccz028>
- Nam, D. (2010). *Productive vocabulary knowledge and evaluation of ESL writing in corpus-based language learning* (Publication No. 3415770) [Doctoral dissertation, University of Alabama]. ProQuest Dissertations and Theses Global. <https://www.proquest.com/docview/734948064>
- Nation, I. P. (2006). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge. <https://doi.org/10.4324/9781315740553>
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130–149. <https://doi.org/10.1017/S0267190512000098>

- Rashtchi, M. (2019). Scaffolding argumentative essay writing via reader-response approach: A case study. *Asian-Pacific Journal of Second and Foreign Language Education*, 4 (1), 1–17. <https://doi.org/10.1186/s40862-019-0078-2>
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press. <https://doi.org/10.1017/9780511609525>
- Schmitt, N. (2010). Formulaic language. In *Researching vocabulary: A vocabulary research manual* (pp. 7–15). Palgrave Macmillan. https://doi.org/10.1057/9780230293977_3
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press. <https://doi.org/10.1093/oso/9780194371445.001.0001>
- Solomon, A. (2019). *Effects of Problem-Based English Language Instruction on Students' Critical Thinking Dispositions, Writing Self-efficacy, and Argumentative Essay Writing Skills* [Unpublished doctoral dissertation]. University of Gondar.
- Supovitz, A., Mayer, P., & Kahle, B. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. *Educational Policy*, 14 (3), 331–356. <https://doi.org/10.1177/0895904800014003001>
- Teshome, T. (2007). *Exploring the effectiveness of the teaching and learning of the writing skills: Asella College of Teachers' Education in focus* [Unpublished master's thesis]. Addis Ababa University.
- Ucar, S. (2017). A corpus-based study on the use of three-word lexical bundles in the academic writing by native English and Turkish nonnative writers. *English Language Teaching*, 10(12), 28–36. <https://doi.org/10.5539/elt.v10n12p28>
- Ungwaraphong, P. (2021). Inductive teaching approaches in Business English writing in an EFL context: Paper-based and product-based instructions. *English Language Teaching Educational Journal*, 4(1), 34–48. <https://eric.ed.gov/?id=ED614608>
- Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1-8. <http://hdl.handle.net/10125/44617>
- Weisser, M. (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Wiley-Blackwell. <https://doi.org/10.1002/9781119180180>
- Wingate, U. (2012). ‘Argument!’ helping students understand what essay writing is about. *Journal of English for Academic Purposes*, 11(2), 145–154. <https://doi.org/10.1016/j.jeap.2012.03.001>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511519772>
- Wright, H. R. (2019). Lexical bundles in stand-alone literature reviews: Sections, frequencies, and functions. *English for Specific Purposes*, 54, 1–14. <https://doi.org/10.1016/j.esp.2018.10.004>
- Yelay, B. (2017). Assessment of the qualities of academic writing in senior essays of English graduates: The case of Dire Dawa University. *International Journal of English and Literature*, 8(8), 102–114. <https://doi.org/10.5897/IJEL2015.0777>