

Abstract

The study reports the findings of a survey conducted in the form of a questionnaire to investigate learner response to the application of data-driven learning (DDL) approach in the context of vocabulary learning. The questionnaire aimed to assess the learner perception of and attitude to corpus use in learning vocabulary, as a part of a larger scale experimental study aimed at teaching spatial prepositions using DDL techniques and Cognitive Linguistics (CL) approach. The participants of the study were 29 second-year undergraduate students majoring in English. The findings revealed both positive feelings and concerns relative to corpus use although benefits, as were indicated by learners, outweighed disadvantages. Overall, the majority of the learners were highly satisfied with the application of corpus to vocabulary learning. Particularly, their concern was related to the hands-on DDL, which, as the learners stated, presented varying degrees of difficulties to them. Also, learners were divided in their views with regard to the need for instructor guidance. The study concluded with the implications and suggestions for future DDL practices in English language teaching.

Keywords: vocabulary, data-driven learning, attitude, perception, survey

Özet

Çalışma, sözcük öğrenimi bağlamında veri-yönlendirmeli öğrenim yaklaşımının kullanımına öğrenci bakış açısını belirlemek için anket yoluyla yürütülen bir araştırmanın sonuçlarını sunmaktadır. Anket, araştırmacı tarafından daha önce yürütülen ve veri-yönlendirmeli öğrenim tekniklerini ve Bilişsel Dilbilim yaklaşımını kullanarak uzamsal edatları öğretmeyi amaçlayan daha geniş ölçekli deneysel bir çalışmanın parçası olarak, sözcük öğreniminde derlem kullanımına öğrenci algı ve tutumlarını değerlendirmeyi amaçlamıştır. Çalışmaya İngilizce öğretmenliği bölümünde okuyan 29 üniversite ikinci sınıf öğrencisi katılmıştır. Çalışmanın sonuçları, derlem kullanımı ile ilgili olumlu görüşlerin ağırlıkta olduğunu belirtmekle birlikte çekinceler olduğunu da göstermiştir. Genel olarak, öğrencilerin büyük bir çoğunluğu, sözcük öğreniminde derlem uygulamalarının kullanımından memnuniyetini belirtmiştir. Öğrenci çekincelerinin özellikle bilgisayar-tabanlı veri-yönlendirmeli öğrenim ile ilişkili olduğu ve bu yöntemin öğrenciler açısından farklı derecelerde zorluklara neden olduğu belirlenmiştir. Ek olarak, öğrenciler, veri-yönlendirmeli öğrenim etkinlikleri uygulanırken eğitimci rehberliğinin gerekliliği konusunda görüş ayrılığı göstermiştir. Çalışmanın sonunda, İngiliz dili eğitiminde gelecekte veri-yönlendirmeli öğrenim uygulamaları konusunda sezdirimler ve öneriler sunulmaktadır.

Anahtar Sözcükler: sözcük bilgisi, veri-yönlendirmeli öğrenim, tutum, algı, anket.

¹ Çukurova University, Faculty of Education, ELT Department, Adana, Turkey, akilimci@cu.edu.tr

Introduction

A considerable amount of research has been undertaken on the classroom application of corpus approach in the form of DDL (Johns, 1991) as a potential tool for teaching English as a foreign/second language (EFL/ESL). Johns (1991) defines DDL as an approach to learning that perceives a language learner as “a research worker whose learning needs to be driven by access to linguistic data” (p. 2) by means of corpora and a concordance program which generates a concordance of a text or corpus, i.e. “a collection of all the contexts in which a word or phrase occurs in a particular text or *corpus* of texts” (Johns, 1994, p. 319). In this sense, corpora and concordancing serve as a valuable tool “to get students to explore regularities of patterning in the target language” (Johns and King, as cited in Mukherjee, 2006, p. 11) by “set[ing] up situations in which students can answer questions about language themselves” (Hunston, 2002, p.170). In other words, DDL foster learning through discovery by exposing learners to rich context, where they can explore various uses of the target word(s).

Compared with traditional methods, DDL has far greater potential for contextualized instruction and enhanced learning experience (Altenberg and Granger, 2002). Similarly, Bondi (2001) points out that working with real data using simple techniques of corpus analysis provides learners with an opportunity to practice the basic skills of language awareness. Furthermore, it is thought that “corpora nicely complement existing reference works and ... may provide information that a dictionary or grammar book may not provide” (Römer, 2011, p. 214).

In DDL, in some cases, the content is derived from corpora by the course instructor and sensibly turned into exercise handouts and such, while in others the student directly uses the concordancer as a point of reference and without much instructor interference, not unlike a dictionary (Johns, 1986). Thus, according to Lenko-Szymanska and Boulton (2015), there are two main modes for using these grand databases as educational tools in the linguistic context, referred to as direct or indirect exposure to corpora.

Direct exposure to corpora involves the removal of the teacher as a solid layer (and perhaps a strong filter) between learner and corpus content and is a method where the student directly accesses corpus and conducts searches on the concordancer via a computerized interface. This has been the actual, initially envisioned model for DDL. It is therefore regarded in the domain of Computer Assisted Language Learning (CALL) (Cobb, 1997). Whereas direct exposure to corpora through classroom concordancing requires learners to have stronger technical skills for conducting database queries (Lenko-Szymanska, 2015), researchers have proven that indirect approaches, such as classroom activities involving paper-based concordance exercises derived from the corpus are effective in increasing student success in ELT courses (Boulton, 2010). The following section briefly reviews the literature on the application of corpus in the form of DDL to different language teaching contexts.

Related Studies

A growing body of evidence suggests that learners, in general, benefited from DDL approach applied to English language teaching pedagogy (Yoon and Hirvela, 2004; Liu and Jiang, 2009; MacArthur and Littlemore, 2008; Cotos, 2014; Geluso and Yamaguchi, 2014; Huang, 2014).

One such study that utilized a corpus-based approach to vocabulary teaching was conducted by MacArthur and Littlemore (2008) to investigate to what extent the use of corpus data would help learners to work out the meanings of the peripheral senses of denominal verbs in English and Spanish. The results indicated that students were able to

work out not only the basic senses but also the different senses of these verbs from the corpus examples. Another study which examined learners' perceptions of and attitude to corpus use regarding the strengths and weaknesses of corpora was conducted by Yoon and Hirvela (2004) in the context of ESL academic writing courses. The findings of the study indicate that, overall, the students found the corpus approach as beneficial to the development of their L2 writing skills. In a similar fashion, Huang (2014) explored the potential of paper-based DDL activities in improving the lexico-grammatical use of abstract nouns in L2 writing. The researcher found that the control group learners 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 an attempt to improve second language graduate students' knowledge of linking adverbials in the context of an advanced academic writing course, Cotos (2014) conducted a study based on two types of DDL activities derived from a native-speaker corpus and both native-speaker and learner corpora. The results of the study indicated that the participants used linking adverbials more frequently, variedly and accurately. The researcher stated that the more significant improvement was observed in the written production of the students who worked on the corpus that their own writings were included in. Geluso and Yamaguchi (2014) investigated the effectiveness of DDL on improving the spoken language. The researchers specifically aimed to find out whether the students would effectively use the newly learned phrases in their spoken production and how they would react to the DDL approach implemented. Based on the findings, the students believed that DDL was a useful and effective resource in the classroom. Another example use of corpus is related to the application of corpus to contextualized lexicogrammar, a study conducted by Liu and Jiang (2009). The researchers found out that the approach adopted produced positive effects on learners' command of lexicogrammar, critical understanding of grammar and discovery learning skills.

The Study

The study reports the findings from a follow-up survey to a larger-scale experimental study conducted by this author on teaching spatial prepositions using DDL techniques and Cognitive Linguistics (CL) approach. The present study partially replicates the previous studies by Yoon and Hirvela (2004) and Huang (2014) with an aim to investigate learners' perspectives towards corpus use in vocabulary learning.

For the purpose of the study, the following research questions were posed.

1. What do learners generally think about corpus use?
2. Which aspects of the corpus application, in general, will present problems to learners?
3. What is learners' perception of corpus-based instruction for vocabulary learning?
4. What are the advantages and disadvantages of corpus use?
5. What do learners think about the instructor's guidance?
6. What do the students think about hands-on corpus practice versus paper-based activities?
7. What is learners' attitude to corpus use in vocabulary learning?
8. What is learners' overall evaluation of corpus use in vocabulary learning?

Methodology

Context and Participants

The study was held in the context of vocabulary instruction. The data collection took place at a vocabulary class at the ELT department of a faculty of education at a university

in Turkey. A total of 29 students took part in the survey, 23 females and 6 males. All the participants were second-year students majoring in English. Their age ranged from 20 to 22 years old, with a mean age of 21. The learners were not administered a placement test to measure their level of proficiency due to resource limitations. Therefore, they were assumed to be advanced learners for the following reasons. First, they all had studied English as a compulsory subject for about 7 years since the sixth-grade at secondary school. In addition, they had to pass the university entrance examination based on English grammar, reading, vocabulary, and translation in order to major in English at a university. Second, after they were admitted to university, they were required to take a proficiency exam based on four skills at the university's Foreign Languages Centre, according to the results of which they either pursued their studies at their department or had to study a year of freshman English and complete it successfully in order to continue their studies at the ELT department. Hence, the participants were assumed to possess a similar level of proficiency in English.

Questionnaire

The questionnaire was adapted from the questionnaires developed by Yoon and Hirvela (2004) and Huang (2014). The questionnaires used by the researchers originally consisted of 42 and 24 Likert-type questions, respectively. For the purpose of the present study, 31 questions were adapted to investigate student perspective towards corpus use in vocabulary learning. The scale consisted of six points such as "strongly disagree" = 1, "disagree" = 2, "somewhat disagree" = 2 "somewhat agree" = 4, "agree" = 5, "strongly agree" = 5. After negatively keyed items were reverse scored, Cronbach's alpha was computed, and the scale was found to have a high level of internal consistency, $r = .936$.

Procedure

The students received a two-week corpus-based instruction of two pairs of spatial prepositions, *over/under* and *above/below*. BNCweb (CQP-Edition) was used for the corpus-based vocabulary instruction and the construction of related exercises. BNCweb is a web-based client program used for searching and retrieving lexical, grammatical and textual data from the British National Corpus (BNC). The students used the web-based interface to query the BNC regarding the uses of the prepositions studied and its concordance to access to the query results in context. BNCweb can be accessed at <http://bncweb.lancs.ac.uk/bncwebSignup/user/login.php>.

The corpus-informed instruction was conducted by the researcher as the classroom instructor and took place in a technology-enhanced seminar hall which was equipped with a projector, a projection screen, wireless connectivity and high-speed internet. The class met once a week for three-hour sessions dedicated to instruction on vertical prepositions over a two-week period. A week before the instruction started, a practice session was held for the purpose of familiarizing the learners with the corpus query techniques. This session consisting of three class periods introduced learners to technical, practical and conceptual aspects of corpus use such as what corpus is, how it could be related to learning and enhancing vocabulary, how BNCweb could be employed to find information on the use of the target pair of words and practice on BNCweb). The following week, students were grouped into pairs and asked to explore the corpus for the use of the first pair of words, *over* and *above* independently. After the students investigated the uses of the words under focus, the students were handed a pre-prepared concordance materials which highlighted the patterns of the target words, i.e. a condensed version of the concordances that they generated. Following their check against the printed concordances, each group presented their findings to the whole class for a final consensus in relation to the uses of the words in question. During the students' both hands-on and paper-based practice on corpus use, the instructor was ready to

provide assistance where necessary. The third week was a three-hour corpus exploration of the use of *under* and *below* and was carried out in the same manner as it had been in the previous week. After the students completed, corpus-based activities, they were administered an anonymous questionnaire.

Analysis

The questionnaire consisting of 31 questions aimed to assess the learners' perspective towards the use of corpus in vocabulary learning. Some items (Table 4) were negatively phrased in order to keep the learners from marking only one side of the questionnaire (Dörnyei and Taguchi 2010, p.43), therefore, these items were reverse scored in order to make them comparable to the other items. Cronbach's alpha was calculated at .936, indicating a high level of internal consistency. A three-stage analysis was applied to the questionnaire data. First, all the responses were recoded into a super-category for an overall evaluation of learner views on corpus use. Next, related items that focus on the same underlying themes were grouped into seven domains, keeping in mind that multi-item scales are more effective to capture the targeted content domain than single items (Dörnyei and Taguchi, 2010). These domains were analyzed in order to highlight which themes, in general, elicited positive and negative responses from learners. Finally, as a fine-grained analysis, the assessment of the data focused on individual items to pinpoint which aspects of corpus use that learners were specifically satisfied or dissatisfied with. For the convenience of the presentation and assessment of the questionnaire data, learner responses were coded into two categories, negative responses being coded into "disagree" and the positive responses into "agree". All the analysis was carried out utilising the Statistical Package for the Social Sciences (SPSS 24.0). The following are the domains emerged after the questionnaire data were restructured.

- The learner perceptions of corpus use in vocabulary learning
- The practice session and searching technique
- The difficulty with hands-on practice
- The facilitative role of instructor
- The advantage of paper-based concordances
- The learner attitude to corpus use in vocabulary learning
- The overall assessment of corpus use

Results and Discussion

The questionnaire findings are presented in the same order as that of the research questions posed at the outset of the study. Table 1 presents the overall evaluation of the learner responses to the corpus use in vocabulary learning, addressing the first research question: "What do learners generally think about corpus use?". The learner responses are presented as two merged categories "agree" and "disagree". The percentages presented refers to the proportion of all positive or negative responses, not the percentages of participants relative to *agree* or *disagree*. The findings reveal that almost all the responses 96.6 % point to the usefulness of the corpus pedagogy, as is indicated by responses clustering around *agree* (Mdn 5).

Table 1

Overall analysis of the questionnaire (n=29)

		Overall Evaluation					
Agree	Disagree	M	SD	Mdn	Min*	Max*	IQR
%	%						
96.6	3.4	5	.756	5	3	6	1

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*1: Strongly disagree 2: Disagree 3: Somewhat disagree 4: Somewhat agree 5: Agree 6: Strongly agree

Table 2 presents the data further broken into the seven aforementioned themes to investigate how learners' generally perceived corpus use relative to the themes concerned and addresses the second research question: "Which aspects of the corpus application, in general, will present problems to learners?". It can be seen that learners almost unanimously found corpus practice useful to learn the use of target words, with the median (Mdn=5.5) suggesting that they strongly agreed. In regard to the *practice session* and *searching technique*, 89.7 % of the learners said that they benefited from the practice session held prior to instruction and found the corpus searching technique easy (Mdn=5.5), while only 10.3 % felt that they did not find the practice session and the searching technique helpful.

As for the *difficulty of corpus practice*, the minimum and maximum values, 2 and 6, respectively, and the larger IQR value of 2 suggest an uneven dispersion of responses and a deviation from the median although most cluster around *somewhat agree* (Mdn=4). Namely, while over half of the respondents (65.5 %) had difficulty with the corpus use, slightly more than a third of them (34.5 %) found it easy. In this respect, Boers and Lindstromberg (2008) point out that corpora may contain language that could be difficult for learners as they are essentially unedited. Therefore, they note that it could be more appropriate for teachers to reduce the amount of text by selecting the concordance lines beforehand.

Concerning the last two dimensions that the learners expressed their agreement, the majority of the learners stated that corpus could be an essential tool when learning vocabulary and they would recommend it to their friends, as is supported by median values of 5.5 and 5, respectively, which means that learners responded *agree* or *strongly agree*. On the other hand, learner opinions differed as to *instructor guidance*, which is also supported by the minimum and maximum values, 2 and 5, respectively and a larger IQR value of 2. While 48.3 % of the learners welcomed the instructor guidance during their practice with the corpus, the remaining 51.7 % found it unnecessary, which suggested that they could manage the hands-on use of corpus data on their own. Finally, with regard to the *advantage of paper-based concordances* (96.6 %), the *inclusion of corpus methodology in vocabulary course* (96.6 %) and *overall evaluation of corpus use* (93.1 %), it can be said that learners were in general satisfied with the corpus application in vocabulary learning. To sum up, despite the difficulties encountered, it seems that learners enjoyed their experience with the corpus and particularly appreciated the rich context that DDL exposed them to.

Table 2
 Analysis of the questionnaire by domains

Domains	A%	D%	M	SD	Mdn	Min*	Max*	IQR
Perceptions of corpus use in vocabulary learning	96.6	3.4	5.4	.801	5.5	3	6	1
Practice session and searching technique	89.7	10.3	5.1	.749	5.5	3	6	1
Difficulty with hands-on practice	65.5	34.5	3.8	1.095	5.5	3	6	2
Facilitative role of the instructor	48.3	51.7	3.7	.938	3.5	2	5	2
Advantage of paper-based concordances	96.6	3.4	5.0	.801	5	3	6	1
Attitude to corpus use in vocabulary learning	96.6	3.4	5.1	.820	5.5	3	6	2
Overall assessment of corpus use	93.1	6.9	5.1	.875	5	3	6	1

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

The following analysis focuses on the individual items which make up each domain examined so far. This approach is considered to provide a more fine-grained analysis of the aspects of corpus use that learners were specifically satisfied or dissatisfied with. Table 3 presents data about the particular items comprising the first domain, *usefulness of corpus for learning vocabulary* and addresses the third research question: “What is learners’ perception of corpus-based instruction for vocabulary learning?”. All the students strongly agreed that corpus was helpful for learning both the usage of vocabulary and phrases (Mdn=6), with responses centering around *agree* and *strongly agree*. In the same manner, nearly all the learners (96.6 %) considered corpus to be a useful resource to enhance their vocabulary. With regard to learning the meaning of words, only a few students (6.9 %) expressed disagreement, while the majority (93.1 %) responded that corpus helped them to find the meaning of the words concerned.

When corpus was compared with a dictionary, learner responses showed variation along the positive side of the scale as is also supported by high standard deviation (SD=1.197). Still, most of the learners (93.1 %) thought that corpus could be more helpful than dictionaries in learning words, while only a few learners (6.9) disagreed in this respect. In response to question 1 *If I had known the corpus earlier, my knowledge of vocabulary would have been better*, the positive response rate was 93.1 %, indicating that only a few learners (6.93 %) were not of the same opinion that earlier acquaintance with the corpus would contribute to their knowledge of vocabulary. To sum up, the learners expressed positive opinions, which can be interpreted that they viewed corpus use as a potential resource to improve their knowledge of vocabulary.

Table 3

Learner perception of corpus use for vocabulary learning

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q8. Using the corpus is helpful for learning the meaning of vocabulary.	27 (93.1)	2 (6.9)	5.3	1.130	6	2	6	1
Q15. Using the corpus is helpful for learning the usage of vocabulary.	29 (100)	-	5.5	.686	6	4	6	1
Q30. Using the corpus is helpful for learning the usage of phrases.	29 (100)	-	5.4	.682	6	4	6	1
Q12. The corpus is more helpful than a dictionary for my vocabulary learning.	27 (93.1)	2 (6.9)	4.8	1.197	5	1	6	2
Q1. If I had known the corpus earlier, my knowledge of vocabulary would have been better	27 (93.1)	2 (6.9)	4.8	1.013	5	2	6	2
Q7. Overall, the corpus is a very useful resource for my vocabulary learning.	28 (96.6)	1 (3.4)	5.5	.688	6	3	6	1

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

Table 4 presents findings as to whether learners experienced any difficulties during hands-on corpus exploration of target vocabulary, addressing the fourth research question: "What are the advantages and disadvantages of corpus use?". In general, the learners provided negative responses about the actual use of corpus. All the items with the exception of Q17 and Q18 in the domain investigated shows varying degrees of concerns, which is also supported by the highest and lowest values along with IQR indicating a wider spread of responses in comparison to the domains examined so far. The most difficult aspect of real corpus use seems to be associated with item 4, which refers to the analysis of collocate outputs. 72.4 % of the learners responded that they had difficulty interpreting the collocations of the target words, whereas 27.6 % said that they could handle it easily.

The next most problematic aspect of corpus use, as rated by learners, is related to the limited number of sentences (Q26) in the concordance output. 65.5 of the learners thought that the number of lines in the concordance output was not sufficient for them to work out the uses of words concerned. However, the positive responses to this item contradict with those provided to item 5, which asks whether learners had met with any difficulties due to too many sentences in concordance output, where 48.3 % of the respondents indicated that there were too many concordance lines to tackle with. However, 51.7 %, of the learners stated that the excessive number of sentences did not present any problems to them. This phenomenon could be explained by the fact that

queries for the given preposition returned a large number of concordance lines, yet most of them were only related to some of the senses of the prepositions investigated. In other words, there were too many sentences, yet they were not sufficient to capture all the senses of that particular preposition.

With regard to the remaining items focusing on time and effort spent (Q6), unfamiliar vocabulary (Q28), cut-off sentences (Q9), analysis of concordance output (Q3), and difficulty of texts (Q2), over half of the respondents reported that they faced a range of difficulties in working with concordance outputs. As far as Q17 and 18 are concerned, which focus on learner views on the practice session for corpus use and the corpus searching technique respectively, majority of the learners (93.1 %) indicated that the practice session helped them to learn the corpus technique, whereas only a small number of respondents (6.9 %) stated that the orientation session did not help them at all. With regard to the searching technique, almost all of them (96.6 %) agreed that they did not experience any difficulty searching for and retrieving target words. The findings presented here also correlate with the findings of the study by Yoon and Hirvela (2004), who found that both intermediate and advanced learners faced difficulties using a corpus. Similarly, Geluso and Yamaguchi (2014) state that the learner group in their study had difficulties related to unfamiliar vocabulary and cut-off concordance lines.

Table 4

Practical and linguistic difficulties experienced by learner

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q2. The real texts in the corpus were too difficult to understand.	16 (55.2)	13 (44.8)	3.9	1.407	4	1	6	2
Q3. I had some difficulty in analyzing concordance output.	15 (51.7)	14 (48.3)	3.8	1.354	4	1	6	2
Q4. I had some difficulty in analyzing collocate output.	21 (72.4)	8 (27.6)	4.2	1.037	4	2	6	2
Q5. I had some difficulty in using the corpus due to too many sentences in concordance output.	14 (48.3)	15 (51.7)	3.4	1.181	3	1	6	2
Q6. I had some difficulty in using the corpus due to time and effort spent on analyzing the data.	17 (58.6)	12 (41.4)	3.8	1.380	4	1	6	3
Q9 I had some difficulty in using the corpus due to cut-off sentences in concordance output.	15 (51.7)	14 (48.3)	3.8	1.405	4	1	6	2
Q17. The searching technique was easy to learn.	27 (93.1)	2 (6.9)	5	1.069	5	1	6	1
Q18. The practice sessions in the computer lab were helpful for learning the technique.	(28) 96.6	1 (3.4)	5	1.013	5	1	6	1

Q26 I had some difficulty in using the corpus due to the limited number of sentences in concordance output.	19 (65.5)	10 (34.5)	4	1.414	5	1	6	2
Q28. I had some difficulty in using the corpus due to unfamiliar vocabulary on concordance or collocate output.	17 (58.6)	12 (41.4)	3.7	1.391	4	1	6	2

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

The questionnaire also asked learners whether the instructor guidance had any facilitative role as they were carrying out corpus-based activities. Table 5 demonstrates the related findings that address the fifth research question: "What do learners think about the instructor's guidance?". The two items received similar responses from learners, indicating a division in learner views. Slightly more than half of the learners (55.2 %) perceived the instructor assistance necessary to find the collocations of the prepositions in the concordance output. Similarly, almost half of the learners (48.3 %) stated that they could have grouped the collocates of the said prepositions according to their semantic categories without the instructor's help. Although the students took part in a three-hour training session, it seems that the duration of the training was not enough for learners to acquire the required skills. This finding also confirms the warning by Lamy and Klarskov Mortensen (2007) that students should be familiarized with the concordances and provided with guidance as to how to draw conclusions from samples of citations.

Table 5

Facilitative role of instructor

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q16. I could have easily found the collocates of prepositions in the concordance output without the instructor's help.	16 (55.2)	13 (44.8)	3.8	1.146	4	1	5	2
Q24. I could have easily grouped the collocates of prepositions in the concordance output according to their semantic categories without the instructor's help.	14 (48.3)	15 (51.7)	3.6	.985	3	2	5	1

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

In line with studies literature reporting findings regarding the usefulness of paper-based concordances (Boulton, 2010), the learners in this study also found printed concordances

useful as opposed to the actual use of corpus, as is presented in Table 6. The analysis of this domain also addresses the sixth research question: “What do the students think about hands-on corpus practice versus paper-based activities?”. The learners stated that the handouts of selected concordance lines helped them both determine the additional meanings of the prepositions they investigated and saved them a lot of time. With respect to the advantage of paper-based concordances, Boulton (2010) points out that much of the tedious labor that might result from hands-on practices can be eliminated with the use of carefully constructed materials. The findings presented here also support those presented in Table 4, where learners expressed their concerns about the online corpus consultation.

Table 6

Advantage of paper-based concordances

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q11. Paper-based handouts of selected concordance lines were useful for me to capture the additional meanings of prepositions.	26 (89.7)	3 (10.3)	5.2	1.037	5.0	2	6	1
Q22. Paper-based handouts of selected concordance lines are necessary to save time.	26 (89.7)	3 (10.3)	4.9	1.047	5.0	2	6	2

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

Table 7 presents findings regarding learners’ attitude to corpus implementation in the vocabulary course such as whether they understood the purpose of the activities, benefited from their experience with the corpus, and would use it on their own in the future. These findings also address the seventh research question: “What is learners’ attitude to corpus use in vocabulary learning?” All the learners indicated they understood why they utilized corpus and how it related to vocabulary learning (Mdn=5). Overall, learners rated corpus-informed activities favorably and pointed to its usefulness in vocabulary courses. Almost all the learners expressed positive feelings with responses centering around the median (Mdn=5), which indicated that they appreciated the purpose of the corpus, enjoyed it and considered using it on their own in the future with the exception of a few learners who expressed disagreement.

Table 7

Learner attitude to corpus use for vocabulary learning

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q29. I understand the purpose of using the corpus in this course.	29 (100)	-	5.4	.677	5	4	6	1
Q27. As I have learned more about the corpus, I have come to like it more.	28 (96.6)	1 (3.4)	5.0	.981	5	2	6	2
Q20. I will use the corpus for my vocabulary learning in the future.	28 (96.6)	1 (3.4)	5.3	.850	5	3	6	1
Q31. I want to use the corpus in my other English courses too.	26 (89.7)	3 (10.3)	5.1	1.026	5	3	6	2
Q13. When I am unsure about the usage of a word, I will search for it in the corpus.	27 (93.1)	2 (6.9)	5.1	1.033	5	2	6	2
Q19. When I am unsure about the meaning of a word, I will search for it in the corpus.	25 (86.2)	4 (13.8)	4.8	1.256	5	1	6	2

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

The last domain that emerged from the merging of related items is learners' overall assessment the corpus-informed pedagogy employed. The findings related to this domain are presented in Table 8, which also answers the last research question: "What is learners' overall evaluation of corpus use in vocabulary learning?". The majority of the learners reported that they were satisfied with the two-week corpus-based instruction implemented, as is indicated by the median values ranging from 5 to 6, indicating that they either agreed or strongly agreed. 96.6 % of the learners reported that corpus use should be introduced to all ESL vocabulary courses (Q25) and that they would recommend it to other students in their department or in other settings where vocabulary learning is concerned (Q21). Responses to Q10 (86.2 %) also echoed similar consensus of opinion reported for Q25 with regard to the inclusion of corpus pedagogy in vocabulary courses. Furthermore, most of the learners (93.1 %) indicated that corpus use should also be introduced to other English classes, suggesting that corpus pedagogy might very well be utilized in other courses for learning and enhancing knowledge of vocabulary. The findings related to this domain are also consistent with Yoon and Hirvela's (2004) study. To sum up, nearly all learners expressed satisfaction with the corpus-based teaching and

learning, adding that they would also use it on their own. Furthermore, almost all learners suggested that it should be integrated into all ESL courses.

Table 8

Learners' overall assessment of corpus-informed pedagogy

Questions	A*	D*	M	SD	Mdn	Min**	Max**	IQR
Q25. Corpus should be introduced in all ESL vocabulary courses.	28 (96.6)	1 (3.4)	5.3	.806	5	3	6	1
Q10. I recommend using the corpus in the same course in the future.	25 (86.2)	4 (13.8)	5	1.035	5	3	6	
Q14 Corpus use should be taught in ESL classes.	27 (93.1)	2 (6.9)	5.3	.996	6	2	6	1
Q21. I will recommend the corpus to other students in my department or elsewhere.	28 (96.6)	1 (3.4)	5.3	.797	5	3	6	1
Q23. I will often use the corpus by my own choice.	24 (82.8)	5 (17.2)	4.6	1.352	5	1	6	2

A=Agree D=Disagree

M = mean SD = standard deviation Mdn = median IQR = interquartile range.

*Raw numbers and percentage in parentheses

**1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.

Conclusion

The study shown that the learners highly appreciated the corpus application in vocabulary learning. Overall, almost all the learners expressed their satisfaction towards the implementation of corpus approach in learning spatial prepositions despite the difficulties they encountered during computer-based concordance analysis. The vast majority of the learners expressed the belief that corpus approach should not only be integrated into ESL vocabulary courses but also be a part of ESL courses. However, the findings suggest that a few points should be noted. First, hands-on DDL presents difficulties to some learners. For instance, over half of the learners stated that practical use of corpus was difficult for them when conducting a hands-on analysis of the target vocabulary. The second issue is related to the learner perception of teacher guidance. The students were divided in their views as to whether teacher guidance was necessary. Nearly half of the learners reported that they could not have tackled with the concordance/collocate output if it had not been for the instructors' assistance, while slightly over half said they did not have any difficulty exploring the concordance outputs.

In this respect, it seems that a few suggestions may be in order with regard to designing and implementing DDL in the classroom, which also echoes three factors pointed out by Cohen (2010, p.161): learning styles, learner strategies, and motivation. Cohen notes that the effectiveness of instruction can be increased if teachers actively address these factors. In this regard, while a small amount of exposure to the technicalities of DDL may be enough for some learners, it may, however, take an extended amount of time for others to

acquire skills required for DLL. Based on the findings presented here and already reported elsewhere in the literature, it is obvious that training sessions should be allocated enough time for learners to completely master DDL techniques and sharpen their skills. This training may be designed in a way to give enough time to slow learners but at the same time not to frustrate the quicker ones, i.e. in a way to hold the attention of those who might grasp these skills more quickly. Another point is that the instructor should always be ready not only to assist learners to overcome technical and linguistic issues but also to provide learners with print-out concordances where they are stuck in order to ensure a smooth flow to DDL instruction, as was the case in this study.

Limitations and Further Research

The integration of DDL into vocabulary teaching mainly elicited positive responses from learners. However, taking into account the small number of students and the resource limitations regarding the selection of the participants, it seems that no firm conclusions could be drawn from the findings presented here, nor might they be generalized to other student populations. Given the hands-on difficulties encountered by learners, the DDL training could be extended. For this reason, future studies might replicate the study with a larger number of students using a more rigorous design and a longer training on DDL.

References

- Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg & S. Granger (Eds.), *Lexis in contrast: Corpus-based approaches* (3-48). Amsterdam: John Benjamins.
- 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. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572.
- Boers, F., & Lindstromberg, S. (2008). From empirical findings to pedagogical practice. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (375-393). Berlin: Mouton de Gruyter.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301-315.
- Cohen A. D. (2010). Focus on the language learner: Styles, strategies and motivation. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 161-178). 2nd ed. London: Hodder Education.
- Cotos, E. (2014). Enhancing writing pedagogy with learner corpus data. *ReCALL*, 26(2), 202-224. <http://doi.org/10.1017/S0958344014000019>
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing* (Second Edi). New York; London: Routledge.
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(2), 225-242. <http://doi.org/10.1017/S0958344014000044>
- Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexicogrammatical patterns in L2 writing. *ReCALL*, 26(2), 163-183. <http://doi.org/10.1017/S0958344014000020>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, 14(2), 151-162.

- Johns, T. (1991). Should you be persuaded.: Two samples of data-driven learning materials. *Classroom Concordancing: ELR Journal*, (4), 1–16.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *ERL Journal*, 4, 27-45
- Lamy, M.-N. & Klarskov Mortensen, J. (2007). Using concordance programs in the modern foreign languages classroom. Module 2.4. In: Davies, G. (ed.), *Information and communications technology for language teachers (ICT4LT)*. Slough: Thames Valley University.
- Lenko-Szymańska, A., & Boulton, A. (2015). Introduction: Data-driven learning in language pedagogy. In A. Lenko-Szymańska & A. Boulton (Eds.), *Multiple Affordances of Language Corpora for Data-driven Learning* (pp. 1-14). Amsterdam: John Benjamins.
- Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *Modern Language Journal*, 93(1), 61–78. <http://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. 159–88). Berlin; New York: Mouton de Gruyter.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art and beyond. In: Braun, S., Kohn, K. and Mukherjee, J. (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5-24). Frankfurt: Peter Lang.
- Römer, U. (2011). Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics*, 31, 205–225.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257–283. <http://doi.org/10.1016/j.jslw.2004.06.002>

Appendix A**Questionnaire about using the BNC in ESL Translation****A1. Background information**

Name: _____ Surname: _____
 Age: _____
 Gender: Male _____ Female _____

B. Reactions to corpus approach

The following questions are regarding your opinions on using the BNC. 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	If I had known the corpus earlier, my vocabulary would have been better	1	2	3	4	5	6
2	The real texts in the corpus were too difficult to understand.	1	2	3	4	5	6
3	I had some difficulty in analyzing concordance output	1	2	3	4	5	6
4	I had some difficulty in analyzing collocate output.	1	2	3	4	5	6
5	I had some difficulty in using the corpus due to too many sentences in concordance output	1	2	3	4	5	6
6	I had some difficulty in using the corpus due to time and effort spent on analyzing the data	1	2	3	4	5	6
7	Overall, the corpus is a very useful resource for my vocabulary learning.	1	2	3	4	5	6
8	Using the corpus is helpful for learning the meaning of vocabulary.	1	2	3	4	5	6
9	I had some difficulty in using the corpus due to cut-off sentences in concordance output	1	2	3	4	5	6
10	I recommend using the corpus in the same course in the future.	1	2	3	4	5	6
11	Paper-based handouts of selected concordance lines were useful for me to capture the additional meanings of prepositions.	1	2	3	4	5	6
12	The corpus is more helpful than a dictionary for my translator training.	1	2	3	4	5	6
13	When I am unsure about the usage of a word, I will search for it in the corpus.	1	2	3	4	5	6
14	Corpus use should be taught in ESL classes.	1	2	3	4	5	6
15	Using the corpus is helpful for learning the usage of vocabulary.	1	2	3	4	5	6
16	I could have easily found the collocates of prepositions in the concordance output without the instructor's help.	1	2	3	4	5	6

17	The searching technique was easy to learn.	1	2	3	4	5	6
18	The practice sessions were helpful for learning the technique.	1	2	3	4	5	6
19	When I am unsure about the meaning of a word, I will search for it in the corpus.	1	2	3	4	5	6
20	I will use the corpus for my vocabulary learning in the future.	1	2	3	4	5	6
21	I will recommend the corpus to other students in my department or elsewhere.	1	2	3	4	5	6
22	Paper-based handouts of selected concordance lines are necessary to save time.	1	2	3	4	5	6
23	I will often use the corpus by my own choice.	1	2	3	4	5	6
24	I could have easily grouped the collocates of prepositions in the concordance output according to their semantic categories without the instructor's help.	1	2	3	4	5	6
25	The corpus should be introduced in all ESL vocabulary courses.	1	2	3	4	5	6
26	I had some difficulty in using the corpus due to the limited number of sentences in the concordance output	1	2	3	4	5	6
27	As I have learned more about the corpus, I have come to like it more.	1	2	3	4	5	6
28	I had some difficulty in using the corpus due to unfamiliar vocabulary on concordance/collocate output.	1	2	3	4	5	6
29	I understand the purpose of using the corpus in this course.	1	2	3	4	5	6
30	Using the corpus is helpful for learning the usage of phrases.	1	2	3	4	5	6
31	I want to use the corpus in my other ESL courses too.	1	2	3	4	5	6