The effectiveness of corpora on Saudi EFL academic writing performance

Assrar Alharbi[1]
Turki Alsolami[2]

English Language Institute, King Abdulaziz University[1,2]
email: tgalsolami@kau.edu.sa[2]

Abstract - This study reports on quantitative research that investigated the effects of the Corpus-Based Approach (CBA) as a pedagogical approach to Academic Writing Instruction in the context of Saudi Arabia. A total of fifteen female preparatory year students at an intermediate language proficiency level at King Abdulaziz University participated in the study. The study aimed at investigating the effect of applying corpora on learners' lexico-grammatical abilities following a pretest-post-test design. Classroom intervention has been conducted and data were collected from written tasks over a period of time. The findings of the study demonstrated that the improvement of lexico-grammatical abilities among students is observed.

Keywords: academic writing, corpus linguistics, EFL, lexico-grammatical patterns, teaching and learning

1. Introduction

The role of Information and Communications Technology (ICT) in the future of Saudi English Language Teaching (ELT) and Teaching English to Speakers of Other Languages (TESOL) is stressed (Picard, 2018). ICT which is defined as the use of technology has provided students with access to information and new learning environments through international online providers. One way to introduce learners to an online linguistics environment is through the implementation of Corpus Linguistics (CL). Research suggests that CL is making remarkable contributions to the fields of language research, acquisition, and education as it provides an enormous amount of almost every-day authentic language input (Friginal, 2018). By using authentic large databases of corpora, it is possible to analyze spoken and written texts to serve different purposes. For example, in education when curriculum designers based their materials on corpus data, teachers as well learners may receive more real-life related language input. As a result, the way learners understand, learn, and teach languages has changed (Paker & Ergül-Özcan, 2017). This change is mainly seen when students will be able to both refer to corpus data and to be able to apply what they have acquired outside the classroom.

In the current study, English as a Foreign Language (EFL) educator is encouraged to develop and apply a more updated curriculum and teaching materials that include corpus data. Corpus Linguistics with its characteristics which are to be authentic, frequent, syntactically stored, and contextualized can offer EFL researchers, materials developers, curriculum designers, teachers, and learners with great potential to better language proficiency. Using corpora can provide EFL learners with the opportunity to acquire authentic language in real contexts that can be transferable to real life. Such valuable knowledge helps EFL learners to be more proficient in the four skills of language which are reading, listening, speaking, and writing. The main focus of this study is on academic writing skills, more precisely the lexico-grammatical patterns.

Recently, a plethora of research studies conducted in various Saudi EFL contexts revealed that Saudi EFL learners seem to lack the competence of performing as well-independent skilled writers (Alharbi, 2019; Obeid, 2017; Alrabai, 2016; Almutairi, 2016; Mohammad and Hazarika, 2016; Alkubaidi, 2014). Writing is considered to be a high-order skill that needs learners to think critically and be aware of the language lexical and grammatical patterns (Obeid, 2017). In the process of academic writing, even competent learners might require further support regarding complex aspects of language. This absence of support for grammatical, lexical, and other linguistic patterns can be the main source of the incompetence for EFL academic writers (Kotamjani et al, 2017). Therefore, to consider including language corpus tools which can enhance students' ability to edit and review the linguistic surface levels of their writing. Consequently, corpus data can help EFL writers to build up their confidence, foster their sense of autonomy and, increase their language awareness in writing skills by being able to check their linguistic hypotheses and going beyond their current linguistic competence and performance.

Another dilemma, as noted by Picard (2018), is that teaching references and materials often neglect linguistic elements that are to be considered frequent in the

native-like language and tend to highlight those that are relatively uncommon. As a result, learners lack the opportunity to deal with the actual use of language that they might encounter outside the classroom. Moreover, Reppen (2010) has affirmed that a key challenge in promoting language teaching materials and resources has always been to offer learners language elements that precisely reflect the way language is used in real-life settings.

Thus, a pedagogical tool to enhance learners' writing abilities by providing them with a source of knowledge to notice linguistic features autonomously based on real-life examples are believed to overlook learners' linguistic problems. Integrating a corpus-based approach to writing instruction can improve learners' academic writing skills by allowing them to search for lexical and grammatical information providing an additional resource of knowledge that can be used inside and outside the classroom which enhances their confidence, improves their learning independent skills, and, increases their language awareness.

This study seeks some answers to provide EFL researchers, material developers, curriculum designers, teachers, and learners with an overview of the potential of applying CBA to writing instruction in EFL classrooms. Within the EFL research field, the study aims to add to the literature of ELT in Saudi Arabia and the integration of technology in classrooms to teach writing to EFL learners. This research is among the first and few studies that seek to investigate both the impact of implementing CBA in Saudi EFL classrooms and the learners' perception of the employment of corpus as a pedagogical tool to improve their academic writing abilities. For EFL material developers, curriculum designers, teachers, and learners, the study enlightens them about the potential of applying CBA to enhance the process of writing, teaching, and learning. Moreover, the results can be utilized by other universities in Saudi Arabia. King Abdulaziz University is considered as a reflected image of other Saudi Arabian universities since they share similar educational visions, administrative systems, and cultural values.

The purpose of the current study is to establish CBA within Saudi EFL classrooms to serve as a pedagogical tool for writing instruction in the writing feedback process. The present study seeks an answer to the following research question namely RQ1: Does the Corpus-Based Approach (CBA) to Writing Instruction have a significant impact on learners' academic writing skills more particularly, their lexical and grammatical accuracy?

A brief introduction to Corpus Linguistics (CL) and its relation to language teaching and learning will help us, at this point, to focus on the foundation of Corpus Application inside the EFL classroom. CL as defined by Liu and Lei (2017) is "the study of language by examining and analyzing naturally- language data, often in the form of texts". Naturally-language data and real language data refer to the language used in every-day by its speakers, in other words; it is not previously prepared language data. CL provides a remarkably powerful tool for the analysis of naturally-language data and can afford enormous insights on how language usages are varied in different contexts such as spoken versus written or formal versus casual situations (Breyer, 2011). CL aims at representing a statistical descriptive interpretation of the speaker's performance. CL studies the actual patterns of language in naturally occurring manners, by relying on a large collection of texts to carry out empirical

studies. CL is based on the idea that language has to be explored as it is performed and not limited to the domain covered by a theoretical paradigm (Breyer, 2011).

In line with the association of Corpus Linguistics, as it is widely known, then a corpus is defined as "a large principled collection of naturally occurring texts (written or spoken) stored electronically" (Reppen, 2010). It is necessary to point out that the term corpus is a singular form while the term corpora is a plural form. Several terms have been used to language learning and teaching practices that make use of corpora, such as Corpus-Driven, Corpus-Aided, Corpus Assisted, and Corpus-Based Approach. Four main characteristics can represent CL which is to be authentic, frequent, syntactically stored, and contextualized. When it comes to frequency, CL intends to investigate the extent to which patterns of language are found in a given body of spoken or written texts in real-life situations. Moreover, it aims to analyze the contextual factors that influence language variation in the texts (Charles, 2018; Cortes, 2018; Vyatkina & Boulton, 2017). The second concept which is authenticity refers to the language input that occurs naturally by its speakers (Lewis, 2000). The third term which is contextualized indicates that a linguistic feature is well thought and learned within an authentic context, where the learners are being exposed to many examples that have been taken from real-life situations (Gilmore, 2008). The fourth and last characteristic which is syntactically stored refers to the capability of corpus tools to technically store language input by its syntactical features.

To illustrate more, corpus data are stored by part-of-speech tagging (POS), that is, each token is tagged by its part-of-speech automatically. Moreover, to be syntactically stored, lexicon and grammar are considered as two connected parts of a single pattern which is known as lexico-grammatical patterns (Liu & Lei, 2017; Benavides, 2015; Liu & Jiang, 2009). Within the lexico-grammatical patterns, a lexical item usually has in nature a grammatical implementation that gives that lexical item a meaning in which it would not be given without a grammatical pattern (McEnery & Gabrielatos, 2006). Corpus Linguistics has confirmed the relationship between grammar and lexis (Hunston & Francis, 1996). The main argument Hunston and Francis (1996) have made is that "a description of a word and its patterns cannot be classified under the heading either of 'lexis' or of 'grammar'". Moreover, Elsherbini and Ali (2017) have stated that "lexical input and grammatical function is of extensive value to someone acquiring English as a second or foreign language, as well as to teachers of English as a second language (ESL) or English for foreign language (EFL)".

Many linguists have argued that teaching EFL based on ready-made language textbooks is not the source of language learning but should include language input that is carried a meaningful context (Lewis, 2000). Yunus et al (2016) have argued that teaching EFL as a product that mainly focuses on the form of a language, giving less or no attention to the contextualized factors cannot provide a full explanation of the complexity of the language aspects. Moreover, Alattar (2014) and Al-Seghayer (2014) have claimed that educational lessons are separated from real-life input might not offer learners valuable learning experience because there is an insufficient application to real-life situations. Since Data-Driven Learning (DDL) concentrates on a genuine and authentic input as compared to the artificial linguistic examples written

in the structural EFL textbooks, it is believed to provide EFL learners with an effective learning experience.

The application of corpora in EFL classrooms can be applied in two different ways, which are indirectly or directly. In terms of indirect use, applying corpora in an educational context is induced by the educators who prepare (e.g., curriculum designers and developers) and by (researchers) who utilize corpus-based tools to track, explore, and investigate different language patterns. While the direct application of corpora is applied by (teachers) who deliver corpus-based instruction and use corpus-based materials. Also, (learners) who make corpus research analyzing various linguistic elements (McEnery & Xiao, 2011).

The indirect use of corpora in ELT as remarked by Leech (1997) has four main focuses which are related to reference publishing, syllabus design, materials development, and language testing. Consequently, the selection of language elements, the order of which to teach, how to teach, and what to test, are all determined by the information corpus data provides (Breyer, 2011). Making decisions based on corpus data has remarkably improved the quality of reference and teaching materials (Römer, 2011). The valuable use of corpus data can significantly contribute to making the description of language more realistic, the sequence of the presented language is more sensible, the examples included are more authentic, and the intended objectives to test are more accurate (Liu & Lei, 2017).

While the indirect application of corpus data is closely associated with what to teach, the direct use of corpus data has also provided valuable insights into how to teach. In a direct application of corpus tools, learners can have direct access to corpus data to search and analyze language patterns (Elsherbini & Ali, 2017). Learners can perform different corpus searches to find out linguistic solutions about what they are looking for. It is important to note that corpus training by applying corpus-based activities in EFL classrooms is essential for learners (Yoon, 2008). This is because learners might not have experience on how to use corpus appropriately and they need to be trained and be familiarized with the way corpus data should be detected and interpreted.

It is recommended that the starting point is with "convergent" tasks that are, tasks in which learners who work with the same data should arrive at the same outcomes. The purpose of these tasks is for the teacher to ensure that essential basic techniques are learned, and a certain track is followed through the processing of corpus consultation. Charles (2007) indicated that teacher guidance can assist learners to succeed in understanding the managing of the arduous quantity of corpus input and support them to progressively develop a better judgment of corpus analyzing without heart-rending the original concept of discovery-learning. Then proceeding to "divergent" tasks, that are tasks in which the same data are alternated by different learner' needs who are expected to arrive at different outcomes (Tono et al, 2014). Moreover, it is important to state that the ability to reach the full potential of corpus tools will vary among learners according to their level of linguistic competence, autonomy, and experience with corpora (Braun, 2006). Therefore, the teacher's supervision and support are essential in facilitating learners to a successful experience in corpus utilization (Yoon & Jo, 2014).

Corpus data have the potential to assist EFL writers in which they can consult corpus researches at any stage of the writing process to check if their written texts are linguistically accurate, to ensure if they carry the intended meaning, to look up for alternatives, to test the linguistic hypotheses they made, and to edit the surface levels of their writing. Kotamjani et al (2017) conducted a study to investigate the effects of corpus consultation on EFL learners. The participants experienced a three-session training to use different corpus resources when they are writing research papers. Moreover, the participants were interviewed. The results revealed that around 604 linguistic problems were solved. Additionally, learners' linguistic awareness was raised. EFL learners are encouraged to explore and analyze corpus data to accomplish various language tasks. Those tasks may include solving their grammatical or lexical problems, checking the appropriate grammatical or lexical usage of their written texts, extending or developing knowledge of linguistic items, differentiating close synonyms, detecting patterns, collocation, and lexicography, as well as other linguistic information (Yunus, 2017).

The number of studies that have applied corpus tools in the EFL Saudi context indirectly as a method of preparing language materials and exploring and investigating specific patterns or values is significant (e.g., Al-Khateeb & Almujaiwel, 2018; Almujaiwel, 2018). Nevertheless, there are few empirical studies that have attempted to apply corpus-based teaching and learning to academic writing skills (e.g., Alharbi, 2012; Alattar, 2014; Alharbi, 2015; Almutairi, 2016; Alshehri, 2018).

Alharbi (2015) performed research to explore the potential of a simplified small corpus in improving Saudi language learners' quality of writing besides an in-depth overview of their evaluation. A sample of twenty-five university-level Saudi students participated in the study. The two instruments used were a classroom intervention in which students' corpus consultation is recorded through a screen-recording software and a survey distributed among the students after each task to determine their attitudes. All participants were competent computer users but had never employed or even heard of corpora before. Hence, the students were trained during a twelve-week period. Small and simplified corpus linguistics was designed for learners to consult and use. The findings of the study have shown that the application of a corpus-related methodology with learners has significantly improved their quality of writing, as well as their attitudes.

In a recent study, Almutairi (2016) aimed at assessing the efficiency of CBA to writing instruction in making corpus-based tasks to teach personal statement writing to EFL learners in the Saudi context. Also, his study sought to evaluate the effectiveness of using corpus-based activities and outline the limitation and potential of corpus tools as a resource in ELT. He used a corpus software called Sketch Engine® to compile a corpus of a sixty-seven-personal statement. After the statements' collection, the lexico-grammatical features were analyzed.

The researcher claimed that the application of CBA in teaching a foreign language has revolutionized the ways in which language is taught since CL is characterized by its authentic appeal that helps linguists and instructors to study the language more closely and to examine characteristics and actual usages of certain genres. The potential of corpora can help these linguists and teachers to create more efficient classroom exercises and even improve their syllabus designs. The researcher

concluded that the application of CBA to writing instruction, more specifically the personal statements teaching in the Saudi EFL context, can help learners to express themselves more effectively.

To sum up, Corpus Linguistics is a methodology used to study, teach, and learn a language. It encompasses large amounts of authentic data in a context. Corpus Linguistics is gradually taking a big part in researching, teaching, and learning practices throughout the world and in Saudi Arabia. Studies have suggested that Corpus Linguistics has a positive impact on ELT. Corpus Linguistics with its characteristics which are to be authentic, frequent, syntactically stored, and contextualized, can offer EFL researchers, materials developers, curriculum designers, teachers, and learners with great potential to better language proficiency. Corpus Linguistics can increase learners' confidence as EFL writers, foster learners' language awareness, and increase learners' autonomy. However, there is still a need for more work to be done in the field to give more clarity and insights into the possible practices of Corpus Linguistics in the educational field.

2. Method

As, according to Yoon (2011), EFL studies involving Corpus Applications to writing instruction "are highly varied in terms of methodology, there is no clearly defined methodology on which one can carry on" (p. 136). Hence, to achieve the purpose of this study and to better answer the research question, the instrument used is believed to adopt a quantitative method. It is a quasi-experimental study.

2.2 Setting
The setting for this study is King Abdulaziz University (KAU), a public university in Jeddah, Saudi Arabia, particularly, at the English Language Institute (ELI). The English Language Program provided by the ELI for KAU Preparatory Year Students (PYS) is aligned and mapped to the Common European Framework of Reference (CEFR) proficiency level descriptors. CEFR is widely known as an international standard for explaining and describing language ability. CEFR describes language ability on a six-point scale, from A1 for beginners, up to C2 for those who are advanced language speakers (Council of Europe, 2001). The English Language Program consists of four courses: 101, 102, 103, and 104, each course targets one of the CEFR levels (101 for A1 learners, 102 for A2 learners, 103 for B1 learners, and 104 for B2 learners).

2.3 Participants
Fifteen Saudi EFL participants took part in this study. The total of participants who started the intervention was 30 learners, however, only 15 participants attended all sessions. As public education in Saudi Arabia is separated by gender, this study takes place with only female learners. To achieve a homogeneous sample, the respondents of the study are to be selected based on The purposive sampling technique- a type of non-probability sampling technique which emphases the sampling techniques where the samples that are investigated are built on the judgment of the researcher as it serves the purpose of the study. In purposive sampling, researchers deliberately select individuals, based on their knowledge of the population to elicit data in which they are interested (Mackey & Gass, 2005). The main purpose of purposive sampling is to

concentrate on specific characteristics of a population that are of interest, which will best help in answering research questions and hypotheses. Moreover, usually, the size sample being investigated is quite small. All the participants as earlier mentioned, are assumed to be at similar, if not the same, levels of English Language proficiency which is at B1 intermediate on CEFR. Participants were fairly homogeneous in terms of sharing the same language proficiency, age, also sharing the same L1 background and educational qualifications.

2.4 Instruments

The data were mainly collected based on classroom intervention. During the intervention, two sets of written tasks were collected. The class meets five times a week during a period of three weeks. According to Hartas (2015), classroom intervention is the investigation of a phenomenon through the manipulation of variables and the relationship between those variables. Moreover, according to Hussain et al. (2013), empirical evidence derived from interventions is central to the quantitative method, those empirical pieces of evidence provide objective, accurate, and generalizable information about the phenomena in an investigation.

2.5 Procedures

2.5.1 Stage One: Pre-test

At this stage, participants completed a written assignment as a first draft. Since the study is designed as One Group Pre and Post-test, the purpose of the written assignment is to be considered as a pre-test to measure the effect of the intervention by comparing it statistically to the second draft (post-test) that students have to write by the implementing of corpus application at stage three. For validity and reliability, four actions have been taken. First, students wrote one essay based on the content of their English Language course because according to Kennedy and Miceli (2001), selecting a range of topics that learners might find relevant is likely to engage them to write. Second, the students were asked to write the essays under test circumstances without being allowed to use any outside resources because the purpose of the study is not to compare the effect of the traditional materials to corpus-based materials but to investigate the potential of corpus application to writing instruction. Third, written essays were graded using a writing rubric. The writing rubric has three main criteria, including Introduction and Conclusion, Body Paragraphs, Essay Structure, and Cohesion, Grammar, and Mechanics. To ensure that the writing is valid and that only grammatical and lexical patterns are the constructs to be measured, all participants wrote the same number of words and followed two body paragraphs form which enabled them to score the same grades. For the criteria of 'Body Paragraphs, Essay Structure, and Cohesion', the participants have a discrepancy of two grades only which is from 6 to 8. Thus, the focus will be mainly on the 'Grammar and Mechanics' criteria. To ensure reliability, a qualified writing rater evaluated the written essays. Intra-rater reliability refers to the consistency that an individual rater has for the same data on different occasions (Zohrabi, 2013). The written tests were graded at different periods of time and no changes have been noticed.

2.5.2 Stage Two: Experimental (Corpus Workshop Consultation)

Following the guidelines addressed earlier for providing students with essential techniques that they need to acquire a successful Corpus Consultation, the workshop training sessions started with "convergent" tasks and then proceeded to "divergent"

tasks in a computer lab. The workshop training was conducted in five main sessions, each session has particular objectives. The workshop content was designed on a PowerPoint file all the handouts and worksheets were designed in Writing Booklet. The tasks of session one were based on 'Convergent Learning' whereas the tasks of sessions two to five were based on 'Divergent Learning'. The workshop content was designed by the researchers based on an analytical reading of corpus literature related to 'Corpora in ELT' (e.g., Lessard-Clouston & Chang, 2014; Cobb, 2016).

In the convergent tasks' sessions, tasks were designed in such a way that students were expected to reach the same outcomes. The corpus website that has been used to complete these tasks was the Compleat Lexical Tutor (see Figure 1). Compleat Lexical Tutor (v.8) was developed by Cobb (2016) of the University of Quebec at Montreal (UQAM) to afford valuable resources for English and French language researchers, teachers, and learners. Due to various resources and technical functions, several researchers have used Compleat Lexical Tutor to conduct experimental studies in different areas, such as lexical bundles, grammatical patterns, writing, and word lists (e.g., Gaskell & Cobb, 2004; Yoon, 2008; Yoon & Jo, 2014). The results of the research have recognized the usefulness of the website and its resources.

Figure1 The compleat lexical tutor, home page

There are three main reasons for selecting this website in particular. First, with its focus on academic discipline, the site seemed to be the most suitable for those working with English for Academic Purposes (EAP) learners. Second, Compleat Lexical Tutor is among the most well-known web-based programs that have various interactive corpus tools. Third, the learners' basic command of the site did not seem to be an issue because Compleat Lexical Tutor: "is quite simple to use" (Yoon & Jo, 2014). Among the Compleat Lexical Tutor tools, the main interactive tool which is of greatest interest to this study is the Corpus Correct tool. This tool offers an effective training practice for the learners, which enabled them to comprehend the central idea behind the employment of corpus data for lexical and grammatical information extraction to correct their written errors. Corpus Correct tool contains sentences in seven lexical and grammatical categories (Prepositions, Singular-Plural, Word Order, Gerund-Infinitive, Simple Past-Present Perfect, Conditionals, and Formulaic). Each sentence holds an error in the designated category. To make the best use of the

Corpus Correct tool, it was applied in the classroom through three steps. First, the students were provided with a link that includes all the Lextutor sentences within only six categories (Prepositions, Singular-Plural, Word Order, Gerund-Infinitive, Simple Past-Present Perfect, Conditionals). In this link, students have to read the sentences and choose if they are either correct or incorrect. The purpose of this step is twofold, to investigate whether the students are aware of the errors or not and most importantly, to engage the students when they will be informed later that all the sentences are wrong. In other words, if learners chose that a sentence is correct, they will be more motivated to examine the concordance lines and find out the correct answer.

In the second step, which is the actual consultation of the Corpus Correct tool, the learners should first look at the error, review it, and click on the CONC option which will show them the concordance lines that have the error used correctly. Then, the concordance lines should enable the learners to build a hypothesis about the correct form of the error. After establishing the hypothesis, the learners should extract the pattern from the corpus data and use it to correct the sentence by examining each hypothesis and writing the form that they think is the correct one inside the 'Check option'. For either the correct or incorrect answer, a mark will appear in the 'Check option' that shows the students the status of their answers (either correct or incorrect). In the last step, which was simultaneously applied with the second step, the learners were provided with a link that has the same sentences but with three options among which the latter should choose one. These options include the following: (1) I found out the correct answer from Corpus, (2) I did not find out the correct answer, and (3) I already know the answer. There are two purposes of the provided link. First, it allows the learners to choose what exactly has happened during the corpus practice. Second, it ensures that all the participants have completed the requirements in the first session as they have to submit their answers through the link.

In the divergent tasks, learners have to work with corpus data by themselves with the teacher's guidelines first and then independently. The Corpus website used the Corpus of Contemporary American English (COCA) because: "it is arguably the most widely used online English corpus and it boasts a very powerful, multifunctional, and user-friendly tool of search engine" (Liu & Lei, 2018). The main purposes of completing the tasks covered from session three to five are to encourage the students to discover new lexical and grammatical rules and, most importantly, to correct the written errors of learners.

2.5.3 Stage Three: Post-Experimental Stage.

At the end of the experiment, students were handed the written essays back with error coded to write the second draft using COCA. All participants were asked to give information about how they corrected their written errors in which they have to choose among three options (1) errors corrected by using COCA, (2) errors corrected without COCA, and (3) errors not corrected.

3. Results and Discussion

As mentioned earlier, 15 students took a pre-test (first draft) at the beginning of the intervention. Then, the students were provided with three weeks workshop on how to use Corpus tools. By the end of week three, the students received a post-test (second

draft) that measures the same constructs of the pre-test, in which they were requested to make changes to their lexico-grammatical patterns using COCA. Table 1 shows the students' grades before the corpus consultation.

Table 1 Students' Pre-Test grades

N	Introduction & Conclusion (5)	Body Paragraphs, Structure & Cohesion (10)	Grammar & Mechanics (5)	Total (20)
1	4	7	3	14
2	4	6	2	12
2	4	5	1	10
4	4	7	4	15
5	4	6	2	12
6	4	7	3	14
7	4	6	2	12
8	4	7	3	14
9	4	5	1	10
10	4	7	4	15
11	4	7	3	14
12	4	6	2	12
13	4	6	2	12
14	4	6	2	12
15	4	7	4	15

Then, the students were provided with three weeks workshop on how to use Corpus tools. By the end of week three, the students received a post-test (second draft) that measures the same constructs of the pre-test, in which they were requested to make changes on their lexico-grammatical patterns using COCA (see Table 2).

Table 2 Students' Post-Test grades

N	Introduction & Conclusion (5)	Body Paragraphs, Structure & Cohesion (10)	Grammar & Mechanics (5)	Total (20)
1	4	8	4	16
2	4	7	4	15
2	4	6	3	13
4	4	8	5	17
5	4	7	4	15
6	4	8	4	16
7	4	7	3	14
8	4	8	4	16
9	4	6	3	13
10	4	8	5	17
11	4	8	4	16
12	4	7	3	14
13	4	7	3	14
14	4	7	3	14
15	4	8	5	17

To determine if there is a significant difference between the pre and the post grades, the Wilcoxon Signed-Rank Test (Table 3) was used to reveal any significant difference between the pre and post-test scores of learners. The Wilcoxon Signed-Rank Test is a type of non-parametric test which is used as an equivalent test to the dependent t-test. It is used when the sample is less than 30, there is no normality in the data, and the use of the dependent t-test is inappropriate. The Wilcoxon Signed-Rank Test is used to compare two sets of scores from the same participants. As the current study followed One-Group Pretest-Posttest Design, The Wilcoxon Signed-Rank Test is appropriate to compare the differences, if any, in the two tests (Woolson, 2007).

Table 3 Wilcoxon signed rank test

		N	Mean Rank	Sum of Ranks
Post-test - Pretest	Negative Ranks	0a	.00	.00
	Positive Ranks	15b	8.00	120.00
	Ties	0c		
	Total	15		

a. Post-test < Pretest
b. Posttest > Pretest
c. Posttest = Pretest

Based on Table 1 and Table 2, one can determine whether the change, due to the intervention, led to a significant difference in the students' overall performance. As shown in Table 4, the p-value, which is (.000), indicates the presence of a statistically significant difference in the learners' performance before and after the classroom intervention.

Table 4 Test statistics for Wilcoxon signed rank test

	Post-test - Pretest
Z	-3.578-b
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

As discussed previously, the results of the post-test have revealed that there is an improvement among the students regarding their lexico-grammatical writing abilities compared to their pre-test results. Moreover, this result is consistent with that of the studies mentioned in the relevant literature (e.g., Yoon, 2008; Alharbi, 2015; Almutairi, 2016). The findings of Alharbi (2015) and Almutairi (2016) are similar to the findings of the current study in terms of the written tasks. Alharbi's (2015) study investigated the effectiveness of corpus tools on Saudi Arabian learners' academic writing skills. The findings confirmed that consultation with corpus data helped the learners to solve language-related problems through the production, analysis, and evaluation of concordance data. Furthermore, the post-test results support the results from Almutairi's (2016) study, which emphasized that in classrooms where EFL

academic writing courses are taught based on the integration of corpus data, learners are more likely to become skilled writers. Based on the results of the current study, it can be concluded that in EFL classrooms, when students are well aided and trained on how to use corpus tools, their writing skills are likely to be improved, especially their lexico-grammatical abilities.

This improvement among learners, according to the post-test results, calls for solutions and interventions. One solution is to evaluate the teaching materials used in the classroom to determine which materials must be more contextualized and to provide the learners with language input that can be used outside the classroom. This supports the arguments of Alattar (2014) and Al-Seghayer's (2014) studies that only the traditional methods of teaching, such as the artificial textbooks, with their emphasis on the direct teaching of artificial language input are not enough. Consequently, the need for a CBA, as a method of teaching that attempts to enhance students' writing abilities by providing them with authentic contextualized data, in EFL classrooms is met.

4. Conclusion

The current study sought to investigate the effect of applying a corpus-based approach to academic writing instruction as a method of teaching on learners' writing performance following a pretest-posttest design. The findings demonstrated that the improvement of lexico-grammatical abilities among students is significant. Several implications were induced to improve the quality of language teaching and learning based on the current findings of the study. Those implications are related to teachers, learners, and material designers. The study findings have revealed some insights for language teachers regarding the application of the corpus-based approach to writing instruction. Teachers should be trained on how to use corpus data and prepare corpus-based materials. The inclusion of authentic and rich corpus-based materials can lead to several advantages for teachers. They will be able to teach the skills of academic writing through real-life examples and enable students to be engaged in exploring different linguistic patterns and usages. Teachers can prepare corpus-based activities and tasks by taking corpus data from concordance lines using (COCA) and establish different objectives for those activities to tackle students' linguistic difficulties. Moreover, the current study provided evidence for curriculum designers and material developers doubting the potential of CBA. Designers and developers may consider using corpus-based data as a source for preparing classroom textbooks and materials with several implications. Including authentic and real life-examples can help learners to use what they have learned outside the classroom. Curriculum designers and material developers are encouraged to provide both teachers and learners with supplementary materials on how to use corpus tools effectively.

References

Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. International Journal of English Linguistics, 4(5), 17. https://doi.org/10.5539/ijel.v4n5p17

Alattar, A. (2014) Content-based instruction and corpus linguistics curriculum for early advanced EFL Saudi students, unpublished thesis (MA), University of San Francisco.

Alharbi, M.A. (2019). Saudi Arabia EFL university students' voice on challenges and solution in learning academic writing. Indonesian Journal of Applied Linguistics, 8(3), 576-587. https://doi.org/10.17509/ijal.v8i3.15276

Alharbi, W.H. (2015) Does Corpus Size Matter? "Exploring the potential of a small simplified corpus in improving language learners' writing quality". Corpus Linguistics (10), UCREL, Lancaster: UK.

Alkubaidi1, M.A. (2014). The Relationship between Saudi English Major University Students' Writing Performance and Their Learning Style and Strategy Use. English Language Teaching, 7(4), 83-95. https://doi.org/10.5539/elt.v7n4p83

Almutairi, N.D. (2016). The Effectiveness of Corpus-Based Approach to Language Description in Creating Corpus-Based Exercises to Teach Writing Personal Statements. English Language Teaching, 9(7), 103. https://doi.org/10.5539/elt.v9n7p103

Alrabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. International Journal of English Linguistics, 6(3), 21. https://doi.org/10.5539/ijel.v6n3p21

Benavides, C. (2015). Using a corpus in a 300-level Spanish grammar course. Foreign Language Annals, 48(2), 218–235.   https://doi.org/10.1111/flan.12136

Breyer, Y. (2011). Corpora in language teaching and learning. Potential, evaluation, challenges. Peter Lang.

Charles, M. (2018). Corpus tools for writing students. The TESOL Encyclopedia of English Language Teaching, 1-7.

Cobb, T. (2016, January 19). The Compleat Web VP. Retrieved from Compleat Lexical Tutor: http://www.lextutor.ca/vp/comp/

Cortes, V. (2018). Corpus tools for writing teachers. In: J. I. Liontas & M. DelliCarpini (Eds.), The TESOL encyclopedia of English language teaching (pp. 1–6). Hoboken, NJ: Wiley-Blackwell.

Dörnyei, Z. (2001). Motivational strategies in the language classroom. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO97805511667343

Elsherbini, S.A. and Ali, A.D. (2017). The Effects of Corpus-based Activities on EFL University Students' Grammar and Vocabulary and Their Attitudes toward Corpus. Journal of Research in Curriculum Instruction and Educational Technology, 3(1), 133-161.

Friginal, E. (2018). *Corpus Linguistics for English Teachers: New Tools, Online Resources, and Classroom Activities.* New York: Routledge.

Gaskell, D. and Cobb, T. (2004). Can learners use concordance feedback for writing errors? System, 32(3), 301–319.

Gilmore, A. (2008). Using online corpora to develop students' writing skills. ELT Journal, 63(4), 363–372. https://doi.org/10.1093/elt/ccn056

Hartas, D. (2010). Educational research and inquiry, qualitative and quantitative approaches. London: Continuum International Publishing Group. https://doi.org/10.1002/9781118784235.eelt0554

Hunston, S. and Francis, G. (1996). *Pattern grammar: A corpus driven approach to the lexical grammar of English.* Amsterdam: Benjamins.

Hussain, M.A., Elyas, T. and Nasseef, O.A. (2013). Research Paradigms: A Slippery Slope for Fresh Researchers. Life Science Journal 10(4), 2374-2381.

Javid, C. and Umer, M. (2014). Saudi EFL learners' writing problems: A move towards solution. Proceeding of the Global   Summit on Education GSE, 4–5.

Kennedy, C. and Miceli, T. (2000). An evaluation of intermediate students' approaches to corpus investigation. Language Learning and Technology, 5(3), 77–90.

Khan, I.A. (2011). Learning difficulties in English: Diagnosis and pedagogy in Saudi Arabia. Educational Research, 2(7), 1248-1257.

Kotamjani, S. S., Razavi, O. F. and Hussin, H. (2017). Online corpus tools in scholarly writing: A case of EFL postgraduate student. English Language Teaching, 10(9), 61–68.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), Teaching and Language Corpora (pp. 11–23). Harlow: Addison Wesley Longman.

Lessard-Clouston, M. and Chang, T. (2014). Corpora and English language teaching: Pedagogy and practical applications for data-driven learning. TESL Reporter, 47(1, 2), 1-20.

Lewis, M. (ed.). (2000). *Teaching collocation*. Hove: Language Teaching Publications.

Liu, D. and Jiang, P. (2009). Using a corpus-based lexico-grammatical approach to grammar instruction in EFL and ESL contexts. Modern Language Journal, 93(1), 61–78.

Liu, D. and Lei, L. (2017). Using corpora for language learning and teaching. Alexandria, VA: TESOL Press.

Mackey, A. and Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.

McEnery, T. and Gabrielatos, C. (2006). "English corpus linguistics". In B. Aarts 8c A. McMahon (Eds.), The Handbook of English Linguistics. Oxford: BlackweU, 33-71.

McEnery, T. and Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), Handbook of research in second language teaching and learning (Vol. 2, pp. 364–380). London: Routledge.

Mohammad, T.,and Hazarika, Z. (2016). Difficulties of Learning EFL in KSA: Writing Skills in Context. International Journal of English Linguistics, 6(3), 105. https://doi.org/10.5539/ijel.v6n3p105

Obeid, R. (2017). Second Language Writing and Assessment: Voices from within the Saudi EFL Context. English Language Teaching, 10(6), 174-181. https://doi.org/10.5539/elt.v10n6p174

Paker, T. and Ergül-Özcan, Y. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. International Journal of Language Academy, 5(1), 62-81.

Pérez-Paredes, P., Sánchez-Tornel, M., Alcaraz Calero, J. M., and Jiménez, P.A. (2011). Tracking learners' actual uses of corpora: Guided vs non-guided corpus consultation. Computer Assisted Language Learning, 24(3), 233–253.

Picard, M. (2018). *The future of EFL and TESOL in Saudi Arabia. In English as a foreign language in Saudi Arabia* (pp. 165–185). Abingdon: Routledge.

Reppen, R. (2010). *Using corpora in the language classroom*. England: Cambridge University Press.

Römer, U. (2011) Corpus research applications in second language teaching. Annual Review of Applied Linguistics, 31: 205–225.

Stubbs, M. and Halbe, D. (2013). Corpus linguistics: Overview. In C. Chapelle (Ed.), The encyclopedia of applied linguistics (pp. 1377-1379). Oxford: Blackwell.

Tono, Y., Satake, Y.and Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. ReCALL, 26(2), 147-162.

Vyatkina, N. and Boulton, A. (2017). Corpora in language learning and teaching. Language Learning & Technology, 213, 1–8.

Woolson, R.F. (2008). Wilcoxon Signed‐Rank Test. Wiley encyclopedia of clinical trials. Vol 6. West Sussex, England: John Wiley & Sons, Inc. 1-3. https://doi.org/10.1002/9780471462422.eoct979

Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. Journal of English for Academic Purposes, 10(3), 130–139. https://doi.org/10.1016/j.jeap.2011.03.003

Yoon, H. and Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. Journal of Second Language Writing, 13, 257–283. https://doi.org/10.1016/j.jslw.2004.06.002

Yoon, H.and Jo, J.W. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 Writing. Language, Learning & Technology, 18(1), 96–117.

Yunus, K. (2017), Corpus Linguistics: Pedagogic application in the 21st century, in International Journal of Academic Research in Progressive Education and Development 6(3): 137–152.

Zohrabi, M. (2013). Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings. Theory & practice in language studies, 3(2).