



## **EFL Learners' Perception of and Attitude to Corpus as a Vocabulary Learning Tool**

**Tara Shankar Sinha**  
 East West University

### **ABSTRACT**

*A considerable amount of research has been undertaken on the use of corpora in teaching English as a second or foreign language. Although there is a growing body of research on the effectiveness of corpus-based vocabulary teaching, there are few studies that have explicitly examined EFL learners' perception of corpus application in vocabulary teaching. The present study explores EFL learners' perception of and attitude to corpus as a vocabulary learning tool. Using a questionnaire, data were collected from thirty-two first-year undergraduate students enrolled in an introductory English language course at a private university in Bangladesh. The findings of the study reveal that EFL learners in general have positive attitudes towards classroom application of corpus. In addition, it has been found that EFL learners have varied opinions about the effectiveness of corpus integration in vocabulary teaching as well as instructor guidance in corpus-based learning activity. The pedagogical implications, including the challenges of using corpora for vocabulary teaching, have been discussed.*

### **INTRODUCTION**

The use of corpora in second and foreign language teaching is steadily growing over the past decades, and it has drawn considerable attention of both researchers and language teachers. Besides, rapid technological developments in recent years have also made corpus integration in language teaching a viable option (Flowerdew, 2012; O'Keeffe, McCarthy & Carter, 2007), and many studies have been undertaken on the classroom application of corpus as a potential tool for teaching English as a second or foreign language. A corpus is "a large and principled collection of natural texts" (Biber, Conrad & Reppen, 1998, p4). However, Johns (1991) first advocated for the use of corpora in language teaching, and introduced the concept of "data-driven learning" (DDL) where learners learn from exploring and analyzing the language data from a corpus.

Although there are arguments both for and against the usefulness of classroom application of corpora, several theoretical arguments have been put forward by some researchers for using DDL activities in second and foreign language teaching. For example, Bondi (2001) argues that working with the corpus data provides learners an opportunity to practice the basic skills of language which ultimately develops language awareness. Moreover, it has also been argued that concordance lines provide authentic language and contexts of language use that pedagogical tasks and materials often lack.

Vocabulary learning, which is an integral part of second language acquisition, can be facilitated by the descriptive analyses of corpora and DDL activities. A corpus can be used as an effective tool for teaching both academic and discipline specific words to adult ESL/EFL learners as it presents lexical items in authentic and meaningful contexts. It has been suggested that corpus-based teaching helps learners to discover the lexico-grammatical patterns of a language (Coxhead & Byrd, 2007; Flowerdew, 2010; Tribble, 2009; Yoon, 2011). But the role of a teacher in corpus-based language teaching is still a highly debated issue. Although DDL, as originally conceptualized, advocated for direct interaction between the corpus and the learners without any intervention from the teacher, there are strong arguments in favor of teacher guidance in corpus-based learning activities. Moreover, studies on learners' perception of DDL and corpus-based teaching have produced mixed results about both the effectiveness of corpus-based language learning and the role of teacher in such learning activities. Therefore, these issues require further investigation. The following section reviews some studies that investigated the effectiveness of corpus use in second or foreign language teaching as well as learners' perception of corpus as a useful tool for language learning.

## **LITERATURE REVIEW**

There are many studies that investigated the effects of corpus use on L2 learners' writing development. While some of these studies (e.g., Gaskell & Cobb, 2004; Yoon, 2008; Yoon & Hirvela, 2004) focused on students' evaluations of corpus use in writing, others mainly addressed the theoretical issues related to corpus integration in L2 writing classes (e.g., Bernardini, 2002; Lee & Swales, 2006; Thurstun & Candlin, 1998; Weber, 2001). There are also some studies that examined the effects of corpus use on vocabulary learning. Though there are variations in the degree to which corpus integration in vocabulary teaching is effective, the general findings of most studies suggest that corpus-based vocabulary teaching can be useful. Using corpora for vocabulary learning helps learners to acquire both vocabulary usage and associated grammar (Coxhead & Byrd, 2007; Tribble, 2002). Balunda (2009) also points out that much of the English language is formulaic in nature and teaching vocabulary as separate from grammar has several limitations. In this case, corpus integration can provide a good platform for teaching vocabulary with grammatical usage in authentic contexts. The following section briefly reviews some studies that found positive effects of corpus on ESL/EFL learners' acquisition of lexico-grammatical patterns.

### **Effectiveness of Corpus**

Koosha and Jafarpour (2006) conducted a study which revealed that using corpus-based materials can be effective in learning collocations of preposition. In another study, MacArthur and Littlemore (2008) investigated whether corpus data help learners to understand the meaning of denominal verbs in English and Spanish. The results indicated that students were able to infer not only the basic senses but also the different senses of the verbs from the corpus examples. Liu and Jiang (2009) examined the effects of integrating corpus and contextualized lexicogrammar in foreign and second language teaching. The findings showed several positive effects of the approach which included improved command of lexicogrammar, increased critical understanding of grammar, and enhanced discovery learning skills. Chao (2010) also found corpus useful for learning the collocation of words. In another study, Huang (2014) explored the effectiveness of paper-based DDL activities in improving the lexico-grammatical use of abstract nouns in L2 writing. The results showed that the learners of the experimental group used

a higher variety of collocational and colligational patterns in their writings and committed fewer linguistic errors in the use of the target abstract nouns. In a more recent study, Ashkan & Seyyedrezaei (2016) compared the relative effectiveness of corpus approach and traditional approach in teaching new vocabulary. The study was quasi-experimental in nature, and data were collected from forty pre-university Iranian female students. A pretest-posttest method was followed for data collection. The results of the study revealed that the corpus approach had a significantly better effect on both learning and retention of new words compared to the traditional approach of vocabulary teaching.

Some meta-analyses (e.g., Boulton & Cobb, 2017; Lee, Warschauer & Lee, 2019) also reported medium to large effect sizes for corpus effectiveness in vocabulary teaching in different contexts. Boulton and Cobb (2017) analyzed sixty-four studies and found large effect sizes for both control/experimental studies ( $d=0.95$ ) and pretest/posttest studies ( $d=1.50$ ). In another recent meta-analysis, Lee et al. (2019) examined twenty-nine studies and reported medium effect sizes for both short-term learning ( $g=0.74$ ) and long-term learning ( $g=0.64$ ) of L2 vocabulary. Moreover, they also found that corpus use had a large effect ( $g=0.92$ ) on improving learners' in-depth knowledge, but it had a relatively small effect on precise knowledge development ( $g=0.46$ ) and productive use ability ( $g=0.53$ ).

The findings of the above studies provide evidence in favor of using corpora in vocabulary teaching, and therefore it can be argued that ESL/EFL learners might benefit from different types of DDL activities. However, before implementing corpus-based teaching techniques in ESL/EFL classrooms, it is necessary to explore learners' attitude to such teaching/learning activities. Learners' perception is an important issue and it requires proper attention as Dörnyei (2005) mentions that learners' perception may considerably impact second language learning. Moreover, different studies on learners' perception of corpus effectiveness as well as teachers' role in corpus-based language teaching have produced mixed results. It has been found that both ESL and EFL learners have varied opinions on these issues. The next section reviews some studies that examined ESL/EFL learners' perception of corpus-based language teaching.

### **Learners' Perception of Corpus**

One of the major studies on learners' perception of and attitude to corpus use in language teaching was conducted by Yoon and Hirvela (2004). They investigated how learners use corpora and how they feel about using it in L2 writing instruction. Data were collected through perception questionnaire and semi-structured interview. It was found that the learners were satisfied with corpus-based learning. The learners reported that corpus use improved their writing skills. In another study, Chujo et al. (2009) found positive attitude of the participants towards the use of DDL tasks for teaching vocabulary and grammar. Simsek (2016) also reported that most of the learners of her study felt motivated and enjoyed the corpus-based activities because they could directly interact with real life data and native speaker language. However, the students also mentioned some drawbacks of corpus data, and their main objection was that some concordance lines were too difficult for them to understand. In another study, Kilimci (2017) found both positive and negative attitudes towards corpus use although most students stated that corpus use had more advantages than disadvantages. While most of the students were satisfied with the application of corpus in vocabulary teaching, their main concern was related to the hands-on DDL. Many of the students pointed out that the DDL tasks presented varying degrees of difficulties to them (e.g., limited number of sentences, difficulty in analyzing the collocation of the target words etc.). Moreover, it was found that the students had two different views about the need for instructor guidance in corpus-based language learning. Slightly more than half of the learners (55.2 %) perceived the instructor's assistance necessary whereas almost half of the learners (48.3 %) stated that they could complete the task without the instructor's help. In a more recent study, Alsolami and Alharbi (2020) explored EFL learners'

perception of corpus use in teaching academic writing. Data were collected through semi-structured interviews, and it was found that the learners had positive attitude towards corpus-based academic writing activities. The researchers also reported that the learners found corpus useful for increasing confidence, providing sufficient input, developing language awareness and promoting learner autonomy.

The above discussion suggests that ESL/EFL learners in general hold positive attitude towards corpus use in second language learning although their opinions on which aspects of language (i.e. vocabulary, grammar, discourse etc.) can be effectively learned by corpus use and the degree of instructor support required in corpus-based learning activities vary considerably. Therefore, these issues require further investigation. Moreover, it is also important to identify the varied levels of challenge that learners face while working with corpus data so that more effective techniques can be developed to make corpus-based teaching/learning activities more fruitful.

## **THE PRESENT STUDY**

The present study was undertaken to explore EFL learners' perception of and attitude to corpus as a vocabulary learning tool. Since there are a good number of studies on corpus effectiveness in vocabulary learning, this study was not designed to directly measure the effectiveness of corpus use in L2 vocabulary learning. It rather addressed what EFL learners think about the usefulness of corpus as a vocabulary learning tool, what challenges they face in using corpus data for vocabulary learning and what they think about teacher support while using a corpus for learning new words.

### **Research Questions**

1. What are EFL learners' views about corpus as a vocabulary learning tool?
2. Which aspect (s) of vocabulary knowledge do EFL learners think can be developed by corpus use?
3. What difficulties do EFL learners face in using corpus data for vocabulary learning?
4. What are EFL learners' opinions about the instructor's guidance?
5. What do EFL learners suggest to enhance the effectiveness of corpus-based vocabulary teaching and learning?

## **METHODOLOGY**

### **Context and Participants**

Data for this study were collected from thirty-two first-year undergraduate students enrolled in an introductory English language course at a private university in Bangladesh. The objective of the course was to develop learners' reading and writing skills along with vocabulary and grammar. The thirty-two participants (18 male and 14 female) were from different departments: Business ( $n=15$ ), Economics ( $n=7$ ), Sociology ( $n=5$ ) and Law ( $n=5$ ). Therefore, the class was heterogeneous in nature. The age range of the participants was 18-20 years. None of the participants had any prior knowledge of corpora.

## Data Collection Tool

Data were collected through a perception questionnaire (see the Appendix). The questionnaire items were adopted from Huang (2014) and Kilimci (2017), and were partially modified to suit the purpose of the current study. Ten Likert-type questions were finally developed to investigate student perspective towards corpus use for L2 vocabulary learning. The scale consisted of six points such as “strongly disagree” = 1, “disagree” = 2, “somewhat disagree” = 3, “somewhat agree” = 4, “agree” = 5, and “strongly agree” = 6. There was also one open-ended question that asked the participants to provide their suggestions about how to improve the efficacy of corpus use in vocabulary teaching. A pilot study was conducted to assess the difficulty and suitability of the questionnaire items and necessary modifications were made after the pilot study.

## Teaching Treatment and Data Collection

This study did not aim at assessing the direct effects of DDL activities on learners’ vocabulary acquisition. The main objective of this study was to find out EFL learners’ views on the use of corpus as a vocabulary learning tool. Therefore, only perception data were collected after teaching certain academic words using a corpus.

Learners’ perception data were elicited after teaching fifteen academic words using the Corpus of Contemporary American English (COCA). The fifteen words were taken from the Academic Word List (AWL) developed by Averil Coxhead (2000). The list consists of 570 words which are divided into ten sub-lists, with the most frequent 60 words in sub-list 1, to the least frequent words in sub-list 10. The list excludes the most frequent 2000 words of English.

Students were first given the 60 words of sub-list 1 and were asked to mark the words in the list that they felt were difficult for them. Only the words that were marked difficult by more than 80% of the students were selected, and thus a list of fifteen words was developed. The words were: approach, assess, assume, constitute, consist, define, derive, evident, estimate, indicate, interpret, legislate, significant, proceed and vary. Then the students were given the list of fifteen words, and the instructor taught them how to search the words in the corpus to identify their collocations, contextual use and grammatical functions. The following figure shows an example of concordance output generated by a search in the COCA.

**Figure 1.** Concordance Output from the COCA for the Word ‘interpret’

SEARCH	CHART	CONTEXT	ACCOUNT
SECTION: ACAD:Education (630) FIND SAMPLE: 100 200 500 PAGE: << < 1/7 > >>			
CLICK FOR MORE CONTEXT <input type="checkbox"/> [?] SAVE LIST CHOOSE LIST <input type="text"/> CREATE NEW LIST <input type="text"/> [?] SHOW DUPLICATES			
1	1997 ACAD Education A B C	drums, sign language, letters and speaking. If those who receive the signs <b>interpret</b> them correctly, ascription of meaning occurs, which can lead to effective learning	
2	1997 ACAD Education A B C	be allowed to disturb the harmony between people. # Language and semantic problems People <b>interpret</b> messages inter alia in terms of their backgrounds and in r	
3	1997 ACAD Education A B C	quite different from the regular' classroom teacher. The substitute teacher is left to <b>interpret</b> lesson plans and implement the use of materials outlined by the abser	
4	1997 ACAD InstrPsych A B C	to set up a study, collect data, select appropriate statistical methods, or <b>interpret</b> the computer-generated reports from conventional statistical solvers. This self rep	
5	1997 ACAD InstrPsych A B C	a strategy, develop a reliable and valid data set, order appropriate analyses. <b>interpret</b> the results, and draw a conclusion? The challenge was to develop a testing	
6	1997 ACAD InstrPsych A B C	of Broudy's three styles of teachers. The " didactic " (which we <b>interpret</b> to be lecture) professor is described as one who " emphasizes memorization, drill	
7	1997 ACAD CommCollegeR A B C	algebra. 5. To represent and solve problems using mathematical techniques. 6. To <b>interpret</b> elementary descriptive statistics. # Critical Thinking 7. To comprehend a	
8	1997 ACAD CommCollegeR A B C	clearly fit under the communication competence. The meanings of other responses were harder to <b>interpret</b> and thus difficult, sometimes even impossible, to codifi	
9	1997 ACAD CommCollegeR A B C	Deal's symbolic frame of leadership in which the leader's primary task is to <b>interpret</b> experience for followers so that what they do has meaning and fits into a large	
10	1997 ACAD CommCollegeR A B C	. As a result, comparisons and trend analyses were difficult to report accurately and <b>interpret</b> with confidence. # Third, some information in this study was gathered	
11	1997 ACAD RoeperReview A B C	likely to use multivariate techniques are also most likely to know how to use. <b>interpret</b> and report them as recommended in the literature. As a result, multivariate	
12	1997 ACAD RoeperReview A B C	estimators that are applicable to the statistical situation in which they are involved and should <b>interpret</b> results accordingly. # Reviewers should make certain that w	
13	1997 ACAD RoeperReview A B C	studies will find the information to be much more valuable if they know how to <b>interpret</b> effect size estimates. # Analysis of effect sizes is often seen as a way	
14	1997 ACAD RoeperReview A B C	subject matter learning in comparison to other first world countries. It is difficult to <b>interpret</b> the data in any way that does not indict American society and its cultur	
15	1996 ACAD Education A B C	Tierney (1992) posits that postmodernist approaches to methodology mean that " we must <b>interpret</b> his words not as if a concrete reality existed irrespective of the	
16	1006 ACAD Education A B C	participation was involuntary and teachers have " thought into " the process, some can <b>interpret</b> their participation as an indirect requirement for maintaining effort et	

Students were instructed to analyze only the academic usage of the words. No specific materials and tasks were developed to teach the words, and the instructor relied on the students' ability to analyze the corpus data. Since the students did not have any prior knowledge of corpus, the instructor had to spend the whole class time (i.e. 90 minutes) on the first day to explain and demonstrate how to use the COCA. There was a large multimedia projector in the class which was used for the demonstration. After the first class, the target words were taught over a period of four weeks when the students met twice a week in the class and the instructor regularly checked their progress. During that period students worked in small groups and analyzed the corpus data generated after searching the target words in the COCA. They used their personal laptops and smartphones for the DDL activity. The instructor monitored each group's progress and provided support only when any group faced difficulty in generating the output or could not interpret the concordance data. The instructor did not provide the meaning of the words, rather encouraged the students to guess the meaning from the contexts and that was found to be the most difficult task for the students. However, the instructor regularly assessed the students' understanding by asking questions and eliciting information about the target words. An example is given below.

*Instructor: Can you give me some words that collocate with 'interpret'?*

*Group-1: to*

*Group-2: correctly*

*Instructor: Ok, can you form a sentence with the word 'interpret'?*

*Group-3: I can interpret*

*Instructor: interpret, what?*

*Group-3: interpret a text*

*Instructor: Ok, then what is the meaning of 'interpret'?*

*Group-4: explain?*

*Group-5: understand?*

*Instructor: 'explain' is more appropriate. Now can you give me a concordance line with the word 'interpret'?*

*(Students identify several concordance lines and tell the teacher)*

However, although the duration of each class was 90 minutes, the instructor could spend only 30 minutes in class for the corpus-based activities as there were other skills that had to be taught. Therefore, students were encouraged to spend at least one hour a week at home to explore the different dimensions of the target words using the COCA. Though each student had a personal computer, it could not be confirmed whether they followed the instruction. As a result, there might be variations in the time spent by each student for the corpus-based vocabulary learning activities.

Finally, after four weeks of corpus activities, the students filled out the perception questionnaire and they took around 15 minutes for the task. For the open-ended question, students were instructed to write two to three suggestions that might help to improve the efficacy of corpus-based vocabulary teaching and learning.

## **Data Analysis**

Data were analyzed to identify the learners' perception of the overall effectiveness of corpus use, the difficulty in corpus use and the area of vocabulary knowledge that can be developed by corpus use. First, the frequencies of preferences on each questionnaire item were calculated and converted into percentage. Then the responses to each question item were summated and the means and standard deviations were calculated using the Statistical Package for the Social Sciences (SPSS 23.0).

In order to enhance the analysis and presentation of the questionnaire data, learners' responses to questions that addressed the effectiveness of corpus use were coded into two categories, "helpful" and "not helpful" by placing all the positive answers (4 "somewhat agree", 5 "agree", and 6 "strongly agree") into the "helpful" category, and all negative answers (1 "strongly disagree", 2 "disagree", and 3 "somewhat disagree") into the "not helpful" category. Similarly, learners' responses to questions that addressed the difficulty in corpus use were coded into two categories, "difficult" and "not difficult" by placing all the positive answers (4 "somewhat agree", 5 "agree", and 6 "strongly agree") into the "difficult" category, and all negative answers (1 "strongly disagree", 2 "disagree", and 3 "somewhat disagree") into the "not difficult" category. Finally, the learners' responses to the open-ended question were paraphrased and analyzed to get an in-depth understanding of the suggestions that the EFL learners provided to improve the effectiveness of corpus use in vocabulary teaching and learning.

## RESULTS AND DISCUSSION

After data analysis, the findings were categorized in order to address the research questions. The first research question of the study was about EFL learners' perception of corpus as a vocabulary learning tool. Questionnaire item one and two elicited learners' responses to address this question. The following table (Table 1) presents EFL learners' views about the effectiveness of corpus-based vocabulary learning.

**Table 1.** Effectiveness of Corpus Use in Vocabulary Learning

Category	Helpful (%)	Not Helpful (%)	Mean	S.D.
Corpus is helpful for learning new words	100	0	5.68	.59
A corpus is more helpful than a dictionary	84.38	15.62	5.16	1.29

The above table shows all students (100%) with the highest mean (5.68) agreed that corpus is an effective tool for vocabulary learning. The standard deviation also shows that all the students had fairly the same view on this matter. The vast majority of the students (84.38%) also opined that a corpus is more helpful than a dictionary for learning new words. Only 15.62 % of the students thought that a corpus is not more helpful than a dictionary. Here the standard deviation shows a relatively wide dispersion of data which means that the learners had varied opinions on this matter. However, it can be claimed from these findings that EFL learners in general have a very positive attitude towards corpus use in vocabulary learning and the majority of them also feel that a corpus is more effective than a traditional dictionary for learning new words.

The second research question was about the aspect(s) of vocabulary knowledge that EFL learners think can be effectively developed by corpus integration. Learners' responses to the third, fourth and fifth questions of the perception questionnaire provided information for this research question. The following table (Table 2) presents the descriptive statistics of learner responses to each question.

**Table 2.** Aspect(s) of Vocabulary Knowledge Developed by Corpus Use

Category	Helpful (%)	Not Helpful (%)	Mean	S.D.
Learning the meaning of words	56.25	43.75	3.78	1.45
Learning collocation	81.25	18.75	5.03	1.33
Learning the grammatical patterns of words	68.75	31.25	4.65	1.49

The average score regarding collocation learning ranks top among the three categories. It can be seen from the table that most students (81.25%) found corpus helpful for learning the collocations of words. The second category that received positive responses from the students is the grammatical patterns of words. Though not as high as collocation, this category received positive responses from a vast majority of the students (68.75%). But it is also important to note that 31.25% of the participants did not find corpus effective for learning grammatical patterns of words. It might have happened due to the nature of the activities they were asked to perform. The instructor asked the students to analyze and identify the grammatical patterns of the words from the concordance lines without explicitly mentioning the forms of the words. Some students, as noticed by the instructor, did not have the necessary metalinguistic knowledge required to perform the task. Consequently, even after providing clues, some students regularly failed to identify the grammatical patterns of the words. However, the last category in the list is meaning of words. It can be seen that almost half of the students (43.75%) did not find corpus data useful for learning the meaning of words. As mentioned earlier, it was noticed during the teaching stage that most of the learners struggled to derive the meaning of the target words from the concordance lines. This is the area where the instructor had to assist the learners very frequently. However, this finding is surprising as some earlier studies (e.g., Huang, 2014) reported positive effects of corpus use for learning the meaning of target words. In this case, the methodological difference between the two studies might provide a plausible explanation. In Huang's (2014) study, paper-based DDL tasks and activities were used, but in the current study no such tasks and activities were used. In the current study, the students directly analyzed the concordance outputs. To sum up, it can be said that the learners found corpus the most useful tool for learning collocation followed by learning the grammatical patterns of words and learning the meaning of words.

The third research question addressed the difficulties that EFL learners face in using corpus data for vocabulary learning. Four questions (six, seven, eight and nine) of the perception questionnaire provided information on this issue. Table 3 summarizes the learners' responses.

**Table 3.** Difficulties Faced in Using the Corpus

Category	Agree (%)	Disagree (%)	Mean	S.D.
Time consuming	78.13	21.87	4.56	1.18
Cut-off sentences	71.88	28.12	4.40	1.38
Unfamiliar vocabulary in data	68.75	31.25	4.37	1.28
Too many sentences in output	84.38	15.62	4.84	1.08



The table shows that most of the students (84.38%) faced difficulty in using the COCA because of the number of sentences generated by a single search. Though the students were instructed to analyze only the academic use of the words, it was still a large number of sentences that posed serious difficulty for the first-year students. Most of the students (78.13%) also opined that the corpus data analysis was time consuming and they had to put a lot of effort in it. However, corpus search generates cut-off sentences and the incomplete texts often create problem in understanding the exact meaning of the words in that context. As can be seen in the table, 71.88 % of the students faced difficulty in analyzing the corpus data because of the cut-off sentences and incomplete texts. Finally, a vast majority of the students (68.75%) mentioned that the number of unfamiliar words in the output also made corpus data analysis very challenging for them. However, probably the last two factors also explain why most of the learners struggled to guess the meaning of the target words from the context. Too many unknown words in the co-text and the incomplete sentences made it difficult for them to derive the meaning of the words successfully from the linguistic context. In other words, contextual support played a very minimal role in inferring the meaning of the target words.

The fourth research question was about instructor's role in corpus-based vocabulary learning. The question was set to find out EFL learners' opinions on the necessity of teacher's guidance in such learning activity. Table 4 summarizes the learners' responses.

**Table 4.** Instructor's Guidance in Corpus-based Vocabulary Learning

Category	Agree (%)	Disagree (%)	Mean	S.D.
Can learn words from concordance lines without any assistance from the teacher	25.00%	75.00%	2.53	1.32

In the current study, the course instructor provided necessary supports whenever the students faced difficulty in generating or analyzing the corpus output. However, as the above table shows, only 25% of the participants felt that they can use corpus on their own for vocabulary learning. 75% of the participants, even after four weeks of corpus-based instruction, felt that they need support and guidance from their teacher. This finding goes against Johns (1991) original idea of DDL that advocated for a direct interaction between learners and corpus data with minimal or no intervention from the teachers. It seems that EFL learners who are using corpus for the first time in their life expect proper and prolonged teacher assistance in order to get the maximum benefit from it.

The last open-ended questionnaire item asked the learners to give some suggestions that could be useful for making corpus-based vocabulary learning more effective and enjoyable. The learner responses address the last research question of the study. The learners' suggestions are summarized and presented in the following points:

- a) More than 80% of the students ( $n=27$ ) suggested to develop specific tasks based on a limited number of concordance lines. In their opinion, specific paper-based tasks designed form corpus output might be more useful than searching a corpus for a word which generates a huge number of sentences and creates confusion for learners.

- b) More than half of the students ( $n=19$ ) mentioned that focusing on one aspect of vocabulary (i.e. meaning, collocation, grammatical pattern) at a time would be more useful than targeting different dimensions of a word at the same time. It would make learning both efficient and manageable.
- c) It has been mentioned in the teaching and data collection section that the students used their personal laptops and smartphones for their class activities. Although there was a computer with multimedia projector, it was mainly used by the instructor for task demonstration. However, those who used smartphones (around 60% of the students) mentioned that they faced some difficulties in completing their tasks because of the screen size of their smartphones. Since most of students did not have any laptop (they had desktops which could not be brought to the class), all of them suggested to conduct the class in a computer lab.

### **PEDAGOGICAL IMPLICATIONS**

The study has several pedagogical implications that could be useful for L2 vocabulary teaching and learning. First, the findings of the study suggest that tertiary level EFL learners have positive attitude towards corpus as a vocabulary learning tool; hence, teachers should integrate corpus in language teaching which facilitate discovery learning and ensures learners' active involvement in the learning process. It promotes learner autonomy. Second, a corpus presents words in authentic contexts and exposes learners to native-like language use. In an EFL context, where exposure to authentic language use is very limited, a corpus can be used as an effective tool for teaching English as a foreign or second language. Third, the findings of the study also reveal that EFL learners do not find corpus useful for developing all aspects of vocabulary knowledge, and corpus output itself can pose different levels of challenge even for tertiary level EFL learners. A language teacher should be aware of these issues while using corpora for vocabulary teaching. Besides, there are other factors (e.g. lack of personal computer, slow internet connection etc.) that can substantially affect the overall success of corpus-based vocabulary learning. A teacher should carefully weigh the factors that can affect learners' success, and take necessary steps to minimize the effects of those factors. Fourth, it has been suggested by most of the learners of the current study that developing specific tasks and activities based on corpus-data might be more useful than random analysis of a large number of sentences which contain many unfamiliar words. Therefore, it might be more useful if teachers can develop hands-on, paper-based tasks based on the proficiency level of the learners and their prior knowledge of corpora. Finally, as it has been found in the current study, teacher's guidance is an important factor that can determine the success or failure of corpus-based vocabulary learning. The course instructor should provide necessary assistance so that learners do not lose motivation and can get the maximum benefit from corpus integration in vocabulary learning.

### **CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH**

Although this study did not measure the actual effects of DDL activities on EFL learners' vocabulary knowledge development, learners' perception data were elicited after teaching some academic words using a corpus, and thus the learners' perception of corpus as a vocabulary learning tool stemmed from their own experience of using a corpus for vocabulary learning. Therefore, the findings of the study provide valuable insights into corpus-based vocabulary teaching in an EFL context.

The results of the study show that EFL learners have both positive and negative attitudes towards corpus-based vocabulary teaching and learning. Overall, they think that corpus is an effective tool for

learning new words, but they also complain about the nature of corpus data which often makes learning difficult for them. Most of them also expect teacher assistance, and hence the absence of teacher guidance might greatly impact the success of corpus-based vocabulary learning, especially if the learners lack any prior experience of using corpora. However, it cannot be suggested that learners' perception is the true indicator of the effectiveness of a language teaching approach because perception might not always represent the reality. But ignoring learners' perceptions altogether might hamper the overall learning process. This does not mean that teachers should blindly follow learners' opinions, but accommodating learners' beliefs about what is good for their learning might prove useful.

The study has certain limitations which should be considered in interpreting the findings. First, the study was conducted to investigate EFL learners' perception of and attitude to corpus-based vocabulary learning, and no test was administered to check the learners' actual performance. An experimental research can find out the real effects of corpus-based tasks on EFL learners' vocabulary knowledge development. Second, learners' perception data were elicited through a questionnaire where only Likert-type questions were used. A major drawback of a Likert scale is that it cannot answer why a participant chose a particular option on the scale. It cannot provide a detailed insight which an open-ended question can. Though there was an open-ended question in the questionnaire, it only elicited learners' suggestions about how to improve the effectiveness of corpus-based vocabulary teaching and learning. Therefore, future researchers should include more open-ended questions and replicate the study to get an in-depth understanding of EFL learners' perception of and attitude to corpus-based vocabulary learning.

## REFERENCES

- Alsolami, T., & Alharbi, A. (2020). Saudi EFL learners' perceptions of the use of corpora in academic writing teaching. *Studies in English Language Teaching*, 8(4), 94-111. <http://dx.doi.org/10.22158/selt.v8n4p94>
- Ashkan, L., & Seyyedrezaei, S. H. (2016). The effect of corpus-based language teaching on Iranian EFL learners' vocabulary learning and retention. *International Journal of English Linguistics*, 6(4), 190-196. <https://doi.org/10.5539/ijel.v6n4p190>
- Balunda, S.A. (2009). *Teaching academic vocabulary with corpora: Student perceptions of data-driven learning*. (Master's thesis, Indiana University, Indiana, USA). Retrieved from <http://hdl.handle.net/1805/2049>
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165–182). New York: Edwin Mellen Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Bondi, M. (2001). Small corpora and language variation: Reflexivity across genres. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small corpus studies and ELT: Theory and practice* (pp. 135–174). Amsterdam: Benjamins.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393. <https://doi.org/10.1111/lang.12224>.
- 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. Retrieved from [http://www2.kuas.edu.tw/edu/afl/20100430Final/Word/2010comp\\_EPCA.pdf](http://www2.kuas.edu.tw/edu/afl/20100430Final/Word/2010comp_EPCA.pdf).
- Chujo, K., 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 concordancer. In M. Mahlberg, V. González-Díaz, & C. Smith (Eds.), *Proceedings of the 5<sup>th</sup> Corpus Linguistics Conference*.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16(3), 129–147. <https://doi.org/10.1016/j.jslw.2007.07.002>
- Dörnyei, Z. (2005). *The Psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Flowerdew, L. (2010) Using corpora for writing instruction. In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 444-457). London: Routledge.
- Flowerdew, L. (2012). *Corpora and language education*. Basingstoke: Palgrave Macmillan.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32(3), 301–319. <https://doi.org/10.1016/j.system.2004.04.001>
- Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, 26(2), 163–183. <https://doi.org/10.1017/S0958344014000020>
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *ELR Journal*, 4, 1-16.
- Kilimci, A. (2017). Learner Perspectives towards corpus use in vocabulary learning. *International Journal of Language Academy*, 5(6), 343-359. <http://dx.doi.org/10.18033/ijla.3765>
- Koosha, M., & Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal*, 8(4), 192-209.
- Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56–75. <https://doi.org/10.1016/j.esp.2005.02.010>
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721-753.
- Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93(1), 61–78. <https://doi.org/10.1111/j.1540-4781.2009.00828.x>
- MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 159–188). Berlin: Mouton de Gruyter.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Simsek, T. (2016). Turkish EFL learners’ reflections on corpus-based language teaching. *Global Journal of Foreign Language Teaching*. 6(1), 21-27. <https://doi.org/10.18844/gjflt.v6i1.806>
- Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267–280. [https://doi.org/10.1016/s0889-4906\(97\)00013-6](https://doi.org/10.1016/s0889-4906(97)00013-6)
- Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew, (Ed.), *Academic discourse* (pp. 131-149). Harlow: Longman.
- Tribble, C. (2009). Writing academic English--a survey review of current published resources. *ELT Journal*, 63(4), 400–417. <https://doi.org/10.1093/elt/ccp073>
- Weber, J. J. (2001). A concordance- and genre-informed approach to ESP essay writing. *ELT Journal*, 55(1), 14–20. <https://doi.org/10.1093/elt/55.1.14>

- Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes*, 10(3), 130–139. <https://doi.org/10.1016/j.jeap.2011.03.003>
- Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology*, 12(2), 31–48. <http://dx.doi.org/10.125/44142>
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257–283. <https://doi.org/10.1016/j.jslw.2004.06.002>

## APPENDIX

### Perception Questionnaire

Name: .....

Age: ..... Gender: .....

Department: ..... Semester: .....

---

The following questions are regarding your opinions on using the COCA. Please use the scale below to circle the response that most closely resembles your perspectives.

	1. strongly disagree	2. disagree	3. somewhat disagree	4. somewhat agree	5. agree	6. strongly agree
1. Corpus is helpful for learning vocabulary	1	2	3	4	5	6
2. A corpus is more helpful than a dictionary for learning new words	1	2	3	4	5	6
3. Using corpus is helpful for learning the meaning of words	1	2	3	4	5	6
4. Using corpus is helpful for learning the collocation of words	1	2	3	4	5	6
5. Using corpus is helpful for learning the grammatical patterns of words	1	2	3	4	5	6
6. I faced difficulty in studying the concordance lines due to time and effort spent on analyzing the data	1	2	3	4	5	6
7. I faced difficulty in using the corpus due to cut-off sentences in concordance output	1	2	3	4	5	6
8. I faced difficulty in studying concordance lines due to unfamiliar vocabulary in the data	1	2	3	4	5	6
9. I faced difficulty in using the corpus due to too many sentences in concordance output	1	2	3	4	5	6
10. I think I can learn the usage of the words from concordance lines without any assistance from the teacher	1	2	3	4	5	6

**What suggestions would you like to give to improve the effectiveness of corpus use in vocabulary teaching and learning?**

a).....

.....

b).....

.....

c).....

.....

*Tara Shankar Sinha is a senior lecturer in English at East West University, Bangladesh. He received his master's degree in applied linguistics from the University of Victoria, BC, Canada. His teaching and research interests include instructed second language acquisition, form-focused instruction, language assessment, and the use of technology in language teaching.*

Email: [tarashankarsinha@gmail.com](mailto:tarashankarsinha@gmail.com)