



**“I WISH COCA HAD...” – LEARNER PERCEPTIONS ON HANDS-ON CORPUS
USE IN EFL CONTEXT**

***"Keşke COCA..." İngilizce'yi Yabancı Dil Olarak Öğrenen Öğrencilerin Sınıfta Uygulamalı
Derlem Araçlarını Kullanımına Yönelik Algıları***

Zehra İTİK¹ ve Hacer Hande UYSAL GÜRDAL²

¹Öğretim Görevlisi,, İzmir Demokrasi Üniversitesi Yabancı Diller Yüksekokulu İzmir, zehrasavran@gmail.com, orcid.org/0000-0002-3827-7676

²Prof. Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, İngiliz Dili Eğitimi Bölümü, Ankara, uysalhande@yahoo.com, orcid.org/0000-0003-2236-4107

Araştırma Makalesi/Research Article

Makale Bilgisi

Geliş/Received:

19.11.2023

Kabul/Accepted:

04.03.2024

DOI:

10.18069/firatsbed.1393134

Keywords

Data-driven learning,
speaking instruction, corpus-
based learning

ABSTRACT

A growing body of research has examined the integration of corpora into foreign language classrooms in the last few decades. A sub-line of research concerning the interaction between the two fields has been student perceptions. Additionally, the growing integration of online learning into curricula, triggered by the COVID-19 pandemic, has aroused a necessity to explore the application of data-driven learning in online learning environments. The current study, therefore, attempted to elicit learner perceptions on learning vocabulary targeted particularly for speaking skill through direct data-driven learning in an online learning context. The data gathered from the participants (N=28) through a questionnaire and a semi-structured interview revealed that the participants found the use of data-driven learning activities in conjunction with the coursebook and oral production tasks helpful to boost their learning. They did not evaluate corpus consultation as a very challenging task to tackle in general. Overall, the study revealed two significant findings: First, it highlighted that collaborative learning might play a significant role in data-driven learning applications. Second, the findings emphasized the necessity for integration of the audio data to the large, publicly available corpora, such as COCA, which could contribute to future research on exploring the efficiency of data-driven learning for speaking skills.

ÖZ

Son yıllarda, derlemin yabancı dil sınıflarında uygulanmasını inceleyen araştırmaların sayısı giderek artmaktadır. İki alan arasındaki etkileşime ilişkin araştırmanın bir alt alanı da öğrencilerin sınıfta derlem kullanımına yönelik algıları olmuştur. Ek olarak, COVID-19 salgınının tetiklediği çevrimiçi öğrenmenin müfredata entegre edilmesi, çevrimiçi öğrenme ortamlarında veriye dayalı öğrenmenin uygulanmasının etkililiğini araştırma ihtiyacını doğurdu. Bu nedenle mevcut çalışma, çevrimiçi öğrenme bağlamında doğrudan veriye dayalı öğrenme yoluyla özellikle İngilizce konuşma becerisini hedefleyen kelime öğrenimine ilişkin öğrenci algılarını ortaya çıkarmayı amaçlamıştır. Katılımcılardan (N=28) bir anket ve yarı yapılandırılmış görüşme yoluyla toplanan veriler, katılımcıların veriye dayalı öğrenme etkinliklerinin, ders kitabı ve üretime dayalı aktivitelerle birlikte kullanımının öğrenmelerini artırmaya yardımcı bulduklarını ortaya çıkardı. Katılımcılar, derlem kullanımı genel olarak üstesinden gelinmesi çok zor bir görev olarak değerlendirmediler. Genel olarak, çalışma iki önemli bulguyu ortaya çıkardı: Birincisi, çalışma, işbirlikçi öğrenmenin veriye dayalı öğrenme uygulamalarında önemli bir rol oynayabileceğini göstermiştir. İkinci olarak bulgular, COCA gibi halka açık geniş bir derlemde sesli verilerin sunulmasının gerekliliğini vurgulamıştır. Çalışmadan elde edilen veriler, konuşma becerileri için veriye dayalı öğrenmenin verimliliğini irdelemeye yönelik gelecekteki araştırmalara katkıda bulunabilir.

Atf/Citation: İtik, Z. ve Uysal Gürdal, H. H. (2024). “I wish coca had...” – Learner perceptions on hands-on corpus use in efl context. *Firat Üniversitesi Sosyal Bilimler Dergisi*, 34, 2, 659-675.

Sorumlu yazar/Corresponding author: Zehra İTİK, zehrasavran@gmail.com

1. Introduction

Data-driven learning (DDL), a hands-on pedagogical approach that intends to integrate corpus tools into second/foreign language teaching, has begun to attract considerable attention over the last decades. Regarding the direct use of corpora, Fligelstone (1993) and Leech (1997) propose that the integration of corpora into educational contexts could be classified under three categories: 1) teaching about, 2) exploiting to teach, and 3) teaching to exploit. The first one refers to teaching about corpora and corpus linguistics as a departmental course content targeted for language and linguistics-related programs at undergraduate/graduate levels. The second one refers to direct corpus consultation by the teacher to teach or create materials. Teaching to exploit, on the other hand, intends to enable learners to enhance their linguistic abilities by giving them opportunities to consult corpora directly.

As a form of “teaching to exploit,” DDL, which is a common term regarding the direct applications of corpora to language learning environments, has been proposed by Johns (1991). Johns (1991) defines the approach as “the attempt to cut out the middleman as far as possible and to give the learner direct access to the data” (p.30). The underlying principle behind DDL is that learners are regarded as researchers, or language detectives, aiming to explore language patterns on their own. In (Cobb & Boulton, 2015) words:

“Getting learners to explore language is nothing new: they are frequently asked to compare example sentences on the blackboard, or identify features of written or spoken texts (Boulton & Tyne, forthcoming). Using corpora merely moves it up a level, increasing the quantity of data available for examination, systematizing the querying procedures and output language, and potentially allowing learners a greater role in the process.”

The corpus-based DDL has been noted as an effective approach that aims to undertake an active role in learning, which in turn, could serve as a means of increasing learner motivation, boosting their autonomy in language learning, and raising learners’ awareness of communicating in real-life contexts (Bernardini, 2004; Chambers, 2005; Geluso & Yamaguchi, 2014; Granger, 2021). Students are exposed to concordance lines on the target language either by directly accessing corpora or data presented by the teachers in the form of concordance lines, as shown in Figure 1.

Line	Year	Source	Text
1	1996	SPOK NPR_Weekend	to get access . And I guess to some degree , access affects legislation but you surely get that . LIANE HA
2	2009	MAG FieldStream	White River National Forest its heavily hu there s food access and healthy elk population . Some bigger bulls live i
3	2012	ACAD TeachLibrar	educational databases and citation of sources ; Students can access and complete these courses anytime , anywhere . De
4	2006	MAG TownCountry	referrals and advocates for research funding , treatment access and insurance coverage . For information , call NEDA
5	1996	ACAD EarNoseThroat	resort to hiring professional " hit men " to find hidden access and perform the central injections for them. (n2 , n7
6	2016	ACAD Human Organization	of NAC access . To test the connection between NAC market access and the number of distributors , a correlation analysi
7	2015	ACAD QuartRevDistanceEd	care provider shortages in the afore mentioned areas . limited access areas include low-income rural areas , inner cities ,
8	2007	MAG Shape	, depending on your weight and metabolism) that you can access as necessary throughout the day . When you eat mor
9	2012	WEB greenpartywatch.org	" the system " is bullshit and purposefully crafted to limit access by anyone other than the corporate parties , and we
10	2004	NEWS NYTimes	acquire top-quality parts at higher cost. # Mr. Benezra said Access Cardiosystems did not contact the F.D.A. or recall dev
11	1995	TV Earth 2	to operate . Well , not if we can crack the access code So , this geo-thingy , it might be used to
12	1994	TV Star Trek: Deep Spac...	them . A couple told me you asked for their financial access codes Helping an elderly couple invest their savings
13	2015	ACAD JAdolAdultLiteracy	the curriculum . A successful IPM activity must incorporate all ACCESS components which are authenticity , collaboration ,
14	1996	MOV Leprechaun 4: In Space	can get out of here . Locked and bolted . : Access denied Shit ! Is there another way in ? -
15	2012	WEB barackobama.com	supports the Ohio GOP 's attempt to eliminate early voting access during the three days before Election Day for all Ohic
16	1990	ACAD CurrentPsych	landmarks , mini-maps of local environments , and a variety of access facilities but this needs to be empirically evaluated v
17	1994	MAG Compute	or mount their creative works on the wall . # Children access features in the Writing Studio by clicking on icons in e
18	1997	MAG USAToday	accurate inventory records and to speed up checkouts ; banks to access financial records more quickly and process transacti
19	2012	BLOG conceptcarz.com	' (Sept. 2) . With its famous free paddock access for all fans ; Historic Festival 30 will be a feast for

Figure 1. KWIC lines for access in COCA

Recently, with the unavoidable integration of distance language education, broadly triggered and accelerated by the global COVID-19 pandemic, corpus consultation in L2 classrooms is a viable option complementary to the tasks and activities offered by the coursebook. Embedding the hard version of DDL into online language learning curricula, rather than the soft version, paper-based DDL, helps eliminate the problem of logistics, which is an oft-cited drawback to using corpora in L2 classrooms. For example, in the case of distance education, all learners need to be equipped with a computer and internet access in order to attend their classes. In this regard, this study intends to contribute to the line of research that aims to investigate the learner attitudes towards the implementation of the hard version of corpora in the case of emergency distance education.

2. Literature Review

A good number of research studies have investigated what learners feel about using the corpora as a language learning tool in class (e.g., Bernardini, 2002; Chambers, 2005; Geluso & Yamaguchi, 2014; Kennedy & Miceli, 2001, 2010; Vannestal & Lindquist, 2007). Most of the research examined corpus use to improve vocabulary acquisition in the written performance of the learners, while only a few studies analyzed spoken performance (Geluso & Yamaguchi, 2014; Aguado-Jiménez, Pérez-Paredes & Sánchez, 2012; Pérez-Paredes & Cantos Gómez, 2004). In general, an extant body of research on DDL use has highlighted positive reactions from learners. Kennedy and Miceli (2001), for example, categorized the learner perceptions of the benefits of DDL as presenting authentic language and opportunities to observe different uses and different functions of a word in different contexts and in particular text forms, all of which helped learners to increase their sense of self-efficacy upon implementing DDL to 17 second-year undergraduate learners with an apprenticeship approach. However, the same group of students reported negative comments about corpus use, centering around the fact that they could not understand all the samples in the dataset and that querying the corpus takes too much time. The time-consuming nature of DDL activities was also reported in other studies, such as Chambers (2005) who implemented DDL as a tool in essay writing with 14 second-year students at the university. In the study, although learners achieved successful queries in the corpus, some of them found corpus consultation "tedious, tiring, and laborious" (p. 118). In Chambers (2005), also the large number of samples found for a word or word sequence and the autonomy given in investigating the lexical and grammatical aspects of the language were appreciated. Another highlighted conclusion by the learners was that corpus activities were appropriate for advanced-level learners, not for beginner-level learners, due to the amount of unknown vocabulary. Comparing the effect of corpus tools and glosses with a control group, Poole (2012) investigated 26 EFL learners' attitudes toward concordance use instead of dictionaries with 5 items. As for the first two items, the author concluded that there was no difference between the perceived difficulty of using a dictionary and concordancing. In addition, the treatment groups did not display any difference in their preexisting vocabulary knowledge. It was reported that the concordance group significantly reported the corpus helped them grasp how to use the words in a sentence. Additionally, there were both positive and negative reactions to too many examples in the corpus. While some participants from the concordance group favored seeing many examples, others found it confusing. The study employed a small number of learners as the analysis of attitudes towards concordancing was a part of the findings from an experimental study. More recently, Wang and Hao (2017) evaluated 76 EFL learners' attitudes after using three types of corpora in vocabulary learning. The findings suggested that the participants reported that the use of corpora was very fruitful, they would recommend it to other learners. However, the participants were not very eager to use the corpus in the future, to use it for other English courses, and they did not think that the more they use it, they would favor it more. In the Turkish EFL learner context, Aşık, Vural, and Akpınar, (2015) implemented DDL to boost lexical awareness using the COCA interface with 126 students majoring in the English Language Teaching department at the university level. Overall, learners were positive about corpus use, yet reported difficulties with the use of the COCA interface. In addition, in contrast to Chambers' (2005) finding, learners found the large number of texts available for one word frustrating. Another highlight of the study was the students' preference towards the use of the corpus in the class with the guidance of a teacher. Concerning spoken language, (Pérez-Paredes & Cantos-Gomez, 2004) conducted a DDL study targeting awareness raising on 25 advanced-level learners' own spoken performance. The study proved effective in increasing learners' noticing of weak points in their own performance, such as the lack of stance adverbials, collocations, and appropriate vocabulary. In a later study, Aguado-Jiménez et al. (2012) examined the direct use of corpora on 47 Spanish EFL first-year university students at C1 level of English according to CEFR. DDL activities were implemented to raise awareness on register and variation. In general, the study was successful at raising learners' awareness of register and variation, and as a result of DDL activities, the learners found the notion of register was easier than they thought before. Finally, Geluso and Yamaguchi (2014) attempted to test DDL with 30 third and fourth-year university students from A2 to B2 levels according to CEFR. The DDL approach was used within a course in which students were also asked to complete speaking journals, prepare student-led lessons, and a behavioral profile study. As a result of the study, learners expressed the usefulness of corpora in improving spoken language and learning new meanings of already-known words. On the negative side, learners found COCA difficult to use and corpus use tedious. All in all, given the range of various learner groups and different educational contexts, it is not really easy to draw common conclusions

from the relevant literature. To this end, the study intends to provide qualitative data for previous research on learner attitudes toward direct corpus use in language classrooms. As Geluso and Yamaguchi (2014) assert, investigating the impact of digital applications of corpus-based learning activities as opposed to paper-based activities that are prepared and tailored by the teacher to meet the objectives of a lesson “could be one exciting avenue”. Distance language learning, therefore, might serve as an ideal learning environment for direct corpus applications to complement the existing curriculum, eliminating the often-cited disadvantages of hands-on corpus use related to logistic problems. In this regard, the study aims to investigate learner attitudes towards the hard version of DDL for vocabulary acquisition targeted for speaking skills and provide evidence to the efficiency of direct corpus use in an online learning environment when used in conjunction with coursebooks in EFL classrooms.

3. Methodology

3.1. Setting and participants

The participants of the study were 28 students attending the English preparatory school at a state university in Turkey. 16 of them were female (%57), and 12 of them were male (%43). The students were placed in intermediate-level classrooms as a result of the placement exam conducted at the beginning of the semester. The exam was prepared by the Testing and Assessment Unit of the institution and aimed to test listening, reading, speaking, and writing skills along with grammar and vocabulary knowledge at B1+ level according to CEFR (Council of Europe, 2001). At the time of the study, the participants were receiving education at B1 - B1+ level courses according to CEFR. The participants were registered in Electric and Electronical Engineering (N=12, %42.9), International Relations (N=10, %35.7), Psychology (N=2, %7.1), Architecture (N=2, %7.1) and Politics and Public Administration (N=2, %7.1) programs.

3.2. Procedure

The study was conducted in a fully online learning environment in an emergency distance learning environment, which was the second academic year after the start of the COVID-19 global pandemic in 2020. DDL activities were integrated into the main course, in which the learners received instruction for four skills in an integrated fashion. The Corpus of Contemporary American English (COCA), as a very large and freely accessible corpus, constructed the primary source for the DDL intervention. COCA involves spoken and written texts totaling over 1.1 billion words gathered between 1990-2015. This study was part of an experimental study with a control/treatment group pretest-posttest design. Therefore, before starting the DDL activities, the learners in the experimental (N=28) and the control group (N=28) sat for a speaking pretest, in which they were asked to participate in pairs and were audio recorded. The three parts of the speaking test required the pairs to provide an answer to an open-ended question (individual task), to describe a picture (individual task), and to come to an agreement on a given prompt (pair task). In addition, the learners and the teacher received two hours of training on how to use COCA functions. In addition, one sample DDL activity was completed in class for four consecutive days of class. The activities aimed at using the Keyword in Context (KWIC) function to search for single words, collocations, and synonyms in general. The DDL vocabulary experiment lasted 10 weeks in total. Each week, the learners received the treatment on a specific day for a two-hour session. The part of the coursebook targeted vocabulary teaching with a specific focus on speaking skills. Therefore, in addition to the in-class speaking activities in which participants completed an individual one-minute speech task, they were given a dialogic task in pairs working on the DDL worksheets every two weeks. As a result, each student submitted 8 monologic tasks, and 4 dialogic tasks. The content of the main course was not completely replaced with DDL activities, but rather, hands-on DDL activities were combined with the coursebook activities. Vocabulary sections in the coursebook were adapted into activities in which the students switched to COCA and completed the same type of activity in the coursebook without changing the objective of the lesson. The control group used only the coursebook. At the end of the treatment, both groups took the posttest. Both the pretest and the posttest were conducted by the researcher online by using Microsoft Teams. The recordings for each pair were then orthographically transcribed, and the transcriptions were subjected to All the participants in the treatment group took the questionnaire and five of them took the interview. Before examining the questionnaires and the interviews, linguistic analyses using Natural Language Processing tools

were run whether the corpus group performed significantly better than the control group. More specifically, the analyses aimed to uncover whether the corpus group produced lexically, morphologically, and phraseologically more complex words and collocations in their oral production. Upon establishing that the corpus group employed more complex words and phrases on the lexical and phraseological level, the questionnaire, and the interviews with the participants from the corpus group were analyzed.

3.3. Data collection

The questionnaire aimed to investigate the participants' attitudes towards a) the general use of COCA, b) the challenges and benefits of using COCA, c) and the use of COCA to improve their L2 lexicon. The questionnaire included 44 items in which the participants could show their level of agreement on a 5-point Likert scale which 1 stands for "definitely disagree", 2 for "disagree", 3 for "not sure", 4 for "agree" and 5 for "definitely agree". It was translated into Turkish from an English version adapted from Geluso and Yamaguchi (2014), and the translation was rechecked by a native Turkish expert in the field of English Language Teaching. It was given in the last week of the study. The questionnaire was prepared using Microsoft Forms, and the links were shared with the students in an online session with the teacher. The second data collection instrument was a semi-structured interview carried out with five learners. The participants who attended the interview were the students who volunteered to do so. The interview involved four questions concerning the benefits and challenges of corpus use, its potential contribution to improving language learning, and the effect, if any, on the participants' views on language. The interviews were conducted with each participant separately in an online meeting in Microsoft Teams. The interviews lasted approximately 15 minutes for each participant. The semi-structured interview attempted to elicit more in-depth data concerning the items in the questionnaire. All participants signed an online consent form before participating in the study.

3.4. Data Analysis

The study conducted both quantitative and qualitative analyses. As for the quantitative analysis, the questionnaire results were analyzed using descriptive statistics for a total of 44 items on a 5 Likert-type scale. The findings were reported as the sum of percentages for the learners who opted for "agree" and "definitely agree", along with the mean and mode values, and the standard deviation for each item. The descriptive statistics were enriched with the insights yielded by the information obtained from semi-structured interviews with 5 volunteering participants. The responses of 5 participants, which form the qualitative data of the study, were examined using content analysis through ATLAS.TI. Content analysis refers to a research method that involves systematically examining and interpreting textual, audio, and visual data to detect patterns or other meaningful information (Hsieh & Shannon, 2005) In parallel, the interviews with 5 participants were first orthographically transcribed by the researcher. In the decontextualization stage, all data for all participants were read by the researcher to get a general sense of the participants comments and to become familiar. In the second step, the transcriptions were coded using the ATLAS.TI. under 5 main themes following a deductive design to code the data as in the questionnaire: the positive effect of the corpus use, the challenges of corpus use, the effectiveness of corpus-based lessons, the effectiveness of speaking portfolios, and their general attitudes towards corpus-based learning approach. The third step was to explore, if any, the emerging themes in the data. As for the last step, the data was rechecked for all the existing and emerging codes to ensure the validity and reliability of the findings. As Friginal (2018) highlights, if the DDL is not effective, there is no meaning in adding it as a major component of the already overburdened curriculum, even if the learners enjoyed using DDL in vocabulary learning. To this end, after conducting a quasi-experimental study following a pretest/posttest design with an experimental and the control group, the current study exploited data elicited from the semi-structured interviews on domains of difficulties and benefits of using corpora for learning vocabulary and the perceived impact of the corpus-based intervention on improving their L2 spoken production. Therefore, the study has potential in contributing to the existing literature by using data not based on subjective observations, but on evidence through quantitative assessment.

3.5. Limitations

The study has a number of limitations. First of all, the small number of participants (N=28) limits the generalizability of the findings. Additionally, the DDL intervention remained limited to only a two hour-session a week. Therefore, further research should consider evaluating perceptions from a larger number of participants and increasing the DDL sessions the learners were exposed to. This study provides support for the necessity of investigating the electronic implementation of DDL, rather than paper-based DDL. However, direct corpus consultation in online learning environments should be more closely investigated.

4. Findings And Discussion

What are the participants' perceptions regarding the use of corpus-based activities for learning vocabulary targeted for speaking skills?

In order to answer this research question, the findings from the questionnaire and the semi-structured interviews were combined. Therefore, both qualitative and quantitative findings were integrated. The findings were examined under five themes. As for each theme, first, the descriptive results from the questionnaires are presented. In addition, the findings were backed up using qualitative data from the interviews and explored using content analysis. The themes are as follows: 1) the positive effect of corpus use, 2) the challenges of corpus use, 3) the effectiveness of corpus-based lessons, 4) the effect of DDL tasks for speaking, and 5) attitudes towards corpus-based learning.

4.1. Positive Effect Of Corpus Use

The first theme investigated was the positive effect of corpus use. 12 items in total (items 3, 13, 14, 15, 18, 19, 20, 21, 22, 23, 26) were used to examine the theme. The findings are presented in Table 1.

Table 1. The positive effect of corpus use

Items	N	Agree	Mean	Mode	SD
3. Corpus was helpful for language learning	28	24 (%86)	4.14	4	.84
12. Corpus was more helpful than a dictionary for finding common phrases	28	24 (%86)	4.17	4	.77
13. Corpus was helpful for learning the usage of vocabulary items	28	24 (%86)	4.17	4	.77
14. Corpus was helpful for learning the usage of phrases	28	25 (%90)	4.32	4	.66
15. Corpus was helpful for learning grammar	28	17 (%61)	3.57	4	1.06
18. Corpus was helpful for improved knowledge of collocations	28	23 (%83)	4.28	5	.76
19. I learned new vocabulary	28	26 (%93)	4.46	5	.74
20. I Improved my English	28	20 (%72)	3.89	4	.87
21. Using the corpus was equally helpful for both speaking and writing	28	16 (%58)	3.50	4	.92
22. Using the corpus was helpful for writing	28	23 (%83)	3.92	4	.85
23. Using the corpus was helpful for speaking	28	22 (%79)	4.07	4	.81
26. Corpus was helpful in finding new ways to use familiar vocabulary	28	23 (%83)	4.28	5	.85

The findings indicated that a vast majority of the participants agreed on the positive effect of corpus on learning new vocabulary (%93) and learning the usage of new phrases (%90), with a mean score of 4.32. This finding finds echo in the relevant literature. In a meta-analysis conducted by (Mizumoto & Chujo, 2015), DDL use has been claimed to have potential, particularly for lexicogrammar. This finding was also supported by the findings from the follow-up interviews. Participant 6, for instance, emphasized expanding vocabulary knowledge as a major benefit of corpus use. The participant stated:

P6: The biggest benefit of the corpus was vocabulary, yes. I believe it was vocabulary for everyone. However, I forget the words I learned very quickly. Actually, this is something about me, but I say here, for example, you give a word. We are researching it, but there are different words in the sentence we chose. For example, words we do not know. We are investigating them. This time, different words come out from there. So we learned word after word. This might seem like a challenge, but it actually turned into an advantage.

The same participant (6) also highlighted the retention effect of corpus on learning phrases. The student commented:

P6: I don't know if it's because it's fun or something, but when we look at the word, but it was the phrases that caught my attention the most. I benefited a lot from this. Especially, it was very easy to remember the phrases I noticed after the class.

Furthermore, participant 7 commented:

P7: In my opinion, the biggest benefit it provides is vocabulary learning, since we see different examples in use, we see some academic content on this site. Apart from that, I can say that since it contains the language that people use in daily life, I can say that it gives a little more familiarity with how people speak in daily use. This comment also leads to another very commonly referred aspect of the corpus, which is authenticity. Learners' exposure to genuine language samples has been asserted to improve vocabulary acquisition, detect linguistic patterns, and construct form-meaning relationships (Römer, 2011; Gilquin & Granger, 2010; Bernardini, 2004). Furthermore, particularly for grammar and vocabulary instruction, the use of authentic language is favored by the current views on communicative approaches. The participants seemed to enjoy examining samples of genuine language use in speech. Participant 23 contributed to this:

P23: As far as I can see from the series I watch or something, we could understand phrases more easily because these contexts are used more frequently. In other words, we learned how the word is actually used in daily use rather than memorizing it. This was more important. You know, since it's a language you need to be able to use in daily life, it's better to have a word together with a context on its own.

In addition, the participants, in general, were certain about the usefulness of corpus in language learning (M=4.14, %86), in learning the usage of vocabulary items (M=4.17, %86), and the usefulness of the corpus in finding the common phrases better than a dictionary (M=4.17, %86). Regarding this topic, Participant 3 commented:

P3: And frankly, the thing I have the most difficulty with in learning a language is memorizing words from a dictionary. In other words, seeing words in a sentence while researching allowed us to learn different words in that sentence. I think that was the biggest benefit. I think it is very different to memorize a word just and see it in a sentence. Because, for example, when I just read the word "connect", very different meanings can come to my mind, but it may mean "having a bond" with someone. In other words, compared to a dictionary, it was much more useful for us to see it in a sentence. We have learned the exact meaning of the words. It also helped us to notice our mistakes in expressing ourselves to other people.

Additionally, participant 14 commented:

P14: It caused us to read many sentences. Therefore, we learned new words, you know, too many of them, I don't know how many of them I gave enough attention to, but I believe that I learned new things because the color-highlighted KWIC lines made it very easy to see the words used together and remember what I worked on.

The benefit of the corpus in improving knowledge of collocations (M=4.28, %83), and in finding new ways to use vocabulary items the learners were already familiar with (M=4.28, %83). The participants also agreed on the effectiveness of the corpus in writing (M=3.92, %83) and speaking (M=4.07, %79). Participant 6 commented:

P6: I felt more comfortable speaking and more like, you know, I made progress in speaking more fluently without getting stuck on grammar.

However, approximately %60 of the students agreed that the corpus was equally useful for writing and speaking. Next, %72 of the people believed that they improved their English in general. Finally, %61 of the participants thought that the corpus helped them learn grammar, with a mean score of 3.57. This could be because the corpus-based lesson focused on vocabulary use integrated with speaking skill, rather than grammar. An additional positive side of the corpus that was appreciated by the participants was free and easy access to the corpus. Participant 6 commented:

P6: So it's free right now, we use it for free, sir, it's free and well, it's easy, you get an email right away, you sign up immediately, it's over, that's it, it's both easy and has a lot of content.

Another benefit of COCA was easy and quick access to the content, which facilitated time management. Participant 14 expanded on that:

P14: Time management is one of the positive aspects of a program that teaches without taking much time because the site offers us many alternatives very quickly.

4.2. Challenges Of Corpus Use

The next theme, which is among the most commonly cited in corpus-based learning, was the perceived difficulties of using corpora on the part of the learners. Particularly, it was a widely claimed notion in corpus research investigating perceptions after using corpora was the level of difficulty of corpus use, especially for lower-level learners. However, the findings imply that only a minority of the participants agreed on the difficulty of using the corpus. Table 2 illustrates the findings.

Table 2. Challenges of corpus use

Items	N	difficult	Mean	Mode	SD
1. Concordancing was difficult.	28	7 (%25)	2.89	2	.87
2. Learning to use COCA was difficult.	28	13 (%47)	3.21	4	1.13
6. Finding phrases around key words was difficult	28	4 (%15)	3.71	4	.89
8. Using the corpus was difficult due to unfamiliar vocabulary	28	11 (%40)	3.14	3	.93
9. Using the corpus was difficult due to cut-off sentences	28	15 (%54)	3.35	4	1.09
10. Using the corpus was difficult due to too many sentences	28	5 (%18)	2.28	2	1.08
11. Using the corpus was difficult due to limited access to internet	28	7 (%25)	2.35	1	1.39
27. It was difficult to understand context of concordance lines	28	11 (%40)	3.10	2	1.06

The item that was agreed by most participants for this theme (%61) was the difficulty of using the corpus due to cut-off sentences, with a mean score of 3.35, which was also found by (Gilquin & Granger, 2010). The next highly agreed item by approximately half of the participants was the difficulty of learning how to use COCA, (%47), with a mean score of 3.21. This finding is not surprising since none of the participants used a corpus or had heard about corpus before. Therefore, they were not familiar with how to use a corpus in technical terms. Concerning this finding, participant 6 commented:

P6: The biggest challenge I had was actually at the very beginning, we tried to figure out how to use the system a little bit. Actually, we tried to solve this problem as a class, it was a bit difficult, making searches and so on. It was a little difficult to register, but the teacher helped us. This problem was solved later.

Participant 14 contributed on this point:

P14: The application is not very easy to use. For me, I had a hard time getting used to it at first. I need to do a search using KWIC. I need to search for a collocation. It was difficult for me to figure out what I should look for, but if we are searching together in class, if I am not doing myself directly, of course, since the teacher described it, I did not have such a problem.

Participant 23, in addition, commented:

P23: I couldn't sign up at the very beginning, so he didn't accept my membership. It did not accept my mail. I had such a hard time. After that, I always procrastinated signing up. So actually, if it was more accessible, I could learn how to use it, even though I attended classes, but I still couldn't learn.

This challenge was attributed to different learning styles and personality traits by participant 23. The student commented:

P23: Corpus was useful, but if it was limited to something, it was not useful if the person was not very good at using technology. For example, I am not good at using technology at all. So maybe that's why I had a hard time.

Finally, participant 3 commented:

P3: The only challenging thing for me was that the system suddenly logged us out of the corpus and asked us to log in again and again too many times.

Depending on the responses, the participants mostly had problems concerning the technical use of the corpus. It is admitted that learners need to receive 'corpus literacy' training before they get their hands on corpus tools (Mukherjee, 2002, p. 179). When learners meet corpora for the first time, it is expected they could feel overwhelmed by the amount of data they see on the screen. However, the findings imply that teacher guidance on the use of the corpus, and training given for each lesson were useful to make up for the challenges.

Next, 11 of the students (%40) believed the corpus use was difficult due to unknown vocabulary, and it was difficult to understand the context of the concordance lines. This finding could be attributed to the fact that, unlike the majority of the studies, the study had participants at the intermediate level, not upper-intermediate or advanced levels, and not with students majoring in language or linguistics-related departments at the university. Römer (2011) posited that complex language representing low-frequency vocabulary items might be a point of concern for beginner-level learners. %40 of the participants (n=11) thought it was difficult to understand the context of concordance lines, with a mean score of 3.10. This was also not interesting as the learners were traditionally exposed to activities in which they were presented with vocabulary items in a reading or a listening text within context. However, not reading horizontally, but vertically, as they are not traditionally familiar, could have served as a challenging skill for the learners. 7 of the participants (%25) found corpus use difficult due to limited access to the internet and found concordancing difficult, with mean scores of 2.35, and 2.89, respectively. Only a small number of participants (n=5, %18) found the corpus challenging due to the comprehensive number of sentences. As Sripicharn (2010) put forward, the use of a large corpus may help learners arrive at accurate conclusions and accurate interpretations of concordance data. This finding was confirmed in the interviews. Participant 6 commented:

P6: There were a lot of sentences, for example, we had the feeling that we would find the best one. In order to find the most beautiful one, for example, we were choosing different words, We were saying that we should choose the most beautiful sentence, for example, everyone was thinking the same way. Then we looked for, for example, different words. I was searching for them as well. In fact, this is not a challenge, but a competition, but let's think of it as a good competition.

Unlike participant 6, participant 7 interestingly commented:

P7: When conducting searches on some phrases, the results were often scarce or nonexistent. I had trouble with that. However, understanding the content of the concordance lines was not very difficult. In general, when I searched for a word, there were too many sample sentences, and the difficulty level of the words in the sentences was not a very big problem.

Therefore, contrary to the widely held belief, the participants did not take the availability of a large number of sample sentences as an advantage.

Finally, only 4 participants (%15) reported difficulties in finding phrases around keywords, not related to finding phrases, but similarly, finding sample sentences for the target keywords or phrases. Participant 14 commented:

P14: Choosing sentences from the corpus is not always easy. Because it was not always easy to find sentences close to the sense that would suit the meaning referred to in class.

4.3. Effectiveness Of Corpus-Based Lessons

The next theme queried by 11 items was related to the effectiveness of corpus-based lessons. It was found that more than half of the participants were positive about the effect of corpus-based lessons. The findings are provided in Table 3.

Table 3. Effectiveness of corpus-based lessons

Items	N	Agree	Mean	Mode	SD
4. The training on how to use the corpus was necessary.	28	20 (%72)	3.78	4	1.03
16. The allocated time in class was sufficient for concordancing	28	19 (%68)	3.78	4	.62
17. I believe I spent enough time concordancing for each speaking portfolio	28	17 (%61)	3.67	4	.90
30. The 1-minute recordings were helpful	28	23 (%83)	4.28	5	.85
31. The 1-minute recordings helped me to understand the context in which to use the key words	28	23 (%83)	4.03	4	.83
34. Having a teacher explain what I find in the corpus was helpful	28	25 (%90)	4.39	5	.99
37. This class afforded me many learning opportunities	28	24 (%86)	4.21	4	.78
39. Using keywords learnt from the corpus in a conversation was an effective way to learn new phrases	28	23 (%83)	4.10	4	.78
40. Using keywords learnt from the corpus in a conversation was a fun way to learn new phrases	28	18 (%65)	3.85	4	.93
41. Using Google docs for corpus-based activities was helpful.	28	21 (%75)	3.96	4	.79
42. This class allowed me to direct my own learning	28	24 (%86)	4.14	4	.75

The item that was agreed upon by the largest number of participants (n=25, %90) was the helpfulness of completing corpus activities with the guidance of a teacher. This finding signifies that the learners may not have felt confident enough about using the corpus individually on their own. Therefore, a ten-week study was not enough for the students to encourage them for autonomous use of corpus. Participant 6 commented:

P6: For example, when we combined our corpus and speaking portfolio, we used what we learned. For example, we found sample sentences from the corpus and pasted it onto a document. After pasting the sentences of our friends there we discussed them one by one with the teacher. That was very effective.

The student added:

P6: In other words, I can say that learning words together in the lesson and looking at the words together actually made them easier to remember. When I learn words on my own, I mostly, I look at the synonyms of some, memorize only the Turkish meaning of some of them. For example, when I look at the English-Turkish dictionary, I find the meaning in Turkish. But when I look at many examples, and study them with the teacher, it becomes much easier to learn these words.

Another important finding was that a substantial majority of the participants (n=24, %86) believed that the class provided them with many learning opportunities, with a mean score of 4.21, and it helped them to direct their own learning, with a mean score of 4.14. This is a critical finding in that it could be interpreted as the corpus-based courses might have provided learners with opportunities to take one step further in improving self-directed learning strategies or autonomous learning although they stated a preference towards the use of

the corpus with a teacher. This finding could be interpreted as the corpus-based activities might be supporting the learners in gaining a sense of responsibility in their own learning. In addition, this finding is also confirmed by (Chambers, 2005), who stated that learners prefer to search for language patterns themselves rather than receiving rule-focused instruction. Participant 14 commented:

P14: We learned both by writing, speaking, and analyzing the sentences in the corpus, searching for adjectives, adverbs and verbs to be used with the keywords. Therefore, it was a three-way learning. We combined what we learned from the corpus with the content from the coursebook. We tried to record our speech, trying to be fluent and to use new words. We also worked on pronunciation. It was a very versatile learning experience.

Although this was not a part of the coursework, participant 3 devised an additional learning strategy for himself and took responsibility for his own learning process. Participant 3 added:

P3: I even have a corpus notebook. I am writing the words I learned on the corpus. It's worked great for me. Now, when I look at the words, I check to see if I know the synonym of the word or not. More precisely, I'm looking, for example, if a second word also meets the meaning of that action, learning it, frankly, would be more advantageous for me to learn a new verb word. Because I can express myself more easily that way. By choosing the words, of course, I wrote the words that would be more useful to me, not all the words I did not know of course, and I was having a look at them 3 4 times a day.

The next three items that more than %80 of the participants agreed on were item 30, "The 1-minute recordings were helpful"; item 31, "The 1-minute recordings helped me to understand the context in which to use the keywords"; and item 39, "Using keywords learned from the corpus in a conversation was an effective way to learn new phrases". A total of 23 participants (%83) agreed on these items, with mean scores of 4.28, 4.03, and 4.10, respectively. These findings were of paramount importance in that corpus-based learning might have been perceived as more useful when integrated into both monologic and dialogic production tasks, rather than when used in isolation. Participant 14 commented:

P14: We used to use basic vocabulary when speaking in class, but corpus lessons helped me look for new, more advanced phrases to express something. For example, we completed a speaking task after using the corpus and I was able to use the words like budget, venue, and sponsor. Normally, I wouldn't use them in a speaking task on a related theme. The words were already on the coursebook, but I don't know, corpus with speaking tasks made them much easier to learn them.

Furthermore, the learners felt using the words learned on corpus in context assisted their learning. Next, three fourth of the participants (n=21) perceived using Google Docs in completing in-class corpus-based activities as useful, with a mean score of 3.96. That might be because using Google Docs was perceived as a fun activity by the students as they had the chance to contribute to the same document simultaneously and use it collaboratively when completing pair or group work activities. %72 of the participants agreed on the item "The training on how to use the corpus was necessary.", with a mean score of 3.78. Therefore, it can be claimed that the finding confirmed the previous findings that training is a crucial part of corpus-based learning, and learners need intensive training before starting to utilize it individually and confidently. However, Sinclair (2004) posits that a few hours of training will suffice for most occasions Participant 23 stated:

P23: Training given was definitely enough. We learned whatever we need to know when using the corpus for an activity especially because the teacher demonstrated every step before activity. So using the corpus was not so hard for us.

Item 16, "The allocated time in class was sufficient for concordancing" (M=3.78) 17, "I believe I spent enough time concordancing for each speaking portfolio", (M=3.67), and 40, "Using keywords learnt from the corpus in a conversation was a fun way to learn new phrases", (M=3.85) was agreed on by 19 (%68), 17 (%61) and 18 (%65) participants respectively. First of all, regarding the fun factor in corpus use, Participant 14 commented:

P14: In online classes, it would be very boring to cover the same activities on the coursebook, as we already studied them through the tutorials. So using COCA was a fun alternative to consolidate what we learned.

Responses to items 4 and 16 indicated that more than half of the participants believed a two-hour session was enough for corpus-based activities, believed they spent enough time engaging with concordance lines, and more than half of them found the use of corpus in class as an enjoyable strategy to enhance their knowledge on co-occurring words.

4.4. Effect Of Speaking Portfolio Tasks

Four items in total investigated the benefits of speaking portfolio tasks. Table 4 elucidates the findings.

The only item that was agreed by the majority of the participants (n=24, %86) was item 36, “Context of planned conversation is important when choosing words and phrases”. Participant 6 commented:

P6: Normally, I would check the meaning of a word in Turkish if I don't know the meaning, and that's it. But in the corpus, we write in English and try to deduce the meaning of that word from the English sentence, you know, we read the sentence, we try to understand the sentence. How did they use the word here? Because sometimes a word doesn't have one meaning, it has more than one meaning. That's why they use it in different ways, for example. How did they use it in the sentence we chose? The corpus showed us the context, and the function of the word.

Table 4. Effect of speaking portfolio tasks

Items	N	Agree	Mean	Mode	SD
5. Choosing a keyword to investigate in the corpus was easy.	28	11 (%40)	2.82	4	1.27
28. It was easy to use my phrases in conversation	28	14 (%50)	3.64	3	.98
29. It was difficult to manipulate the speaking portfolio tasks to use my keywords	28	7 (%25)	2.71	2	.97
36. Context of planned conversation is important when choosing words and phrases	28	24 (%86)	4.21	4	.68

Therefore, it could be inferred that speaking portfolio tasks integrated with corpus-based activities helped learners to comprehend the important role of context in learning vocabulary items and employ them in oral performance. It also means that the corpus-based activities, along with speaking portfolio tasks, helped learners to raise awareness of register. It was confirmed in the literature that the use of corpora could serve as a complementary component to the invented content in readily available materials in language classrooms (Cobb & Boulton, 2015). Therefore, DDL might be more fruitful when used in conjunction with more traditional resources (Geluso & Yamaguchi, 2014). Regarding item 28, the participants proved indecisive about whether they found it easy to use the phrases they learned from the corpus in their speaking tasks. Participant 6 commented on this stating:

P6: I mean, for me it was not difficult because when we combine corpus with speaking portfolio, we tried to use the words we learned from the corpus. For example, we had a task about guests, I remembered 4 of the five words that were about the guest theme in the exam. I mean, because the speaking tasks and corpus activities were about the coursebook activities, it wasn't difficult, but fun.

However, it is interesting to find that in the next item, only 7 participants (%25) found it difficult to manipulate the tasks in their speaking portfolio to use the keywords learned in the corpus-based lessons, with a mean score of 2.71. On the contrary, the participants found it useful. Participant 7 commented:

P7: Corpus, on its own, was not an application to help students do a lot of practice, so combining what we learned from corpus with speaking portfolio tasks was the main thing that improved our speaking skill. Corpus is great to learn different words, but I think trying to use them in speech had a serious effect on my speaking skill.

Participant 14 added:

P14: We really cared about using words. You know, in every task, we tried to use a minimum of three words per person, We even took notes of the words we don't know there, in a blue box. We especially tried to use them, you know, you had to design a context using our imagination as well. After that, we corrected each other's sentences. Before we record a task, we checked what words sounded more natural and more grammatical.

Therefore, the data presents contradictory findings. The reason could be attributed to the fact that speaking portfolio tasks were already adapted from the coursebook activities. Finally, less than half of the participants (n=11, %40) thought that it was easy to choose a keyword to investigate in the corpus. It is an interesting

finding that in most of the activities, the learners were given the keywords to explore, as the target vocabulary items were based on the coursebook itself. However, it might have been due to the fact that they found learning to use COCA difficult in general.

Another interesting comment regarding the integration of the speaking portfolio was by Participant 14:

P14: I think it was the portfolio part that brought the corpus one step ahead, as I said. The Corpus is a very nice app on its own, but it needs to be supported somehow. I think I understand, if I think about it not for speaking skills, but only for learning vocabulary, then of course, I believe that it will be more useful than the coursebook.

Finally, participant 3, referred to a major problem of L2 learners. The student commented:

P3: Since we are all in Turkey, we don't need to speak English much outside of class. In other words, speaking portfolio and one-minute or 2-3-minute recordings with our pair worked great for me to improve my speaking skill. But I think it would be good if the number of tasks were increased a little more.

Overall, with a focus on speaking skill, the participants seemed to agree that the corpus activities needed to be supported with other speaking tasks to be effective in enhancing oral performance. However, when integrated with speaking activities, corpus use could serve as a very fruitful resource for expanding vocabulary knowledge.

4.5. Attitudes Towards Corpus-Based Learning

The last theme that the current study examined was the general attitudes toward and beliefs about corpus-based learning. All items, except one, were agreed on by more than half of the participants. The findings are provided in Table 5.

Table 5. Attitudes toward corpus-based learning

Items	N	Agree	Mean	Mode	SD
24. Corpus use should be taught in English classes more regularly	28	19 (%68)	3.82	4	.90
25. I will use the corpus in future English classes	28	16 (%58)	3.42	4	1.19
32. Using a corpus is best in combination with a dictionary	28	16 (%58)	3.60	4	1.03
33. I trust the phrases I find in the corpus to be grammatically correct	28	19 (%68)	3.78	4	.83
35. After taking this class I believe that grammar and vocabulary are more closely related	28	21 (%76)	3.78	4	1.06
38. I would recommend this type of class to other English learners in Turkey	28	23 (%83)	3.96	4	.96
43. I enjoyed being able to direct my own learning	28	22 (%79)	4.14	4	.75
44. A class of this nature is better suited to advanced learners than beginners.	28	12 (%43)	3.39	3	1.03

The only item that less than half of the participants was item 44. It was a surprising finding since, in related research, it was a recurrent claim that corpus-based lessons are better suited to advanced learners. However, this finding implies that the learners did not perceive corpus-based activities as too challenging or too complex, though %43 of them still find it better for advanced learners. There was one item that more than %80 of the students agreed on (item 38), with a mean score of 3.96. It is a very encouraging finding that the learners found corpus as an effective tool for learning English. In addition, concerning item 38, the majority stated that they would recommend corpus-based lessons to other L2 learners in Turkey. It was a highly convincing finding that all interviewees agreed that they would recommend it to other L2 learners in Turkey, however, in different ways. Participant 23 stressed the importance of learners' aptitude to use technology when learning English. Participant 3 expressed that as learners, they always struggle to find a fun way of learning, and corpus might

be definitely one of them. Participant 7 regarded the use of corpus as the recipe for the ones who desire to learn vocabulary. The student commented:

P7: The collection of texts taken from real use of language, such as from blogs, news, movies, etc. literally gets my attention, and keeps me concentrated and facilitates learning.

Next, 22 participants (%79) expressed that they enjoyed being able to direct their own learning, which implies that they favored corpus activities, giving them opportunities to boost autonomous learning in a way. %76 of the participants (n=21) raised awareness of the close interrelation between lexis and grammar. 19 of the students (%68) counted on the grammatical accuracy of the input in the corpus. This might be owing to the fact that they engaged with the spoken component of COCA, a kind of input that they were not used to “read”. Moreover, 19 participants (%68) reported that corpus use should be taught more regularly in English classes, with a mean score of 3.82. This finding indicates that the participants of the study welcomed the use of corpus as a part of the curriculum. Finally, item 25, and item 32 were agreed on by 16 participants (%58). More than half of the participants, though not a very high number, reported they would use corpus in future English classes. That might be because they do not seem to feel confident enough to conduct corpus searches on their own. Finally, more than half of the participants agreed that they would benefit best from the corpus in combination with a dictionary. This might be attributed to the fact that nearly %40 of the participants found corpus use difficult due to unknown vocabulary and found it difficult to understand the context of the concordance lines.

In addition to the themes available, the findings uncovered two issues of paramount importance. First of all, a very important finding was that there was an emerging theme from the interviews, which was collaboration. There were no items that dealt with this issue on the questionnaire. However, all interviewees made comments concerning the impact of collaboration on corpus use. Participant 6 commented:

P6: My partner came up with very interesting ideas. It was really fun to work with a partner. I think that if I had done it alone, maybe these works would not have been so permanent. Because it was like learning while having fun.

She went on to add:

P6: Because, for example, I used to make a lot of sentences out of one word and paste a sentence of everyone there, for example, I would look at a sentence or 2 sentences and I would get bored the most, but when everyone there, for example, finds a sentence and pastes it there. You know, other people see my sentence. For example, we can say that such a word also has this usage. Together it is easier to make sense of what we are working on.

Participant 23 contributed to the theme by adding that:

P23: When we had problems using the corpus, there was always someone in the group who knew how to solve it. When we needed to complete a task as a group work, one person definitely knew how to solve the problem and help us, so we completed the task together more easily.

Participant 3 added:

P3: As we were working as a group, or as a whole class activity on Google docs, it was very useful to work on the words no one knew of. It was like word pooling. Everybody contributed on it, and when I missed something in the corpus, the group helped me understand it.

This was a very important finding since all the participants highlighted the positive impact of collaborating with others, although the theme was not queried in the questionnaire or the interviews. However, it seems that there is a need to add “the impact of collaboration on corpus use”. They stated it as a major factor that made corpus activities more entertaining and informative. Therefore, it could be concluded that, collaboration, as one of the 21st-century skills, needs to be considered as a part of corpus-based learning approaches.

The second theme that emerged from the interviews was a very important factor that could affect the use of corpus in learning an L2, particularly with a focus on speaking skill. The finding highlighted a major limitation of COCA. The participants expressed a clear need to listen to the concordance lines, particularly if they were to engage in speaking activities. Participant 7 commented:

P7: Ok, COCA was useful to learn new vocabulary to use when speaking English, but I don't think corpus on its own, is very useful in improving spoken language because we couldn't hear the sentences in the corpus. I

mean, I can't say that it improved my speaking, since there was no robot or something to voice the sample sentences, but I believe it improved my vocabulary knowledge.

Participant 14 contributed on that stating:

P14: Yes, in terms of vocabulary learning and so on, corpus use was effective, but there was not even a button in the corpus where you could click on and listen to the sentences aloud. I wish COCA had had such a feature. I think if we want to listen to that sentence so that we can listen and repeat it, for example, we need to copy it, or we need to listen to that sentence from a translation or another place that will read it to us aloud.

Participant 3 commented:

P3: I mean if we could listen to the sentences we found on the corpus, it could be a very rich learning resource.

Overall, the findings indicate that the participants were quite positive about the use of corpus. A number of technical issues were reported; however, the technical issues were not regarded as major. The solution to these kinds of problems, such as difficulties in learning how to use COCA, was attributed to conducting corpus-based lessons with the guidance of a teacher or receiving assistance from the group members. In general, corpus use was found to be useful for learning words and phrases. One particular finding was that the corpus use integrated with a speaking portfolio helped learners improve their speaking skills. In addition, the content of the corpus was not found very challenging by the participants on contrary to the common belief. Finally, the participants reported they would recommend it to other L2 learners in Turkey, particularly for expanding vocabulary knowledge as an effective tool of language learning.

5. Conclusion

The current study attempted to evaluate Turkish L2 learners' perceptions regarding the integration of DDL activities into the curriculum within an emergency distance education context. The study supports the three main findings in the meta-analysis conducted by Boulton and Cobb (2017). First, the study confirmed that DDL use is not to be used only with advanced learners, yet lower-level learners may also benefit from it to improve oral performance. Second, it was established that corpus consultation may yield effective findings for English for general purposes context, contrary to the belief that DDL is more successful in the ESP context. Finally, although the current study did not aim to compare learner perceptions for hard and soft versions of corpora, it provides support to the finding that hands-on corpora use was reported to be more effective than paper-based DDL. The study established that the majority of the learners found corpus use effective in improving their spoken performance, particularly concerning phrases, appreciated the authenticity in the corpus, and found it more helpful than a dictionary in finding phrases. The learners seemed to find corpus use difficult more owing to the unfinished sentences, unlike in Geluso & Yamaguchi (2014). However, unfamiliar vocabulary did not seem to present a very big challenge for the participants. The learners seemed to have the most difficulty with technical issues in using the corpus, such as registering, and logging in. Therefore, in addition to easy access to user-friendly tools, providing learners with general computer literacy training might diminish the problems faced when using a corpus (Daskalovska, 2015). Furthermore, personality traits or learning preferences of the learners should be considered as a crucial factor, since if the majority of the learners feel intimidated by using technology in the classroom, the DDL approach may fail to meet the objectives of a course. Although Vannestal and Lindquist (2007) reported that lower-level learners tend to assess corpus use as boring and difficult, the participants of the study did not report the approach in that way despite not being advanced-level EFL learners, as in most corpus research. Regarding the speaking tasks assigned, the learners found it helpful to use the corpus in conjunction with oral production tasks, particularly if the focus is the speaking skill. Regarding general attitudes, the learners stated they would recommend it to other learners, and it helped raise their awareness of lexicogrammatical relations. Moreover, the study reveals two crucial findings. First, the qualitative data revealed that collaboration, which is among the 20th-century skills, plays a vital role in conducting DDL activities (Chambers, 2005; Mackey & Ziegler, 2017), particularly with lower-level learners. Following the Interaction hypothesis (Long, 1981), collaborative and communicative environments 1) to receive input, 2) to receive positive and negative evidence to notice the gaps in their interlanguage and to reinforce the correct forms used, and 3) to produce language. In that regard, the study appears to provide evidence that collaboration plays an important role in boosting performance in oral performance. Therefore, further research might add a collaboration domain to the questionnaires, if relevant, and investigate the effect

of collaborative tasks in corpus consultation. Secondly, the data uncovered a major limitation of publicly available corpora such as COCA. In particular, when the DDL intervention is targeted at vocabulary learning integrated with speaking skills, the learners highlighted the need for audio data, as well as written transcriptions. Consequently, the addition of the audio files to the spoken component of COCA could take implementations of direct corpus consultation one step further.

In conclusion, the present study revealed that the incorporation of corpus implementations into EFL classrooms could serve as a fruitful approach, particularly for online learning environments, which help remove logistic problems involved in the hard version of DDL. The findings of the study suggested that EFL learners agreed that direct corpus consultation used in conjunction with coursebook activities especially facilitated noticing and learning lexical and phraseological units in spoken language. Therefore, the study implies that corpus-based activities adapted from the EFL coursebooks could effectively complement the existing language learning curricula in the context of online education. Regarding spoken language, this study illustrates that using corpus tools in vocabulary learning could improve EFL learners' spoken language, especially when applied in collaborative settings, as well as written language. Considering that spoken competence entails the use of accurate and a diverse set of vocabulary items (Milton, 2013; Uchihara & Saito, 2018), DDL seems to serve as a useful tool in boosting learners' lexicon. To this end, the study has several implications for foreign language teaching and corpus design. The study provided evidence that the corpus-based activities integrated into the language learning curriculum in EFL contexts of higher education might help learners benefit from this approach in learning general vocabulary to improve the complexity level of their spoken performance. The study also contributed to the existent corpus-based research as corpus studies drawing on the spoken learner corpus seem to be rare (Yoon, 2020). In addition, the study provided evidence that large corpora such as COCA may need to be revised in terms of available resources for online language learning environments, especially for EFL speaking instruction.

Acknowledgments

This study is supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) within the scope of 2211-A National Scholarship Program for PhD Students. I would like to thank them for their support.

References

- Aguado-Jiménez, P., Pérez-Paredes, P., & Sánchez, P. (2012). Exploring the use of multidimensional analysis of learner language to promote register awareness. *System*, 40(1), 90–103. <https://doi.org/10.1016/j.system.2012.01.008>
- Aşık, A., Vural, A. Ş., & Akpınar, K. D. (2015). Lexical Awareness and Development through Data Driven Learning: Attitudes and Beliefs of EFL Learners. *Journal of Education and Training Studies*, 4(3), 87–96. <https://doi.org/10.11114/jets.v4i3.1223>
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis. Proceedings from the Fourth International Conference on Teaching and Language Corpora, Graz 19-24 July, 2000* (pp.165-182). Amsterdam: Rodopi.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. McH. Sinclair (Ed.), *How to use corpora in language teaching* (pp.15-36). 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>
- Chambers, A. (2005). Integrating Corpus Consultation in Language Studies. *Language Learning & Technology*, 9(2), 111–125.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 478–497). Cambridge University Press.
- Fligelstone, S. (1993). Some reflections on teaching from a corpus linguistics perspective. *ICAME Journal*, 17, 97-109.
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130–144. doi:10.1080/09588221.2013.8039827
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(2), 225–242. <https://doi.org/10.1017/S0958344014000044>

- Gilquin, G., & Granger, S. (2010). How can DDL be used in language teaching? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359-369). London, England: Routledge.
- Granger, S. (2021). Phraseology, corpora and L2 research. In *Perspectives on the L2 Phrasicon: The View from Learner Corpora* (pp. 3–22). Channel View Publications. <https://doi.org/10.21832/9781788924863-002>
- Johns, T. (1991). Should you be persuaded: Two examples of data driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing*. ELR Journal 4 (pp. 1-16). Birmingham: University of Birmingham
- Kennedy, C., & Miceli, T. (2001). an Evaluation of Intermediate Students’ Approaches To Corpus Investigation. *Language Learning & Technology*, 5(September), 77–90. <http://lt.msu.edu/vol5num3/kennedymiceli/>
- Kennedy, C., & Miceli, T. (2010). *CORPUS-ASSISTED CREATIVE WRITING: INTRODUCING INTERMEDIATE ITALIAN LEARNERS TO A CORPUS AS A REFERENCE RESOURCE* (Vol. 14, Issue 1). <http://lt.msu.edu/vol14num1/kennedymiceli.pdf>
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McNery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). New York: Addison Wesley Longman.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills, in Bardel, C., Lindqvist, C. & Laufer, B. (eds.), *L2 Vocabulary Acquisition, Knowledge and Use: New Perspectives on Assessment and Corpus Analysis*. Eurosla Monograph Series 2. Amsterdam: Eurosla. 57–78.
- Mizumoto, A., & Chujo, K. (n.d.). A Meta-analysis of Data-driven Learning Approach in the Japanese EFL Classroom URL 「論文」 A Meta-analysis of Data-driven Learning Approach in the Japanese EFL Classroom 1. <http://hdl.handle.net/10112/13009>
- Mukherjee, J. (2002): *Korpuslinguistik und Englischunterricht: Eine Einführung*. Frankfurt am Main: Peter Lang.
- Pérez-Paredes, P., & Cantos-Gomez, P. (2004). Some Lessons Students Learn: Self-discovery and Corpora (pp. 247–257). <https://doi.org/10.1075/scl.17.18per>
- Poole, R. (2012). Concordance-based glosses for academic vocabulary acquisition. *CALICO Journal*, 29(4), 679–693.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205–225. <https://doi.org/10.1017/S0267190511000055>
- Sripicharn, P. (2010). How can we prepare learners for using language corpora? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 371- 384). Abingdon/New York: Routledge
- Uchihara, T., & Saito, K. (2019). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal*, 47, 64 – 75.
- Vannestal, M.E. and Lindquist, H., (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 3 (329-350). doi:10.1017/S0958344007000638
- Wang, X., & Hao, T. (2017). Designing interactive exercises for corpus-based English learning with Hot Potatoes software. In *Emerging Technologies for Education: Second International Symposium, SETE 2017, Held in Conjunction with ICWL 2017, Cape Town, South Africa, September 20–22, 2017, Revised Selected Papers 2* (pp. 485-494). Springer International Publishing.
- Yoon, S. (2020). The Learner Corpora of Spoken English: What Has Been Done and What Should Be Done? *Lanaguage Research*, 56(1), 29–51. <https://doi.org/10.30961/lr.2020.56.1.29>
- Ziegler, N., & Mackey, A. (2017). Interactional feedback in synchronous computer-mediated communication: A review of the state of the art. *Corrective feedback in second language teaching and learning*, 80-94.

Etik, Beyan ve Açıklamalar

1. Bu çalışmanın yazar/yazarları, Hacettepe Üniversitesi Etik Kurulu’nun 30.04.2021 tarih ve E-51944218-101.02.02-00001561321 sayılı yazısı ile etik kurul izin belgesi almış olduklarını beyan etmektedir.
 2. Bu çalışmanın yazar/yazarları, araştırma ve yayın etiği ilkelerine uyduklarını kabul etmektedir.
 3. Bu çalışmanın yazar/yazarları kullanmış oldukları resim, şekil, fotoğraf ve benzeri belgelerin kullanımında tüm sorumlulukları kabul etmektedir.
 4. Bu çalışmanın benzerlik raporu bulunmaktadır.
-

