

**THE EFFECTIVENESS OF USING
CORPUS-BASED MATERIALS IN VOCABULARY
TEACHING**

Bütüncü Temelli Etkinliklerin İngilizce

Kelime Öğretiminde Etkililiği

Turan PAKER¹ & Yeliz Ergül ÖZCAN²

Abstract

Our study aimed at finding out the effectiveness of corpus-based vocabulary teaching activities as well as students' attitudes towards concordance-based materials when corpus-based tasks in English vocabulary learning are used. The study was conducted in a preparatory school in a private university. The participants were 28 intermediate level students whose native language was Turkish and they were between the age of 17 and 20. They were placed into two intact groups as experimental and control group as a result of a pre-test. While the participants in the experimental group were taught through corpus-based vocabulary teaching activities, those in the control group were taught vocabulary activities through the tasks in their text book and dictionary. The data were collected through a proficiency exam, pre- and post-tests and a 12 item-questionnaire based on a 5-point Likert scale as quantitative instruments and through interviews and learner diaries as qualitative instruments. The quantitative data were analyzed through descriptive analysis and independent samples t-tests, and qualitative data were analyzed through content analysis by observing themes in students' responses during the interviews and learner diaries. The statistical analysis indicated that using corpus-based vocabulary tasks were more effective than the tasks in the textbook. Additionally, our findings indicated that the attitudes of the students were positive in the use of corpus-based vocabulary tasks.

Key words: Corpus-based tasks, concordance lines, corpora, vocabulary learning, learner attitude.

Özet

Bu çalışma, bütüncü temelli aktivitelerin İngilizce kelime öğretiminde etkinliğini incelemiştir. Çalışmanın ana amacı ders kitabı veya sözlükle kelime öğretiminin yerine bütüncü temelli etkinliklerle kelime öğretiminin daha etkili olup olmadığını araştırmaktır. Bu çalışmanın diğer bir amacı ise öğrencilerin bu tür bütüncü temelli kelime etkinliklerine karşı olan tutumlarını anlamaktır. Bu çalışmada, Türkiye'de özel bir üniversitede eğitim gören, orta seviyede İngilizce bilgisine sahip 28 öğrenci yer almıştır. Katılımcılar kendilerine uygulanan ön-test sonucunda deney ve kontrol gruplarına ayrılmışlar, deney grubundaki katılımcılar bütüncü temelli etkinliklerle kelime öğrenimi çalışmaları yaparken kontrol grup ise sadece ders kitabı ve sözlük kullanımına dayalı çalışmalar yapmıştır. Bu çalışma, nitel ve nicel veri toplama araçları kullanılan yarı-deneysel bir çalışmadır. Bu çalışmadaki nitel veri toplama araçlarını bir yeterlilik testi, ön ve son testler ve 12 maddelik 5'li likert tipi öğrenci tutum anketi oluşturmaktadır. Nicel veri toplama araçları ise öğrenci görüşmeleri ve günlükleridir. Nicel veriler betimsel analiz ve bağımsız örneklem t-testleri uygulanarak analiz edilmiş, nitel veriler ise içerik analizi yapılarak incelenmiştir. Verilerin istatistiksel analizi göstermiştir ki bütüncü temelli kelime etkinlikleri ders kitabındaki etkinliklerden daha etkilidir.

¹ Doç. Dr., Pamukkale Üniversitesi, Eğitim Fakültesi, Yabancı Diller Eğitimi Bölümü, tpaker@gmail.com

² İstanbul Teknik Üniversitesi, Yabancı Diller Yüksekokulu, yeliz_ergul@hotmail.com

Buna ek olarak, tutum anketine verilen yanıtların analizi öğrencilerin bütünce temelli kelime etkinliklerine karşı olumlu tutumlara sahip olduklarını göstermiştir. Görüşme verileri ve öğrenci günlükleri de anket sonuçlarını desteklemektedir.

Anahtar Kelimeler: Bütünce temelli etkinlikler, konkordans satırları, korpora, kelime öğrenimi, öğrenci tutumları.

Introduction

The integration of corpus in the field of English as a Foreign Language (EFL) has resulted in a big shift in both 'methodology and ideology' (Binkai, 2012). Before corpora, linguistic descriptions have relied on what native-speakers know about language, or what they perceive language to be, rather than how language is used. Thanks to corpora, we have gained a better understanding of the contextual use of language, and we have developed new insights into language structure (Tsui, 2004). Corpora provide a number of examples of an item searched in its context of use, but it does not give the meaning of the word or phrase on the screen. The meaning of new item has to be deduced through the examples generated (Q'Keefe, McCarthy and Carter, 2007). Teachers can also use concordance to begin discussion on words, phrases and collocations or guide students to discover patterns of language use (Reppen, 2011). Data gathered from a corpus allow determining the criteria of word associations in collocations (Donesch-Jezo, 2013). The richness that authentic texts provide regarding their cultural and linguistic content, having the opportunity while selecting relevant, suitable, and interesting materials for the specific groups of learners, the motivational aspects of learning through authentic materials rather than artificial ones are some of the advantages of using corpora (Mishan, 2004).

In the past, teachers used some examples made up by them such as isolated sentences to exemplify a grammatical pattern while teaching. These examples were non-authentic and limited in number. Most of the time, students had difficulty exploring these poor sentences (Sun and Wang, 2003). On the other hand, using corpora has altered the roles of learners and teachers in language teaching. It has not only reinforced learner-centered tasks, but also changed the conception of teachers as sources of knowledge, and thus, teachers have assumed new roles as 'guides and facilitators, or even co-researchers' (Gabrielatos, 2005). Such student-centered discovery learning is very beneficial for the students when they are given access to the authentic language use. They are challenged to make generalizations and observe patterns of language behavior through the text they are exposed to. In addition, studying concordances can create more awareness in language use (Krieger, 2003). Data Driven Learning (DDL) is considered as an important breakthrough in ELT history because concordances give students access to an unlimited database of authentic materials (Basanta & Martin, 2005). The analysis activities (corpus-based activities) help learners to process text better and to systematize their knowledge. Learners are likely to notice more, and reflect on further occurrences of the language items when they pay attention to the meanings, uses, and functions of common words in the target language. This process leads to interlanguage development (Willis, 2011). Gavioli and Aston (2001) have emphasized that learning activities including the analysis of corpus data are in line with the current principles of language learning theory. Thus, learners can develop more autonomy provided that they are guided properly regarding how they should observe language and make generalizations about the discourse or patterns in the language. According to Reppen (2011), teachers usually integrate corpus into their teaching in three ways: first, they gather data from the corpus searches; then, prepare the materials depending on the level, and finally, they have students work with

these materials. In addition, teachers can also use online available corpora while teaching a specific language pattern or they can create specialized corpora from authentic texts or student papers and have students analyze the data

The corpus approach (data-driven learning approach)

Data-driven learning (DDL) refers to the use of a corpus of texts through the concordancing software when trying to find answers to linguistic questions by observing different uses of language. Although it is not a communicative approach, it has benefits such as being learner-centered, using authentic language, and promoting learner discovery (Allan, 2009). Data-driven learning (DDL) was defined by Johns and King (1991, p.iii) as “the use of computer-generated concordances in the classroom to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output.” Johns (1991) lays out the principles of DDL as follows:

- 1) The learners discover the language through their own questions and the language learner is a researcher who has access to linguistic data.
- 2) The computer acts as an informant answering the questions that the learners ask themselves.
- 3) The basic computer tool is the concordancer. Thus, the concordances on a computer screen show all instances of a word or phrase in a key word in context (KWIC) format.
- 4) The teacher is in the role of a facilitator in a student-centered classroom and concordancers are the important focus for learning.

Corpus approach or data-driven learning approach (DDL) includes hypothesis formation while carrying out inductive corpus-based exercises, teacher’s explicit explanations either to confirm or correct these hypotheses, and thus, hypothesis testing in terms of follow-up exercises regarding the task and output by learners (Chujo, Anthony & Oghigian, 2009). One of the central principles of the corpus approach is that vocabulary and grammar are interrelated rather than distinctive from each other (Halliday, 1992; Sinclair, 1991 cited in Yoon & Hirvela, 2004). Vocabulary instruction is moving away from teaching words in isolation, and placing a greater emphasis on exposing learners to lexical items in authentic and meaningful contexts. Furthermore, there is a growing amount of evidence that much of the English language is formulaic which suggests that teaching vocabulary as separate from grammar has limitations (Balunda, 2009).

On the other hand, corpus studies have promising results. Koosha and Jafarpour’s study (2006) revealed that using corpus-based materials were effective in both vocabulary teaching and learning. Anaç’s study (2009) also indicated that DDL has a marked effect on the learning of language items presented in the materials. Similarly, the current study revealed that using corpus-based materials were effective and helpful in vocabulary teaching. In Chao’s study (2009), it was revealed that concordancing was really effective in learning vocabulary through collocation. Chujo, Utiyama and Miura’s study (2006) revealed that most of the participants had positive attitude towards using a DDL tasks in learning vocabulary.

Current study

This study aims at examining the effectiveness of corpus-based tasks/activities in vocabulary teaching instead of using textbook activities and dictionaries in vocabulary teaching. Another purpose is to identify students' attitudes towards corpus-based vocabulary activities.

Method

Participants

The participants were 34 intermediate level preparation class students at a private university in Turkey. Their native language was Turkish and they were between the age of 17 and 20. In our school, all incoming students are given a proficiency test at the onset of the academic year. The test includes reading, listening and writing sections. If students get 60 marks or above out of 100, they are exempted from the preparatory program, and they can start studying in their departments as they are considered proficient enough to carry out their academic studies in their departments in which the medium of instruction is English. However; if students score less than 60 marks, they have to study in the preparatory program for one academic term or year according to their proficiency level. Students are divided into four levels, elementary, pre-intermediate, intermediate and upper-intermediate depending on the results of the scores they have in the proficiency test.

The participants of the current study participated in classes four hours a day, twenty hours a week for a semester. One academic year has two terms or thirty-two weeks overall, and students attend classes four or five hours a day, twenty or twenty five hours a week in general. They were placed into two groups using a convenience sampling technique. They were divided into two groups as the experimental and the control groups. There were eighteen students (ten female and eight male) in the experimental group and sixteen students (eight female and eight male) in the control group. In both groups, the students were placed at an intermediate level as a result of the placement test administered by the testing office of the school. The attendance was mandatory, so the majority of the students attended classes regularly. However, two students from the experimental group and four students from the control group dropped out of the program because of several personal reasons. As a result of SPSS calculations, the means and standard deviations of the proficiency test of the experimental group (1) and control group (2) indicated that the mean of the experimental group ($\chi=8.31$, $sd= 2.21$) was slightly higher than the control group ($\chi=8.00$, $sd= 2.21$). As the results were parametric, the independent samples t-test was used to see whether the control group and the experimental groups were homogenous. Although the mean score for the experimental group was slightly higher, our calculations indicated that there was no statistically significant difference between the two groups ($p = .73$, $p>0.05$) in their English proficiency levels. The pre-test results indicated that the mean of the experimental group ($\chi=8.06$, $sd = 2.81$) was higher than the control group ($\chi=6.83$, $sd= 1.80$). Even though the mean score of the experimental group was higher, there was no statistically significant difference between the two groups ($p = .19$ ($p>0.05$)) in their English vocabulary knowledge.

Instruments and materials

Proficiency Exam: In the first phase of this study, a proficiency test was given to the participants to find out whether two groups were homogenous or not. The researchers adopted Cambridge Preliminary English Test (PET) and Pearson Practice Reading Test as

proficiency test. The proficiency test included 20 reading, vocabulary and grammar questions. Items were analyzed by two experienced EFL teachers and an expert so as to judge their level of difficulty. After the pilot study, it was revised and used for the study as a proficiency test. The Cronbach-alpha value of the proficiency test was 0.71.

Pre- and Post-Tests: After having sampled target words from the textbook used by the school for this level, two tests (pre-and post tests) were prepared by the researchers. The tests consisted of multiple choice gap filling based on concordance lines. In each item, four to five concordance lines were given and students were asked to choose the best word that matches all the blanks. A couple of examples were given to the students before the pre-test because the test format was new to the students. In order to test reliability of the pre- and post tests, a pilot study was conducted. The Cronbach-alpha value of the test was 0.67.

Questionnaire: The students in the experimental group completed a questionnaire in 5-point Likert Scale format after the treatment. The items in the questionnaire were taken and adopted from Farr (2008), Yoon and Hirvela (2004), and Vannestal and Lindquist (2007). The questionnaire was translated into Turkish by the researcher to make learners feel more comfortable and to gather more reliable data. The researcher consulted an English instructor and an expert to find out whether the English version and the translated version were similar. A few changes were made on the wording of the questionnaire items after getting feedback and one item was omitted after the pilot study to increase the reliability. The Cronbach-alpha value of the final version of the questionnaire was 0.72.

Interview Questions: In addition to the administration of the questionnaire, semi-structured interviews were conducted to explore students' views and attitudes in detail towards using corpus-based vocabulary tasks. The interviews were conducted with all of the students (16) in the experimental group, and the interviews were administered in their native language to obtain more reliable and detailed data as they can easily reflect in their native tongue. The interviews were tape-recorded and later transcribed by the researcher for analysis. While analyzing the data descriptively, each student was given a number like A, B, C, D, etc. in order to ease data analysis and keep their identity anonymous. Some of the direct quotations from the interviews were included in data analysis to increase credibility.

Materials

The Textbook and Corpus-based Materials: North Star 3 Reading and Writing Student's book (Barton & Dupaquir-Sardinos, 2008) was used as a text book in the course. There was a vocabulary section at the beginning of each unit, and there were approximately eight to ten target words. These sections contained mostly matching and fill-in-the-blanks type of vocabulary tasks including non-authentic materials. For the experimental group, the vocabulary tasks in the textbook were replaced with corpus-based vocabulary tasks. Corpus of Contemporary American English (COCA) (Davies, 2012) was the resource to choose appropriate concordance lines for the target words and prepare corpus-based vocabulary materials. The researcher chose COCA because it is a very rich corpus with 450 million words of texts and it offers authentic examples from spoken language, fiction, magazines, newspapers and academic texts. Corpus-based vocabulary activities included analyzing concordance lines and answering questions, matching activities, and fill-in-the-blank exercises. These vocabulary activities were new to the students as they had not done them at school before.

Procedure

The researchers administered a pre-test to experimental and control groups at the beginning of the experiment. The results of pre-test revealed that there was no statistically significant difference between the groups ($p = .19$ ($p > 0.05$)) in their English vocabulary knowledge. Thus, we started the experiment. Target words in the weekly lesson plans were studied in the experimental group through the corpus-based activities. Each group had 20 hours English classes. Overall, both groups had 5 hours of vocabulary learning and practice. Even though both groups used the same course book, the vocabulary sections in the course book were skipped while teaching experimental group. On the other hand, the control group learned the target words using textbook and dictionary-based activities. After the treatment, the researchers administered a post-test that included the target vocabulary to both groups. Both groups were considered equal because they had been taught the same words in the same time period. At the end of the term, all the students from each group filled in an attitude questionnaire in order to find out their attitudes regarding the use of corpus-based activities in English vocabulary teaching. After the questionnaire had been analyzed using SPSS, audio-taped interviews with each participant from the experimental group were conducted to back up and clarify the questionnaire data. The researchers transcribed the interview data and analyzed them descriptively. The researchers also asked 5 students from the experimental group to keep learning diaries to reflect their opinions and attitudes about the process.

Results and Discussion

In order to investigate the first research question, which aims to find the effects of using corpus-based activities on intermediate level EFL learners' performance in vocabulary learning, the post-test scores of the students in both experimental and the control group were analyzed to identify whether there was a significant difference between the scores of the two groups. As shown in Table 1, the results showed that the mean score of the experimental group ($\chi = 12.88$, $sd = 2.41$) was higher than that of the control group ($\chi = 9.33$, $sd = 2.67$). As a result of the independent samples t-test, the p-value ($p = .001$, $p < 0.05$) showed that there was a statistically significant difference between the experimental and the control group. Namely, using corpus-based activities in vocabulary teaching had a marked effect on the experimental group students' performance in this study.

Table 1. The results of the post test

Groups	N	χ	Std. Deviation	P
Experimental group	16	12.88	2.41	.001*
Control group	12	9.33	2.67	

* $p < 0.05$

The comparison of gain scores for the groups showed that all students demonstrated some learning of the target vocabulary. However, the students in the experimental group learned the target vocabulary more than the students in the control group (see Table 2). Thus, it can be pointed out that the results of using corpus-based activities differed when compared to using a course book vocabulary activities only.

Table 2. The comparison of the groups' gain scores

	Groups	N	χ	Std. Deviation
before training	Experimental group	16	8.06	2.82
	Control group	12	6.83	1.80
after training	Experimental group	16	12.88	2.42
	Control group	12	9.33	2.67

In order to answer the second research question of the study, which aims to explore the attitudes of the intermediate level EFL learners regarding the use of the corpus-based vocabulary activities, the responses of the students to the questionnaire were analyzed by using SPSS. Descriptive statistics were also carried out and the means of the questionnaire items were calculated to find out the students' attitudes towards corpus-based vocabulary activities. The following data in Table 3 presents the mean values for the students' attitudes towards using corpus-based materials in vocabulary teaching.

Table 3. Results of the questionnaire

Items	N	Mean	Std. Deviation	Level of Participation
Item 4	16	4.38	.71	Strongly agree
Item 1	16	4.31	.60	Strongly agree
Item 3	16	4.31	.70	Strongly agree
Item 5	16	4.31	.79	Strongly agree
Item 12	16	4.25	.85	Strongly agree
Item 2	16	4.19	.75	Agree
Item 11	16	4.06	.92	Agree
Item 6	16	3.94	.77	Agree
Item 8	16	3.75	1.06	Agree
Item 9	16	3.62	1.14	Agree
Item 10	16	3.62	.88	Agree
Item 7	16	3.06	1.34	Somewhat agree

The mean values of the items presented in Table 3 revealed that the students' overall attitudes were positive towards corpus-based vocabulary activities. Most of the students strongly agreed that corpus-based vocabulary activities were helpful for learning the meaning of vocabulary (Item 4; $\chi = 4.38$). They all agreed that corpus-based vocabulary activities were beneficial to learn the target words when asked to the students in the interviews. Student D expressed that:

Corpus-based activities were very useful for me. We first found the part of speech of the word and then guessed the meaning of it, which was very beneficial. I have learned many new vocabularies apart from the highlighted words (target words). (Student D, experimental group)

In the interviews, most of the students stated that these activities fostered their strategy to guess the meaning of new words from the context and their consciousness about the different usages of the words when they analyzed the concordance tables.

Most of the students strongly agreed that integrating corpus-based vocabulary activities was beneficial for them in learning the target language, (Item 1; $\chi= 4.31$). They also agreed that corpus-based vocabulary activities were beneficial for their education when asked to the students in the interviews. Student F noted that:

While doing corpus-based activities, I was using some strategies like analyzing the structures and then getting the meaning by guessing. I tried to do this in quizzes and exams, too. (Student F, experimental group).

Student F's response reveals that she benefited from corpus-based vocabulary activities and transferred the strategies she used while doing corpus-based activities to reading. In addition, the students were asked to compare whether using corpus-based activities were more helpful than using a dictionary (Item 3; $\chi= 4.31$), and the majority of the students strongly agreed that corpus-based vocabulary activities were more helpful. Student K mentioned that:

I used to check the meaning of each unknown word in a book before, but now, I do not feel like I need to use dictionary all the time. With the help of corpus-based activities, I learned how to understand the part of speech of a word and guess the meaning. Sometimes it gets a little bit difficult but it is much more fun to discover them (Student K, experimental group).

Student K's response reveals that corpus-based activities lessened his/her dictionary dependence and promoted his strategy of analyzing the words and guessing the meaning. Student K also found corpus-based materials fun as they gave the opportunity to discover the part of speech and meaning of the target words.

The mean score of 4.31 for Item 5 in Table 5 points to the fact that the majority of the students also agreed that corpus-based vocabulary activities were valuable for learning the usage of vocabulary. Most of the students agreed that corpus-based activities should be taught in English classes in their country (Item 12; $\chi= 4.25$). They were reluctant to welcome the idea of using corpus-based materials in English classes. They also agreed on recommending using the corpus-based activities in the same course in future quarters (Item 11; $\chi=4.06$). The following account is from one student in the experimental group.

At first, corpus-based vocabulary activities were challenging for me but now I am doing OK. I think that they are more durable (than learning words from the vocabulary activities in textbook). When we talk about the words and discuss the meaning and parts of speech, I understand better. I hope that we will have corpus-based vocabulary activities in the following module, too (Student A, experimental group).

When the students were asked whether their understanding of language systems benefited from corpus-based vocabulary activities, the majority agreed that integrating corpus-based materials raised their awareness of language systems (Item 2; $\chi=4.19$). The mean score 3.94 of item 6 shows that the majority of the students agreed that they had some difficulty in understanding the concordance lines due to the limited number of sentences. Student C mentioned that:

Sometimes, it's not easy to understand because it's not complete and a little difficult. Sometimes, I can't guess the meaning of the word. Other times, it's nice. I like the new style. We do it like there are some definitions and we try to match it. This kind is easy to remember (Student C, experimental group).

The statement indicates that incomplete concordance lines made it difficult to comprehend the whole idea. The reason of this might be that these kinds of vocabulary activities were new to the students. It was their first time doing vocabulary activities using corpus-based materials.

Most of the students agreed that they understood the purpose of using corpus-based vocabulary activities (Item 8; $\chi = 3.75$). Few students disagreed that they understood the purpose behind the corpus-based activities.

Most students agreed that corpus-based activities improved their guessing the meaning from the context strategy (Item 9; $\chi = 3.62$). Few of them thought that such activities did not help them to improve their guessing strategy when asked in the interviews. The mean score 3.62 for Item 10 shows that some of the students thought they actively participated in the corpus-based vocabulary activities while others were neutral about this item. The mean score 3.06 for Item 7 shows that the majority of the students were neutral about the difficulty level of the real texts in the corpus. Some of them agreed that the real texts were too difficult to understand while others disagreed.

All in all, the results of the questionnaire revealed mostly the same results when they were repeated in the interviews. The students' responses to the seventh item in the questionnaire showed that the majority of the students agreed that the real texts in corpus were too difficult to understand. However, in the interviews, students stated that the cut-off sentences made it difficult to understand not the real texts. Apart from the seventh item, the results and responses revealed that the qualitative data supported most of the quantitative data in the study. The findings revealed that students from the experimental group had positive attitudes towards corpus-based vocabulary activities.

Consistent with previous research (Chan & Liou, 2005; Kobelinski, 2005; Koosha and Jafarpour, 2006; Chao, 2010; Chujo, Utiyama & Miura, 2006; Balunda's, 2009), the current study revealed that using corpus-based materials were effective and helpful in vocabulary teaching. The current study also confirms the findings of some previous studies that examined students' attitudes towards concordancing in the context of collocation learning (Chao, 2010; Chujo, Utiyama & Miura, 2006).

The study has revealed some insights for language teachers about the use of corpus-based materials in vocabulary teaching. We believe that integrating authentic and rich corpus-based vocabulary activities will bring a lot of benefits to teachers. Such integration will enable teachers to teach target words through real texts and give students opportunities to engage in exploring different uses of the target words in intermediate level and above. Teachers can prepare corpus-based vocabulary activities by taking concordance lines from the corpus (COCA) and preparing different types of activities like lead-in questions to help the students guess the meaning from the context, gap-filling exercises, and matching activities. As Yoon and Hirvela (2004) state, it may be pedagogically unwise to simply let students work on corpora without giving proper guidelines. Teachers should guide students while using corpora. Otherwise, they can easily get lost and give up learning in this way.

The present study provided evidence for material designers questioning the effectiveness of using corpus-based vocabulary activities. Hence, material designers may consider incorporating such materials and activities into their classes to make the learning process more interesting and effective. They can easily access free online corpora and learn how to use them effectively. Using corpora and preparing corpus-based activities are easy to

grasp but it may be time-consuming to find appropriate concordance lines for different proficiency levels. It is easy and quick to find rich and authentic concordance lines for high proficiency level students while simplification and adaptation of the concordance lines are necessary for lower proficiency levels. Therefore, materials designers can adopt concordance lines according to the level of the students. Intermediate and high level students can manage to do these activities; however, students with low proficiency levels may need easier and more comprehensible activities.

As the questionnaire and interview data suggested, students held positive attitudes towards corpus-based vocabulary activities. One of the reasons is that they were new and interesting for them. In addition, they also enjoyed deducing the meaning by themselves with the help of the lead-in questions and engaging in the activities actively. Moreover, they became more confident as they did not use dictionary while doing the activities. Students can find corpus-based vocabulary activities on the Internet and consult online corpora while writing or studying grammar and vocabulary. Online corpora can be a very helpful resource for students if they learn how to use them effectively. Their teachers can assist them or they can watch video tutorials on how to use these online corpora. Students can refer to COCA especially in writing for appropriacy of work.

Corpus-based vocabulary tests used in the study were similar to corpus-based vocabulary activities in terms of rich and authentic context. These types of tests will have a practical use for testing units in English preparatory programs in universities. Instead of preparing artificial test items, testing office members can prepare rich and authentic test items to assess students' vocabulary knowledge. Alderson (1996, cited in McEnery & Xiao, 2007) emphasize that 'the potential advantages of basing our tests on real language data, of making data-based judgments about candidates' abilities, knowledge and performance are clear enough.' A crucial question is whether the possible advantages are born out in practice. Therefore, integrating corpus-based approach into teaching and testing should go hand in hand.

Conclusion

This study aimed to find out the effectiveness of corpus-based activities in vocabulary teaching. Another aim of the study was to investigate students' attitudes towards corpus-based vocabulary activities. Comparison of the post test scores for experimental and control groups indicated that there was a statistically significant difference between the groups; the experimental group scored higher than the control group. The results of the attitude questionnaire revealed that the students' attitudes were positive towards using corpus-based vocabulary activities. The interview data revealed the same result that the students had positive attitudes towards using corpus-based materials in vocabulary teaching. The post test, questionnaire and interview data indicated that using corpus-based activities in vocabulary teaching was fruitful for students. Thus, it can be claimed that using corpus-based vocabulary materials were more effective than using textbooks or dictionaries.

Limitations of the Study

The study was limited to 34 intermediate level EFL students in a preparatory program. It could be replicated with more participants so as to decide whether the conclusions drawn from the present study are similar. Next, conducting corpus-based studies with different proficiency levels (elementary, pre-intermediate and high) would be helpful to gain more insights about the effectiveness of these materials. Therefore, further research may

investigate the effectiveness of corpus-based vocabulary activities on students with lower or higher proficiency levels. Furthermore, the study was conducted in a semester. A further study may be conducted in a longer treatment period to understand the long-term effects of using corpus-based materials. Another limitation of the study is that we are unable to provide data for students who dropped the program. We also have no information (attendance rates or being exposed to online corpus-based materials) which may affect achievement.

References

- Alderson, C. (1996). Do corpora have a role in language assessment? In J. Thomas and M. Short (Eds.), *Using Corpora for Language Research* (pp. 248-259). London: Longman.
- Allan, R. (2009). Can a graded reader corpus provide 'authentic' input? *ELT Journal* Volume, 63 (1), 23-32. doi: 10.1093/elt/ccn011.
- Anaç, G. (2009). *Effects of data-driven learning on language learning: a study of Turkish learners of English as a foreign language*. Unpublished master's thesis, Abant İzzet Baysal University, Bolu, Turkey
- Balunda, S.A. (2009). *Teaching academic vocabulary with corpora: student perceptions of data-driven Learning*. Unpublished master's thesis, Indiana University, Indiana, USA
- Basanta, C.P., & Martin, M.E.R. (2005). The application of data-driven learning to a small-scale corpus of conversational texts from the BNC-British National Corpus. *International Journal of Learning*, 12 (8), 183-192.
- Barton, L., & Dupaquir-Sardinos, C. (2008). *North Star 3 Reading and Writing*. New York: Pearson.
- Bennett, G.R. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. Ann Arbor, MI: University of Michigan Press.
- Binkai, J. (2012). An empirical study on corpus-driven English vocabulary learning in China. *English Language Teaching*, 5 (4), 131-137. doi: 10.5539/elt.v5n4p131.
- Chan, T., & Liou, H. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18 (3), 231-250. doi:10.1080/09588220500185769.
- Chao, P. (2010). *A study of collocation learning of junior high students in Taiwan via*

concordance. Paper presented at the 2010 International Conference on English Teaching, Kaohsiung, Taiwan.

Chujo, K., Utiyama, M., & Miura, S. (2006). Using a Japanese-English parallel corpus for teaching English vocabulary to beginning-level students. *English Corpus Studies*, 13, 153-172. Retrieved from http://www2.nict.go.jp/univ-com/multi_trans/member/mutiyama/pdf/06_English_Corpus_Studies.pdf

Chujo, K.L., Anthony, L., & Oghigian, K. (2009). DDL for the EFL classroom: effective uses of a Japanese- English parallel corpus and the development of a learner friendly,online parallel concordance. In M. Mahlberg, V. Gonzales-Diaz, & C. Smith (Eds.), *Proceedings of the Corpus Linguistics Conference, CL 2009*. Liverpool University of Liverpool. Available online at

<http://ucrel.lancs.ac.uk/publications/cl2009/>.

Davies, M. (2012). The Corpus of Contemporary American English (COCA): 450 million words (1990-2012). Available online at <http://www.americancorpus.org>.

Donesch-Jezo, E. (2013). Using language corpus in teaching foreign language vocabulary. *International Multidisciplinary e-Journal*, 2 (1), 11-25.Retrieved from <http://www.shreeprakashan.com/Documents/2013128181020978.2.%20Ewa%20Donesch-Jezo.pdf>

Ewa%20Donesch-Jezo.pdf

Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: perspectives from the users, *Language Awareness*,17(1), 25-43. doi: 10.2167/la414.0.

Gabrielatos, C. (2005). Corpora and language teaching: just a fling or wedding bells? *TESL-EJ Teaching English as a Second or Foreign Language*, 8 (4), 1-37.

Retrieved from <http://tesl-ej.org/ej32/a1.html>

Gavioli, L. and G. Aston. (2001). Enriching reality:language corpora in language pedagogy. *ELTJournal* 55/3, 238-46.

Johns, T. (1991). Should you be persuaded – two samples of data-driven learning materials. In T. Johns & P. King, (eds.), *Classroom Concordancing*. Birmingham University: *English Language Research Journal*, 4,1-16.

Johns, T. & King, P. (eds.) (1991). Classroom concordancing. *English Language Research Journal*, 4, 27-45.

Kobelinski, S.G. (2005). *A Corpus-based approach to ESP: EST vocabulary in*

- information technology*. Unpublished master's thesis, Federal University of Santa Catarina, Santa Catarina, Brasil
- Koosha, M., & Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal Quarterly*, 8 (4), 192-209. Retrieved from http://rua.ua.es/dspace/bitstream/10045/17423/1/RAEI_23_04.pdf
- Krieger, D. (2003). Corpus Linguistics: what it is and how it can be applied to teaching. *The Internet TESL Journal*, 9 (3). Retrieved from iteslj.org/Articles/Krieger-Corpus.html
- McEnery, A.M. & Xiao, R.Z. (2007). 'Parallel and comparable corpora: What are they up to?' In G. M. Anderman, & M. Rogers (Eds.), *Incorporating Corpora: Translation and the Linguist*. Clevedon: Multilingual Matters.
- Mishan, F. (2004). Authenticating corpora for language learning: a problem and its resolution. *ELT Journal*, 58 (3), 219-227. doi: 10.1093/elt/58.3.219.
- Q'Keeffe, A., McCarthy, M. J. & Carter, R. (2007). *From corpus to classroom: language use and language teaching*. Cambridge: Cambridge University Press.
- Reppen, R. (2011). Using corpora in the language classroom. In B. Tomlinson (Ed.), *Materials Development in Language Teaching (second edition)*. Cambridge: Cambridge University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sun, Y.C. & Wang, L.Y. (2003). Concordancers in the EFL classroom: cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16 (1), 83-94. doi: 10.1076/call.16.1.83.15528.
- Tian, S. (2005). Data-driven learning: Do learning tasks and proficiency make a difference? *Proceedings of the 9th Conference of the Pan-Pacific Association of Applied Linguistics*. Retrieved from <http://www.paaljapan.org/resources/proceedings/PAAL9/pdf/TianShiaup.pdf>
- Tsui, A.B.M. (2004). What teachers have always wanted to know- and how corpora can help. In J. M., Sinclair (Ed.), *How to use corpora in language teaching* (pp.39-65). Amsterdam, Philadelphia: John Benjamins.

Vannestal, M.E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 19 (3), 329-350. doi:10.1017/S0958344007000638.

Willis, J. (2011). Concordances in the classroom without a computer: assembling and exploiting concordances of common words. In B. Tomlinson (Ed.), *Materials Development in Language Teaching (Second edition)*. Cambridge: Cambridge University Press.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13 (4), 257–283. doi:10.1016/j.jslw.2004.06.002.

APPENDIX 1

ATTITUDE QUESTIONNAIRE

Name: Age: Email address:
 Gender: Male Female: Major subject:
 How long have you been learning English?

Please circle the most appropriate number of each statement which correspond most closely to your desired response.

1: Strongly disagree 2: Disagree 3:Somewhat agree 4: Agree 5: Strongly agree

1-I feel my education, in general, benefited from the integration of corpus-based activities.	1	2	3	4	5
2- I feel my understanding of language systems benefited from the integration of corpus-based activities.	1	2	3	4	5
3- The corpus is more helpful than a dictionary for my vocabulary learning.	1	2	3	4	5
4- Corpus-based activities are helpful for learning the meaning of vocabulary.	1	2	3	4	5

5- Corpus-based activities are helpful for learning the usage of vocabulary.	1	2	3	4	5
6- I have some difficulty in understanding the concordance lines due to the limited number of sentences.	1	2	3	4	5
7- The real texts in the corpus are too difficult to understand.	1	2	3	4	5
8- I understand the purpose of using corpus-based vocabulary activities in this course.	1	2	3	4	5
9- Corpus-based activities have increased my guessing the meaning from the content strategy.	1	2	3	4	5
10- I actively participated in corpus-based activities during the sessions.	1	2	3	4	5
11- I recommend using the corpus-based activities in the same course in future quarters.	1	2	3	4	5
12- Corpus-based activities should be taught in English classes in my country.	1	2	3	4	5

APPENDIX 2

CORPUS-BASED VOCABULARY ACTIVITIES

A. Read the concordance lines below and answer the following questions.

Campaign:

1 America is deep in the throes of its presidential **campaign** season, never a time noted for its

2 The Department of Health is launching a **campaign** to raise awareness of the symptoms of

3 will target people over the age of 55. The **campaign**, which will run from 8 May to 30 June,

4 e-mails, the FBI raids. Solyndra investors raised **campaign** money for Obama. The

1. Is '**campaign**' a(n) verb, noun or adjective?
2. What does '**campaign**' mean?
3. Can you write one sentence using '**campaign**'?

Competition:

1 incorporate some form of class competition between groups and therefore, the students are

2 rise as tuition will also increase the competition for students among institutions, including

3 between men; drive for success, power, and competition; conflict between work and family),

4 using computers. The winner of one recent competition was a pair of amateurs using three

1. Is 'competition' a(n) verb, noun or adjective?
2. What does 'competition' mean?
3. Can you write one sentence using 'competition'?

B. Read the following sentences. What do the underlined words have in common? How do they differ from each other?

Technology and engineering education continues to evolve as it becomes more apparent that students need this information to become more successful in college and careers.

Mrs. Garcia shared that she felt she was part of the team and knew that the school cared and wanted her son to succeed.

Difference:.....
Similarity:
.....

C. What can you observe from concordance lines below?

Did you miss out on some key fact that influenced their decision to turn you down? In the educational setting, academic performance is influenced by teachers, peers, or family members.

Lady Gaga has admitted to being influenced by Madonna, but when " Born This Way " sounded more than a little similar to " Express Yourself, " some wondered if influence was plain-out copying?

Influence:
.....
.....

APPENDIX 3

PRE-TEST & POST TEST

Choose the most appropriate word or phrase that fits all the blanks in each set of lines.

have been a problem. After this, they will always be _____ of seafood, but will credit it with
took her mom to the dermatologist to examine a _____ spot on her fair skin, she
planned
watch volunteer who thought the youth looked _____ The shooter, George
Zimmerman,

- a) mysterious c) curious
b) courageous d) suspicious

7.

coldly indifferent, while animal suffering provokes _____ feelings of pity and
compassion?
The grief is like a lightning strike. It is so powerful. It is so _____. " She said, " This
individual Her tattoo had to be broken into several sessions, due to the _____ pain. In the
words of
deal maker in the fall elections. Instead, both parties are using _____ feelings about the
law

- a) personal c) intense
b) deep d) great

8.

health decisions that might help them _____ those goals. This may involve decisions
about
and Google s SPDY both of which can _____ better performance, reduce the total number
of
different interventions to _____ improvements in health and well-being at the individual
States). If children really need to _____ optimal blood levels of 30 ng/mL (- 75 nmol/L),
then

- a) achieve c) observe
b) reach d) arrive

9.

While McInturff says he sees the house as a(n) _____ from his job, he didn't completely
jazz and good food to audiences looking for a(n) _____ from war. As close to a musical of
prison and house arrest, Chen made a daring _____ - from his rural village in political in
narrative fiction have a new place to _____ to: the recently unveiled Jal Mahal

- a) leave c) return
b) escape d) move

10.

White Oak; 713-862-3838. Phoebe Hunt: Fiddle phenom is _____ well known in these
parts for
style. Leading him on by discreet questions, I soon had a _____ clear idea of the Wells
outfit,
when I was thirty five, then thirty six. Till then I'd lived a _____ normal life, if normal
includes
of our getaway vehicle, not far from the palace. So I'm _____ certain we'll have no
difficulty

- a) clearly c) slightly
b) absolutely d) fairly

16.

life, many people thought that a woman's job was only to ____ children and take care of her
 was a young mother, newly arrived in West End, struggling to ____ her son alone and
 make her way that fresh air and room to run around. What a great place to ____ a family.
 "I signaled he'd stuck on the ceiling swaying in the air. He didn't have to ____ him. Boys
 needed mothers

- a) grow c) raise
 b) rise d) increase

17.

champion, the Secretary of Sorrow Services. Her _____, she knew, was the reason the
 He then repeated the importance of American _____ toward significant climate-change
 have shared with us that early service and _____ experiences while still in high school
 played
 peace deal and Turkey's determination to play a _____ role in the Arab world at the
 expense

- a) unionship c) friendship
 b) leadership d) relationship

18.

been sent to find. Caleb Bahame was a terrorist and a _____ so cruel that it was difficult
 to Killer had terrified San Francisco in the '90s, a _____ who took eight innocent lives,
 left
 never pull off. Too bad he was a cold-blooded _____. Or that was the going theory. One
 she
 angle. " John, Dennis Rader is the BTK killer. He's a _____. They asked for a great
 figure, not

- a) torturer c) murderer
 b) admirer d) robber

19.

I'm a thirty-year-old woman, a doctor who deserves some privacy and _____. "I spoke
 calmly
 back then. Lucille scowled at me. "You should teach those boys some _____. In my day
 standing find love. For herself, she hoped only to earn her husband's _____ and to give
 it in
 day when I try to keep my eyelids shuttered. # But I have enormous _____ for the work
 she

- a) manner c) behaviour
 b) respect d) custom

20.

it didn't matter. He was determined to set things right and keep _____ people from dying.
 claiming the lives of a number of your neighbors, including _____ children, as a result.
 The
 is murdered, or if any harm comes to his wife or any one of his five _____ children -- then
 in
 Laden before they did their worst. If you killed them, you'd be killing an _____ man,

- a) innocent c) cruel
 b) pure d) smart