The Effectiveness of Corpus-Based Approach on Vocabulary Learning Gains and Retention in Saudi Tertiary EFL Context

Ahmed Fathy Farag-Allah Youssef[1]* and Dr. Mahmoud Ali Ahmed Omar[2]

[1]*English Department, Sudan University of Science and Technology, Sudan/Department of English Language, Faculty of Humanities and Administrative Sciences, Buraydah Colleges, Saudi Arabia*

[2]*Associate Professor of Literature, Department of English Language and Literature, Faculty of Languages, Sudan University of Science and Technology, Sudan*

Corresponding Author: Ahmed Fathy Farag-Allah Youssef, E-mail: concordafg@gmail.com

ARTICLE INFO	ABSTRACT
	Corpus Linguistics has been noted to be a trendy approach in language research in the last few decades as it is closely related to technological advancements. The rising number of available online corpora and the advanced concordancing software have been urging and boosting researches on potentials and uses of corpora in language analysis and teaching. In this study the researcher tried to fill in the gap apparent in the Saudi EFL context of poor or lack of corpora use. Therefore, the researcher has investigated the effect size of using corpus-based approach on vocabulary gains and retention. He has employed quasi – experimental study to thoroughly investigate the study questions. A diagnostic pre-test, posttest and delayed posttests have been employed with both groups, the control which has been taught using conventional approach and the experimental which has been taught using corpus-based approach. At the end of the treatment test scores for both groups have been statistically analyzed to find out the differences. The results of posttests and delayed posttests showed a statistically significant differences in favor of the experimental group, which indicated the effectiveness of using corpus-based approach on vocabulary gains and retention rates. Consequently, the researcher has recommended integrating corpus pre-service and in-service training for EFL teachers. He has also suggested enhancing friendly user interfaces for available corpora to be easily accessed and used by teachers and learners. Finally, combining conventional tools alongside with corpus-based method is recommended in some instances to meet some emerging challenges with corpora use.

1. Introduction

It can be seen that there is a high degree of agreement that vocabulary learning is an integral part of mastering a second language. The best way to attain solid and successful vocabulary learning, however, is still controversial, partially because it depends on a wide range of elements (de Groot, 2006). Consequently, it is not surprising to find teachers and learners sometimes confused about the right way to achieve it, especially as common coursebooks and teaching materials have not offered straightforward sufficient guidelines and descriptions.

Undoubtedly, vocabulary is fundamental to language and vital to the average language learner (Zimmerman,1997); while teaching and learning are considered major concerns for both EFL teachers and learners, solid acquisition of second language vocabulary is of a particular necessity for English as a second or as a foreign language learner (EFL/ ESL) who used to acquire dead abstract lexicons throughout long years of formal study (Hunt & Beglar, 2005).

Zhan (2008) ascertained a profound notion in this regard through emphasizing the impact of adopting convenient teaching methods and materials on getting learners acquire much deeper impression and richer information about the target words and making it easier to be saved and retained in a long time period. In the same regard, Laufer (2001) claimed that most learners acquire new vocabulary effectively through exposure to language input especially reading input, rather than the deliberate sophisticated process of memorizing words. It is been believed that one of the essential and vital resources for providing this language input with distinct features of authenticity and abundance is corpora-based approach which considered a revolutionary teaching trend in EFL classrooms (Sinclair, 2004). Corpus has been defined by Crystal (1997) as *"a large collection of linguistic data, either written texts or transcription of recorded speech, which can be used as a starting point of linguistic description or as a means of verifying hypotheses about a language".* This collection of massive authentic data is electronically stored in readable and searchable forms which enable users doing a variety of linguistic analysis and description. The term corpus linguistics as a discipline has been branching into sub-disciplines and methods that mainly pertinent to language teaching, like corpus-based approach, Data Driven Learning, Corpus-based materials, concordancing, concordancers, …. etc. (Barbieri & Eckhardt, 2007).

In a classroom, Corpora can be utilized directly or indirectly in teaching English; i.e. indirectly in designing teaching materials and in syllabus design, while directly through the direct use of corpora inside or outside the classroom (Gavioli & Aston, 2001, Römer, 2011). In employing corpus-based approach learners are expected to be independent researcher and autonomous language learners to explore language patterns and cooccurrences, rather than passive knowledge recipients (Conrad, 2005). This study has been intended to investigate the effectiveness of employing this brand-new approach in vocabulary classrooms with university level students in Saudi Arabia, hoping to participate in innovating and enriching EFL teaching context in Saudi Arabia and coin an effective model for teaching and learning vocabulary. The first spark that initiated the researcher's interest in this topic is attributed to his successful and rich experience with teaching legal translation to languages department four years back using parallel corpora to teach legal texts and terminology.

1.1 Problem Statement

The researcher has noticed that a great bulk of recent studies on ELT have agreed upon the importance of learners' involvement, authenticity of language input, and contextualized vocabulary learning. Corpus linguistics has already claimed to provide authentic data input, enhance learners' autonomy, and promote learner-centered classroom. The first incident that initiated the researcher's interest in the study dated back to 2016 when he was teaching legal translation course to students of Languages Department and he was first using parallel corpora as a tool for teaching legal texts and terminology. Then the researcher started to preview some prominent studies on corpora utilization in teaching vocabulary and surprisingly found no studies (at that time 2016) conducted on the Saudi context except very few ones handled the idea of forming specialized corpora or checking or using corpora for linguistic analysis.

To fill in this gap, the researcher has designed this study to thoroughly investigate the effectiveness of this approach, corpus-based approach, on Saudi learners vocabulary learning and retention, hoping that this study contribute to raising awareness of Saudi EFL teachers and professional of exploiting this new approach in their classrooms, coursebooks and teaching materials.

1.2 Questions of the Study

The following research questions have been formulated:
1. What is the effect of using corpus-based approach on improving vocabulary learning in Saudi EFL classroom?
2. How far is corpus-based approach effective in improving vocabulary retention in Saudi EFL classroom?
3.

1.3 Hypotheses of the Study

The researcher has made the following hypotheses:
1- There are statistically significant differences, at the level P≤0.05, between the mean scores of post-tests for control group' learners that received conventional vocabulary instruction and those of the experimental group that received corpus-based instruction in favor of the experimental group.
2- There are statistically significant differences, at the level P≤0.05, between the mean scores of delayed post-tests for control group's learners that received conventional vocabulary instruction and those of the experimental group that received corpus-based instruction in favor of the experimental group.

1.4 Significance of the Study

To the researcher's knowledge this study is considered one of the very few studies to be conducted on corpora integration in the Saudi EFL classrooms. It could contribute to raise awareness of both ELT professionals and learners of the feasibility and efficacy of corpora integration in EFL classrooms. It has also suggested practical classroom activities and hands-on for corpus-based approach in vocabulary teaching that may benefit interested teachers.

2. Literature Review and Previous Studies

Learners should have opportunities to develop a feel for language and enhance their discovery learning through given direct access to corpora use as stated by Huang (2011). One of the recent corpus studies is Boontam and Phoocharoensil (2018) which was conducted on 30 Thai learners to investigate the effect size of employing corpus-based approach. The researcher applied a treatment for six weeks (one 50-minute class every week). Results of pre and posttests illustrated some improvements in vocabulary learning as a result of using corpus-based activities and exercises. Moreover, results of questionnaires and interviews revealed learners' positive attitudes towards using corpus approach. Another study, Al-Mahbashi et al (2016) which has been applied on 77 female university students in Sana'a University, Faculty of Education to investigate the effectiveness of DDL approach on writing skills. They have employed COCA Corpus (*Corpus of Contemporary American English*) (COCA) in creating the corpus-based hands-on and tasks for the experimental group. The results of pre and posttests analysis for both groups, experimental and control groups, revealed significant differences between the mean scores of their post test scores in favor of the experimental group which applied corpus-based approach. Similarly, the differences in delayed post test scores were significant in favor of the experimental group. the researcher concluded that corpus-based group outperformed its counterpart in both learning gains and retention rates. They concluded stating *"The long and short-term impact of DDL method could be attributed to learning the target words in rich contexts through numerous exposures to them that would subsequently lead to vocabulary acquisition (through what is called discovery learning".*

On the other hand, some studies investigated the use of corpora in learning other languages, rather than English, like Çelik and Elkatmiş (2013). They investigated the effect of corpus-based approach on learning Turkish punctuation marks compared to traditional lecture-based approach. They conducted an experimental study on 171 university students at Kirikkale University, in Turkey. They used an achievement test on punctuation marks. Results of post-tests have revealed significance differences between both groups in favor of the experimental group. Furthermore, based on the questionnaire responses and interviews the majority of experimental group's learners in revealed positive perceptions on studying punctuation marks through obtained concordances lines.

3. Methodology

The current study has followed an experimental method that is designed to offer answers to the main research questions and provide more information related to the potential vocabulary gains and retention rate after applying corpus-based approach. This type of research, quasi-experimental or experimental research designs essentially aim at investigating the casual relationship between two variables in the treatment, independent and dependent variables. For more specifics, the independent variable is the influential variable which has a direct impact on the dependent one (Loewen 2018). Thus, adopting the independent variable will inevitably cause a change or variation in the dependent variable. In this study, the independent variable was considered the use of corpus-based approach which was assumed to have direct impact on the dependent variable, experimental group's learners' post-tests scores.

Therefore, the design has included a placement test, a pretest, a posttest, and a delayed posttest in which all learners of both groups had to do. The purpose of both the placement test and the pretest is to ascertain and prove that the two randomly assigned groups were comparable before the treatment, while the immediate posttest measures the immediate vocabulary gains of the treatment. In addition to these tests, a two-week delayed posttest has been included to examine the effects of the treatment over the longer term. The researcher has substantially employed quantitative analysis on the collected data.

3.1 Research Procedures

The researcher has designed the procedures of this study in a way that provide reliable and validated data for testing the study hypotheses and answering its questions. These procedures can be summarized in the following steps. First, the researcher has randomly assigned two groups as the participants of the study; the experimental group, and the control group. The experimental group has been taught vocabulary lessons through using corpus-based approach and utilizing COCA

Corpus (Corpus of Contemporary American English) to enhance their learning experience, whereas the control group has been taught the same target vocabulary via traditional approach. Second, the researcher has designed and applied Cambridge General Placement Test, vocabulary pre-tests, post-tests, and two-week delayed post-tests on all learners of experimental and control groups to assess their vocabulary knowledge and retention before and after the treatment. Lastly, the quantitative data, taken from tests results, has been statistically analyzed using Statistical Program for Social Sciences Package (V.25).
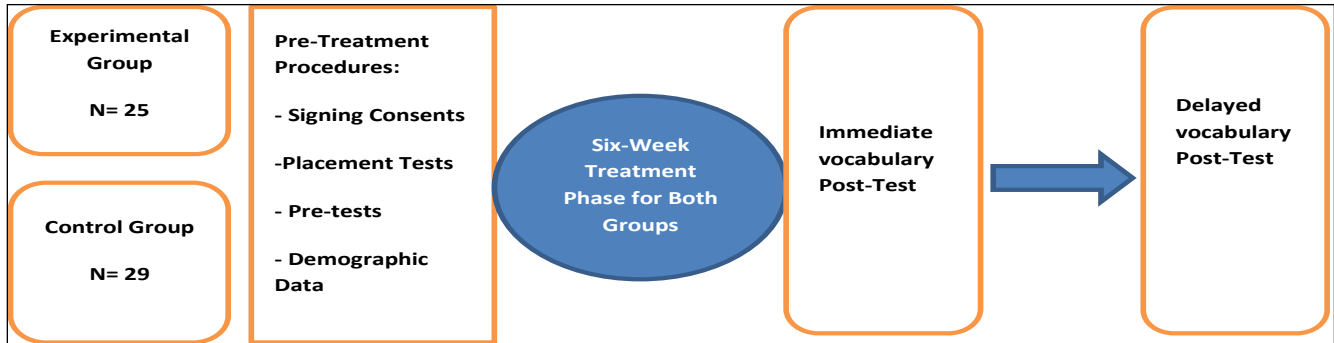
(See Figure 1: An Overall Research Design).

3.2 Study Population and Sampling

The study sample consists of 54 students studying foundation courses at Humanities and Administrative Sciences College, at Buraydah Colleges in Saudi Arabia. They are mostly 18 to 22 years enrolled in foundation courses to pursue their studies at one of four available programs, Accounting, Law, Human Resources Management, and Business Administration at Buraydah Colleges. Participants have been randomly placed into the control and the experimental condition following a chance procedure (Gravetter & Forzano, 2018; Hatch & Lazaraton, 1991; Kirk, 2009; Loewen & Plonsky, 2016; Nunan, D. & N. David, 1992). One class has become 29 students and has been assigned as the control group (traditional approach for teaching vocabulary), and the other one (*n=25*) has been assigned as the experimental group that received a different treatment (corpus-based teaching for vocabulary). Both groups have been taught by the researcher who has been teaching general English course, reading and writing sections, for 10 hours a week for each, at least 5 hours of them have been devoted to teaching vocabulary. Students enrolled in the experimental group have been informed about the research and signed a consent form to participate in the experiment. In the meantime, they have been told that they have the right to participate in this experimental research or to withdraw from the experiment if they feel uncomfortable at any time without any consequences.

3.3 Data Collection Tools

The substantial assigned tool to measure the effectiveness of using corpus-based approach vs traditional approach on vocabulary learning and retention is vocabulary achievement tests. The researcher has applied four tests on participants of both groups, 54 students. At the first week, the researcher has applied two tests, CPT and vocabulary pre-test, to measure their proficiency levels and vocabulary background respectively. After six weeks of study for both groups, an immediate post-test has been applied to measure their gains after the treatment. The main aim for this test is identify whether or not there are statistically significant differences in gains between the two groups after going through six –week treatment and having been exposed to different teaching approaches, traditional approach with the control group vs corpus-based approach with the experimental group. After a two-week interval, another delayed post-test has been applied on both groups to measure the differences, if any, in retention rate for both groups.

The tests were organized into two main equal categories, receptive or definitional knowledge and productive or transferrable knowledge; each part has 25 points out of 50. Both tests were identical, except that the items in the post-test and delayed post-test were re-arranged to dismiss the effect of memorizing the answers. However, both groups' students have not been told that the pre-test will be taken again as a post test and a delayed post-test. Tests are mainly based on the assigned target vocabulary in the three chapters in *Interactions 1 Reading Book*. The main themes of these tests were assessing students' perceptive and productive knowledge of the target vocabulary. (Figure 2 and Figure 3 for samples of vocabulary tests):

A. Match each word or phrase in Column (A) with its definition or meaning in Column (B):

NOTE: EACH WORD equals ONE POINT and *there are TWO EXTRA answers in Column (B)*.

COLUMN (A)	COLUMN (B)
1. Excuses [ ]	a. Blood vessel attack in the brain.
2. Opportunities [ ]	b. The air around the earth.
3. Miss [ ]	c. Of enough value for the money.
4. Memorable [ ]	d. Explanations that aren't good reasons.
5. Worth the cost [ ]	e. Fail to have.
6. Miserable [ ]	f. Worth remembering.
7. Attend college [ ]	g. Chances to do something.
8. Get along [ ]	h. Be friendly with someone.
9. Atmosphere [ ]	i. Go to university.
10. bio meteorologist [ ]	j. The researcher who studies health and emotions in response to atmospheric conditions.
11. a stroke [ ]	k. Very unhappy or uncomfortable.
12. moods [ ]	l. Illnesses like pneumonia. Asthma, headaches, etc.
13. extreme weather [ ]	m. Examples: tornadoes, storms, floods, hurricanes.
14. diseases [ ]	n. A synonym for the word feelings
15. universal [ ]	o. Another word for "worldwide".
	p. A blood sugar disorder.
	q. Prepared items from an inexpensive restaurants, snack bars, or food stands.

Figure 2: Samples of Receptive Vocabulary Question.

A. Read the following concordance lines (sentences) and use one of the following words to fill in the gaps in each, according to the provided context: *1 Point for Each*

(Health disorders - blizzards – weather - in contrast to - bio meteorologists – nutritionists – diabetes – probably – influence – privacy – interests – slowly - academically – attend – worth the cost)

1. In that private schools, students were supposed to do a little better, and their parents to feel that this was ——————————————————, there's a business to be had there. #
2. violence and insecurity had become " normalized " and that " mental ———————————————————— such as anxiety, depression, or post-traumatic stress " could manifest themselves once the..
3. Though preliminary, scientific studies back up the notion that weather affects your health, says a ————————————————————, ,professor at Utah State University in Logan
4. Patient lose weight and hand him a diet plan – or refer him to a ———————————————— . At the Family Health clinics,
5. I agree to the beliefnet.com terms of service, rules of conduct and ——————————— —————————— policy (the " agreements ").
6. ice has led to unusual warming of the Arctic atmosphere, that in turn impacts ———— —————————————conditions in the northern hemisphere
7. the day began cold and blustery, ————————————————almost two weeks of uninterrupted sunshine
8. around the world will be forced to face flooding and drought in the summer and ————————————————— in the winter
9. for funding from the American Government to get a free ride to ————————————

Figure 3: Samples of Productive Vocabulary Questions.

3.4 Tests Reliability and Validity

The topics and the relative weight of the questions in the vocabulary achievement tests have been equally selected from the three chapters. Collocations, contexts, and language chunks have been received much focus. The reliability of the vocabulary achievement test has been tested through applying the test on another sample of 15 male and female students early before the treatment. Test-retest method has been applied on this sample then reliability coefficient factor has been calculated. According to the tests and retests scores of the pilot sample, the reliability coefficient was 0.828; which indicates a good level of reliability. In addition, the researcher found that the Pearson inter-rater reliability was relatively high, 0.906.

For validity issues, the research has consulted 7 experts in the fields of ELT, Linguistics, Curricula Design and Research Methodologies to check the validity of all the instructional materials, assessment procedures, learning objectives, lessons plan, and tests before getting started in the application process. First, the researcher received some valuable suggestions and feedback from the jury and accordingly he did some changes to the tests and the activities. Second, another draft has been sent back to the jury for their final assessment and remarks. As it has been shown in the following table (Table 6 Summary of Validation Results), there was a high percentage of satisfaction among the jury on the appropriateness of the proposed teaching tools, assessment tools, teaching plans, assigned coursebook and assigned corpus, tests, and designed activities and tasks for both groups.

Table 1: Teaching Plans and Materials, and Tests Validation Results

Study Tools and Criteria of Assessment	Minimum	Maximum	Mean	St. Deviation
1. Teaching Plans: Criteria	**-0.5**	**1**	**0.6428**	**0.5149**
A. Fit the Time Framework	-0.5	1	0.5714	0.612
B. Appropriate to the Levels	0	1	0.5	0.651
C. Equal Content for both Groups	0	1	0.5714	0.606
D. Measurable Objectives	0.5	1	0.9285	0.188
2. Assessment Plans: Criteria	**-0.5**	**1**	**0.7321**	**0.2672**
A. Convenient in Respect to the Time Table	-0.5	1	0.5	0.4082
B. Appropriateness of Assessment Tools	1	0.5	0.7142	0.2672
C. Representativeness for The Proposed Objectives	1	1	1	0
D. Equality between the Two Groups	0	1	0.7142	0.3933
3. Activities and Tasks: Criteria	**-0.5**	**1**	**0.75**	**0.2868**
A. Achieving the planned Objectives	0.5	1	0.8571	0.2439
B. Fit the Adopted Approach for Each Group	0.5	1	0.7857	0.2672
C. The Variety of Types and Forms	0.5	1	0.9285	0.1889
D. Promoting Students' Participation and Passion	-0.5	1	0.4285	0.447
4. Pre-tests and Post-tests: Criteria	**-0.5**	**1**	**0.7714**	**0.3296**
A. Diversity of Used Questions	1	1	0.7857	0.3933
B. Representativeness of the Target Vocabulary	-0.5	1	0.4285	0.6726
C. Suitable Length and Time for the Test	-0.5	1	0.7142	0.3933
D. Meeting Different Achievement Levels	0.5	1	0.9285	0.1889
E. Convenient Distribution of Marks.	1	1	1	0

N.B.: *Strongly Agree = 1   Neutral = 0   Disagree = -0.5*
 Agree = 0.5   Strongly Disagree = -1

The mean scores of the Jury assessment for the four illustrated elements were relatively high as it ranges from *Agree* (0.5) to *Strongly Agree* (1). The highest validity score (0.77) was given to the criteria of Pre-tests and Post-tests then immediately came the validity of designed activities and tasks (0.75), while assessment plans and teaching plans were (0.73) and (0.64) respectively. It is noteworthy that the majority of the jury (6 out of 7) highly valued the appropriateness of the planned objectives in the teaching lesson plans and the diversity and purposefulness of the proposed activities and tasks especially for corpus-based group. On the other hand, three of them negatively rated the appropriateness of corpus-based activities in promoting student's passion and participation. Some of them justified that most of these activities were not familiar to the students and they might not be enthusiastic to try new things. Overall, the research has already applied some changes based on the initial feedback from the jurors in lesson plans, activities and tasks designs, and tests. Therefore, the overall ratings of the jury for the final drafts of instructional tools and plans were relatively high and positive.

3.5 Data Analysis

At the end of the treatment, in the ninth week, the researcher has analyzed the tests scores for placement, pre-, post-, and

delayed post-tests for both groups, experimental and control to test the main hypothesis of the study. To get a detailed and authentic answer for the first and main study's question: First the researcher measured the difference between pre-test and post-test scores for each individual group and ran a t-test for each group separately, One Sample T-test, to measure the gains of each group after the six-week experiment. Second, the researcher ran Independent Samples T-test to compare the mean differences of corpus-based group's post-test scores to the conventional group's scores to test if there were significant differences in the final gains between them. Finally, he ran Independent Samples T-test to compare the mean differences of delayed post-test scores to detect any significant differences in retention rates.

3.6 Instructional Procedures

For experimental group, the instructor has used corpus-based instruction and corpora-based materials to present target vocabulary in each assigned reading passage and to reinforce previous learnt ones. Teaching has been happening in English Lab to enable learners to instantly access the assigned online corpus and be participating in corpora activities prepared by the researcher, whereas for the other control group, teaching has been happening in the traditional classroom using a whiteboard and sometimes slideshows projection. All computers in the English Lab are equipped with internet connection and wireless network for smartphones and personal laptops to help learners' instant access to online corpora. In addition, the researcher has employed Edmodo online platform for homework assignments, class announcements, quizzes and teachers' feedback on group work assignments submissions. Both Classroom instruction or feedback and follow-up have been available equally and fairly for participants of both groups on a daily basis.

Early orientation and training sessions have been held for participants of the experimental group to be familiarized with using COCA Corpus tools and searching options and be introduced to concordancing skills and reading obtained concordances lines. As participants not only received training on the technical use of the corpus but were also given guidelines on how to read and analyze the concordance lines, following Sinclair's (*Reading Concordances*) model. This enhanced their performance in the answering activities. As seen in previous research, when participants receive technical guidelines without training on how to read and analyze the concordance lines, or if they undergo a short training period, the result is either insignificant performance or less positive attitudes (e.g., Pérez-Paredes et al., 2012; Cobb, 1999). Lastly, at the end of the course, participants of both groups had to acquire the same target vocabulary for each chapter whatever the employed teaching approach is (See tables 2&3 for an overview of the plan).

Table 2: An Overview of the Classroom Plan.

Domain	Course Learning Outcomes (CLOs)	Teaching Strategies and Learning Resources	
		Control Group	**Experimental Group**
1. Knowledge	1.1 Identify the meaning, definitions, and synonyms for the target words. 1.2 Identify the appropriate collocates for some target vocabulary in real contexts. 1.3 Recognize the correct spelling, pronunciation, part of speech, and use of the target vocabulary in each chapter	• Using the Traditional Classroom with its normal facilities to deliver lessons. • Traditional lecturing approach and other context and definitions strategies will be used for words teaching	• Using the Department's Computer Lab to deliver all the vocabulary classes using available PCs and internet connection. • Discovery learning and autonomous problem-solving approaches will be used to learn target words.
2. Cognitive	2.1 Utilize new learnt vocabulary in authentic contexts and personal situations. 2.2 Demonstrate high level of awareness for words frequency in different genres. 2.3 Utilize word clusters rather than single isolated words.	• Using the assigned Course book, Interactions 1, Reading Section, by *E. Hirn and Hartmann*, to teach the target vocabulary in the first three chapters. • Traditional lecturing approach and other	• Using the assigned course book for teaching reading strategies only, and using COCA Corpus to learn the target vocabulary in each chapter. • Using printouts and hands-on mainly derived from COCA Corpus Concordance

	2.4 Employ words collocation for producing spoken or written utterances.	context and definitions strategies will be used for words teaching	lines and queries data. • Discovery learning and autonomous problem-solving approaches will be used to learn target words.
3. Interpersonal Skills and Responsibility	3.1 Students are able to accomplish the assigned tasks and appropriately fulfill their role in their team, contribute to others' work, and solve their team problems.	• Using work sheets, slides and activities mainly based on the course book activities and exercises. • Students are given pair work and group works tasks based on the course book activities and exercises.	• Using printouts and hands-on mainly derived from COCA Corpus Concordance lines and queries data. • Students are given pair work and group works tasks based on the online use of corpora searching tools and concordancers.
4. **IT and Communication Skills**	4.1 Students are able to use different forms of IT tools appropriately to assist learning and researching, and to give presentations. 4.2 Students are able to use online dictionaries and online	• The teachers will be using the available projector and personal lab top in the class to give them presentations and demos in using online dictionaries.	• All students will be using the available PCs or their own smart phones to get direct access to online corpora use and doing queries about the target vocab.

Table 3: Covered Chapters and Target Vocabulary in Each.

No.	Topics to be Covered	Number of Assigned Vocabulary/ per week	Week No.	Number of Hours
1.	Orientation for the Study, Placement Tests, and Pretests	---	1st	4 hr.
2.	Chapter 1: Academic Life around the World.	20	1st	4 hr.
3.	Chapter 1: Academic Life around the World.	41	2nd	6 hr.
4.	Chapter 2 Experiencing Nature.	16	3rd	6 hr.
5.	Chapter 2 Experiencing Nature.	20	4th	6 hr.
6.	Chapter 3: Living to Eat, or Eating to Live.	17	5th	6 hr.
7.	Chapter 3: Living to Eat, or Eating to Live.	12	6th	4 hr.
8.	**IMMEDIATE POSTTEST**	----	6th	2 hr.
9.	**INTERVAL**	---	7th	--
10.	**INTERVAL**	---	8th	--
11.	**DELAYED POSTTEST**	---	9th	2 hr.
12.	**Total Weeks and the Teaching Actual Hours**	**116 words**	**6 weeks**	**40 hours**

The researcher adopted a gradual approach for getting students, of experimental group, engaged in corpus-based learning, starting with detailed workshops on using corpora in vocabulary learning, specifically COCA corpus, then moving a step forward through using worksheets, hands-on, and materials based mainly on concordances lines from COCA on the target vocabulary, and the last step is learners' direct access to COCA. The teacher will be the key in planning and designing the individualized classroom tasks and the appropriate mode of corpora integration, basically depending upon the target objectives and content of each chapter and in the light of his students' language capacity to handle these objectives. As the researcher mentioned in the previous paragraphs that it will be a simplified gradual process in which all learners will be smoothly introduced to the corpora, COCA Corpus, through three stages process. These stages can be broadly named introduction and orientation, Teacher's corpora-based activities, and lastly learners' direct access to corpora and individualized tasks. (See figures 4&5 for sample corpus-based exercises and tasks)

A. **Read the concordance lines below and answer the following questions.**
Attitudes:

approaches to difficult problems , focused specifically on	attitudes	about	crime	. the findings can be used to understand the
you . It 's just couches ! " # On her	attitude	about	money	. One of the big things I loved about growing
n't really make that much difference in terms of their own	attitudes	about	the nominee	. MR-WARREN : It made a great deal of
' use of homophobic epithets . Group-level homophobic	attitudes	and	bullying behavior	were included at Level 2 to predict the
. " We 've seen remarkable changes in the kids '	attitudes	and	confidence levels	, " says Pam . " Our introvert ,
Quick to support . Fine range in zone coverage . Great	attitude	and	effort	. Can get separated from a man-to-man coverage .
urge to relieve suffering toward revising the static , habitual	attitudes	and	norms that	created it . Zen koan study , contingency "
knowledge of the harmful effects of noise actually affect the	attitudes	and	practices	. Study design # This is a cross-sectional
mid-1980s , and some people assert that spouses fulfilled their	attitudes	and	responsibilities for	decades before that . However , the

1. What does the word "attitude" mean?
2. What part of speech is it?
3. What are the prominent collocates in these sentences?
4. Tick the appropriate synonyms for the word attitudes? View, stance, opinion, feeling, richness, academic.
5. Write a sentence on your own using this word?

Figure 4: Samples of Printouts for Corpus-based Experimental Group

GROUP WORK TASKS – HOME ASSIGNMENTS – SAMPLE A

Complete the table below, following the stated instructions:

- Work in GROUPS, you and your PARTNERS together to fill in the table.
- Go to the website (www.english-corpora.org/COCA/
- Search for the words (low carb diet – nutritionists – diabetes – fast food – desserts)
- Use Key Word In Context (KWIC) Type of search.
- AS A LAST OPTION, you can use the available links for bilingual dictionaries on the website

Words/ elements	low carb diet	nutritionists	diabetes	Fast food	desserts
Part of speech					
meaning					
Collocations before					
Collocations after					
synonyms					

Figure 5: Samples of Printouts for Corpus-based Experimental Group

On the other hand, learners of the control group have their assigned context for the target vocabulary in each chapter, two reading passages of about two pages each, besides using traditional dictionaries and exercises set by the course book designers who basically set one single topic for a set of words. A detailed two course descriptions for the specific content of reading course has been prepared by the researcher to be followed during the study, one for the experimental group and the other for the control group. Each description is completely aligned with the required learning objectives and content set by the department of Foundation Courses in Burayadah Colleges, whereas there are some differences in methods of teaching and learning tools (See *Table 4 Similarities and Differences Between the Two Sample Groups*).

Table 4: Similarities and Differences between the two Sample Groups

Similarities in the Experiment	Variances in The Conditions of The Experiment		
1. The Teacher	**Features**	**Experimental group**	**Control group**
2. The Syllabus and Lessons	**Method**	Corpus-based approach	Conventional Approach
3. Target Vocabulary	**Classroom**	Language laboratory	Traditional classroom

4. Allocated Time 5. Comparable Numbers and Levels 6. Placement Test 7. Pretest/Posttest &Delayed test	**Materials**	Handouts with concordance lines	Worksheets based on coursebook activities
	Activities	Hands-on & paper-based	Paper-based only
	Teacher's roles	Supporter and facilitator Monitor Tasks-planner	Knower Lecturer Administrator Presenter

4. Results and Discussions

### *4.1 Comparability of the Assigned Groups*

It was crucial to make sure that any variation in research results between groups could not be attributed to variations in the participants' English proficiency levels or vocabulary knowledge backgrounds. To address these two issues, two commonly used and freely available tests were administered to the participants: Cambridge General Placement Test (CPT) and Vocabulary achievement pre-test prepared by the researcher based on the target vocabulary. It was thus necessary to have empirical evidence that the two assigned groups are comparable and there are not any statistically significant differences between the mean scores of their placement test and pre-test scores.

In order to validate this claim, all participants in each of the two groups have been given a placement test CPT (Cambridge General Placement Test) at an early stage of the study. The scores of both groups, experimental and control, have revealed diversity in learners' proficiency levels. On one hand, 64.2% of control group's learners are categorized as intermediate, upper-intermediate and advanced levels and only 35.6% are categorized as elementary and pre-intermediate levels. On the other hand, 60% of experimental group students are categorized as intermediate, upper-intermediate and advanced levels and only 40% are categorized as elementary and pre-intermediate levels. The mean score for the experimental group is 28.2 out of 50 while it is 25.8 out of 50 for the control group as it is shown in details in *table (5)*.

Table 5: Placement Test Scores Descriptive Statistics for Both Groups.

Test	Group	N. of Sts.	Mean	Std. Deviation	Std. Error Mean
PLACEMENT TEST	Control	25	28.20	8.441	1.688
	Experimental	28	25.79	7.997	1.511

Though there have been slight differences between the scores of both groups as it has been just shown, we don't know whether they accordingly affect the treatment or don't. Therefore, the researcher has run independent samples T-Test to prove whether these differences in proficiency levels, placement test scores, statistically significant, at the level $P \leq 0.05$, that may have direct impact on the treatment results or NOT. It has been shown in table (3.7) that P value is 0.29 which is much greater than the significance level ≤ 0.05. The above shown results proved that the slight differences in proficiency levels are not statistically significant and thus warranted the comparability of the randomly assigned two groups, the control and the experimental and supported dismissing the variable of groups' proficiency levels from affecting the scores of vocabulary achievement post-test at groups level. (*see figure 6* for learners' categorization)

1
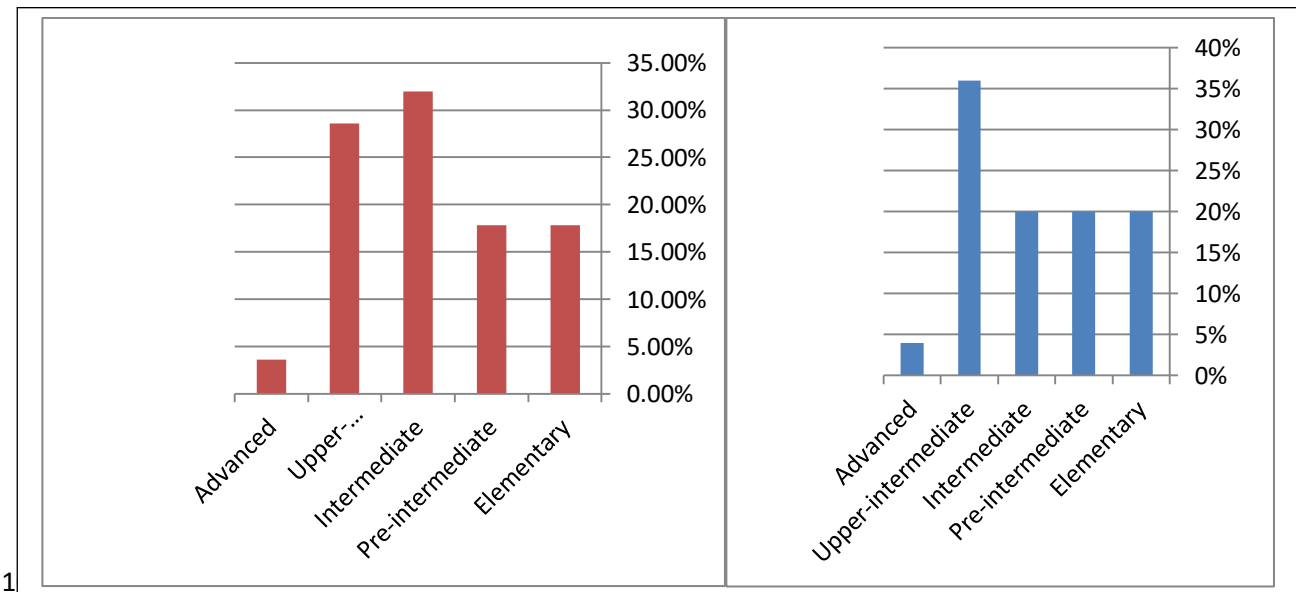
4.1.1 Vocabulary Achievement Pre-Test

An achievement pre-test has been given to students of both groups, experimental and control groups, as a second tool to test the comparability of both groups besides being a baseline for measuring the resulting gains after going through the treatment. For the experimental group, the students achieved the mean score of 23.48 out of 50 whereas the control one achieved the mean score of 20.14 out of 50. Though they are not identical means, it has been proved to be NOT statistically significant through running independent samples T-Test. As the P value was 0.17 which is greater than the significance level ≤ 0.05. (*See Table 6*)

Table 6: Vocabulary Pre-Test Scores Descriptive Statistics

Group Statistics for Pre-test Scores					
TEST	**group**	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
PRETEST	Control	25	23.48	8.377	1.675
	Experimental	28	20.14	9.272	1.752

4.2 QUESTION ONE: What is the effect size of using corpus-based approach on vocabulary learning gains?

According to the means of the pre-test results, both groups performed at a very similar rate (Experimental pretest mean score = 23.48, Control pretest mean score= 20.14, and *P value=0.177*). These results were meaningful, as they revealed that there was no significant discrepancy between the groups pre-experimentally in terms of their pre-learning backgrounds. After confirming and proving the comparability of the two assigned groups and setting an equal baseline for any further measurements of vocabulary gains, the researcher has measured the gains of each individual group by analyzing the immediate posttests scores. Data collected from this test have been analyzed by running One Paired Sample T-Test on pre and immediate post-tests scores of each group separately. For the control group, Table 9 reveals that the pre-test mean score is 20.14 whereas the posttest mean score is 32.21 which indicates and average of 12.07 points of gains. To test the significance of these gains One Paired Samples T-test has been run and revealed that *P =0.00* which indicated a statistically significant difference at the level of (*P ≤ 0.05*) between pre and post-tests scores (*See Table 7*).

Table 7: Control Group's Pre and Post Tests Scores Descriptive Analysis

TEST		Mean	Mean Gains	N	Std. Deviation	Std. Error Mean
Control Group	PRETEST	20.14	**12.07**	28	8.936	1.752
	POSTTEST	32.21		28	9.272	1.601

Similarly, For the experimental group, descriptive statistics in *Table 8* reveals that the pre-test mean score is 23.48 whereas the posttest mean score is 40.28 which indicates and average of 16.8 points of gains. To test the significance of these gains One Paired Samples T-test has been run on pre and post -tests scores and revealed that *P =0.00* which indicated a statistically significant difference at the level of (*P ≤ 0.05*) between pre and post-tests scores.

Table 8: Experimental Group's Pre and Post Tests Scores Descriptive Analysis

TEST		Mean	Mean Gains	N	Std. Deviation	Std. Error Mean
Experimental Group	PRETEST	23.48	**16.8**	25	8.377	1.675
	POSTTEST	40.28		25	6.374	1.275

To sum up, these results proved that vocabulary learning gains of both teaching methods, corpus-based and conventional approaches, were significant and clear in learners' immediate posttests scores. Regarding the last and the main step of this analysis, which is measuring the differences in gains between the two groups and identifying the effectiveness of using corpus-based approach in vocabulary learning with the experimental group learners. As it can be seen in *Table 9*, the mean scores of immediate posttests are 40.28 and 32.21 for the experimental group and the control group respectively whereas the mean gains scores are 16.8 and 12.07 for the Experimental and the control groups respectively. Although these mean posttest scores and gains reveal differences of vocabulary learning gains in favor of the experimental group (corpus-based approach), we still need further variances analysis to identify their statistical significance. For this reason, the researcher has run an Independent Samples T-test on immediate posttests scores of both groups to figure out the significance of these differences at the level (*P = 0.05*). As it is shown in *table 10* below, the P value is 0.000 which indicates a high level of significance for the differences in vocabulary learning gains in favor of the experimental group (corpus-based approach). This result has answered the first and main question and met the first hypothesis of the study which stated *"there are statistically*

significant differences in vocabulary learning between the control group which received conventional vocabulary learning and the experimental group which received corpus-based vocabulary learning in favor of the experimental group."

Table 9: Experimental and Control Groups Posttests Scores Descriptive Statistics Analysis

Tests	group	N	Mean	Std. Deviation	Std. Error Mean
POSTTEST	Experimental	25	40.28	6.374	1.275
	Control	28	32.21	8.469	1.601

Table 10: Independent Samples T-test for Post-test Scores of Experimental and Control Groups

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
								95% Confidence Interval of the Difference	
TEST / PARAMETER		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
POSTTEST	Equal variances assumed	1.219	.275	3.879	51	.000	8.066	2.079	3.892	12.240
	Equal variances not assumed			3.942	49.645	.000	8.066	2.046	3.955	12.176

4.3 QUESTION TWO: How far is corpus-based approach effective in improving vocabulary retention in Saudi EFL classroom?
In answering the second question of the study the researcher has hypothesized that there are statistically significant differences, at the level P≤0.05, between the mean scores of delayed post-tests of control group that received conventional vocabulary instruction and those of the experimental group that received corpus-based instruction in favor of the experimental group. To test this hypothesis the researcher has analyzed the collected data of delayed posttests scores (that measured vocabulary retention of learners) for both groups. As it's shown in *Table 11*, the mean scores of delayed posttests are 32.82 and 41.52 for the control and experimental groups respectively with a standard deviation 10.883 for the control group and 10.564 for the experimental one. Even though there are clear differences in mean scores of delayed posttests between both group in favor of the experimental group, another variance analysis, Independent Samples T-test, has been run to prove this claim. *Table 12* shows that *P value = 0.005* which indicates that the differences in mean scores of delayed posttests are statistically significant in favor of the experimental group. Therefore, the second hypothesis proposed by the researcher has been met and proved that suggested the effectiveness of corpus-based approach over the conventional approach in enhancing vocabulary retention rate.

Table 11: Descriptive Statistics Analysis of Delayed Posttest Scores for Both Groups

Tests	Group	N	Mean	Std. Deviation	Std. Error Mean
DELAYED POSTTEST	Experimental	25	41.52	10.564	2.113
	Control	28	32.82	10.883	2.057

Table 12: Comparing Variances of Delayed Posttest Mean Scores (T-Test)

TEST / PARAMETER		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
DELAYED POSTTEST	Equal variances assumed	.186	.668	2.945	51	.005	8.699	2.954	2.769	14.628
	Equal variances not assumed			2.950	50.627	.005	8.699	2.949	2.778	14.619

To recap, the above shown results confirmed and proved that learners of the experimental group who received corpus-based instruction and learning activities outperformed their counterparts in the control group in vocabulary learning and retention rates. These results are aligned with many other recent studies (Alruwaili, 2018; Elsherbini, S., and Ali, A., 2017; Alharbi. R., 2017; Almutairi, N., 2016; Almahbashi, A. et al, 2015; Edo-Marzá, 2014; Varley, 2008; and others) that proved the effectiveness of integrating corpora in enhancing vocabulary learning and retention.

5. Implications and Suggestions

The findings indicated that there were statistically significant differences at the level (p≤0.05) between the mean scores of post-test and delayed post-test scores for both groups in favor of the experimental group (corpus-based group). This finding is in alignment with the majority of corpora literature which proved the effectiveness of corpora utilization in vocabulary learning and retention.

This conclusion could be elaborately justified in the light of Cobb's (1999) phenomenal study on corpora use. Cobb (1999) concluded that differences in the definitional knowledge of vocabulary were marginal and nonsignificant, while the experimental groups' learners who received concordances-based hands-on and tasks showed higher ability in transferring this knowledge into new contexts and retaining it to much longer periods. On the other hand, students who used wordlists or dictionaries could have strong definitional knowledge of vocabulary items and even this knowledge was not well retained in a longer time span (Cobb, 1999, P.30).

Consistently, another prominent study was Nation's (2001) and Nagy's (2007), they concluded that the abundance of contexts and word encounters could facilitate and reinforce word's acquisition and retention. Consequently, vocabulary is best learnt and retained when being learnt through abundant authentic contexts with or without deliberate word-meaning teaching. This can be attributed to the notion of multifaceted nature of words learning that can be covered through wordlists or dictionaries. Knowing a word is not just identifying its meaning, it also includes many linguistic and metalinguistic aspects, such as, collocations, colligation, part of speech, synonyms, prosody, morphology, phonology, spelling, frequency, register, genre, and semantics or connotation. Most of these facets are less likely to be manifested and grasped through conventional tools, though using rich authentic concordance data and a variety of query tools and options could help in recognizing and understanding these aspects inductively or deductively.

Going beyond the benefits of providing abundant authentic data and recurring words encounters, some researchers highlighted the benefit of learner's autonomy when using corpus-based approach. This approach encourages learner's active participation and student-centered learning style. Chen (2004) claimed that a learning process in which learners have to learn knowledge prepared only by teachers, is considered a passive way of learning. Conversely, corpus linguistics enables learners to take control of their own learning to achieve the required tasks through identifying query types, choosing relevant concordances, noticing words in their contexts, identifying frequencies, word collocations, clusters, inferring patterns, inferring grammatical rules and so on.

13

To recap, the greater vocabulary gains and the longer retention of the experimental group's learners have been proved to be as a result of the use of corpus-based approach. This effect size of corpora utilization could be interpreted by its features of providing abundant authentic contexts, plenty of word encounters, enhancing learner's autonomy and active participation in content selection and learning processes, and manifesting linguistics and meta-linguistic aspects of a word. All these features could by far consolidate vocabulary knowledge acquisition and retention.

Given the results of this study, the researcher has formulated the following suggestions for an effective corpus use in the classroom:

1. Krashen (2009) stressed on the providing comprehensible input as an essential condition for effective learning. However, it occasionally happens with learners, especially low levels, to encounter unfamiliar contexts or incomprehensible outputs that may cause frustration and distract learning process. Thus, the researcher suggested that teachers could prepare paper-based hands-on and materials including comprehensible inputs and simplified samples that fit learners' proficiency level. This comprehensible and proficiency-fit input are suggested to be one level higher than the learner's level.

2. In traditional strict corpus-based method or DDL, teachers' role is restricted to monitoring and planning, though, in this study the researcher has adopted a more flexible mode which ensured teacher's essential role in facilitating the learning process of students and making sure that they are able to effectively use the corpus and infer language patterns from obtained concordance lines before letting them work alone and have direct access to corpora. In this proposed model, teacher's involvement in every stage and every procedure is fundamental till students gain the required skills for corpus use and concordances reading and be able to screen out irrelevant contexts.

3. The researcher assumed that corpus-based approach can be used solely or in conjunction with other conventional methods according to the intended learning outcomes for the lesson. Consequently. A model that combine both concordances-based and conventional approach in an integrated way could be the best fit for the Saudi EFL context and more most applicable for various learning objectives. Adopting such method could maximize corpus benefits, fit a variety of lesson types, and meet potential challenges in one-way corpora integration.

4. It's evidenced that learners' sufficient and effective training on corpus use and concordances reading is fundamental and indispensable for a successful corpus-based application. The researcher has noticed, in this treatment, that acquiring the technical skills of online corpora use happened easily and quickly, while the more sophisticated and complex concordancing skills of noticing, determining, grouping, analyzing, interpreting, and inferring took much longer time, carefully planned training, and adequate practice. Accordingly, learners' corpus training has to be carefully prepared in a gradual enriched way to cover both technical skills and the most importantly concordancing skills.

5. The cornerstone of the effective use of corpus-based approach is a well-trained teacher that has the required skills for introducing and implementing such brand-new approach in his classroom. Thus, equipping teachers with adequate training on all aspects of corpus use is inevitable for effective adoption of this approach in EFL context. The proposed model of training, whether pre-service or in-service, would include technical skills of online corpora use, pedagogical skills of dealing with obtained concordances lines, pedagogical skills of corpus-based lesson plans, pedagogical skills of designing corpus-based hands-on and designing corpus-based teaching materials and tests.

6. Recommendations for Further Study

A highly recommended corpus study could be for investigating the changing roles of teachers in a corpus-based classroom compared to the conventional roles in traditional classrooms. This study could specifically identify teachers' roles and responsibilities in each stage of a corpus-based classroom. Another aspect of corpora that can be probed is designing a proposed model for learners training on both technical skills on online corpora use and concordancing skills required for analyzing and interpreting obtained query data. This study would aim at designing training plans, teaching materials, instructional guidelines, corpus-based hands-on, and practice exercise. The third recommended is designing a curriculum for teacher' training that achieve all embedded skills in corpus-based approach application. This study could propose a framework for designing a teacher training curriculum that fill in this gap of teachers' lack of training in EFL context. A fourth suggested study can investigate the effect of using a mixed approach of both concordances-based and conventional dictionaries like GOOGLE TRANSLATE and MERRIAM WEBSTER, ONLINE OXFORD DICTIONARY and others to check unfamiliar words in concordances lines.

References

[1]   Alharbi, R. M. S. (2017). Acquisition of lexical collocations: a corpus-assisted contrastive analysis and translation approach (Doctoral dissertation, Newcastle University).

[2]   Aljuhani, H. S. (2016). Using corpora in a lexicalized stylistics approach to teaching English-as-a-foreign-language literature.

[3]   Al-Mahbashi, A., Noor, N. M., & Amir, Z. (2015). The Effect of Data Driven Learning on Receptive Vocabulary Knowledge of Yemeni University Learners. 3L: *Language, Linguistics, Literature®,* 21(3).

[4]   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), 103-111.

[5]   Alruwaili, A. (2018). Integrating corpus linguistics in second language vocabulary acquisition (Doctoral dissertation, University of Nottingham).

[6]   Barbieri, F., & Eckhardt, S. E. (2007). Applying corpus-based findings to form-focused instruction: The case of reported speech. *Language Teaching Research*, 11(3), 319-346.

[7]   Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English Preposition Learning through Data-Driven Learning (DDL). 3L: *Language, Linguistics, Literature®,* 24(3).

[8]   Celik, S., & Elkatmis, M. (2013). The Effect of Corpus Assisted Language Teaching on the Learners' Proper Use of Punctuation Marks. Educational Sciences: *Theory and Practice, 13(2),* 1090-1094.

[9]   Chen, J., Chau, R., & Yeh, C. H. (2004, January). Discovering parallel text from the World Wide Web. In *Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalization-Volume 32* (pp. 157-161).

[10] Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning, 12(4),* 345-360.

[11] Conrad, S. (2005). Corpus linguistics and L2 teaching. *Handbook of research in second language teaching and learning*, Routledge: 417-434.

[12] Davies, M. (2008). The Corpus of Contemporary American English (COCA): 900+ million words, 1990-present. Available online at http://www.americancorpus.org.

[13] De Groot, A. M. B. (2006). Effects of Stimulus Characteristics and Background Music on Foreign Language Vocabulary Learning and Forgetting. *Language Learning, Vol. 56(3)*, PP 463–506.

[14] Edo-Marzá, N. (2014). "A practical corpus-based approach to teaching English for tourism."

[15] Elsherbini, S. A. H., & 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.

[16] Gavioli, L., & Aston, G. (2001). Enriching reality: language corpora in language pedagogy. *ELT journal, 55(3),* 238-246.

[17] Gravetter, F.J & Forzano, L. B. (2018). Research Methods for the Behavioral Sciences. *Cengage Learning*, Inc. 978-1337-613316

[18] Hatch, E. M. and A. Lazaraton (1991). The Research Manual: Design and Statistics for Applied Linguistics, Newbury House Publishers.

[19] Huang, L. (2011). Corpus-aided language learning. *ELT Journal, 65(4),* 481-484.

[20] Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal,* 91(4), 663-667.

[21] Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly, 37(3),* 467-487.

[22] Kirk, D. (2009). Physical education futures, Routledge.

[23] Krashen, S. (2009). Does intensive decoding instruction contribute to reading comprehension? *Knowledge Quest, 37(4*), 72.

[24] Laufer, B. (2001). Reading word-focused activities and incidental vocabulary acquisition in a second language. *Prospect*,16,44-54.

[25] Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51(3),* 285-329.

[26] Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. Vocabulary acquisition: *Implications for reading comprehension*, 52-77.

[27] Nation I.S.P. (2001). Learning Vocabulary in Another Language. *Cambridge University Press*: Cambridge

[28] Nunan, D. and N. David (1992). *Research methods in language learning,* Cambridge University Press.

[29] Pérez-Paredes, P., Sánchez-Tornel, M., & Calero, J. M. A. (2012). Learners' search patterns during corpus-based focus-on-form activities: A study on hands-on concordancing. *International Journal of Corpus Linguistics,* 17(4), 482-515.

[30] Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics, 31,* 205-225. http://dx.doi.org/10.1017/S0267190511000055

[31] Sinclair, J. M. (2004). Corpus, Concordance, Collocation. UK: Oxford University Press.

[32] Varley, S. (2008). I'll just look that up in the concordance: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning, 22(2),* 133-152.

[33] Zhang, W. (2008). In search of English as a foreign language (EFL) teachers' knowledge of vocabulary instruction.

[34] Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. *Second language vocabulary acquisition, 6(1),* 5-19.