

Language Teaching and Educational Research

e-ISSN 2636-8102

Volume 6, Issue 2 | 2023

EFL Instructors' Corpus Literacy and Their Perceptions of Using Corpora to Teach L2 Vocabulary

Fulya Eslek-Onur
Sibel Tosun

To cite this article:

Eslek-Onur, F., & Tosun, S. (2023). EFL instructors' corpus literacy and their perceptions of using corpora to teach L2 vocabulary. *Language Teaching and Educational Research (LATER)*, 6(2), 110-123. <https://doi.org/10.35207/later.1284574>

[View the journal website](#)



[Submit your article to LATER](#)



[Contact editor](#)



Copyright (c) 2023 *LATER* and the author(s). This is an open access article under CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Research Article

EFL instructors' corpus literacy and their perceptions of using corpora to teach L2 vocabulary

Fulya Eslek-Onur*  ORCID

EFL Instructor, Department of Foreign Languages, Munzur University, TURKIYE

Sibel Tosun  ORCID

EFL Instructor, School of Foreign Languages, Firat University, TURKIYE

Abstract

Corpus literacy is "the ability to use the technology of corpus linguistics to investigate language and enhance the language development of students" (Heather & Helt, 2012, p.417). The recent literature focused primarily on the perceptions of students and teachers about corpus use in learning and teaching vocabulary or investigated corpus literacy alone, but EFL teachers' corpus literacy is not explored. This study employed an explanatory sequential mixed-method design to investigate the corpus literacy of EFL instructors and their perceptions of using corpora to teach second-language vocabulary. A total of 41 EFL instructors working in 17 different state universities in Turkiye participated in the study. Data were gathered through a questionnaire and semi-structured interviews. The results showed that the EFL instructors had low to medium levels of corpus literacy. In addition, there was no correlation between corpus literacy and the teaching experience of the participants. The results also showed that despite their familiarity with corpora, the instructors had never used one to teach vocabulary. Furthermore, the majority of the participants learned about corpora and corpus tools during their Ph.D. and MA studies in English language teaching and linguistics, mostly through coursework and publications.

Received

17 April 2023

Accepted

05 July 2023

Keywords

Corpus literacy
Corpora
Vocabulary teaching

Suggested APA citation: Eslek-Onur, F., & Tosun, S. (2023). EFL instructors' corpus literacy and their perceptions of using corpora to teach L2 vocabulary. *Language Teaching and Educational Research (LATER)*, 6(2), 110-123. <https://doi.org/10.35207/later.1284574>

Note(s)

*Corresponding author

**The ethics committee approval was obtained from Munzur University with a document number of 75395 on 28/11/2022.

Author(s)' statements on ethics and conflict of interest

Ethics statement: We hereby declare that research/publication ethics and citing principles have been considered in all the stages of the study. We take full responsibility for the content of the paper in case of dispute.

Statement of interest: We have no conflict of interest to declare.

Author contribution disclosure: Both authors contributed equally to this study.

Funding: None

Acknowledgements: We would like to thank 41 instructors who accepted to participate in the study. We are also grateful to the anonymous reviewers for their comments and suggestions.

Introduction

Accessing information in the present era is relatively simple, and the centrality of technology in our lives makes it a vital component of education. The rapid changes in technologies have led to innovations not only in our daily lives but also in schools and classrooms, allowing the use of new tools, particularly in language education. A corpus (plural corpora) is one of these tools and is defined as a compilation of written and spoken language in computerized databases. According to Conrad (2005), a corpus is an extensive, systematic collection of naturally occurring texts stored and presented in electronic form. Similarly, Vyatkina and Boulton (2017) define corpora as "systematically organized electronic collections of texts" (p. 1). As the researchers stated, corpora represent a natural and authentic source of language, and the texts are from real life, namely from academic journals, newspapers, magazines, TV shows, movies, and so on. However, corpora are not merely regarded as collections of texts but as "genuine theoretical resources used in a number of applied research areas" (Oğat, 2016, p. 41). Various types of corpora can be utilized for different areas of study, including general and specialized corpora, written and spoken corpora, and native and learner corpora.

Linguistics, as the scientific study of language, intersects with corpus linguistics in every subfield. Corpus linguistics focuses on how languages are used in practice and examines language as it is used in real life. Therefore, it is a method of language analysis in which a collection of texts comprising authentic language is stored in a corpus. According to Conrad (2000), corpus linguistics is "the empirical study of language relying on computer-assisted techniques to analyze large, principled databases of naturally occurring language" (p. 548). Corpora enable users to analyze word frequency and examine examples of word usage within specific contexts. On the other hand, corpus literacy refers to "the ability to use the technology of corpus linguistics to investigate language and enhance the language development of students" (Heather and Helt, 2012, p. 417). In his research, Callies (2019) outlined several aspects of corpus literacy, including:

- a. Understanding basic concepts in corpus linguistics: What is a corpus, and what types of corpora are available and how? What can you do – and cannot do – with a corpus?
- b. Searching corpora and analyzing corpus data using corpus software tools, e.g., concordancers: What is corpus software, and how can it be used to search a corpus? How can corpus output be analyzed?
- c. Interpreting corpus data: How can general trends in language use or change be inferred from corpus data?
- d. Using corpus output to generate teaching materials and activities: How can corpus materials be utilized for teaching purposes? (p. 247)

The use of corpora in language education has focused on various domains, including vocabulary, grammar, reading, and writing. Being used for multiple purposes and being a rich source of real-life data, corpora have greatly increased in popularity more recently (see Abdel Latif, 2020; Aşık, 2017; Çalışkan & Kuru Gönen, 2018; Frankenberg-Garcia, 2012; Hirata & Hirata, 2019; Lee, 2011; Poole, 2020; Şimşek, 2020; Xodabande & Nazari, 2022) in second language pedagogy, especially in vocabulary teaching and learning. A myriad of corpus-based and corpus-related studies (see Barabadi & Khajavi, 2017; Belkhir, 2013; Çilak, 2017; Kazaz, 2015; Koçak, 2020; Tosun, 2017) have been undertaken recently to teach vocabulary, and their

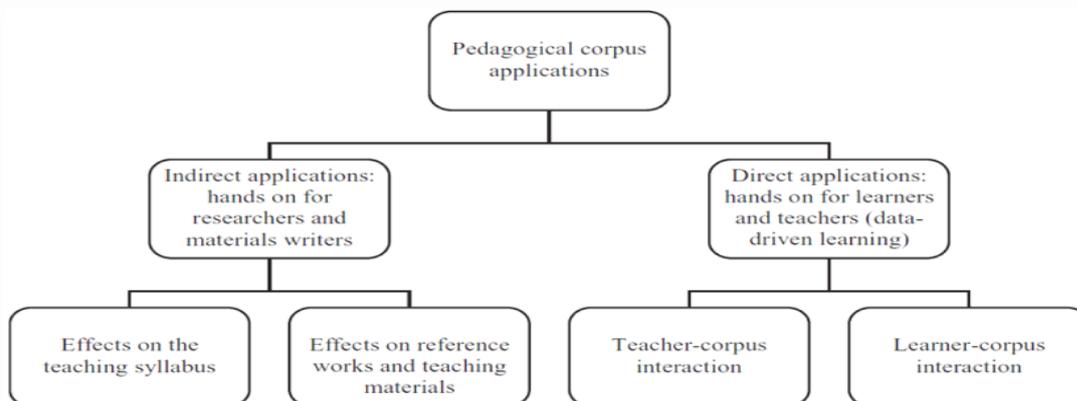
findings indicate that using corpora to expand students' vocabulary is more effective than traditional methods.

Vocabulary is an essential component of language learning and teaching. Second language learners need to acquire a sufficient vocabulary in order to use the target language effectively. Lewis (1993) emphasized the importance of vocabulary in a language, claiming that "lexis is the core or heart of language" (p. 89). According to Schmidt's noticing hypothesis (1990, 1993), learners learn new vocabulary or grammatical features when they are aware of them, and they are unlikely to learn grammatical forms unless they consciously pay attention to them. Retention improves when learners are given more opportunities to hear and use the target language and when they focus their attention on the vocabulary being learned (Schmitt, 2008). Therefore, EFL instructors and teachers should place special emphasis on vocabulary teaching and utilize novel techniques and methods to teach the target vocabulary. It should be kept in mind that 21st-century EFL learners are technologically competent digital natives (See Akayoğlu et al., 2020; Farhadi & Öztürk, 2023; Solak & Recep, 2014). Hence, as EFL teachers and instructors, we can integrate technology into our teaching to make use of corpora and concordancers to teach vocabulary. At this juncture, having corpus literacy would help EFL instructors and teachers develop corpus-based syllabi and materials for vocabulary instruction. Therefore, the present study is significant as it explores the extent of English instructors' knowledge, and aims to raise awareness and knowledge of corpora among EFL instructors. It offers significant implications for second language teacher education (SLTE), including pre- and in-service training, and fills the gap in the limited number of studies on this issue.

Review of the Literature

Over a few decades, researchers in the field of language education have focused on corpus-related studies, and both indirect and direct uses of corpora are common pedagogical applications in teaching and learning English as a foreign language. In indirect applications, researchers and teachers use corpora to create curricula, syllabi, and materials, which can lead to the use of actual language samples in textbooks instead of invented examples. Direct corpus applications for language teaching and learning, on the other hand, usually involve students accessing a corpus directly (Römer, 2011). Figure 1 summarizes the direct and indirect applications of corpora. As depicted in Figure 1, different kinds of corpus methods and tools can have varied effects on different people and things (Römer, 2011).

Figure 1. The use of corpora in second language learning and teaching by Römer (2011)



There is an increasing number of corpus-related studies in the existing literature. Most of these studies were conducted with EFL learners or pre-service EFL teachers. For instance, Belkhir (2013) conducted a study to illustrate how computer corpus data mitigates the challenges associated with EFL vocabulary teaching and learning. Belkhir (2013) carried out this study to determine the extent to which EFL teachers are familiar with the idea of computer corpus data and to explore EFL teachers' views on using computer corpus analysis as a language source for EFL vocabulary selection and training. Utilizing a mixed-method research design, the participants of the study were 10 EFL teachers working at the Department of Foreign Languages, University of Tlemcen in Algeria. Data were obtained from the participants through a semi-structured interview and a five-point Likert scale. It was reported that all participating teachers used the Oxford Advanced Learner's Dictionary and other vocabulary teaching materials, such as English Vocabulary in Use, to improve their students' word knowledge and vocabulary skills. The results showed that the majority of participants expressed dissatisfaction with the teaching sources and materials they were using at the time. Furthermore, the results of the semi-structured interview revealed that almost all the participants were unfamiliar with computer corpus data, but they had positive attitudes towards using computer corpus data as a source for EFL vocabulary selection in particular and English teaching in general. In conclusion, Belkhir (2013) emphasized the importance of vocabulary in strengthening students' four language skills. She suggested that curriculum designers and language teachers use computer corpus data to teach vocabulary. She also advised EFL teachers to update their resources and materials to teach more effectively.

As for the utility of data-driven methods to teach vocabulary, Barabadi and Khajavi (2017) conducted a study to compare the data-driven learning approach and traditional methods of teaching vocabulary. The study involved 62 Iranian students at two English institutes in Iran. The researchers formed three intact groups, two of which were experimental groups. Both the experimental and control groups were administered a vocabulary test that included the essential words encountered in their textbooks after the 7-week treatment. The results showed no significant difference between the experimental and control groups in this vocabulary test. The results of the vocabulary size test indicated that the vocabulary size of the two groups was the same in terms of reading comprehension. Furthermore, the results showed that the corpus-based, data-driven approach to teaching and learning vocabulary was more effective than traditional methods. Learners in the experimental group performed better than learners in the control group. Similarly, Çilak (2017) studied effects of corpus-based materials on the vocabulary knowledge of EFL learners. The purpose of the study was to determine whether corpus-based vocabulary activities have an impact on the learning and retention of target vocabulary. For this experimental study, which involved 41 EFL learners, quantitative data were collected through pre-tests and post-tests, while qualitative data were obtained through interviews with 10 learners in the experimental group. Two pre-tests were designed: one based on corpus-based material and the other on coursebook material. Each pre-test contained 50 new vocabulary items that would be taught during the study. The items were selected from corpus-based and coursebook resources and included nouns, verbs, adjectives, adverbs, and phrasal verbs. The results of this study demonstrated statistically significant differences in the impact of corpus-based materials and the traditional teaching approach on learners' average vocabulary scores. Furthermore, the paired sample t-test revealed no statistically significant difference between the coursebook-based test and the corpus-based test.

Analysis of the interviews revealed participants' positive attitude toward corpus-based vocabulary. The study also indicated the positive effects of using corpus-based vocabulary teaching resources on learners.

The studies on perceptions were centered around both teachers and students. For example, Çalışkan and Kuru Gönen (2018) conducted a qualitative study to investigate university instructors' perspectives on the use of concordance lines in vocabulary teaching and their attitudes toward corpus-based materials after receiving training. Three EFL instructors participated in a four-week training program that focused on teaching about corpora, using corpora to teach language, and incorporating corpus-based language pedagogy principles into the classroom. Data were collected using various instruments, including semi-structured interviews, reflective logs, and an open-ended questionnaire. The results of the study revealed the limited effectiveness of corpus-based materials for vocabulary instruction. Participants identified the challenges of developing corpus-based materials and highlighted technology issues as difficulties they faced.

As well as exploring EFL instructors' perceptions on using corpus-based materials in vocabulary instruction, EFL learners' attitudes to corpus have been studied. Sinha (2021), for instance, investigated EFL students' perceptions and attitudes toward corpus as a vocabulary learning tool. Data were gathered from 32 first-year undergraduate students enrolled in an introductory English language course at a private university in Bangladesh. It was reported that most of the participants found the corpus a useful tool for learning new words, but they also complained that the nature of the corpus data made learning challenging for them.

To sum up, the studies in the existing literature mostly focused on the perceptions of students and teachers about corpus use in learning and teaching vocabulary or investigated corpus literacy alone. For this reason, exploring EFL teachers' corpus literacy with a particular emphasis on vocabulary is an under-researched area. In addition, the exploration of the association between corpus literacy and teaching experience has received scant attention in the research literature. This study, therefore, sets out to examine EFL teachers' corpus literacy, their perceptions of teaching vocabulary, and its relationship with teaching experience. The present study aims to explore EFL instructors' corpus literacy levels and how EFL teachers perceive the use of corpora for teaching second language vocabulary. With this in mind, this study was guided by the following three research questions:

1. What is the corpus literacy level of EFL instructors?
2. Is there an association between the corpus literacy levels of EFL instructors and their teaching experience?
3. What are EFL instructors' overall perceptions towards the use of corpora to teach vocabulary?

Methodology

Research design

The current study employed an explanatory sequential mixed-method design to explore EFL instructors' perceptions of using corpora to teach second language vocabulary. Creswell (2014) describes mixed method research design as "a combination of qualitative and quantitative approaches which provides a more complete understanding of a research problem than either approach alone" (p. 4). Teddlie and Tashakkori (2009) also asserted that mixed-methodologists primarily follow the pragmatic paradigm and are interested in both qualitative

and quantitative data and their analysis. Exploring the perceptions of EFL instructors regarding the use of corpora through the combination of these methodologies would therefore provide a comprehensive understanding of the phenomenon. In this research design, a Quan > Qual sequence was followed (Creswell, 2014). Firstly, quantitative data were collected and analyzed, and then the researchers started to collect qualitative data. Three research questions were formed, and a purposive sampling method was utilized to select the participants. Table 1 provides the overall plan of the study.

Table 1. Research design

Research Questions	Sampling Strategy	Participants	Data Collection Tools	Data Analysis
1. What is the corpus literacy level of EFL instructors?	Purposive Sampling	41 EFL Instructors	Questionnaire	Descriptive Statistics
2. Is there an association between the corpus literacy levels of EFL instructors and their teaching experience?	Purposive Sampling	41 EFL Instructors	Questionnaire	Chi-Square Analysis
3. What are EFL instructors' overall perceptions towards the use of corpora to teach vocabulary?	Purposive Sampling	6 EFL Instructors	Semi-structured Interviews	Content Analysis

Participants

The participants in this study were 41 EFL instructors from 17 different state universities in Türkiye. The participants in the study were selected using a procedure called purposive sampling. The inclusion criterion was as follow: work as an EFL instructor at a university. 29 of the 41 participants were female and 12 were male. The participants were from 17 different state universities. [Adana Alparslan Türkeş Science and Technology University (N=2), Adıyaman University (N=2), Anadolu University (N=1), Atatürk University (N=1), Bursa Technical University (N=1), Bingöl University (N=1), Fırat University (N=18), İzmir Demokrasi University (N=2), Kırşehir Ahi Evran University (N=2), Middle East Technical University (N=1), Munzur University (N=4), Muğla Sıtkı Koçman University (N=1), Niğde Ömer Halis Demir University (N=1), Osmaniye Korkut Ata University (N=1), Samsun University (N=1), Turkish Aeronautical Association University (N=1), and Yalova University (N=1)] in Türkiye. While 40 of them were working at the School of Foreign Languages, only one of the participants was working at the Applied English and Translation Programme. The demographic information of the participants is presented in Table 2.

Table 2. Demographic information of EFL instructors

Educational Status	Number	Teaching Experience	Total
Bachelor degree (BA)	10		
Master degree (MA)	10		
Phd Students/Candidates	18	1-20+ years	41
PhD degree (PhD)	3		

Data collection tools

In this study, quantitative data were collected using a questionnaire to ask for descriptive and background information about the participants and their prior knowledge of corpora. The questionnaire was adopted from Çalışkan (2020), who also adapted Bunting's (2013) questionnaire. This questionnaire consisted of three sections: a. Demographic Information, b. Your Knowledge about Corpus Tools and c. Experience in the Use of Corpora. Given the scope of the current study, only one question (Would you like to attend extensive workshop sessions on the use of corpus tools in EFL classrooms?) was deleted. A colleague with a Ph.D. in the field of ELT was consulted to determine the validity of the questionnaire, and she approved it and did not make any suggestions. In addition, a semi-structured interview was used for the qualitative data. The questions of the semi-structured interview were adapted from Çalışkan (2020) and revised, and two questions were modified within the scope of the current study.

Data collection procedure

The data were collected using a questionnaire and semi-structured interviews. The questionnaire was designed using a free online tool, and the link was shared with the participants. Since the participants were from different state universities in different cities, collecting the data through this online tool was time-saving and convenient. The data were gathered from a total of 41 EFL instructors. Similarly, the semi-structured interviews were conducted via Zoom meetings, and six participants volunteered to attend. Each session was recorded with the participants' knowledge and consent.

Data analysis

The data analysis stage consisted of two phases due to the mixed qualitative and quantitative data collection methods utilized in the study. The participants' corpus literacy levels were explored through descriptive and frequency analysis using SPSS. The mean scores and standard deviation were presented in tables. The quantitative analysis also included examining the association between corpus literacy levels of EFL instructors and their teaching experience using Chi-Square analysis. Additionally, participants' perceptions were analyzed qualitatively through content analysis. The semi-structured interviews were transcribed verbatim. The researcher initially reviewed the data to gain a general understanding. Each transcription was then emailed to the participants for verification. The transcriptions were analyzed through iterative readings, with the researchers manually coding the raw data by highlighting relevant phrases or sentences. In vivo codes derived from participants' own speech were used.

Results

Corpus literacy levels of EFL instructors and its association with teaching experience

In an attempt to answer the first research question, which explores the corpus literacy level among EFL instructors, the quantitative data, gathered through the questionnaire, was analyzed on SPSS. The results of the descriptives analysis are presented in Table 3.

Table 3. Descriptives on the knowledge of corpora

Items	<i>N</i>	<i>M</i>	<i>STD</i>
1. Corpus software programs (e.g., Antconc, WordSmith Tools)	41	1.92	0.93
2. Online corpora (e.g., COCA, Sketch Engine, Lextutor)	41	2.46	0.95

As seen in Table 3, the participants have low awareness of corpus software programs such as Antconc, WordSmith, etc. ($M = 1.92$, $SD = 0.93$). On the other hand, it has been found that EFL instructors have a moderate-high familiarity with online corpora such as COCA, Sketch engine, Lextutor, etc. ($M = 2.46$, $SD = 0.95$). The participants were also surveyed on their knowledge of using corpus tools in various areas. Similarly, frequency analysis was run for mean scores. The findings are displayed in Table 4 below.

Table 4. Descriptives on the knowledge of the use of corpus tools

Items	<i>N</i>	<i>M</i>	<i>STD</i>
1. Using corpus tools directly with students (e.g., analyzing concordance lines)	41	2.12	1.02
2. Using corpus tools to create corpus-informed classroom materials	41	2.19	0.98
3. Using corpus tools to create vocabulary activities	41	2.12	0.92
4. Using corpus tools to check students' vocabulary knowledge	41	2.19	0.98
5. Using corpus tools to build my own language knowledge	41	2.39	1.09

As seen in Table 4; overall, the participants demonstrate low-moderate knowledge using corpus tools. The highest mean score belongs to item 5 ($M = 2.39$, $SD = 1.09$), indicating that EFL instructors use corpus tools to build their language knowledge the most among other purposes. Personal academic development purpose is followed by the creation of corpus-informed materials ($M = 2.19$, $SD = 0.98$) and checking students' vocabulary knowledge ($M = 2.19$, $SD = 0.98$) with the same mean score. The least resorted purpose of corpus use belongs to the direct use of corpus with students ($M = 2.12$, $SD = 1.02$) and the creation of vocabulary activities ($M = 2.12$, $SD = 0.92$). In the light of these findings, it can be concluded that EFL instructors tend to use corpus tools for self-improvement, but they seem to have little information on how to use corpus tools directly with students or how to create corpus-informed materials.

The EFL instructors' corpus literacy was also explored through open-ended questions, inquiring about the source of their corpus awareness. First, regarding the question about the educational source of corpus and corpus use, 17 participants indicated that they were informed on the use of corpora as part of BA, MA, or Ph.D. education. 11 out of 17 participants reported that they have learnt about the corpus and some corpus tools during their Ph.D. studies. While six participants mentioned MA courses as the source of their corpus knowledge, only two mentioned they were informed about corpus during their BA.

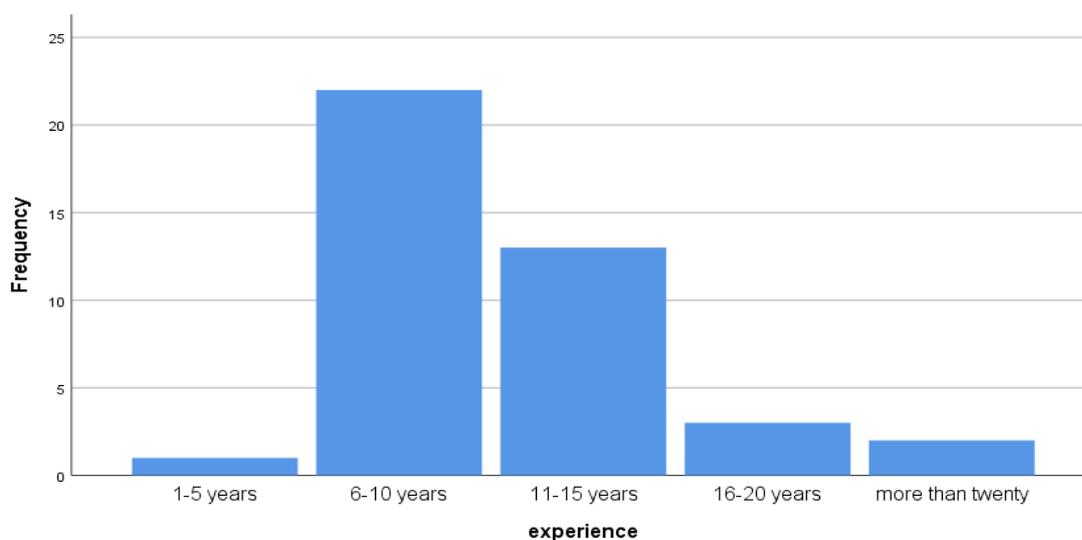
The participants were also asked whether they had received training in the use of corpora as part of a continuing education program (e.g., before or during training). It was found that only one female participant had received training on corpus and corpus use as part of her ICALT training. This finding is noteworthy in that the participants either had not received training or the use of corpora and corpus was not part of their training programs.

With respect to the question investigating other academic sources of corpus knowledge such as conferences, journals, and so on, six participants indicated that they read about corpus in journals and articles. In contrast, two instructors stated that they had been informed about corpus through conferences and one through online workshops. One participant also mentioned that his MA thesis was on the corpus, so he had studied it by himself. Another participant stated that she was informed about corpus when she attended her colleague's corpus training.

Lastly, the EFL instructors were asked which corpus tools they had used before. Only one-third of the participants (N = 14) reported using a corpus. COCA was the most frequently used corpus tool, with a frequency of 11, followed by BNC and Sketch Engine. Only one participant reported using the Antconc concordance as part of her MA thesis. As for the purpose, most participants who used corpus tools stated that they used them to teach vocabulary and collocations. Some also indicated that they used the corpus for writing; a minority of the participants mentioned that they used the corpus for grammar and linguistics. Only one participant reported the use of corpus to prepare course materials and exam questions.

To answer research question 2, which examines the association between corpus knowledge and teaching experience, the Chi-square test was calculated. First, descriptive information about the teaching experience of the EFL teachers was obtained, and the results are presented in Figure 2.

Figure 2. Teaching experience of the participants



As illustrated in Figure 2, the teaching experience of EFL instructors varies between 1-5 years and more than twenty years. However, the teaching experience of the majority of the participants ranges between 6 and 15 years. Instructors with 6-10 years of experience hold the highest percentage (53.7%, N = 22), followed by instructors with 11-15 years of experience (31.7%, N = 13) and 16-20 years of experience (7.3%, N = 3). Only two participants stated that they had more than twenty years of experience (4.9%), and one reported 1-5 years of experience (2.4%). Besides the descriptive analysis of teaching experience, its association with

corpus literacy was further explored with Chi-Square analysis, whose results are presented in Table 5.

Table 5. The Chi-Square test for the association between corpus literacy and teaching experience

<i>Pearson chi-Square Items</i>	Value	df	Asymp. Sig. (2-sided)	Phi
1. Knowledge of corpus software programs	5.99	12	.91	.38
2. Knowledge of online corpora	9.56	12	.65	.48
3. Using corpus tools directly with students (e.g., analysing concordance lines)	7.75	12	.84	.43
4. Using corpus tools to create corpus-informed classroom materials	9.35	12	.67	.47
5. Using corpus tools to create vocabulary activities	9.63	12	.64	.48
6. Using corpus tools to check students' vocabulary knowledge	7.93	12	.79	.44
7. Using corpus tools to build my own language knowledge	10.0	12	.61	.49

The Chi-square test for independence indicated no significant association between the knowledge of corpus and corpus tools and teaching experience. The non-significant difference was found in the knowledge of software programs $\chi^2(12) = 5.99$, $p = .96$, $phi = .38$; knowledge of online corpora, $\chi^2(12) = 9.56$, $p = .65$, $phi = -.48$; using corpus tools directly with students $\chi^2(12) = 7.75$, $p = .84$, $phi = .43$, using corpus tools to create corpus informed classroom materials $\chi^2(12) = 9.35$, $p = .67$, $phi = .47$, using a corpus to create vocabulary activities $\chi^2(12) = 9.63$, $p = .64$, $phi = .48$, using corpus tools to check students' vocabulary knowledge $\chi^2(12) = 7.93$, $p = .79$, $phi = .44$, and lastly, using corpus tools to build their own knowledge $\chi^2(12) = 10$, $p = .61$, $phi = .49$. These findings indicate that EFL instructors' knowledge on corpus and corpus tools is not dependent on their teaching experience.

The findings of the qualitative data obtained through semi-structured interviews with six EFL instructors revealed that three participants had knowledge of corpus but never used them to teach vocabulary. On the other hand, two participants stated that they had to use corpora such as COCA and BNC in their Ph.D. assignments but never used them in their classrooms. They shared their ideas in the following words:

I know COCA. I took a course called Web-based language teaching, and we covered it, but not in detail. We can teach vocabulary via corpora, but I think native corpora are not appropriate for our students' language level. A classroom corpus might be used for them (Interviewee III).

They also reported that textbooks, prescribed word lists, and collocation dictionaries could be more effective and user-friendly in teaching vocabulary. On the issue, Interviewee I and Interviewee II expressed their ideas as follows:

I do not find corpora user-friendly. Their interfaces are complex and should be improved. It is also not suitable for the level of my students. Textbooks have become standardized, and some publishers are very successful in this business. Corpora might be used while teaching English to immigrant students or while teaching English for

academic purposes. I think collocation dictionaries or some other websites such as Ludwig are more useful and practical (Interviewee I).

I heard the term corpus but I have never used it in my classes. I think they are not user-friendly and they are for researchers and professionals. Their interfaces are not attractive and for me they are time consuming. Using online dictionaries, for example collocation dictionaries are more practical (Interviewee II).

On the other hand, Interviewee VI emphasized the importance of corpora in language learning, and she expressed that corpora provide authentic materials and they are beneficial for students to see the different structures and usage of a word, to learn collocations of words. She highlighted the importance of corpora as below:

I first heard the term "corpus" when I was a student. At that time, my lecturers were dealing with the subject of corpus. Then I did research and learned a lot. I love to use corpora. I developed a corpus for my master's thesis. I used both written and spoken corpora to teach vocabulary and speaking. I also use them for my studies and they are part of my life (Interviewee VI).

Discussion

This study was designed to investigate the corpus literacy of EFL instructors, their perceptions of the use of corpora, and the relationship between corpus literacy and teaching experience. A main finding of the study is that the level of corpus literacy among the participating EFL instructors was low to moderate. Most of them had little to some knowledge of corpora and corpus tools. The participants had some knowledge of online corpora but were less knowledgeable about corpus software and programs. These findings are in line with those of Aşık (2015) and Callies (2019), who reported that only a limited number of participants were familiar with corpora and frequently used them in their classes. The findings are also consistent with those of Belkhir (2013), who sought to determine the extent to which EFL teachers were familiar with the idea of computer-corpus data and to explore the views of EFL teachers on the use of computer-corpus analysis as a language source for EFL vocabulary selection and training. The results of that study showed that almost all of the participants were unfamiliar with computer-corpus data but had positive attitudes towards using computer-corpus data as a source for EFL vocabulary selection in general and for English teaching in particular. Even so, it is worth noting that in the current study, the instructors who were familiar with corpora and corpus tools tended to use them for self-improvement rather than for instructional purposes.

There are many studies in the literature which focused on the use of hands-on corpora with students and the development of instructional materials, and these studies have shown many of the benefits of using corpora in language pedagogy. However, as the results of the present study suggest, it seems that the corpus has not yet found a place for itself in language education.

Several factors may contribute to the limited incorporation of corpora, such as tight teaching schedules or the lack of readily available teaching materials. Considering these factors, the findings have implications for program and curriculum developers, as well as materials developers. Including corpus-informed materials and tools in the curriculum could provide

instructors with greater benefits in terms of planning and implementing corpus use. This approach can help instructors gain more knowledge about corpora and reduce time spent on course syllabi and pre-made materials.

Another obvious finding of the study is that there was no association between the level of corpus literacy and teaching experience, suggesting that familiarity with the corpus and corpus tools does not depend on teaching experience. The responses to the open-ended questions about the sources of corpus knowledge gave the rationale for the result and a more likely determinant. The responses of the instructors showed that the majority of them had learned about corpora and corpus tools during their Ph.D. and MA studies in ELT and linguistics, mainly through courses and articles. Experienced instructors without MA and Ph.D. degrees and with graduate degrees in various departments reported little or no familiarity with the corpus. This result suggests that corpus literacy may be more attributed to academic development than teaching experience. There are no studies which have compared the academic development of EFL instructors with their usage of corpora. The current study is, therefore, crucial in this regard.

The findings discussed above provide useful implications for BA, pre-service and in-service training programs as well. The fact that only EFL instructors who had studied for an MA or a Ph.D. in ELT and linguistics majors were familiar with the corpus and its use points out the necessity of including corpus literacy courses as part of BA curriculums. For other disciplines than ELT and linguistics, training could be provided through pre-service and in-service programs. Römer (2011), however, commented that developments in corpus research had had little impact on the practice of English-language instruction since comparably few teachers and students are aware of valuable resources and use corpora or corpus software.

Conclusion and Implications

This study has provided a deeper insight into the corpus literacy of EFL instructors and has contributed to our understanding of instructors' knowledge of corpus and corpus tools and their relationship to experience. In the light of the findings, it can be implied that the potential of the corpus is not well received by EFL instructors for a variety of reasons. Future studies could focus on improving the corpus literacy of EFL instructors, as well as training them in the use of corpus tools for pedagogical purposes. In addition, the results suggest that there may be a stronger association between academic development and corpus literacy. Further studies regarding the role of academic development in corpus literacy would be worthwhile. On the other hand, second language teacher education programs could revise their curriculum, and program developers could offer prospective teachers a course called "Corpus Literacy". Such a course could develop the TPACK (Technological Pedagogical Content Knowledge) of prospective teachers.

While the study provides useful insights, the present study is not without limitations. The data were collected from the participants at 17 different universities. However, the generalizability of the results is subject to certain limitations, such as the number of participants. In addition, due to time constraints and the reluctance of participants, semi-structured interviews could only be done with six participants. The interviews could be conducted with more participants. All in all, it is recommended that future research be conducted with a larger sample and a more diverse population.

References

- Abdel Latif, M.M.M. (2020). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices. *ReCALL*, 33(1), 34-48. <https://doi.org/10.1017/S0958344020000129>
- Akayoglu, S., Satar, H. M., Dikilitas, K., Cirit, N. C., & Korkmazgil, S. (2020). Digital literacy practices of Turkish pre-service EFL teachers. *Australasian Journal of Educational Technology*, 36(1), 85-97. <https://doi.org/10.14742/ajet.4711>
- Aşık, A. (2015). The use of corpora in EFL classrooms: What do Turkish EFL teachers know and offer to promote pragmatic competence? *The Asian EFL Journal*, 17(4), 194-221.
- Aşık, A. (2017). A sample corpus integration in language teacher education through coursebook evaluation. *Journal of Language and Linguistic Studies*, 13(2), 728-740.
- Barabadi, E. & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary. *Cogent Education*, 4(1), 1-13. doi.org/10.1080/2331186X.2017.1283876
- Belkhir, F. Z. (2013). A survey on teachers' awareness and attitudes on computer-corpus data: An assisted technology-based EFL vocabulary selection and instruction source. *Procedia- Social and Behavioral Sciences*, 103, 77-85.
- Callies, M. (2019). Integrating corpus literacy into language teacher education. In Götz, Z. & Mukherjee, J. (Eds.), *Learner Corpora and Language Teaching*, (pp. 245 – 264). John Benjamins Publishing Company.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34(3), 548 - 560.
- Conrad, S. (2005). Corpus linguistics and L2 teaching. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 393-409). Mahwah, NJ: Lawrence Erlbaum.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.)*. Thousand Oaks, CA: Sage.
- Çalışkan, G., & Kuru Gönen, S. İ. (2018). Training teachers on corpus-based language pedagogy: Perceptions on vocabulary instruction. *Journal of Language and Linguistic Studies*, 14(4), 190-210.
- Çalışkan, G. (2020). *An investigation into EFL university instructors' perspectives on the use of corpora in language teaching*. [Master's thesis, Anadolu University]. Ulusal Tez Merkezi. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Çilak, E. (2017). *The effects of using corpus-based materials in vocabulary instruction*. [Master's thesis, Ufuk University]. Ulusal Tez Merkezi. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Farhadi, S. & Öztürk, G. (2023). Technological pedagogical content knowledge (TPACK) level and needs of pre-service English as a foreign language (EFL) Teachers: Evidence from Turkey. *Revista Educación*, 47(1). <http://doi.org/10.15517/revedu.v47i1.51920>
- Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, 45, 475-489. [doi:10.1017/S0261444810000480](https://doi.org/10.1017/S0261444810000480)
- Gonzalez-Martinez, A. & Gandon-Chapela, E. (2021). Boosting English vocabulary knowledge through corpus-aided word formation practice. *RAEL*, 20(1), 49-70.
- Heather, J. & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, 20(4), 415- 440.
- Hirata, Y. & Hirata, Y. (2019). Applying "Sketch Engine for language learning" in the Japanese English classroom. *Journal of Computing in Higher Education*, 1, 233-248. <https://doi.org/10.1007/s12528-019-09208-z>

- Kazaz, İ. (2015). *Corpus-aided language pedagogy: The use of concordance lines in vocabulary instruction*. [Master's thesis, Bilkent University]. Ulusal Tez Merkezi. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Koçak, A. (2020). EFL students' perceptions of using COCA to develop their vocabulary. *Journal of Language Research*, 4(1), 12-26.
- Lee, S. (2011). Challenges of using corpora in language teaching and learning. *Linguistic Research*, 28(1), 159-178.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove, UK: Language Teaching Publications.
- Oğat, D. (2016). Corpus linguistics outcomes and applications in the digital era. In Dejica, D., Hansen, G., Sandrini, P. & Para, I. (Eds.), *Language in the Digital Era. Challenges and Perspectives*, (pp. 37-47). De Gruyter Open Poland. <https://doi.org/10.1515/9783110472059-006>
- Poole, R. (2020). "Corpus can be tricky": Revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 1-22. <https://doi.org/10.1080/09588221.2020.1825095>
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11 (2), 129-158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics* 13, 206-226.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language teaching research*, 12(3), 329-363.
- Sinha, T. S. (2021). EFL learners' perception of and attitude to corpus as a vocabulary learning tool. *The Reading Matrix: An International Online Journal*, (21)2.
- Solak, E. & Recep, Ç. (2014). Examining preservice EFL teachers' TPACK competencies in Turkey. *Journal of Educators Online*, (11)12.
- Şimşek, T. (2020). *Corpora in foreign language teacher education: Introducing a corpus literacy course to ELT pre-service teachers*. [Doctoral dissertation, Çukurova University]. Ulusal Tez Merkezi. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage, London.
- Tosun, S. (2017). *Data-driven language learning: The use of concordance lines in vocabulary learning*. [Master's thesis, Çukurova University]. Ulusal Tez Merkezi. <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1-8.
- Xodabande, I. & Nazari, M. (2022). Impacts of a corpus linguistics course on in-service EFL teachers' corpus literacy. *Computer Assisted Language Learning Electronic Journal*, 23(1), 318-346.