

## EFL Students' Perceptions of Using COCA to Develop Their Vocabulary

Arzu Koçak<sup>1</sup>, Hacettepe University, kocaka@hacettepe.edu.tr

**Recommended citation:** Koçak, A (2020). EFL Students' Perceptions of Using COCA to Develop Their Vocabulary. *Journal of Language Research (JLR)*, 4(1), 12-26.

**Abstract:** This paper presents a research study investigating students' perceptions about the use of COCA (Corpus of Contemporary American English) to expand their English vocabulary. The purpose of the study is to determine perceived benefits of using online corpora for improving EFL students' vocabulary from the perspective of students. In the study, a group of EFL students in the Preparatory School at a state university in Turkey were introduced what a corpus is and how to use it to develop their vocabulary. The study was carried out for four weeks and the participants attended three hours of instruction each week. After each instruction, students were required to do some corpus-based vocabulary activities. The data for this research were collected via a questionnaire. At the end of the fourth week, the students reflected their perceptions about using the online corpus (COCA) in the questionnaire given. The data gathered from the study were analyzed quantitatively. The results indicated that most of the students think that using COCA as an online corpus is highly beneficial to enriching their English vocabulary. The results will supply EFL teachers with planning information about how corpora and corpus-based vocabulary activities might best be integrated into English language curriculum.

**Keywords:** *corpus, corpora, corpus-based vocabulary activities, vocabulary, perceptions*

### INTRODUCTION

In recent years, corpus-informed approaches have greatly affected foreign language teaching and second language acquisition research. Language corpora are used as an important language learning tool in second language learning today (Al Saeed & Waly, 2009; Dash, 2003; Varley, 2008). It is believed that corpora is a crucial source as it enables us to access reliable, authentic and various information that can hardly be found in other sources. Since corpora provides learners with a variety of opportunities to see language properties from various directions, second language learning (SLL) that lacks of reference to corpora does not seem scientific (Varley, 2008). Basically, the data gathered from corpora presents important complementary sides for traditional linguistic principles of SLL (Biber, 1996 as cited in Varley, 2008).

In language teaching, there are two basic types of using corpora: "(a) as a primary resource and (b) as a secondary resource" (Varley, 2008, p.3). Using corpus as a main resource refers to the direct use of corpora in the second language learning (SSL) classrooms or language laboratories. Therefore, students have an opportunity to access corpora directly in order to get the information they search and improve their language skills. The second type of the use of corpora addresses to benefit from corpora in the design of foreign language materials such as syllabuses, grammar books and bilingual dictionaries (Varley, 2008).

Besides different areas like genre analysis, semantics, pragmatics and sociolinguistics, today, corpus linguistics applies to language teaching. Moreover, it is used as an instructional tool to teach vocabulary, grammar, reading, writing, listening and speaking (Fligelstone, 1993). In this sense, considering the effect of corpus linguistics on foreign language teaching and learning, this study investigated students' perceptions of using an online corpus (COCA) to develop their vocabulary.

<sup>1</sup> ORCID: <https://orcid.org/0000-0002-3641-7102>

Submitted: 04.04.2020

Accepted: 09.07.2020



## THEORETICAL FRAMEWORK

Vocabulary is an important part of foreign language learning since it is necessary for learners to know a wide range of vocabulary in order to develop their language skills. Moreover, the master of the vocabulary of the foreign language has a great effect on learners' expression in communication (Chao, 2008). Thus, in foreign language instruction, much more attention should be paid to the importance of vocabulary learning both by teachers and students. There are different ways of teaching and learning vocabulary. However, it is still difficult for most of the learners to enlarge their vocabulary of the target language (Feng-xia, 2009). Therefore, new techniques of teaching and learning vocabulary should be integrated into foreign language curriculum.

In recent year years, computers have facilitated language learning to a great extent. With regard to foreign language learning and teaching, one of the significant outcomes of computer technology is corpus linguistics. Corpus linguistics can be applied to teaching vocabulary in English language instruction as a new way of helping students to widen their English vocabulary.

As this study focuses on using corpora in English language teaching, the following section will give information about what a corpus is and what kind of corpora teachers can use for teaching English.

### Corpus-based Lessons

It is not possible to say exactly how many words English contains. Global Language Monitor has reported that there are nearly a million words (988,968) in English. Researchers also attempted to find out the number of words native speakers use in order to evaluate how many words learners must learn. The results indicate that the number of words native speakers know changes between 12,000 and 20,000 according to their level of education (Goulden, Nation, and Read, 1990 as cited in McCarten, 2007). Another result shows that almost 20,000 word families, except expressions and phrases, are known by a native speaker university graduate. When these numbers are taken into consideration, it seems that it is impossible for foreign language learners and even for many native speakers to know all of these words (McCarten, 2007).

Fortunately, learners can operate in English using fewer than 20,000 words. Determining the number of various words used in an average spoken or written text also enables us to decide how many words learners need. Because of the repetition of some high frequency words, learners are said to understand many texts through a relatively small vocabulary. Therefore, for instance, nearly 80 percent of words in an average text should be understood by learners who know the most frequent 2000 words. In addition, learners' understanding is enhanced to 88.7 percent with a knowledge of 5000 words (Francis and Kucera 1982 as cited in McCarten, 2007). With regard to spoken language, almost 1800 words constitute over 80 percent of the spoken corpus (McCarthy, 2004; O'Keeffe, McCarthy & Carter, 2007 as cited in McCarten, 2007). Although learning up to 5000 words is still difficult for learners, it is possible for many learners to accomplish this learning goal instead of learning 20,000 words.

As a result, it is essential that the most frequent 2000 to 5000 vocabulary items be found and they be given priority in teaching. In addition, students must be autonomous learners. It is impossible for teachers to teach the vast number of the vocabulary items students need to know. Therefore, students need help about how to learn vocabulary and what to learn (McCarten, 2007).

### *What can we learn from the corpus about vocabulary?*

A corpus basically consists of a variety of texts stored in a computer (McCarten, 2007). Crystal (1992) defines the corpus as "a collection of linguistic data, either compiled as written texts or as a transcription of recorded speech" (p.85). Crystal (1992) also states that a corpus aims to prove a hypothesis on language - for instance, to identify how the usage of a specific sound, word or syntactic construction changes. The collection of texts in a corpus can include written or spoken language. Written texts such as magazines and newspapers can be entered into the computer by means of a scanner, a CD



or the Internet. Spoken texts, like conversations, can be loaded into the computer database after the conversations are recorded and transcribed. As a result, a corpus enables us to analyze the language in the corpus and see how people really speak or write through corpus software tools (McCarten, 2007).

A large corpus contains different sections, or subcorpora, which includes various types of English such as “subcorpora of different varieties like North American English and British English, or different types of language like conversation, newspapers, business English, and academic English” (McCarten, 2007, p.2). In order to integrate the corpus in to the design of the syllabus, the first step is to determine the type of English that will be used for the course material because different corpora provide the teachers with various words and different uses of words to teach. For instance, although the word “nice” is in the top fifteen words in conversation, it is not frequent in written academic English. Another example is the word “see”. The corpus analysis shows that it has the same frequency both in written academic English and conversation; however, it has a variety of uses. In written academic English, the word “see” is generally used to guide the reader to another article or a book such as *see McCarthy, 2004*.

In conversation, “see” has different uses. For example, “*I see*”, which means “I understand,” and “*See*” and “*You see*”, which introduce what the speaker feels is new information for the listener (McCarten, 2007, p.2).

Consequently, the kind of the corpus the teachers use can influence which words will be used in the course material and which meanings of those words will be taught to learners

McCarten (2007) summarizes what can be learnt from the corpus about vocabulary:

**“Frequency:** Which words and expressions are most frequent and which are rare

**Differences in speaking and writing:** Which vocabulary is more often spoken and which is more often written

**Contexts of use:** The situations in which people use certain vocabulary

**Collocation:** Which words are often used together

**Grammatical patterns:** How words and grammar combine to form patterns

**Strategic use of vocabulary:** Which words and expressions are used to organize and manage discourse” (p. 3).

According to McCarten (2007), the huge amount of data in the corpus, which can contain millions of words, can be analyzed through corpus tools. The corpus enables us to reach not only “statistical kind of information (a quantitative analysis)” but also a great number of texts which offer an opportunity to analyze the use of vocabulary in context – “a qualitative analysis” (p.3). Nevertheless, McCarten (2007) states that it is impossible for corpus to tell the teachers what is the best way of teaching and learning vocabulary as it is only an instructional tool.

### *Aim of the study*

The use of technology in the classroom is gaining importance day by day. Moreover, in the following years, it is likely to become a usual a part of English Language Teaching (ELT) practice. One of the technology-based instructional tools for English language instruction is the corpora. Many studies have been conducted on the effectiveness of using corpora as an instructional tool in teaching vocabulary, grammar and language use in the English language classroom (Barlow, 2002; Conrad, 2000; Neff et al., 2001; Willis, 1998; Varley, 2008).

However, little is known about students’ perceptions of using online corpora to develop their vocabulary in foreign language teaching and learning. In order to help students to enlarge their vocabulary, in addition to traditional techniques and activities, new ones can be integrated into our ELT classes. Therefore, an investigation into students’ perceptions of using online corpora to improve their vocabulary will give ELT teachers information about whether online corpora can be used as an



instructional tool to help students develop their vocabulary. In this sense, this study aims to answer the following research question:

- What are EFL students' perceptions of using COCA as an online corpus to develop their vocabulary?

## **METHODOLOGY**

This paper presents research that investigates EFL students' perceptions of using an online corpus (COCA) in English courses to expand their vocabulary. The participants are EFL learners at a state university in Turkey. In this study “*perception*” refers to what students think and feel about the use of COCA to improve their vocabulary knowledge. Therefore, with this meaning, in this article the word “*perception*” is used interchangeably with its synonyms “*view*”, “*opinion*” and “*attitude*”.

### **Sample / Participants**

This study was conducted at the Preparatory school at a state university in Turkey. The aim of the Preparatory School is to enhance students' language skills and enable them to follow their undergraduate studies in their own department where the medium of instruction is English. Students attend English courses for 25 hours per week in order to reach the required level. According to the results of proficiency exam students take at the beginning of the each academic year, they are placed in classes considering their English proficiency level.

The EFL students in the preparatory school attend English courses for a year to enhance their language skills and to get prepared for their subsequent four year education. These students have the instruction in their own department in English (% 30, % 50 or % 100). Therefore; they need English not only for operating successfully in the target language but also for learning English for a Specific Purpose (ESP) or for Academic Purposes (EAP). They have to access, read and analyze various English-language academic texts when they start to study in their own departments. As a result, developing their English vocabulary is very important for these students.

The participants in this study are 27 students (9 male and 18 female) at the Preparatory School. The students are at A2 level. The participants, aged between 18- 21, study in the same class but they are from different departments.

The main reason for choosing these students for the study is that they have not known what a corpus is and they have not done corpus-based activities in English courses before. In addition, perception of these students about corpus-based vocabulary activities gave information about how corpora can best be integrated into the curriculum of English courses.

### **Instrument(s)**

The data for this research were collected via a questionnaire for students. In addition, COCA (Corpus of Contemporary American English) was used by the teacher and students to do some corpus-based vocabulary activities.

### ***Student Questionnaire***

The researcher prepared a questionnaire for students to gather data for this research. The questionnaire was designed to get information about students' perception of using online corpus (COCA) to develop their vocabulary. It was prepared in Turkish in order to enable students to completely understand the questions and express themselves well. The questionnaire consisted of only one section which included 16 questions eliciting responses on 5 points Likert scales. Some questions



were adapted from the questionnaires used in a previous study investigating students' attitudes towards a technology-based instructional tool (Koçak, 2010). Before the questionnaire was administered, preliminary questions in the questionnaire were checked by five English instructors working in Testing Unit in the Preparatory School to assess their validity and to get feedback about the questions and after these processes, necessary changes were made in the questionnaire. In addition, the data obtained from the study indicated that the reliability of the questionnaire, as measured by internal consistency, was found to be satisfactory (Cronbach  $\alpha = 0.862$ ).

### ***The Corpus of Contemporary American English (COCA)***

The Corpus of Contemporary American English (COCA) is accepted as the first big and varied corpus of American English. It was made available online in early 2008. Moreover, it contained more than 385 million words from 1990-2008 taken from spoken, fiction, newspapers, popular magazines and academic journals (Davies, 2009). The current corpus includes "more than 600 million words of text (20 million words each year 1990-2019)" (<https://corpus.byu.edu/coca/>). In this study, all corpus-based activities were done via COCA.

### **Data collection procedures and analysis**

The study was carried out for five weeks. During the four weeks of the study, the participants were introduced what a corpus is and how to use it to develop their vocabulary. They attended three hours of instruction each week. After each instruction, students were required to do some corpus-based vocabulary activities. The corpus-based activities used in the first week of the study were adapted from the activities presented at an ELT conference (Walters, 2010). The activities done by the participants in the second and third week of the study were designed by the author according to the participants' needs and language level. The fourth week's vocabulary activities which required students to use various functions of COCA they learnt till then were taken from the Internet. After students experienced corpus-based vocabulary activities, they reflected their experiences in the questionnaire in the fifth week of the study.

In the first week of the study, first of all, the students were informed about the study. They were also informed about what a corpus is, what COCA is and how to access and use it on the Internet through a power point presentation and Internet connection. The topic of the instruction in the first week of the study was frequencies and collocations. Students were taught to get information about which words are frequently used and which are rare by using the corpus. For example, students were shown how many times the word 'cause' occurs in the corpus (52908). In addition to frequencies, students are given basic instructions about "what word usually goes with X" through collocates function of the corpus. For instance, students are shown how to find on the corpus what words follow the word "incredibly" (the words that the corpus found: important, difficult, powerful, strong ...). After giving instructions for using COCA to get information about frequencies and collocations, students were required to do some corpus-based activities out of the class as an assignment (shown in Appendix A).

In the second week of the study, at the beginning of the course, the answers for the students' assignments were checked and a collocation game about collocates of "do and make" were played as a whole class. In this activity, students were required to use the information they got from the corpus for their assignment about the most frequent collocates of "make and do" (shown in Appendix B). Then, the students were taught how to get data on the corpus about which vocabulary is often written and which is more spoken in English. For example, students are shown how to compare the use of "probably and however" in the written and spoken language (The data gathered from the corpus showed that "probably" is more frequent in spoken language than the written one (newspapers & academic texts). In addition, "however" is more frequent in written language (newspapers & academic texts) than the spoken one. After the activity, students were given a new assignment about topic of the instruction in the second week of the study (shown in Appendix B).



In the third week of the study students were required to do contextual analysis. First of all, they were given a number of sentences obtained from COCA and asked to analyze the context of each sentence to see why one modal verb is used rather than the other. For example, students were given some sentences and asked why “must” is used instead of “have to”. Secondly, they were given some sentences taken from COCA but this time, the modal verbs in the sentences were deleted. Students were required to fill in the blanks with a modal verb they think is most appropriate to use by working with a partner (I told you. Didn't I tell you? You \_\_\_\_\_ fill the name! I told you in the hospital) (shown in Appendix C).

In the fourth week of the study, students were given instruction about how to find synonyms and prepositions (What preposition comes after X) on the corpus. In addition, they were given a worksheet that includes various corpus-based activities taken from the Internet (<http://writingcenter.unc.edu/files/2011/12/corpus.pdf>). The activities on the worksheet required students to use all the instructions they learnt during the study on using COCA. The activities were done in the class as a whole class activity by connecting to the Internet and using COCA. Therefore, students had an opportunity to revise all the basic instructions on how to use the corpus to develop their vocabulary (shown in Appendix D).

Finally, in the last week of the study (5 May, 2011), the questionnaire was conducted with the students to investigate their opinions about using the corpus to develop their vocabulary. The data gathered from the student questionnaire were analyzed quantitatively using Statistical Packages in Social Sciences (SPSS) Version 11.5. The answers to Likert type questions were analyzed quantitatively using percentages and frequencies. The results obtained from the questionnaire provided information about students' perception of using online corpus to develop their vocabulary.

## FINDINGS

This study investigated tertiary level EFL students' perceptions about the use of online corpus (COCA) to develop their English vocabulary. The study explored students' perceptions about the benefits of using online corpus (COCA) in English courses in order to enlarge students' vocabulary.

The results concerning students' perceptions of using online corpus (COCA) are presented in Table 1 respectively. These questions required responses on 5-point Likert scales (ranging from “strongly disagree=1” to “strongly agree=5”). In the table, the responses “agree” and “strongly agree” are combined “A/SA”; “disagree” and “strongly disagree” are combined “D/SD”.

Table-1 Students' perceptions of using online corpus to develop their vocabulary

	A/SA	N	D/SD	M
	%	%	%	
Q 1 I believe online corpus (COCA) is useful for developing my English vocabulary	96.3	3.7	-	4.33
Q 2 Online corpus (COCA) can be used to develop my English vocabulary to support my learning	100	-	-	4.56
Q 3 I generally have positive attitudes towards using online corpus (COCA) in English courses	92.6	7.4	-	4.44
Q 4 I believe that online corpus (COCA) offers a good source to develop my English vocabulary	100	-	-	4.67
Q 5 Online corpus (COCA) enables me to access English words easily and quickly	100	-	-	4.59
Q 6 Online corpus (COCA) helps me to solve the problem of learning collocations	100	-	-	4.70
Q 7 Online corpus (COCA) helps me to see the differences between written and spoken English	92.6	7.4	-	4.59
Q 8 Online Corpus helps me to develop the skill of guessing the unknown words from the context	81.5	18.5	-	4.11



Q 9	I think that I need training to use online corpus (COCA) effectively to develop my English vocabulary (Students need training about how to use online corpus in order to use it effectively in English courses).	96.3	-	3.7	4.52
Q 10	I think using online corpus (COCA) is interesting	85.2	14.8	-	4.11
Q 11	I think using online corpus (COCA) is enjoyable	66.6	25.9	7.4	3.89
Q 12	I believe my motivation to learning new words increases by using online corpus (COCA)	81.5	14.8	3.7	3.96
Q 13	I have difficulty in using online corpus (COCA) as I do not have enough computer operating skills	11.1	-	88.9	1.78
Q 14	I liked corpus-based vocabulary activities	88.9	11.1	-	4.11
Q 15	I want to do corpus-based vocabulary activities in English courses	85.1	11.1	3.7	4.33
Q 16	I use / will use online corpus (COCA) out of the class while studying and doing homework	74.1	25.9	-	4.56

Note: M:Mean Number of Students:27

A: Agree SA: Strongly Agree N: Neutral D: Disagree SD: Strongly Disagree

The mean scores of responses to the questions about the benefits of using online corpus (all questions except the questions 9 and 13) were between 3.89 and 4.70, indicating that students perceive the use of corpus (COCA) to be beneficial to their English vocabulary development. It seems that a majority of the students (96.3 %) believe that online corpus offers them many opportunities to improve their vocabulary. Moreover, all of the participants (100 %) stated that it can be used to improve their vocabulary as an instructional tool and it offers a good source to learn vocabulary.

The results also show that most of the students (92.6 %) have positive attitudes towards using online corpus in order to enlarge their vocabulary. As for the benefits of using COCA, the results indicate that the corpus enables students to access English words easily and quickly (100 %) and helps them to learn collocations without having difficulty in learning. In addition, it enables students to learn the differences between the vocabulary of spoken and written English (92.6 %) and do contextual analysis by guessing the meanings of unknown words from the context (81.5 %). With regard to the students' feelings about using the corpus to develop their vocabulary, the results show that most of the students found using corpus interesting (85.2%) and enjoyable (66.6). In addition 88.9 % of the students stated that they liked corpus-based vocabulary activities. The results also show that 81.5 % of the students believe that their motivation to learn new English words increases through using the corpus. It is noteworthy that a vast majority of the students (88.9) do not have difficulty in using COCA on the Internet. Another result gathered from the study indicated that students need training about how to use the corpus effectively to develop their vocabulary (96.3 %). The results also show that most of the students (85.1) want to do corpus-based vocabulary activities in English courses and they are using or will use COCA (74.1) as an instructional tool out of the class while studying or doing homework.

As Table 1 shows, the overall mean scores of the items indicate that students do not have difficulty in doing corpus-based activities and most of the students have positive attitudes overall towards doing corpus-based vocabulary activities in order to improve their vocabulary in English instruction.

## DISCUSSION

The results of this study indicated that students perceive using online corpus (COCA) to be beneficial for their English vocabulary development. The data gathered from the study also show that students have positive viewpoints about the use of online corpus (COCA) to develop their English vocabulary. This confirms previous research, as reviewed by Varley (2008), which has indicated that



learners believe that corpora are very beneficial for their language learning and they enable learners to improve their English proficiency. A majority of the students who took part in this study think that COCA provides learners with many opportunities to enhance their vocabulary by offering a good source including a wide range of vocabulary. As a result, students can see the differences between the written and spoken language and they might use this information for speaking and writing activities.

The results show that the most important benefit of using COCA is that it helps students to find “what word usually goes with a specific word” by offering collocates function. Students generally have difficulty in using collocates and they make several mistakes while speaking and writing. The corpus helps students to solve this problem as it enables them to see easily what word comes before or follows the target word. Previous studies have also shown that giving a corpus-based vocabulary instruction designed to teach L2 learners collocational pairs is effective in developing their knowledge of collocations (Daskalovska, 2015; Jafarpour et. al., 2013; Walker, 2011). In this study, the participants learnt how to use collocates function of COCA which presents English collocations with frequency numbers. The findings of the study indicate that most of the participants think that corpus enables them to learn English collocations with ease and it seems that they have a positive perception of using COCA to expand their vocabulary with collocations; therefore, they also develop a positive attitude towards using the corpus to improve their competence of collocations. In the same way, previous studies indicate that corpora are regarded as a useful instructional tool by foreign language learners to learn collocations. (Akıncı, 2009; Chao, 2010).

These findings parallel those in the literature. Many researchers have stated that corpora are effective in language teaching (Barlow, 1992; Simpson & Mendis, 2003; Varley, 2008; Willis, 1998). Willis (1998), for example, states that in a corpus students might find out:

- “the potential different meanings and uses of common words
- useful phrases and typical collocations they might use themselves
- the structure and nature of both written and spoken discourse
- that certain language features are more typical of some kinds of text than others” (as cited in Krieger, 2003, p. 3).

In the study, the participants also did contextual analysis through analyzing the sentences taken from COCA and tried to figure out the use of particular grammatical structures. They made an effort to understand the communicative meaning of the target structures used in the authentic sentences obtained from the corpus. The data gathered from this study indicated that most of the students think that the corpus help them to improve their skill of guessing the unknown word from the context since it provides the forms of authentic language use coming from various real contexts. These findings tally with Gabrielatos (2005), who mentions that corpus samples enable learners to make lexical inference, creates a good opportunity to revise target grammar structures and develop an analytical approach to assess the use of the structures critically.

Gabrielatos (2005) also states that authentic language samples taken from corpora help learners to develop their reading skills, especially their skills of guessing the meaning of unfamiliar words from the context. In addition, Gabrielatos (2005) emphasizes that language samples taken from corpora provides learners more opportunities than the ones offered by texts to improve their ability in using contextual guessing technique as corpora include various language samples that show multiple uses of language in different contexts. Similarly, Barlow (2002) suggests that the corpus can be used to analyze the language in a context. It seems that the corpus can assist foreign language learners in doing contextual analysis through presenting authentic and real language samples and so it can serve to enhance their vocabulary.

As regards the benefits of using corpus, the participants also stated that students need training to be able to use the corpus effectively for their vocabulary development. As using online corpora is a





new technique in foreign language teaching and learning, it is essential that students be taught what corpora are, how to access them on the Internet and how to use them to enhance their English vocabulary. Previous research has also emphasized the necessity of giving training to learners on how to use corpora (Cheng et al., 2003; Chambers & O'Sullivan, 2004). During this study, the participants had various training courses on the use of COCA and learnt how to make search in the corpus for vocabulary through using its different functions (frequency, collocations, synonyms, contexts of use, differences in speaking and writing). It can be concluded that learners believe that they cannot use COCA effectively to develop their English vocabulary without receiving practical instruction in corpora.

The results also indicated that most of the participants stated that they did not have difficulty in using online corpus. This result shows that using online corpus does not require students to have advanced computer operating skills. The students who have basic computer operating skills can easily use online corpus (COCA). It is clear from the findings that it is easy for learners to use COCA. As Durand (2018) emphasizes, "COCA has a very user-friendly interface" (p.137). Durand (2018) also states that COCA allows learners to do various searches without any difficulty since it provides tutorials and search commands which facilitates navigating the website of corpus. In addition, a variety of informative videos on using different functions of COCA are available for learners on the Internet. Therefore, learners can benefit from instructional videos to learn how to use corpora and after a short training course, they can integrate it into their learning process as an online reference tool.

A number of studies have indicated that foreign language learners have positive attitudes towards using corpora to develop their vocabulary in the target language (Akıncı, 2009; Chao, 2010; Chujo, Utiyama & Miura, 2006, Paker & Ergül-Özcan, 2017). Likewise, the findings of the study suggest that a majority of the students think that using corpus to learn vocabulary is both enjoyable and interesting. In addition, they stated that using corpus enhanced their motivation to learn new English words. Moreover, a vast majority of the students stated that they liked the corpus and they would use it while studying English and doing assignments. It is clear from the results that the participants are highly likely to integrate corpus consultancy into their learning process to enhance their vocabulary development.

## CONCLUSION

The results of this study suggest that COCA is widely perceived as a useful reference tool by students and they are of the opinion that it helps to develop their English vocabulary in foreign language instruction. Thus, considering their favorable opinions of using online corpora for enlarging their vocabulary, corpus-based vocabulary activities can be integrated in foreign language instruction and the COCA can be used as an instructional tool for students' vocabulary development. However, it is seen from the results that having positive views about using the corpus for developing vocabulary is not enough to integrate corpus-based vocabulary activities in language instruction. The results of this study suggest that there is a need for training for students to be able to use online corpus effectively to enhance their English vocabulary. In fact, although corpora have started to be used in foreign language learning, most of the students and teacher are not aware of this teaching technique. They do not know what a corpus is and how it can be used as a learning and teaching tool to develop their language skills. The corpus is a valuable tool as it shows how native speakers of English use particular vocabulary items, language forms and expressions. It contains millions of words taken from different genres such as newspapers, popular magazines and academic texts. Therefore, both the students and teachers should be introduced what a corpus is and how to use it in foreign language learning. In addition, they should be informed about the benefits of using corpus to improve their English vocabulary.

## PEDAGOGICAL IMPLICATIONS

Vocabulary instruction in a foreign language involves teaching various aspects of vocabulary (the meaning(s) of the words, its spoken and written forms, word parts (prefix, suffix, root), word class, its collocations, its register, what associations it has, what connotation it has and its frequency). A



corpus, such as COCA and British National Corpus (BNC) serves like a native speaker and provides both teachers and learners with accurate and up-to-date information about the use of English through providing users with many functions which enables them to conduct various searches for vocabulary. When students, teachers and curriculum designers are confused about the use of particular vocabulary or grammar item, they can obtain relevant information through doing contextual analysis in corpora. Similarly, in the class, when students look puzzled after the teachers' explanation about the meaning and use of some abstract difficult words, they can benefit from the direct use of COCA as a reference tool and authentic language samples presented through concordance lines can be used to clarify the points students are confused about. Therefore, besides established and conventional vocabulary teaching techniques, corpora can be incorporated into foreign language instruction as an alternative supplementary source and corpus-based vocabulary activities can be designed for EFL students to enable them develop their vocabulary.

On the other hand, teachers should be careful about the design of corpus-based vocabulary activities. The concordance lines taken from corpora to teach vocabulary should be chosen through considering students' language level and language samples gathered from corpora should be meaningful and serve to pedagogical methods used in teaching vocabulary. Moreover, lack of knowledge about how to design and use corpus-based vocabulary activities in foreign language instruction can negatively affect teacher's attitudes towards applying these activities in English courses. Therefore, teachers should be given pedagogical information about how to prepare and use corpus-based vocabulary activities in English instruction.

The corpus offers a good source including a wide range of vocabulary. It can be integrated into foreign language instruction as an effective alternative reference tool for learners' vocabulary development. Moreover, corpora can be used to teach students of different levels. Thus, the results of this study might supply teachers with planning information about how corpora might best be integrated into English language curriculum to develop learners' vocabulary and how learners can benefit from it both in and outside the ELT classroom.

## **LIMITATIONS**

The main limitation of this study is the number of the participants (27 EFL learners) who took part in this study. If the study had been carried out with more students, it could have yielded more reliable results. In addition, in order to get in-depth information about the students' opinions on the use of COCA to develop their English vocabulary, some interviews could have been carried out with some participants. However, unfortunately, because of the time limit problem, the interviews could not be conducted with the participants.

## **RECOMMENDATIONS FOR FURTHER RESEARCH**

This study examined EFL learners' perceptions of using an online corpus (COCA) to expand their vocabulary. Although the participants were at A2 level, they thought that it was not difficult to learn how to use COCA and they were very eager to do corpus-based vocabulary activities. Therefore, further research that investigates higher level EFL learners' (language learners at B2 or C1 level) opinions about the use of online corpora to develop their vocabulary in greater details needs to be conducted.

## **ACKNOWLEDGEMENT**

1. The author would like to thank the participants, EFL students, for their contributions to the study.
2. This study was not financially supported by any institutions.
3. This study was only presented at 15th International INGED ELT Conference: "Taking it to the limits", at Hacettepe University, Ankara, Turkey and the paper has not been published previously anywhere.



## REFERENCES

- Al Saeed, N., & Waly, S. (2009). *Corpus for classrooms: Ideas for material design*. Proceedings of the 10th METU ELT Convention. Ankara, Turkey.
- Akıncı, M. (2009). *Effectiveness of corpus consultancy in teaching verb+noun collocations to first year ELT students*. Unpublished master's thesis. Bogazici University, Istanbul, Turkey.
- Barlow, M. (2002). *Corpora, concordancing, and language teaching*. Proceedings of the 2002 KAMALL International Conference. Daejeon, Korea.
- Chambers, A., & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL*, 16(1), 158-172.
- Chao, Z. (2008). Study on English vocabulary teaching in college. *US-China Foreign Language*. 6 (3), 39-42.
- 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.
- Cheng, W., Warren, M. and Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173-186.
- 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.
- Crystal, D. (1992). *An encyclopedic dictionary of language and languages*. Oxford, 85.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*. 34, 548-560.
- Dash, N. S. (2003). Use of language corpora in second language learning. *South Asian Language Review*, 13 (2), 1-27.
- Daskalovska, N. (2015). Corpus-based versus Traditional Learning of Collocation. *Computer Assisted Language Learning*, 28(2), 130-144
- Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990-2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, 14 (2), 159-190.
- Durand, S. (2018). Embracing computer corpora in the language learning classroom and using it in your classroom. In J. Perren, K. Kelch, , J.Byun, S.Cervantes, & S.Safavi (Eds.), *Applications of CALL theory in ESL and EFL environments*. United States of America: IGI Global.
- Feng-xia, Z. (2009). Using Lexical Approach to teach vocabulary. *US-China Foreign Language*. 7 (8), 44-47.
- Fligelstone, S. (1993). Some reflections on teaching from a corpus linguistics perspective. *ICAME Journal*. 17, 97-109.
- Gabrielatos, C. (2005). 'Corpora and language teaching: Just a fling, or wedding bells?'. *TESL-EJ*, 8 (4), 1-37.
- Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A Corpus-Based Approach toward Teaching Collocation of Synonyms. *Theory and Practice in Language Studies*, 3(1), 51- 60.  
doi:10.4304/tpls.3.1.51-60



- Krieger, D. (2003). Corpus linguistics: What it is and how it can be applied to teaching. *The Internet TESOL Journal*, 9 (3). Available: <http://iteslj.org/Articles/Krieger-Corpus.html>
- Koçak, A. (2010). *The attitudes of one teacher and her students towards using internet sources to develop students' reading skills*. Bilkent University, Ankara.
- McCarthy, M. J. (2004). *Touchstone: From corpus to course Book*. Cambridge: Cambridge University Press
- McCarten, J. (2007). *Teaching vocabulary. Lessons from the corpus, lessons for the classroom*. Cambridge, Cambridge University Press.
- Neff, J., Martinez, F. & Rica, J. P. (2001). *A contrastive study of qualification devices in native and nonnative argumentative texts in English*. Paper presented at the annual meeting of the American Association of Applied Linguistics, St. Louis, MO.
- Paker, T., & Ergül-Özcan, Y. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. *International Journal of Language Academy*, 5(1), 62-81.
- Simpson, R. & Mendis D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 27 (3), 419-441.
- Varley, S. (2008). I'll just look that up in the concordancer: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22 (2), 133-152.
- Walker, C. (2011). How a corpus-based study of the factors which influence collocation can help in the teaching of business English. *English for Specific Purposes*, 30(2), 101- 112.  
doi:10.1016/j.esp.2010.12.003
- Walters, J. (2010). The native speakers on your computer. Paper presented at the 1st Black Sea International ELT conference, Zonguldak, Turkey.
- Willis, J. (1998). 'Concordances in the classroom without a computer', in B. Tomlinson (ed.), *Materials Development in Language Teaching*. Cambridge: Cambridge University Press, pp. 44–66.

## Appendix A.

### Corpus Tasks (week 1): Assignment

1. How many times does the word “affect / accept / change / think” occur in the corpus? Find the frequency number?
2. What word usually goes with the words “extremely / almost /badly / certainly”? Use the COLLOCATES function.
3. What kind of adverbs go with the words “happy /kind / nervous / perfect”? Use the COLLOCATES and POS LIST function.
4. What are the most frequent collocates of “make” and “do”?

## Appendix B.

### Corpus Tasks (week 2)

- Even if house prices are going down, doesn't it still \_\_\_\_\_ sense for to you buy?
- You recommend a fund that one can own in their IRA that would \_\_\_\_\_ money from China's growth
- This is the time where he's got to \_\_\_\_\_ decisions
- I mean, this is an opportunity, opportunity to \_\_\_\_\_ things better
- \_\_\_\_\_ room for your daughter's expression of her feelings about having a car
- If you do, then you know that people \_\_\_\_\_ mistakes



- Up next, women \_\_\_\_\_ fun of how men act when they're sick, but maybe were not just whiners.
- In an interview posted on YouTube, the President called for Mubarak to \_\_\_\_\_ changes, but refused to take sides
- But for now, states across the country are in fiscal crises, facing difficult cuts to \_\_\_\_\_ ends meet
- To \_\_\_\_\_ matters worse, the region's political leaders rarely have the patience or the skill
- I was too excited to play or sleep, much less \_\_\_\_\_ homework

### Assingment

Compare the use of look into & investigate, go up & rise in the written language (newspapers & academic texts) and in conversations (spoken language) in terms of frequency. Take one sentence for each word from corpus.

### Appendix C.

#### Corpus Tasks (week 3)

##### Contextual Analysis

*Analyze the context of each to see why one modal verb is used rather than the other*

1. At most tourist sites, you **have to** wait in line, purchase a ticket, and see things under the supervision
2. with a higher atmospheric CO2 concentration, plants **have to** open their stomata less in order to get the same amount of CO2
3. You want to live in and around a luxury product? You **have to** pay more
4. What I presume to guess about others **must** come from what I have read about and been told
5. I use a dictionary, even though I have been told that I **must** begin to think in Italian
6. Farmers and agricultural managers **must** consider the variety of ways that diversification can occur within the system and develop methods that best meet their specific needs of crop production and resilience

##### Cloze activities

*Fill in the blanks with "must / have to / should / be able to". Make the necessary changes*

1. My disadvantage as a historian is that, for the most part, I lack objectivity. For my information I \_\_\_\_\_ be self-reliant
2. I told you. Didn't I tell you? You \_\_\_\_\_ fill the name! I told you in the hospital
3. In principle, changes in climate \_\_\_\_\_ show up in rainfall statistics, hurricane frequency, temperature records, and so fort
4. He didn't look at his watch, it \_\_\_\_\_ be after midnight  
you may not \_\_\_\_\_ sell your book if you lack a solid platform, says Gini Graham Scott
5. What three goals do you want to accomplish in the next year? What challenges do you \_\_\_\_\_ overcome to reach them?



## Appendix D.

### Corpus Tasks (week 4)

Sample Searches of the Corpus of Contemporary American English (COCA)

[www.americancorpus.org](http://www.americancorpus.org)

1. Is “corpus” a common word? Search for CORPUS using the CHART option. Click on the bars to see examples.

2. What adjectives can I use to describe a corpus? Use the LIST option to search for [j\*] corpus

Use the POS (part of speech) list to find the correct code: [j\*] = any adjective

Select the ACADEMIC section.

3. What prepositions come after research? Search for research [i\*]

Click on the entries to see examples of the different prepositions..

5. But what verbs can you use with the noun RESEARCH? You don't want to repeat “do research” all the time.

RESEARCH

CONTEXT: any verb [v\*]

2 / 0 (=2 words to the left of research; 0 to the right)

ACADEMIC section

6. But what if you're bored with the word RESEARCH. What synonyms can you use? Search for [=RESEARCH]

in the ACADEMIC section

7. Comparing two words: – search and seek.

COMPARE search seek

8. So, how do we use *seek*? Use the CHART option and enter seek

CONTEXT [nn\*] 0 3

This looks for nouns used after the verb “seek”.

(Taken from <http://writingcenter.unc.edu/files/2011/12/corpus.pdf>)

