

The Effects of Corpus-based Activities on EFL University Students' Grammar and Vocabulary and Their Attitudes toward Corpus

Dr. Saneia Abdel Hameed Elsherbini

Sadat Academy for Management Sciences

Dr. Amira Desouky Ali

Sadat Academy for Management Sciences

Abstract

The study investigated the impact of corpus-based activities on enhancing EFL students' grammar and vocabulary and explored their attitudes toward using corpus. Two classes of 104 freshmen taking business English course at Sadat Academy for Management Sciences in Egypt participated in the study; with 54 in the experimental group and 50 in the control group. The mixed-methods design was used to gather data from a grammar test, a vocabulary test, a questionnaire, and semi-structure interviews. Students in the experimental group were trained to use corpus and were taught through corpus-based activities for 11 weeks; whereas control group students were instructed using the course book only. The differences in students' performance in grammar and vocabulary were examined between their pre- and post- tests. The results of the t-test indicated that there was a significant difference between the mean scores of the pre- and post-tests of grammar and vocabulary of the experimental group in favor of the posttests. Also, experimental group students outperformed their counterparts in the control group in the mean scores of the grammar and vocabulary posttests. After the treatment, the attitudes of experimental group students toward corpus were explored through administering the questionnaire and interviews. Data analysis of the questionnaire revealed that respondents were generally positive toward using corpus-based activities in learning both grammar and vocabulary. Analysis of the interviewees' responses supported the findings of the questionnaire and offered some insights into using the corpus. Suggestions for pedagogical implications and recommendations for further research were provided in this study.

Key words: *Corpus-based Activities, Grammar, Vocabulary, Attitudes toward Corpus, Mixed-Methods, EFL University Students.*

Introduction and Background

O'Keefe, McCarthy, and Carter (2007) define Corpus as “a collection of texts written or spoken, which is stored on a computer and can be analyzed qualitatively or quantitatively using analytical software” (p.1). Analyzing qualitatively means looking at lines of concordance to see how a word or phrase is used, and analyzing quantitatively means focusing on the number of times that a particular word is used.

Corpus can be used in the language classrooms in two ways: Students' direct use of concordancing software, i.e. computer-based activities, and indirect use of corpora i.e. the presentation of teacher-prepared concordance data in handouts (paper-based activities). In the direct use of concordance, learners find language rules by themselves through directly accessing concordance. They find some difficulties using this approach because of the intimidate amount of concordance examples which can be meaningless. In the indirect use of corpora, the teacher accesses a concordance directly and prints out some examples. He edits the concordance to be simple and clear for the learners (Yoon & Jo, 2014).As for this study, both corpus direct and indirect techniques were implemented.

The use of corpus-based activities in EFL classrooms is essential for learners to get acquainted with real reliable language. Going through the concordance lines of a searched word, learners can easily analyze language patterns and collocations (Hadley, 2002; cited in Chang & Sun, 2009). Viewing the incorporation of corpus research into language teaching and learning from a register-specific and lexico-grammatical perspective, Conrad (2000) argues that three major influences of corpus-based linguistics on the teaching of grammar can be seen in the 21st century. These three significant impacts are:

1. Detailed descriptions of grammar rules will be replaced by register-specific descriptions.
2. Grammar instruction will be closely integrated with the teaching of vocabulary. Lexico-grammar patterns will be central to language description and language learning.

3. Emphasis on structural accuracy will be shifted to the appropriate use of alternative grammatical constructions.

However, researchers assert that students face lots of difficulties in the direct use of corpora; especially lower-level students who are overwhelmed with too much concordance examples which become too numerous and meaningless. On the other hand, teacher guidance can help students manage this huge amount of data and gradually develop a better sense of corpus use (St.John, 2001; Charles, 2007).

One of the key basis of the corpus approach characterizations is that vocabulary and grammar are integral rather than distinctive from each other (Sinclair, 1991; Halliday, 1992). Within this association of vocabulary and grammar, also known as lexico-grammar, there is a priority on most constant combinations of words, i.e., “collocation” (Conrad, 2000; Biber & Conrad, 2001). For instance, the noun “location” is often followed by the prepositions “of” and “for”. This focus in 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 ESL or EFL. Recent attempts to connect the corpus approach with genre analysis have been particularly beneficial in such domains as English for Academic Purposes (EAP) or English for Specific Purposes (ESP) (Yoon & Hirvela, 2004).

In light of such benefits of corpus use in second or foreign language, many studies have argued for the inclusion of corpus in teaching EFL as follows:

Studies conducted in the context of grammar teaching through corpus-based sources

Vannestal and Lindquist (2007) investigated advanced proficiency level EFL learners’ attitudes toward using concordancing in grammar learning and the effects of corpora on learners’ motivation to learn grammar. The results of the study revealed that the control group improved as much as the experimental group, in contrast to what the study hypothesized. According to the results of the questionnaire, the students in the

experimental group felt positively toward using corpora in grammar learning.

Boulton (2009) conducted a study with 132 first-year intermediate and lower levels of English students, looking at the use of corpora to deal with linking adverbials in English. The participants were randomly divided into four groups, two groups dealt with traditional sources and the other two groups used KWIC (Key Word in Context) concordances or short contexts (SC). Results indicated that the corpus groups (KW and SC) used information more effectively than the traditional groups (GU and BD). The researcher concluded that corpus data could be used by learners for reference purposes, and lower level students could also benefit from data-driven learning.

Perez-Liantada (2009) investigated a corpus-based approach in teaching and learning spoken grammar of English for Academic Purposes. Data resulting from the assessment process and student production suggested that corpus-informed instruction grounded in Bhatia's multi-perspective model can constitute a pedagogical approach in order to: i) obtain positive students' responses from input and authentic samples of grammar use, ii) help students identify and understand the textual, genre and social aspects of grammar in real contexts of use, and therefore iii) help develop students' ability to use grammar accurately and appropriately.

Rapti (2010) examined the impact of Data-driven learning (DDL) on a group of adolescent students in Greek, the degree of motivation to learn grammar, and the effectiveness of DDL in teaching grammar. Concordance-based grammar materials were introduced to the experimental group, whereas the control group used a conventional grammar book. Results showed that most learners expressed their preference for concordance based learning and further access to corpora. The findings pointed out the importance of using corpus-based grammar teaching to EFL learners.

Girgin (2011) investigated the effectiveness of using corpus-based activities on lower level EFL students' learning of English grammar and their attitudes towards using corpus-based activities in English grammar learning. Six intact lower level EFL classes at Erciyes University School of Foreign languages participated in the study. Results showed that the students were able to use corpus-based activities in the learning of the target grammar structure produced similar results when compared to using a course book. Results also revealed students' positive attitudes toward using corpus-based activities in learning grammar.

Hanafiyeh and Keshi (2013) investigated the effects of the corpus-based language learning on Iranian EFL learners' grammatical knowledge by analyzing the learners' writing samples. The sample was 60 students in Azad University and they were assigned randomly into an experimental group of concordance (n=30) and a control group of thesaurus (no=30). The subjects were analyzed using corpus linguistic analysis to investigate their writing quality. The results indicated that there were noticeable differences in the EFL writing quality between the two groups. Also, there were significant differences where the concordance group gained more grammatical skills than the thesaurus group.

Chujo, Kobayashi, Mizumoto, and Oghigian (2016) examined the effectiveness of combining two newly developed web-based tools for the foreign language DDL classroom. The tools were freeware based on the same parallel corpus and Para News which consists of newspaper text in English with translation in Japanese. The results asserted that the group which used the combined tools outperformed the group used the single tool. Also it was indicated that using different types of corpus tools were helpful in understanding the targeted grammar than the traditional method.

Studies examined the effectiveness of corpus-based sources on EFL learners' performance in vocabulary

Chan and Liou (2005) investigated the influence of using five web-based practice units on English verb-noun collocation with the design of a web-based Chinese-English bilingual concordance on collocation learning. Thirty-two EFL college students participated by taking a pre-test and two post-tests, and responding to background questionnaire and an evaluation questionnaire. Results indicated that learners made significant collocation improvement immediately after the online practice. Both the online instructional units and the concordancer were acceptable to most participants.

Jafarpour and Koosha (2006) conducted a study which tried to define the role of DDL (data-driven learning) in the teaching of collocation of English prepositions to Iranian EFL adult learners by dividing them into two groups. One received data-driven instruction whereas the other received instruction in the conventional approach. The study showed that the participants who received data driven instruction outperformed those who received conventional instruction in the learning of collocation of prepositions.

Varely (2008) carried out a study on 19 EAL students to find out if corpora improved students' vocabulary and grammar. Findings revealed that most of the participants agreed that corpora were very beneficial for their language learning. Also corpora helped students to learn about certain language uses that were not available in any of the traditional tools. Another benefit was the fact that corpora helped students get broader view of language (i.e. corpora could compare spoken and written language).

To ascertain the effectiveness of corpus-based approach in teaching collocations of synonyms pairs, Jafarpour, Hashemian, and Alipour(2013) conducted a study comparing the effectiveness of corpus-based approach with the effect of the traditional approach. Participants with equal competence on

learning collocations were selected and were then divided into two groups, experimental and control groups. The experimental group was taught using concordancing materials of BNC (British National Corpus). The control group, on the other hand, was taught collocations in a traditional way. Findings revealed that a difference between the two approaches and the corpus-based approach had an effect on the comprehension and production of collocations of synonyms in favor of the experimental group.

Kayaoglu (2013) conducted a quasi-experimental research to explore the feasibility of a corpus by using the corpus of Contemporary American English (COCA) to distinguish between close synonyms which have parallel meanings but cannot be replaced one for the other. Students were given 40-minutes training to use COCA and were expected to make their judgment of choosing the appropriate English words from a collection of real language sample. The results of this study showed that the use of corpus was potential for determining word choice. Also, it was observed that corpora can be an effective tool in solving grammatical and related problems.

Unaldi, Bardakci, and Akpinar(2013) investigated the effectiveness of corpus-informed, contextualized vocabulary instruction. In this study, one of the experimental groups was exposed to corpus-informed vocabulary instruction, and the other group was treated with decontextualised vocabulary instruction. The results showed that the group treated through decontextualized learning activities scored statistically high mean compared to the other group. The group treated with corpus-informed instruction showed some progress while the control group made the least progress.

Ucar and Yukselir (2015) sought to reveal the impacts of corpus-based activities on verb-noun collocation learning in EFL classes. The study was carried out on one experimental group and one control group; each of which consisted of 15 students. The students were preparatory class students at school of foreign languages. Throughout the study, the experimental group was taught verb-noun collocation through a conventional

method. The results demonstrated that there was a statistically significant difference between experimental and control group which signifies that corpus-based activities had a significant impact on verb-noun collocations in EFL classes.

Statement of the Problem

Based on the review of literature and the results of a pilot study conducted by the researchers on a sample of freshmen, other than that of the present study, to examine their proficiency in grammar and vocabulary; findings showed that freshmen had difficulty in learning grammar and vocabulary. These difficulties might be a result of ineffective grammar and vocabulary teaching strategies and low motivation to learn these sub-skills; which led to their low level in grammar and vocabulary.

Aims of the Study

The main aims of this study were to determine the effectiveness of corpus-based activities in enhancing students' grammar and vocabulary achievement and also to explore their attitudes toward using corpus-based activities in EFL learning.

Questions

By conducting this study, the researchers sought to find answers to these questions:

1. What is the effect of using corpus-based activities on improving students' grammar?
2. What is the effect of using corpus-based activities on improving students' vocabulary?
3. What are the attitudes of the experimental group students toward using corpus in learning grammar and vocabulary in EFL classes?

Hypotheses

The present study tested the following hypotheses:

1. There is a significant difference at the 0.05 level between the mean scores of the grammar pretest of the experimental group and their mean scores of the grammar posttest in favor of the posttest.

2. There is a significant difference at the 0.05 level between the mean scores of the grammar posttest of the experimental group and the mean scores of the grammar posttest of the control group in favor of the experimental group.
3. There is a significant difference at the 0.05 level between the mean scores of the vocabulary pretest of the experimental group and their mean scores of the vocabulary posttest in favor of the posttest.
4. There is a significant difference at the 0.05 level between the mean scores of the vocabulary posttest of the experimental group and the mean scores of the vocabulary posttest of the control group in favor of the experimental group.

Significance of the Study

Few studies have examined the effects of using corpus in teaching both grammar and vocabulary in EFL context. This study will open new avenues for researchers interested in investigating how corpus-based sources affect grammar and vocabulary achievement. The study suggests effective techniques for teachers to use corpus-based activities as a valuable approach in teaching grammar and vocabulary. This will develop learners' autonomy in using tools that enhance their grammatical as well as lexical competencies. The results of the study will also provide guidance for EFL specialists and developers to integrate corpus-based activities in courses and programs as one of the most important approaches used recently in learning EFL/ ESL.

Method

Participants

The participants were 104 freshmen taking business English course at Sadat Academy for Management Sciences in Cairo; Egypt (academic year 2015-2016). They were chosen from two intact classes; aged from 18 to 20. The classes were assigned randomly into two groups; a control group (50 students) and an experimental group (54 students). Students of the two groups

were heterogeneous in terms of gender and language proficiency. Out of the 54 students of the experimental group, 21 were males and 33 were females, while the control group contained 24 males and 26 females. All students had adequate computer proficiency as indicated by their responses to the demographic part of the questionnaire.

Instruments

The instruments used in this study were consisted of:

1. Tests of grammar and vocabulary.
2. A questionnaire of attitude toward corpus use.
3. Semi-structured interviews.

Grammar and vocabulary tests.

One form of grammar test and vocabulary test were created and administered for both pre-test and post-test. The grammar test consists of 6 questions while the vocabulary test is composed of 9 questions. The grammar test examines students' knowledge of: singular and plural, word order, infinitive and gerund, conditionals, tenses (simple past, present perfect, and past perfect), and modal verbs. Questions in the vocabulary test were divided into three parts: a) phrasal verbs, b) prefix/suffix, and c) collocations. The questions in the vocabulary test were adapted from (Kim & Chun, 2008). The types of questions in both tests include: completion, matching, odd-one-out, fill-in-the-blanks, MCQs, and put in order.

The topics and the relative weight of the questions in the grammar and vocabulary tests were chosen according to the course the students studied. Collocations were given more weight because it is the focus of the course and they contain different types. The questions in collocations part respectively assess students' knowledge of: adjective+noun, noun+verb, verb+noun, noun+verb, adverb+verb, and adverb+adjective collocations.

The reliability of the tests was examined by administering them to a sample rather than that of the present study. Test-

retest reliability coefficient for the grammar test was 0.7 and for the vocabulary test was 0.8; which indicated good reliability. The tests were validated by specialists in TEFL. The tests were conducted on the experimental and control groups before the experiment on February 19th and after the experiment on May 19th. The researchers scored each test out of 50; a mark for each correct answer. The Pearson inter-rater reliability was relatively high; (0.99).

The attitude questionnaire.

The questionnaire is composed of two parts: the demographic information part and a likert scale part. The first part was created to have general information about the age and gender; computer use (frequency of use, having computer at home, and accessing the internet); and knowledge about corpus. The questionnaire was adapted from (Yoon & Hirvela, 2004). The questionnaire was organized in a five-point likert scale: 5 strongly agree; 4 agree; 3 neutral; 2 disagree; and 1 strongly disagree. This section consisted of 13 items and covers 3 categories: benefits of corpus use in learning grammar and vocabulary, problems and challenges in using corpus, and overall perception of corpus use. Cronbach Alpha was used to examine the reliability; which was found to be(0.93), that is relatively high. Therefore, the questionnaire was considerably reliable. The validity was checked by jurors from TEFL. The demographic part of the questionnaire was administered to the experimental group before the experiment while the likert scale part was conducted a day after the administration of the post-tests.

Semi-structured interviews.

The second researcher conducted the interviews at the end of the study to further explore students' views of corpus use in learning grammar and vocabulary, the challenges they faced when using corpus, their overall perception of corpus, and whether they plan to use it for further work. The interview is made up of 5 main questions. Fifteen students were selected beforehand according to their proficiency levels in English (5 lower-intermediate, 5 intermediate, and 5 upper-intermediate).

The interviews were conducted the same day of the questionnaire. In order to get a comprehensive picture of corpus use, students were interviewed individually and each interview lasted for 10-15 minutes. The interviews were tape-recorded and transcribed verbatim for analysis. The validity of the interview questions were checked by submitting them to the jurors specialized in TEFL.

Procedures

The experiment took place during the second semester of February/ May 2016. The corpus used in this study was Contemporary American English (COCA); a database of 450million entries from spoken and written sources developed by Brigham Young University. COCA was chosen for its large size as well as free availability online. The concept of corpus and its use were introduced to students in the experimental group. Students created accounts on COCA. They were trained for two sessions to use the basic functions of the corpus like using tags and codes for POs (parts of speech), conduct search, and interpret the output. Students were asked to search in BWC written and Brown+BNC written, as they suit writing, grammar, and lexis. Students in the control group studied the course book with the traditional way of teaching grammar and vocabulary.

While the first researcher taught the control group, the second researcher taught the experimental group, prepared the corpus-based activities, proposed a tutorial and a manual and uploaded them to Edmodo, and reviewed students' quizzes, and commented on their answers and corpus work. Students in the control and the experimental groups were instructed 11 sessions. The sessions were given in lecture rooms where students have access to the internet and used their smart phones and laptops to use online concordances and do activities on grammar and vocabulary. Before asking experimental group students to generate their own corpus queries, they were introduced to different worksheets and hands-on exercises of web-based corpus exercises on grammar as well as vocabulary (indirect use of corpus). The students did these activities in pairs,

individually or in groups. These exercises include: collocations, affixes and prefixes, phrasal verbs, word order, singular and plural, conditional, modal verbs, tenses, and infinitive and gerund. When students mastered these functions and skills, they were asked to independently conduct searches, identify patterns of language, read the concordance lines, and access the original context (direct use of corpus). Most of the corpus work were done outside classroom; either at home or university labs. Students were given homework tasks to do as quizzes on Edmodo. The quizzes consisted of a series of corpus queries for students to type in to complete or draw conclusions about grammatical rules or collocation use in the light of the concordance lines and output to check their comprehension and mastery of the functions taught in each session.

Study Design

The study is mixed-methods; both qualitative and quantitative approaches were used. Students of the two groups were chosen randomly and were assigned to a control group and an experimental group. Grammar and vocabulary tests were created by the researchers and carried out to the two groups before and after the experiment. The questionnaire was administered only to the experimental group after the study. The interviews were used to get deeper insights into the experimental group students' responses to the questionnaire.

Statistics Used

SPSS was used to analyze the data of the tests and the questionnaire. Paired sample t-test and independent sample t-test were utilized to explore the effects of using corpus on students' achievement of both grammar and vocabulary. Descriptive statistics including percentages and frequencies were used to examine the attitude of the experimental group students toward using corpus in EFL learning classes. In addition, Nvivo 10 was used to analyze the interviews and help in creating patterns in the qualitative data. Cohen Kappa coefficient was carried out to calculate a proportion of corresponding codes. The reliability index was (0.82); which

shows high inter-coder reliability. Therefore, themes of above (0.70) were retained. Data were coded under four themes:

- a. Using corpus in learning grammar
- b. Using corpus in learning vocabulary
- c. Difficulties in using corpus
- d. Overall attitude toward using corpus.

Results

The results of the current study were analyzed in light of the research questions. The first and second questions were answered through implementing t-tests. The third question was answered by analyzing the data collected from the questionnaire and the interviews.

Effect of Corpus-based Activities on Students' Grammar

A paired sample t-test was administrated to examine the difference between the mean scores of the grammar pre- and post-test results of the experimental group students before and after using corpus-based activities in learning grammar (see Table (1)).

Table (1) Paired Sample T-Test for Grammar Pre- and Post-Tests

	N	Mean	SD	t	df	p-value	Sign
pair 1 gram. pre test	54	13.66	4.61	-21.21	53	.000	0.05
pair 1 gram. post test	54	25.05	6.14				

As shown in table (1), data analysis of the paired sample t-test in this study revealed that the mean scores of the pretest (M=13.66) and the posttest (M= 25.05) were significantly different ($t = -21.21, p < .001$). It can be strongly confirmed that there was significant difference at the 0.05 level between the mean scores of the grammar pretest of the experimental group students and their mean scores of the grammar posttest in favor of the posttest. Therefore, the first hypothesis was confirmed.

An independent sample t-test was performed in order to compare the grammar posttest mean scores of the control and experimental groups to determine if there was a significant difference between the groups. Table (2) presented the results of the difference between the groups.

Table (2) Independent Sample T-Test for the Difference between the Grammar Posttest of the Control and Experimental Groups

	Levene's Test for Equality of Variances				t-test for Equality of Means					
	N	M	SD	t	df	p	Mean Difference	Std. Error Difference	95% of Confidence Interval of the Difference	
Control Group	50	17.10	4.00						Lower	Upper
				-10.62	87.06	.000	-11.511	1.083	-13.664	-9.357
Experimental Group	54	28.61	6.78							

Table (2) indicated that there was a statistically significant difference between the mean scores of the grammar posttest of the control group (M=17.10, SD=4.00) and the experimental group (M=28.61, SD=6.78) in favor of the experimental group. This means that hypothesis two was retained.

Effect of Corpus-based Activities on Students' Vocabulary

A paired sample t-test was administered to examine the difference between the mean scores of the vocabulary pre- and post-test results of the experimental group as shown in table (3).

Table (3) Paired Sample T-Test for Vocabulary Pre- and Post-Tests

	N	Mean	SD	t	df	p-value	Sign.
pair 1 vocab. pre test	54	17.55	3.86	-20.78	53	.000	0.05
pair 1 vocab. post test	54	32.83	4.95				

Considering the data and results illustrated in Table (3), corpus use showed significantly positive effects as the mean scores of the pretest (M= 17.55) was significantly different from

the posttest ($M= 32.83$) ($t=-20.78$, $p<0.01$). So, it could be asserted that there was a significant difference at the 0.05 level between the mean scores of the vocabulary pre- and post- tests of the experimental group students in favor of the vocabulary posttest. Thus, hypothesis three was maintained.

An independent sample t-test was performed in order to compare the vocabulary posttest mean scores of the control and experimental groups to determine if there was a significant difference between the two groups. Table (4) illustrated the difference between the groups.

Table (4) Independent Sample T-Test for the Difference between the Vocabulary Posttest of the Control and Experimental Groups

	Levene's Test for Equality of Variances					t-test for Equality of Means				
	N	M	SD	t	df	p	Mean Difference	Std. Error Difference	95%of Confidence Interval of the Difference	
Control Group	50	34.04	3.92						Lower	Upper
				-35.62	97.20	.000	-32.552	.913	-34.366	-30.738
Experimental Group	54	66.59	5.33							

Table (4) indicated that there was a statistically significant difference between the mean scores of the vocabulary posttest of the control group ($M=34.04$, $SD=3.92$) and the experimental group ($M=66.59$, $SD=5.33$) in favor of the experimental group. As a result, hypothesis four was accepted.

Attitudes toward Using Corpus

As for the demographic section of the questionnaire, (1.9%) of the students stated that they used computer for school work 5 times a week; (51.9%) of the students used it twice a week; (31.5%) of the students used it once a week; and (14.7%) stated that they used it from three to four times. With respect to the availability of internet at students' homes, (94.4%) of the students indicated that they have internet access at home while (5.6%) of the students didn't have internet. In addition, (18.5%) of the students use English while using computer; whereas (27.8%) use Arabic and (53.7%) use both languages. None of the students were familiar with corpora or have used it before. Therefore, the respondents used computer frequently for school work and the majority had internet access at home.

Data gathered from the likert scale section of the questionnaire as well as the semi-structured interviews were analyzed and presented below under three themes: Attitudes toward using corpus in grammar and vocabulary learning, challenges faced in using corpus, and general perception of using corpus.

Using corpus in learning grammar and vocabulary.

As illustrated in table (5), 70.4% ($M=3.93$; $SD=1.226$) of the students agreed that using corpus is easy while 13.2% of the students found it difficult to use and 16.7% were neutral. While 59.3% ($M=3.57$; $SD=1.092$) of the students signified that using corpus in learning grammar was helpful, 18.5% disagreed. 56.3% ($M=2.91$; $SD=1.391$) of the students indicated that using corpus in learning grammar is not more difficult than leaning it through a course book while 38.9% of students found learning grammar using corpus is more difficult than learning it through a course book. Students who found corpus helpful in learning vocabulary were 62.9% ($M=3.85$; $SD=1.106$); whereas 13% perceived it as unhelpful. 70.4% ($M=3.83$; $SD=1.194$) asserted that corpus was more helpful than a dictionary while 16.7% of the students disagreed.

These results were supported by students' responses to the first and second questions in the interviews concerning how students used corpus in learning grammar as well as vocabulary. As for using corpus in learning grammar, seven interviewees indicated that they read the concordancer examples to deduce rules. They found it interesting than checking grammar books with their general explanations. They stated that having background knowledge about the rules helped them to infer rules and understand the concordance lines. A student said: "*I think corpus is good for learning grammar than grammar books. It is easy to guess the rules from extracts or search for grammatical usage than to study it from grammar books; where the presentation of these rules is boring*".

However, five interviewees reported that corpus was not helpful in learning grammar and believed that grammar books are more direct and convenient justified this by stating that using corpus to learn grammar was not easy and need more time to interpret the overwhelming information. They found that it was difficult to analyze the corpus results to interpret rules. One of the interviewees viewed this as a matter of preference to grammar books; as it was easy to use and not time-consuming. Three interviewees were neutral in their attitudes.

In addition, interviewees reported their use of corpus to learn about synonyms, collocations, phrasal verbs, prepositions, and suffixes and prefixes. They admitted the resourcefulness of corpus and the concordancer examples regarding the target words they looked up. A student said: "*learning about synonyms, collocations, and phrasal verbs are things I used corpus for in learning vocabulary*". Students stated that corpus gave them more accurate and relevant data with a big number of examples when compared to dictionary. With the word frequency feature in corpus, students reported being able to check the most common words that collocate with the target words. Reading the target words in authentic contexts facilitate understanding of the usage of these words. While dictionaries are useful in looking up the meaning of words, corpora are beneficial in learning the

usage and place of these words in context. Words in COCA are underlined and bold; which make them easier to identify from the great amount of information. An interviewee considered corpus good for her; though she prefers printed dictionary. However, one student complained of his inability to use corpus effectively in finding collocations and lost time searching; which made him upset.

Table (5) Descriptive Statistics for the Questionnaire

	<i>Items</i>	<i>M</i>	<i>SD</i>
1	The corpus is easy to use.	3.93	1.226
2	Using the corpus is helpful for learning grammar.	3.57	1.092
3	I think that learning grammar using corpus is more difficult than learning it through a course book.	2.91	1.391
4	Using the corpus is helpful for learning vocabulary.	3.85	1.106
5	The corpus is more helpful than a dictionary.	3.83	1.194
6	I have some difficulties in using the corpus due to the time and effort spent on analyzing the data.	2.91	1.470
7	I have some difficulty in analyzing concordance output.	2.70	1.341
8	I have some difficulties in performing the search technique.	2.41	1.367
9	When I search for information in the corpus, I usually get the information I need.	3.89	1.076
10	Learning English using corpus has increased my confidence.	3.44	1.176
11	I will use the corpus in learning grammar and vocabulary in the future.	3.43	1.312
12	I really felt positively toward using corpus in learning English.	3.83	1.060
13	Teacher should use corpus in EFL classes.	3.85	1.246

Difficulties in using corpus.

Students who found some difficulties in using corpus due to the time and the effort spent on analyzing the data were 35.2% ($M=2.91$; $SD=1.470$) and student who didn't find any difficulties were 54.4%. While 25.9% ($M=2.70$; $SD=1.341$) of the students found some difficulties in analyzing concordance output, 58.1% didn't find any difficulties. Most students (68.5%; $M=2.41$; $SD=1.367$) found no difficulties in performing the search

technique and some students (27.8%) found some difficulties in performing the search technique.

Analysis of students' responses to the third interview question, asking about the difficulties they encountered when using corpus, showed that they perceived corpus difficult for being: a) time-consuming; b) difficult in interpreting concordance results and understanding the cut-off examples; and c) challenging in dealing with technical problems such as registering to COCA. The interviewees said that corpus use was not difficult but needs more practice and training to be able to search and interpret concordance output effectively; a result which agrees with responses to item 7 in the questionnaire. Three interviewees indicated having technical problems in logging and getting search results; which goes in line with responses to item 8 in the questionnaire. This was supported by this comment of one of the interviewees: *"Signing-in to COCA was difficult as after ten minutes the site asks you to log in again; which I found difficult and sometimes boring. It would be better if it is offline"*. Three students reported having difficulty in search techniques as they got no results if words are typed incorrectly or if the number of words before or after the target word was inaccurate. A student said: *"Well. It is not bad but it took me much time to search correctly especially for collocations and I had problem in specifying the number of words before the target word. Sometimes, I set two words but I found nothing"*. Seven students encountered difficulty in interpreting data. A student said: *"I followed the steps but I found no results. The examples were a little bit confusing for me. Many examples and I found them sometimes difficult to understand"*. In addition, two interviewees reported that corpus use is time-consuming.

Overall attitude toward using corpus.

Students who agreed that they usually got the information they needed when searching for it in the corpus were 74.21% ($M=3.89$; $SD=1.076$), while 13 % of the students disagreed. 63.5% ($M=3.44$; $SD=1.176$) of the students affirmed that learning English using corpus has increased their confidence and 24.3%

of the students said it didn't. Students reported their willingness to use corpus in learning grammar and vocabulary in the future were 51.8% ($M=3.43$; $SD=1.312$) and students who had no intention to use it were 24.1%. Most students (68.5%; $M=3.83$; $SD=1.060$) felt positively toward using corpus in learning English while a few number of students (11.1%) felt negatively toward using corpus in learning English. Finally, students who indicated that teachers should use corpus in EFL classes were 74% ($M=3.85$; $SD=1.246$) and students who didn't agree were 14.9%.

In responding to the fourth question in the interview inquiring about overall attitude toward using corpus, interviewees felt generally positive about using corpus in learning English. One of them stated: *"The most interesting thing about using corpus is that it contains all the things I need to look up somewhere else. It is easy to check for vocabulary or grammar. All in one place"*. Another student said: *"Generally, I feel positive in learning vocabulary than grammar. For me, it is faster and convenient than dictionary or grammar books"*.

When asked about using corpus for further learning, seven students recommended using it for learning grammar and vocabulary in the future; which is in the same vein of their responses to item 11 in the questionnaire. They suggested using corpus in other EFL courses; as showed in responses to item 13 in the questionnaire. A student commented: *"Yes. I feel positive and I believe that the corpus is useful especially in citing many examples on a word which is not in the dictionary or grammar books"*. On the other hand, five students were negative as they believed that corpus is difficult to use than dictionary and grammar books. Three students were neutral in their attitudes.

Discussion

Regarding the results of the achievement tests on both grammar and vocabulary, the progress in students' achievement can be attributed to corpus-based sources that supported learners' grammatical structures and vocabulary. The results of the present study corroborated by Jafarpour and Koosha(2006); Jafarpour et al.(2013); Kayaoglu (2013); Unaldi et al. (2013); and

Ucar and Yukselir (2015), who found that learners made significant usage of grammar and vocabulary after using corpus-based activities. Similarly, the findings of this study was in line with Varely (2008), who concluded that corpora helped students to learn certain language uses; grammar and vocabulary that were not available in any of the traditional tools.

The findings demonstrated in tables (1) and (2) revealed that using corpus in learning grammar were found to be significantly affecting students' grammar learning. This result is consistent with the previous research that have indicated the positive effect of corpus-based activities on learning grammar over the traditional methods (Vannestal & Lindquist, 2007; Boulton, 2009; Perez-Liantada, 2009; Girgin, 2011). Likewise, Perez-Liantada (2009) asserted that corpus-based approach developed students' ability to use grammar accurately and appropriately.

Similarly, Rapit (2010) found out that most learners expressed their preference for concordance-based learning and pointed out the importance of using corpus-based grammar teaching to EFL learners. Hanafiyeh and Keshi (2013) indicated that there were noticeable differences in the EFL writing quality between the two groups and the concordance group gained more grammatical skills than the thesaurus group. In addition, Chujo et al. (2016) asserted that the group which used the combined web-based tools outperformed the group used the single tool; therefore, using different types of corpus tools were helpful in understanding the targeted grammar than the traditional method. Overall, the results revealed that corpus-based approach enhanced students' ability to use grammar accurately and appropriately.

Moreover, the data collected and figured out in tables (3) and (4) revealed that using corpus in learning vocabulary had a great effect on students' progress and achievement. These results were corroborated by Varely(2008); Jafarpour et al. (2013); Kayaoglu(2013); Unaldi et al.(2013); and Ucar and Yukselir(2015), who found that learners made significant usage

of grammar and vocabulary after using corpus-based activities. The findings of the current study also echoes that of Jafarpour and Koosha (2006); who defined the role of data-driven learning in teaching collocation and indicated that participants who received data-driven instruction outperformed those who received conventional instruction in the learning of collocation.

Like the researchers in this study, Jafarpour et al. (2013) compared the effectiveness of corpus-based approach with the effect of the traditional approach. Findings revealed that corpus-based approach had a positive effect on the comprehension and production of collocations. Kayaoglu (2013) conducted a quasi-experimental research to explore the feasibility of a corpus to distinguish between close synonyms which have parallel meanings. The results showed that corpora can be effective tool in solving grammatical and lexical related problems. Unaldi et al. (2013) found that the group treated through decontextualized learning activities scored statistically high means compared to the other group. Likewise, the results of the present study goes in line with Ucar and Yukselir (2015)'s results that corpus-based activities has a significant impact on verb-noun collocation.

Furthermore, the results of the questionnaire indicated that the students were generally positive toward the use of corpus activities in learning vocabulary and grammar. They reported that corpus use was particularly beneficial for gaining and improving their vocabulary and grammar. As was found by Vannestal and Lindquist (2007) when investigating how the use of concordancing would affect students' attitudes toward grammar, using corpora was perceived as a helpful approach in grammar learning. Thus, having received more direct training in corpus use, learners could develop more adequate attitudes toward using corpus in their EFL learning. This result was also coincided with Yoon and Hirvela (2004); Chan and Liou (2005);Varely (2008); Girgin (2011); and Kayaoglu (2013); who detected that most of the learners agreed that corpus was very helpful for their language learning.

In addition, it is important to note that students identified some problems in using corpus. For example, the shortage of time available to analyze the data provided by corpora and the difficulty of analyzing concordance lines. The present study suggests offering sufficient time and adequate training in using corpus for language learning. This is consistent with the literature that has pointed out the need to train students for effective use of corpus to overcome any of these problems (St. John, 2001; Charles, 2007; Fuster & Clavel, 2010; Marinow, 2013; Yoon & Jo, 2014).

In conclusion, students who received instruction through corpus-based activities displayed higher performance in both grammar and vocabulary than those who studied using the traditional way. The study was significant in terms of using corpus-based activities in learning grammar and vocabulary and had a positive effect on student's attitude toward using corpus.

Conclusion

In this mixed-methods study, the researchers investigated the effects of using corpus-based activities on EFL tertiary students' grammar and vocabulary development. They also explored the attitudes of the students who received the treatment toward using corpus in learning grammar and vocabulary. The findings have revealed that the achievement of the students who were instructed grammar and vocabulary using corpus-based activities outperformed their counterparts who received traditional instruction; as indicated by the results of the grammar and vocabulary posttests. The data obtained from the questionnaire and the interviews have shown that the majority of the students hold positive attitudes toward using corpus in grammar and vocabulary instruction. The interviewees found corpus more helpful than dictionary in learning the usage and places of words in context. They stated that using corpus in learning grammar in interpreting rules is more interesting than checking grammar books. However, they experienced some challenges in logging in COCA, using proper search techniques to

get data, sorting and interpreting corpus results, and spending much time trying to understand the complexity of the language.

The study is limited to the small sample size, 104 participants, with a short-term treatment; one semester with a two-hour session per week, which makes the results of this study not generalizable to all EFL learners in tertiary level. In addition, the study focused only on teaching grammar and vocabulary through corpus-based activities. The corpus used in the study is COCA; as students searched BWC written and the Brown+BNC written to complete some corpus-based activities. Moreover, the grammar test examined students' knowledge of: singular and plural, word order, infinitive and gerund, conditionals, tenses (simple past, present perfect, and past perfect), and modal verbs; whereas the vocabulary test checked students' mastery of phrasal verbs, prefix/suffix, and only six types of collocations (adjective+noun, noun+verb, verb+noun, noun+verb, adverb+verb, and adverb+adjective). The topics and structures provided to the students in the experimental and control groups were restricted to these rules and structures; the only difference was in the instruction method. Finally, the study is limited to one post-administration of the questionnaire and the interviews as students indicated in the demographic part of the questionnaire having no idea or experience of corpus. The potential data gathered from more implementation of the questionnaire or interviews would enrich the study with differences in the students' attitudes toward corpus in different phases of the treatment.

This study provides some pedagogical concerns to EFL educators. First, in teaching vocabulary, it is important to familiarize students with various types of collocations rather than the ones they learn from the textbooks. Second, teachers should give students chances to recycle the newly learned collocations by creating follow-up activities or quizzes after each corpus-based lesson or ask them to use these collocations in their writing. Third, before asking students to independently generate their own corpus queries or interpret grammatical

rules from corpus data, they should be trained to formulate queries and interpret results under the guidance of the teacher. Adequate time and practice should be allocated to use corpus consultation by students of various levels of language proficiency. To achieve this goal, teachers should be trained to adequately use corpus and scaffold students to use different features of corpus. In addition, teachers should provide modeling, tutorials, mini-lessons, handouts or booklets, or multimedia digital files like podcasts. Finally, material designers should incorporate corpus-based learning in EFL curriculum and language programs. They should consider the challenges that face students when using corpus for lexico-grammar learning and provide strategies or alternatives to overcome them.

Further studies could explore the effectiveness of using corpus-based activities on students' autonomy and inductive learning not only in grammar and vocabulary instruction but also in teaching written and spoken language. More empirical studies are needed to investigate the long-term effects of learning grammar and vocabulary through corpus-based activities. It is also recommended to examine which grammatical structures and collocations can be best taught by corpus-based instruction. Future studies are required to investigate the correlation between students' attitudes towards corpus use and their language proficiency levels. It is also noteworthy to examine the effect of using corpus on students' motivation, self-confidence, writing apprehension, or self-efficacy. It would be informative to conduct research that compare the effect of using corpus to dictionaries or grammar books on students' recycling of language structures and expressions, correcting errors in written assignments, or solving their linguistic and writing problems.

References

- Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than been counting. *TESOL Quarterly*, 35 (5), 331-336.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and a training. *RECALL*, 21(1), 37-54.

- Chan, T.,& Liou, H. (2005). Effects of web-based concordancing instruction of EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231-251.
- Chang, W.,& Sun, Y. (2009). Scaffolding and web-concordancers as support for language learning. *Computer assisted Language Learning*, 22(4), 283-330.
- Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, 6, 289-302.
- Chujo, K., Kobayashi, Y., Mizumoto, A.,& Oghigian, K. (2016). Exploring the effectiveness of combined web-based corpus tools for beginner EFL DDL. *Linguistics and Literature Studies*, 4(4), 262-274.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34, 548-560.
- Fuster, M., & Clavel, B. (2010). Corpus linguistics and its applications in higher education. *Revista Alicantina de Estudios Ingless*, 23, 51-67.
- Girgin, U. (2011). *Corpus-based activities at lower levels of EFL proficiency: The effectiveness of using concordance lines on grammar learning* (Unpublished master's thesis). University of Bilkent, Ankara, Turkey.
- Halliday, M.A. (1992). Language as system and language as instance: The corpus as a theoretical construct. In J.Savartivik (Ed.), *Proceedings of Nobel Symposium 82 on Directions in Corpus Linguistics*, Stockholm, 4-8 August 1991 (pp. 61-77). Berlin: Mouton de Gruyter.
- Hanafiyeh, M.,& Keshi, A. K. (2013). The effect of corpus-based instruction and thesaurus-based teaching on Iranians EFL learners' grammatical knowledge. *Journal of Basic and Applied Scientific Research*, 3(2), 167-179.
- Jafarpour, A.,& Koosha, M. (2006). Data-driven learning and teaching collocation of prepositions: The case study of Iranian EFL adult learners. *Asian EFL Journal*, 8(4), 192-209.

- Jafarpour, A., Hashemian, M., & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies*, 3(1), 51-60.
- Kim, H., & Chun, S. (2008). Fostering lexis awareness and autonomy by corpus-based data driven learning. *English Teaching*, 63(2), 213-235.
- Kyaoglu, M. N. (2013). The use of corpus for close synonyms. *The Journal of Language and Linguistics Studies*, 9(1), 128-144.
- Marinow, S. (2013). Training ESP students in corpus use-challenges of using corpus-based exercises with students of non-philological studies. *Teaching English with Technology*, 13(4), 49-76.
- O' Keefee, A., Mc Carthy, M. & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press, Cambridge, UK.
- Perez-Liantada, C. (2009). Textual, genre and social features of spoken grammar: A corpus-based approach. *Language learning & Technology*, 13(1), 40-58.
- Rapti, N. (2010). *A study of classroom concordancing in the Greek context: Data-driven grammar teaching and adolescent EFL learners* (Unpublished doctoral dissertation). University of Nottingham, United Kingdom.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- St, John. E. (2001). A case for using a parallel corpus and concordance for beginners of a foreign language. *Language Learning & Technology*, 5(3), 185-203.
- Ucar, S., & Yukselir, C. (2015). The effects of corpus-based activities on verb-noun collocations in EFL classes. *Turkish online Journal of Educational Technology*, 14 (2), 195-205.
- Unaldi, I., Bardakci, M., & Akpınar, K. D. (2013). A comparison of contextualized, decontextualized and corpus-informed vocabulary instruction: A quasi experimental study. *Journal of Language and Literature Education*, 8, 78-95.
- Vannestal, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordance in a university grammar course. *RECALL*, 19(3), 329-350.

- Varley, S. (2008). I'll just look that up in concordance: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22(2), 133-152.
- Yoon, H., & Jo, J.W. (2014). Direct and indirect access to corpora: An exploratory case study comparing student's error correction and learning strategy use in L2 writing. *Language Learning & Technology*, 18(1), 96-117.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.

