



**DESIGNING AND IMPLEMENTING A CORPUS-BASED LANGUAGE
TEACHING (CBLT) TRAINING PROGRAM IN A PRE-SERVICE EFL
TEACHER EDUCATION CONTEXT
MA THESIS**

Berkcan KÖSE

Eskişehir 2024

**DESIGNING AND IMPLEMENTING A CORPUS-BASED LANGUAGE
TEACHING (CBLT) TRAINING PROGRAM IN A PRE-SERVICE EFL
TEACHER EDUCATION CONTEXT**

Berkcan KÖSE

MA THESIS

Department of Foreign Language Education

MA in English Language Education Program

Supervisor: Assoc. Prof. Dr. Safiye İpek KURU GÖNEN

Eskişehir

Anadolu University

Graduate School

July 2024

JÜRİ VE ENSTİTÜ ONAYI



ABSTRACT
**DESIGNING AND IMPLEMENTING A CORPUS-BASED LANGUAGE
TEACHING (CBLT) TRAINING PROGRAM IN A PRE-SERVICE EFL TEACHER
EDUCATION CONTEXT**

Berkcan KÖSE

Department of Foreign Language Education
MA Program in English Language Teaching
Anadolu University, Graduate School, July 2024

Supervisor: Assoc. Prof. Dr. Safiye İpek KURU GÖNEN

The present study implemented a 20-hour, 9-week systematic Corpus-based Language Teaching (CBLT) Training involving pre-service English teacher candidates with the use of corpora and investigated their perceptions before and after the training regarding the use of corpora in EFL classrooms and the CBLT training. Initially, quantitative and qualitative data were obtained from 102 participants through the Corpus Background Survey on familiarity and awareness of available corpora. Following the training, qualitative data were collected from the 20 participants of the CBLT Training via reflective logs, post-teaching reflective logs, and semi-structured interviews. Descriptive analysis was employed for the quantitative data, while the qualitative data were analyzed qualitatively by using the Constant Comparative Method (CCM). The findings revealed that the participants had very low prior familiarity and knowledge regarding corpora. The results highlighted the benefits (e.g., enhancing vocabulary and grammar instruction and error correction) and drawbacks (e.g., difficulties in finding appropriate language items, technical problems with different concordancers) of corpora use for language teaching. The participants also pointed out some benefits (e.g., raising their language awareness, creating a reflective and collaborative learning environment) and weaknesses (e.g., too much workload on production, need for more corpora practice) of the CBLT training. Finally, the results were discussed in the context of relevant literature, presenting pedagogical implications and suggestions for future research.

Keywords: Corpora use in language teaching, Corpus literacy, Teacher education

ÖZET

HİZMET ÖNCESİ İNGİLİZCE ÖĞRETMEN EĞİTİMİ BAĞLAMINDA BİR DERLEM TABANLI DİL ÖĞRETİMİ (DTDÖ) EĞİTİM PROGRAMININ TASARLANMASI VE UYGULANMASI

Berkcan KÖSE

Yabancı Diller Eğitimi Anabilim Dalı

İngilizce Öğretmenliği Programı

Anadolu Üniversitesi, Lisansüstü Eğitim Enstitüsü, Temmuz 2024

Danışman: Doç. Dr. Safiye İpek KURU GÖNEN

Bu çalışmada, İngilizce öğretmen adaylarının derlem araçlarının kullanımını kapsayan, 20 saatlik ve 9 haftalık, sistematik bir Derlem Tabanlı Dil Öğretimi (DTDÖ) Eğitim programı uygulanmış ve öğretmen adaylarının derlem araçlarının İngilizce öğretimine entegrasyonuna ilişkin eğitim öncesindeki ve sonrasındaki algılarını ve DTDÖ eğitimine ilişkin görüşlerini araştırmak amaçlanmıştır. 102 İngilizce öğretmen adayının derlem araçlarına yönelik farkındalıklarını araştırmak amacıyla Derlem Geçmişi Anketi ile nicel ve nitel veriler elde edilmiştir. Eğitim sonrasında, eğitimin 20 katılımcısından eğitime yönelik yansıtıcı günlükler, öğretim sonrası yansıtıcı günlükler ve yarı yapılandırılmış görüşmeler yoluyla nitel veri toplanmıştır. Nicel analiz için betimsel analiz kullanılırken, nitel veriler Sürekli Karşılaştırma Metodu kullanılarak analiz edilmiştir. Bulgular katılımcıların derlem farkındalık seviyelerinin düşük olduğunu, dil öğretimi için derlem araçlarının kullanımının faydalarını (örneğin, kelime bilgisi ve dil bilgisi öğretimi ve hata düzeltilmeyi geliştirmek) ve eksikliklerini (örneğin, uygun dil öğelerini bulmaktaki zorluklar, farklı derlem araçlarıyla ilgili teknik sorunlar) vurgulamaktadır. Ayrıca katılımcılar DTDÖ eğitiminin bazı faydalarına (örneğin, dil farkındalıklarının artması, yansıtıcı ve işbirlikçi bir öğrenme ortamı yaratılması) ve zayıf yönlerine (örneğin, üretime yönelik aşırı iş yükü, daha fazla derlem pratiğine ihtiyaç duyulması) dikkat çekmişlerdir. Son olarak, bulgular alanyazın bağlamında tartışılmıştır ve pedagojik çıkarımlar ile gelecek çalışmalar yönelik öneriler sunulmuştur.

Anahtar Kelimeler: Dil öğretiminde derlem kullanımı, Derlem okuryazarlığı, Öğretmen eğitimi

ACKNOWLEDGEMENTS

First and foremost, I would love to express my heartfelt and deepest gratitude to my dearest supervisor, Assoc. Prof. Dr. S. İpek KURU GÖNEN, as the completion of this thesis would not have been possible without her constant support, never-ending motivation, constructive feedback, and invaluable guidance. Her dedication to everything she does has been one of the major sources of my academic motivation and ethical mindset since the first days of my academic journey. Having been fascinated by her mindful, caring, and encouraging mentorship, I owe her the biggest thanks and gratitude for all her contributions to my academic and personal development. It is her diligent efforts, sincere support, and great company that I am indebted to the most for all I have achieved as a teacher and academician.

I would like to express my gratitude to my committee members, Prof. Dr. Ali MERÇ and Assoc. Prof. Dr. Hale IŞIK-GÜLER, for their insightful suggestions, comments, and contributions to my study. I would like to extend my deepest gratitude to Assoc. Prof. Dr. Hale IŞIK-GÜLER for her constant encouragement, constructive feedback, and significant contributions to the study, especially the designing process of the CBLT training.

I owe special thanks to Assoc. Prof. Dr. Erdem AKBAŞ, Assoc. Prof. Dr. Betül BAL-GEZEGİN, Assoc. Prof. Dr. Ali Şükrü ÖZBAY and Asst. Prof. Dr. Ufuk GİRGIN for their contribution and insightful suggestion to my study and the development process of the data collection instruments. I am also deeply indebted to my dearest colleagues and friends, Asuman ŞİMŞEK-TONTUŞ, Gözde KIRMIZI, and Zeynep Yadigar TOSUN, for their encouragement, everlasting assistance, and constant support, which made this study possible.

I would like to thank my dearest students for taking the time to participate in the training and showing their enthusiasm regarding the training content and data collection procedures throughout the training.

I would love to express my very special gratitude to my dear family for their unconditional love, faith, and trust in me, not only during the completion of this thesis but my whole life.

I would love to extend the most special thanks to my dearest cats, Turşu and Fıstık, for making this journey and whole life bearable. I am grateful for their company and most innocent love.

Last but not least, my heartfelt gratitude and eternal thanks to my beloved wife, Gamze KILIÇ KÖSE, for her endless and unconditional love, encouragement, and patience since we first met. Without her company, not only this thesis but my whole life would have been incomplete.

This study has been funded by the Scientific and Technical Research Council of Türkiye (TÜBİTAK) by providing a scholarship, “2210-A Genel Yurt İçi Yüksek Lisans Burs Programı”.

Berkcan KÖSE
Eskişehir, 2024



STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Anadolu University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Berkcan KÖSE

TABLE OF CONTENTS

	<u>Page</u>
COVER PAGE	i
FINAL APPROVAL FOR THESIS	ii
ABSTRACT	iii
ÖZET.....	iv
ACKNOWLEDGEMENTS	v
STATEMENT OF COMPLIANCE WITH ETHICAL PRINCIPLES AND RULES	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiii
1. INTRODUCTION	1
1.1. Background to the Study	1
1.2. Statement of the Problem	5
1.3. Purpose of the Study	8
1.4. Significance of the Study	8
1.5. Scope and Limitations of the Study.....	10
1.6. Definitions of Key Terms.....	11
2. REVIEW OF LITERATURE	13
2.1. Corpus Linguistics in Language Pedagogy	13
2.2. Consulting Corpora in Language Education	19
2.2.1. Indirect corpus applications in language education	20
2.2.2. Direct corpus applications in language education	23
2.2.3. Using corpora for language skill development	25
2.3. Corpora in Language Teacher Education	32
3. METHODOLOGY	40
3.1. Research Design	40
3.2 Setting and Participants	41
3.2.1. Setting	41
3.2.2. Participants	42
3.3. Corpus-based Language Teaching Training	44

3.3.1. Procedure of the Corpus-based Language Teaching Training	47
3.3.2 Tools Used in the Corpus-based Language Teaching Training	54
3.3.2.1 <i>ODTÜClass</i>	54
3.3.2.2 <i>Doodle</i>	55
3.3.2.3 <i>WhatsApp</i>	56
3.4. Data Collection Instruments	57
3.4.1. Corpus Background Survey	58
3.4.2. Reflective Logs	59
3.4.3. Post-teaching Reflective Logs	59
3.4.4. Semi-structured Interviews	60
3.5. Data Collection Procedures	61
3.6. Data Analysis	63
4. FINDINGS	66
4.1. Overview of the Study	66
4.2. RQ1: Corpus Familiarity, Knowledge, and Experiences of Pre-service EFL Teachers in EFL Contexts	67
4.3. RQ2: Perceptions of Pre-service EFL Teachers after the Training ...	75
4.3.1. RQ2.1: Perceptions on the integration of corpora into actual teaching contexts	76
4.3.1.1. <i>Benefits of corpora use for language teaching</i>	76
4.3.1.2. <i>Drawbacks of corpora use for language teaching</i> ...	88
4.3.2. RQ2.2: Perceptions regarding the CBLT training	95
4.3.2.1. <i>Strengths of the CBLT training</i>	95
4.3.2.2. <i>Weaknesses of the CBLT training</i>	100
5. DISCUSSION	102
6. CONCLUSION	116
6.1. Summary of the Findings	116
6.2. Conclusions	119
6.3. Implications	123
6.4. Suggestions for Further Research	127
REFERENCES	129

APPENDICES
CURRICULUM VITAE



LIST OF TABLES

	<u>Page</u>
Table 2.1. The list of available online corpora used in the EFL context	15
Table 3.1. Demographic information of the participants	43
Table 3.2. Demographic information of the participants in the training sessions	44
Table 3.3. The list of corpus software and online corpora demonstrated in the training	47
Table 3.4. The syllabus and schedule of the Corpus-based Language Teaching Training	48
Table 3.5. Research Questions and Data Collection Instruments	57
Table 3.6. Data Collection Procedure	63
Table 3.7. Data Analysis Procedure	65
Table 4.1. The pre-service EFL teachers' knowledge of online corpora and corpus software	68
Table 4.2. The results related to the ways of being informed about corpora in EFL	69
Table 4.3. The pre-service EFL teachers' knowledge of using corpora for different pedagogical purposes	70
Table 4.4. The main categories related to the integration of corpora into actual language teaching	76
Table 4.5. The sub-categories related to the benefits of corpora use for language teaching	77
Table 4.6. The sub-categories related to the drawbacks of corpora use for language teaching	88
Table 4.7. The distribution of the codes regarding the drawbacks of corpora use for language teaching	93
Table 4.8. The main categories related to the systematic corpus-based language teaching training.....	95
Table 4.9. The sub-categories related to the strengths of the systematic corpus-based language teaching training	96
Table 4.10. The sub-categories related to the weaknesses of the systematic corpus-based language teaching training	100

LIST OF FIGURES

	<u>Page</u>
Figure 2.1. The interface of COCA	16
Figure 2.2. Some different features of COCA	18
Figure 2.3. Sample concordance lines retrieved from COCA	19
Figure 2.4. The use of corpora in second language learning and teaching (Römer, 2011)	20
Figure 2.5. Sample published corpus-informed dictionaries and textbooks	22
Figure 2.6. A sample corpus-informed grammar teaching activity (Girgin, 2011)	24
Figure 3.1. Outline of the present study's research design (visual model adapted and revised from Ivankova et al., 2006)	40
Figure 3.2. Overview of the experiential learning cycle by Kolb (1984)	46
Figure 3.3. A screenshot of a sample corpus search on AntConc used in session 3	51
Figure 3.4. Overview of CBLT Training course page on ODTÜClass	54
Figure 3.5. Screenshots of the resources to be used in EFL Classrooms and the Podcasts post	55
Figure 3.6. The interface of the Doodle poll for semi-structured interviews	56
Figure 6.1. The outline of the implications of the study	126

LIST OF ABBREVIATIONS

BAWE	: British Academic Written English
BNC	: British National Corpus
CBLP	: Corpus-Based Language Pedagogy
CBLT	: Corpus-Based Language Teaching
CELTA	: Certificate in Teaching English to Speakers of Other Languages
CL	: Corpus Linguistics
COCA	: Contemporary Corpus of American English
DDL	: Data-Driven Learning
EAP	: English for Academic Purposes
EFL	: English as a Foreign Language
ELT	: English Language Teaching
ESP	: English for Specific Purposes
FLE	: Foreign Language Education
HEC	: Higher Education Council of Türkiye
KWIC	: Keyword in Context
LMS	: Learning Management System

MICASE : Michigan Corpus of Academic Spoken English

MoNE : Ministry of National Education of Türkiye

TESOL : Teaching English to Speakers of Other Languages



1. INTRODUCTION

This chapter aims to thoroughly outline the background of the study, the statement of the problem, the purpose of the study, the significance of the study, the scope and limitations of the study, and the definitions of the key terms, respectively.

1.1. Background to the Study

Recent developments and rapid progress of computer technologies have revolutionized language education as much as it has affected all aspects of our daily lives. These breakthroughs highlight the importance of the introduction and integration of new, innovative tools for enhancing language instruction and practice in language teaching and learning environments, as well as providing language learners, teachers, and researchers with significant opportunities to enhance language education (Başal et al., 2024; Eslek-Onur & Tosun, 2023; Ma et al., 2023). Along with many other recent advancements in computer technologies that call for the need to keep language teaching environments up-to-date, artificial intelligence (AI) has also entered our lives at an unprecedented pace through various applications of AI-based generative language models (e.g., ChatGPT, Claude, Gemini, Pi, etc.) with ground-breaking possibilities like more accessible education, learning management and (r)evolutionary changes regarding teaching methods that might deeply affect the current nature of language education and teacher training (Alharbi, 2023; Ninaus & Sailer, 2022; Pokrivcakova, 2023).

Leveraging the integration of corpora into language teaching pedagogy as a “computer-assisted technique” for enriching educational practices has received a considerable amount of attention as a result of such advancements in technology (Farr & Leńko-Szymańska, 2023; Ma et al., 2023; Özbay & Gürsoy, 2023; Vyatkina & Boulton, 2017; Tsai, 2019). Referred to as the systematic collection and compilation of naturally occurring, authentic language data from texts in digital forms (Bennett & Uí Dhonnchadha, 2023; O’Keeffe et al., 2007; Römer, 2011), a corpus (plural corpora) is consulted for analyzing the language used in real life by identifying and interpreting patterns in language data (Başal et al., 2024; Köse & Kuru Gönen, 2023; Schmidt, 2023; Tsai, 2019; Xodabande & Nazari, 2022). The collection of language data as a corpus includes the compilation of written texts and the transcriptions of spoken language in

written form (Farr & O’Keeffe, 2019; Huang, 2011; Poole, 2022; Timmis, 2015). Based on the purposes and criteria for the inclusion and exclusion of each corpus, language data for the majority of corpora are obtained from written materials like articles, blogs, books, newspapers, and magazines, and the recorded and transcribed scripts of TV shows, movies, radio programs, weather forecasts, and conversation (Hunston, 2002; McCarthy & O’Keeffe, 2010; Pérez-Paredes, 2022; Poole, 2022). Each corpus represents a different aspect of a language as they differ in terms of their types, vocabulary size, and types of texts and genres collected (Huang, 2018).

The utilization of corpora in the forms of online-accessible corpora and corpus tools and applications in language pedagogy holds significant potential for English language teaching as corpus use not only enables linguistics analysis of a language but also provides access to each instance language use of a specific language item across various contexts with multiple features (Abdel Latif, 2021; Chen et al., 2019; O’Keeffe et al., 2007; Poole, 2022). A specific query related to a word or a phrase in a corpus unveils valuable insights into the language item, such as its genre, usage patterns and frequency, collocations, phrasal verb usage, idiomatic expressions, register, and use across various contexts (Ebrahimi & Faghih, 2017; Farr & O’Keeffe, 2019; Hunston, 2002; O’Keeffe et al., 2007; Özbay & Gürsoy, 2023; Şen Bartan, 2023). The use of corpora for English language teaching and learning has received an emerging, increasing popularity as access to authentic language is vital to language education, especially since conventional teaching materials are criticized for being based on subjective intuitions and judgments of material writers (Abdel Latif, 2021; Aşık, 2015; Gilmore, 2019; Leńko-Szymańska, 2017; Meunier & Reppen, 2015; Naismith, 2017; Sun & Hu, 2023; Şimşek, 2020).

In order to utilize the benefits of corpora for language education, Johns (1991) coined the term “data-driven learning” (DDL), the involvement and engagement of language learners with corpora for recognizing and identifying language patterns inferring grammar rules and drawing insights from concordance lines autonomously at their own pace (Bennett & Uí Dhonnchadha, 2023; Boulton, 2010; Can, 2023; Chen et al., 2019; Flowerdew, 2015; Ma et al., 2023). A concordance line is defined as the display of the distribution of the searched target word or phrase in a number of lines where each occurrence of that item with other words that surround it across texts (Boulton & Cobb, 2017; Özbay & Gürsoy, 2023; Şimşek & Can, 2023). The consultation of corpora by language learners is believed to foster discovery learning (Bernardini, 2004; Boulton &

Cobb, 2017; Michali & Patsala, 2020), as it puts learners in the form of a researcher, and they engage with the linguistic analysis of the language by analyzing concordance lines (Bennett & Uí Dhonnchadha, 2023; Chambers, 2019; Çalışkan, 2020; Gilquin & Granger, 2010). Moreover, with the integration of DDL into language pedagogy, language learners become more motivated to take charge of their own learning, and it enhances their autonomy and language awareness as they work on real language data on their own (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Charles, 2022; Çalışkan, 2020; Farr, 2008; Gilquin & Granger, 2010; Zareva, 2017).

The use of corpora for language learning and teaching purposes in language classrooms is known as corpus-based language pedagogy (CBLP) (Farr & Leńko-Szymańska, 2023; Ma et al., 2021), and it is regarded as the integration of corpus linguistics (CL) into language education. This integration works in two distinct ways. The first one is the *indirect* application of corpora in language education. The indirect use of corpora in language education covers a vast number of applications of corpora, including corpus consultation and corpus-based language analysis, and it enables the creation and development of language teaching materials like textbooks and grammar books, reference publishing, and dictionaries, syllabus design, and language test development and assessment (Cushing, 2017; Đurović & Bauk, 2022; Götz, 2022; Kaya et al., 2022; McEnery & Xiao, 2011; Park, 2014). The second way of integrating corpora is the *direct* application of corpora by language teachers or learners in language education environments (Abdel Latif, 2021; Bennett & Uí Dhonnchadha, 2023; McEnery & Xiao, 2011; Sinha, 2021). The direct applications of corpora in language education refer to the hands-on use of corpora by language learners and teachers for practicing the naturally occurring, actual language data in concordance lines (Bennett & Uí Dhonnchadha, 2023; Bernardini, 2004; Boulton & Cobb, 2017; Köse & Kuru Gönen, 2023; Şimşek, 2020).

As DDL stands out from conventional methods in language education contexts with its significant benefits in language instruction on various language skills (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Çalışkan, 2020; Gilquin & Granger, 2010; Godwin-Jones, 2017; Lee et al., 2019; Mizumoto & Chujo, 2015; Sinha, 2021; Timmis, 2015; Xodabande & Nazari, 2022), there has been a growing number of studies on the integration of corpora and DDL into vocabulary instruction and learning in recent years as it provides language learners with the actual usage of vocabulary items across different contexts (Alsahafi, 2022; Barabadi & Khajavi, 2017; Can, 2023; Chen, 2017;

Daskalovska, 2015; Eslek-Onur & Tosun, 2023; Karras, 2016; Koçak, 2020; Li, 2017; Lu & Dang, 2022; Saddhono et al., 2023; Sinha, 2021; Tsai, 2019; Uçar & Yükselir, 2015; Vyatkina, 2016; Zhang, 2022). Likewise, studies examining corpora use for grammar teaching also highlight benefits as DDL offers authentic examples of grammar usage across different contexts with corpus-informed grammar teaching activities for learners (Girgin, 2011, 2019; Emir & Yangın-Ekşi, 2023; Hughes, 2010; Khan et al., 2021; Lin & Lee, 2019; Lakew et al., 2021; Lin, 2019; Özer & Özbay, 2022; Rodríguez-Fuentes & Swatek, 2022; Römer, 2004; Smart, 2014; Vannestal & Lindquist, 2007; Zhong & Wakat, 2023). Similarly, empirical research on the integration of corpora into language education has also highlighted the value of utilizing corpora in writing instruction and error correction (Anthony, 2016; Chang, 2014; Emir, 2019; Gaskell & Cobb, 2004; Gilmore, 2008; Luo & Liao, 2015; O’Sullivan & Chambers, 2006; Quinn, 2015; Yoon & Hirvela, 2004).

Despite its emerging popularity in relevant research, the integration of corpora in language classrooms, as well as the proficiency of language teachers in corpus-based language teaching pedagogy, remains limited (Başal et al., 2024; Callies, 2019; Chambers, 2019; Ebrahimi & Faghih, 2017; Leńko-Szymańska & Boulton, 2015; Ma et al., 2022). The limited use of corpora in language education contexts largely stems from language teachers’ lack of awareness and requisite knowledge on how to effectively consult and utilize corpora (Breyer, 2009; Chen et al., 2019; Frankenberg-Garcia, 2012; Gilquin & Granger, 2010; Naismith, 2017; Tribble, 2015). Aşık (2015) surveyed 42 EFL instructors in the Turkish EFL context, revealing that nearly half the instructors lacked familiarity with corpora, with only a few having prior experience using corpora. Similarly, Bunting (2013) conducted a study with EFL teachers to assess their familiarity with corpus tools, and it was revealed that most of the EFL teachers lacked knowledge of how to integrate corpora into their teaching despite their familiarity with CL from their university courses. In a more recent investigation by Pérez-Paredes (2022), the findings revealed that teachers do not utilize corpora in their classrooms, and these results stressed the importance of training language teachers in the pedagogical applications of corpora by devoting more training and courses on the *know-how* of corpus consultation that connects the use of corpora with classroom practices and activities (Bennett & Uí Dhonnchadha, 2023; Leńko-Szymańska, 2017; Mukherjee, 2004; Römer, 2011; Vyatkina, 2020). Therefore, the integration of corpora consultation into initial teacher training for

English teacher candidates is an area open to investigation (Breyer, 2009; Chen et al., 2019; Lin, 2019; McCarthy, 2008; Mukherjee, 2006; O’Keeffe & Farr, 2003; Şimşek & Can, 2023).

Researchers have underlined the significance and necessity of providing courses and training on features, tools, and utilization of corpora for prospective EFL teachers as such training equips them with the ability to consult corpora for language analysis by concordancing and create corpus-informed language teaching materials to utilize in their future teaching practices (Abdel Latif, 2021; Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Çalışkan, 2020; Heather & Helt, 2012; Köse & Kuru Gönen, 2023; Leńko-Szymańska, 2017; Naismith, 2017; Tribble, 2000). Leńko-Szymańska (2014) points out the significance of obtaining skills and expertise in consulting corpora effectively for pre-service EFL teachers during teacher training. Similarly, it was emphasized in Naismith’s (2017) study that the integration of corpus tools in language teaching indicated potential benefits for improving language awareness and facilitating lesson and teaching material creation. Zareva (2017) also highlights language teachers’ low familiarity with and limited prior experience with corpora consultation and the need for more hands-on corpora practice opportunities for language teachers. Şimşek (2020) and Köse & Kuru Gönen (2023) also emphasized a need for more courses to expand the corpus literacy and practice of corpora consultation for pre-service teachers for their future language teaching practices after graduation.

All in all, the crucial role of offering training on using corpora and its pedagogical implications in language education is highlighted as empirical studies stress the need for structured, systematic, and comprehensive training initiatives to equip EFL teachers with the necessary skills for the effective integration of corpus-informed English teaching materials into their classroom practice (e.g., Başal et al., 2024; Farr & Leńko-Szymańska, 2023; Köse & Kuru Gönen, 2023; Schmidt, 2023; Şimşek, 2020). Corpora can enhance language education across various dimensions and aspects if prospective teachers are introduced to corpora and different corpus applications in language teaching during initial pre-service teacher training.

1.2. Statement of the Problem

According to research studies in relevant literature, it is evident that language teachers must be exposed to corpora and need instruction on integrating its use into their

teaching practices such as utilizing concordancing applications and that there is a need for the inclusion of corpora into both initial teacher training for prospective pre-service EFL teachers and the professional development of in-service language teachers (Chambers et al., 2011; Chen et al., 2019; Çalışkan & Kuru Gönen, 2018; Frankenberg-Garcia, 2012; Lin, 2019; McCarthy, 2008; Mukherjee, 2004; O’Keeffe & Farr, 2003; Römer, 2011; Şimşek & Can, 2023). However, in the Turkish EFL context, the integration of corpora into language teaching remains limited due to EFL teachers’ lack of awareness and knowledge regarding corpus utilization as they are unfamiliar with what is available in corpora, as well as how to access corpora and use concordance lines for language teaching purposes (Aşık, 2015; Başal et al., 2024; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Şimşek, 2020). Moreover, the involvement and integration of corpora into teacher training programs is still not a common practice despite the benefits in the relevant literature, and there is not a systematic, comprehensive training program approach to the integration of corpora into ELT contexts for the professional development of prospective teachers other than individual efforts of some researchers in the Turkish context (Aşık, 2017; Köse & Kuru Gönen, 2023; Özbay, 2017; Şimşek, 2020).

Recently, the use of AI has been gaining an increasing amount of attention in every aspect of life with various applications of AI-based generative language models like ChatGPT, Claude, and Gemini, and many more with endless possibilities of what can be generated with access to any type of information with one click at lightning speed (Alharbi, 2023; Ninaus & Sailer, 2022; Pokrivcakova, 2023). While these AI models are capable of generating human-like texts for any specific prompt given, providing personalized learning experiences according to the needs and expectations of the user, and even creating language teaching and learning materials, the use of AI in education has brought about critical concerns regarding the reliability and credibility of AI-generated content (AlAfnan et al., 2023; Akgün & Greenhow, 2022; Balta, 2023; Kung et al., 2023; Perkins, 2023; Pokrivcakova, 2023).

Unlike the use of corpora and similar data-driven approaches that are based on authentic language usage with high reliability and credibility dating way back in the literature, it is a well-known fact that AI models can sometimes generate inaccurate or misleading information, thereby posing critical concerns regarding ethical considerations while creating language teaching materials for language instruction (Akgün & Greenhow, 2022; Balta, 2023). While AI-generated materials can also be a valuable supplement to

language education, there is a need for language teachers and learners to use such models with caution and evaluate the AI-generated content critically (AlAfnan et al., 2023; Balta, 2023; Perkins, 2023; Pokrivcakova, 2023). Therefore, this calls for the need to the integration of corpora into language teaching environments for accessing reliable and authentic language data, as well as systematic and comprehensive training programs for the professional development of prospective EFL teachers to equip them with required skills to consult corpora for language teaching purposes (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Şimşek, 2020). Nonetheless, it should be noted that the recent advancements in AI technology hold the potential to foster the integration of corpora into language teaching environments in the future in terms of assisting corpora consultation efficiently for the analysis of large corpora, identifying language patterns, and generating insights to further enrich teaching practices as language teachers and learners can benefit from the combination of the strengths of AI-generated content with the reliability and authenticity of language data retrieved from corpora.

It is crucial to raise the language awareness of pre-service EFL teachers to improve their future teaching practices with opportunities to learn about corpora, how to search for a language item, analyze concordance lines to interpret language data, and use corpora for teaching practices via corpus applications and creating corpus-informed language teaching materials for an actual classroom environment (Çalışkan, 2020; Frankenberg-Garcia, 2012; Leńko-Szymańska, 2017; Lin, 2019; Mukherjee, 2004; Römer, 2011; Şimşek, 2020). Despite such different applications of corpora with the aforementioned benefits in the literature, there is a lack of studies that encourage the consultation of corpora and creating corpus-informed language teaching materials and trace their use in actual language teaching classrooms after training pre-service EFL teachers on corpora and corpus literacy. Within the Turkish context, although some research has explored the incorporation of corpus literacy for pre-service EFL teachers (Aşık, 2017; Köse & Kuru Gönen, 2023; Özbay, 2017; Şimşek, 2020), only limited insight exists regarding English teacher candidates' views and experiences related to the use of available corpus tools and corpus-informed language teaching materials into their teaching practices. In order to meet this need and to assist pre-service EFL teachers in the effective integration of recent advancements such as corpora use, this study aimed at designing a systematic and comprehensive training program on the integration of corpora

into EFL teaching environments and offer various suggestions for EFL teacher educators, instructors, administrators, and language practitioners.

1.3. Purpose of the Study

The primary purposes of the present study were to investigate the perceptions and reflections of Turkish pre-service EFL teachers regarding corpus and corpus-based language teaching and its practical implementation in English language teaching contexts before and after receiving the 9-week systematic and comprehensive Corpus-based Language Teaching (CBLT) Training tailored for the specific context of the current study. Based on the relevant need for providing training opportunities for pre-service EFL teacher education, this study aimed at designing a corpus training program with a specific focus on hands-on practices of corpora integration into actual teaching contexts aligned with the needs of the learners and EFL teaching environments along with the necessary theoretical information on corpus-based language teaching. Moreover, this study explores the perceptions and views of the English teacher candidates about the CBLT Training program, which aimed to engage the participants with corpus consultation for designing corpus-informed language teaching materials and integrating these materials into their actual teaching experiences in language classrooms in their practicum schools.

Taking into account the scope and the aforementioned purposes, the present study seeks to answer the following research questions:

- 1) What do Turkish pre-service EFL teachers think about corpora and the integration of corpora into language teaching contexts?
- 2) What are the perceptions of Turkish pre-service EFL teachers regarding the
 - a. integration of corpora into actual teaching contexts?
 - b. systematic Corpus-Based Language Teaching Training?

1.4. Significance of the Study

In the relevant literature, empirical studies aimed to investigate the effectiveness of corpus applications in language teaching for skill development and the perceptions of EFL learners on the implementation of corpus-informed EFL activities and materials (Crosthwaite, 2017; Girgin, 2019; Lin & Lee, 2019; Tono et al., 2014). Regarding EFL teacher training, the primary focus was on training pre-service and in-service EFL

teachers about corpora and the use of corpus tools, as well as investigating their views and perspectives on incorporating corpora into language teaching and learning (Çalışkan & Kuru Gönen, 2018; Ebrahimi & Faghih, 2017; Farr, 2008; Tribble, 2015; Zareva, 2017). However, unlike these studies, the present study focuses on the interplay between equipping prospective pre-service EFL teachers with corpus literacy and providing opportunities to create DDL language teaching materials for actual teaching practices in their practicum schools.

In the Turkish EFL context, there is a lack of empirical studies that encourage the integration of an integrative, interactive, and systematic model of training programs on the training of pre-service EFL teachers on corpus literacy and the consultation and utilization of different tools for EFL teaching with a focus on monitoring their developing knowledge and skills, and material design development. This study highlights the significance of the integration of corpus-based language teaching into actual teaching practices with systematic and comprehensive training that covers a wide range of strategies for prospective EFL teachers so that they could experience corpora integration both as a learner and a teacher by first learning about corpora and then, exploiting corpora to teach (Breyer, 2009; Fligelstone, 1993; Schmidt, 2023). This study may shed light on senior English teacher candidates' perceptions and views about the use of various corpus tools and corpus-informed teaching materials in their teaching practices, as well as insights into the efficacy and implementation of such training on corpus literacy in teacher training programs.

Considering the initial low familiarity with corpora and lack of research and training on corpora in the Turkish EFL context, there is a need to investigate Turkish pre-service EFL teachers' corpus knowledge, their views on corpora, and their experience in its practical applications in language instruction to gain deeper insights and develop tailored systematic, comprehensive training programs on the integration of corpora into language teaching environments. The relevant research suggests the need for pre-service training to familiarize EFL teachers with corpus use and corpus-based language teaching pedagogy, equip them with practical skills for consulting corpora, demonstrate how corpora can be integrated into English language instruction, and provide opportunities for gaining hands-on experience for implementing corpus-informed teaching activities into their EFL teaching practices (Başal et al., 2024; Schmidt, 2023). In order to meet this need and to assist prospective EFL teachers in the effective integration of corpora, this

study aimed at designing a systematic and comprehensive training program on the integration of corpora into EFL teaching environments and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms. The CBLT Training was tailored for the specific context of the study and implemented into the practicum processes of senior pre-service EFL teachers as an attempt to equip them with the theoretical knowledge and practical skills of consulting and utilizing corpora, demonstrating how corpus use can be integrated across various language domains and skills, and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms.

The present study suggests pedagogical implications regarding integrating the use of corpora for language teacher training programs to equip EFL teacher candidates with corpus literacy and the integration of corpora into actual teaching practices of EFL teachers. The study also contributes to the teacher training program curricula in the Faculty of Education of universities in the Turkish context for training English teacher candidates. Moreover, it can be noted that the integration of corpora into language classrooms would also contribute to the teacher training curricula for the teacher education of teachers and learners of other languages.

1.5. Scope and Limitations of the Study

The present study was conducted at the Middle East Technical University in Türkiye during the Fall Term of the 2023-2024 Academic Year to explore the perceptions and views of Turkish English teacher candidates on corpora, corpus-based teaching pedagogy and the implementation of teaching materials created by the utilization of various available corpora in language teaching before and after a systematic and comprehensive training on the theoretical and practical aspects of corpus software and online corpora. While 102 senior Turkish pre-service EFL teachers studying at the Department of Foreign Language Education (FLE) participated in the study's initial data collection phase, 20 voluntary and available pre-service EFL teachers took part in the Corpus-Based Language Teaching (CBLT) Training and the second phase of the study. Therefore, the findings of the present study are limited to its context and the 20 pre-service EFL teachers, and this could potentially limit the generalizability of the findings to a broader population of pre-service EFL teachers in all EFL teaching contexts. Hence,

future studies might be conducted in contexts with larger sample sizes to generalize the effects of integrating corpora into language teacher training programs and actual teaching experiences.

In the present study, pre-service EFL teachers designed at least three corpus-informed language teaching materials to integrate their language teaching experiences during their practicum at various high schools in Ankara, Türkiye. However, due to the nature of the practicum, they could only teach two lessons throughout the semester and the training. Additional research could explore the adaptation of more extended implementation and long-term effects of such training on corpora use for pre-service EFL teachers in their actual classroom practices. As the participant student-teachers were assigned to various schools and grades for their practicum process, the specific classroom, students, and the topic they were assigned to teach might have an influence on their perceptions and views regarding the implementation of corpora into language teaching. Therefore, future studies may consider incorporating different grade levels to generalize the findings. Moreover, as the CBLT training was held in a regular classroom without computer access, participants had to bring their laptops to practice corpora. For future research, providing computer facilities to participants would help to surpass this limitation.

1.6. Definitions of Key Terms

Corpus (plural corpora) is the systematic, electronic collection of authentic language data derived from texts (Römer, 2011).

Corpus linguistics (CL) refers to the empirical examination of systematic collections of authentic language data using computer-based methods (McCarthy & O’Keeffe, 2010).

Corpus literacy refers to the knowledge, skills, and competencies required to employ corpora in language teaching to investigate the language and enrich learners’ language development by analyzing concordance lines (Heather & Helt, 2012).

Concordance line refers to the display of the distribution of a specific word or phrase across various texts and contexts (Hunston, 2002; O’Keeffe et al., 2007).

Corpus-based language pedagogy (CBLP) refers to the capacity to integrate corpus linguistics into language teaching practices in the classroom (Boulton, 2009; Römer, 2011).

Corpus-informed teaching materials refer to language teaching materials and exercises designed by consulting concordance lines derived from a corpus.

Data-driven learning (DDL) is a language teaching approach involving language learners actively engaging with corpora to explore language patterns by concordancing independently (Johns, 1991).

The direct application of corpora is the use of corpora and corpus data by language learners and teachers for practicing the naturally occurring, authentic language data in concordance lines in actual classroom practices (Römer, 2011).

The indirect application of corpora is the use of corpora and corpus data by material writers, curriculum developers, and researchers to integrate corpora use into teaching curricula, dictionaries, and textbooks (Römer, 2011).

2. LITERATURE REVIEW

The primary purpose of this chapter is to offer a comprehensive outline of the relevant research in CL under three main sections. The first section begins by introducing corpus linguistics along with detailed information about different corpora and the role of *Corpus Linguistics in Language Pedagogy*. The next section presents the *indirect* and *direct applications* of corpora in language education as well as the use of corpora for *language skill development* as the ways of *Consulting Corpora in Language Education*. The last section delves into the integration of *Corpora in Language Teacher Education* to equip foreign language teachers with the ability to utilize corpora for language teaching purposes.

2.1. Corpus Linguistics in Language Pedagogy

The emergence of the first electronic, computerized corpora dates back to 1961 as the Brown Corpus was compiled by Henry Kučera and W. Nelson Francis for the sake of linguistics research by systematically collecting a large amount of naturally occurring language data and analyzing that large, principled data with reliable, computerized results (Conrad, 2000; Kučera et al., 1967; Schmidt, 2023; Tognini-Bonelli, 2010). Technological progress and the availability of the internet paved the way for the ease of compiling corpora, and there has been a substantial amount of increase in the number of empirical studies on “computer-assisted techniques” for analyzing language ever since (Ma et al., 2023; Özbay & Gürsoy, 2023; Sinclair, 2004; Vyatkina & Boulton, 2017).

Conceptualized as analyzing and investigating the language in a short time with little effort, the first attempts of corpus linguistics were shaped around running language analysis for research purposes, which was followed by the formation and compilation of corpus-informed dictionaries (Huang, 2018; O’Keeffe et al., 2007; Şimşek, 2020). John Sinclair’s COBUILD monolingual learner’s dictionary project (1987) initiated one of the earliest indirect applications of corpus use for creating corpus-based dictionaries and grammar reference books like the “Longman Grammar of Spoken and Written English” (Biber et al., 1999) for learners of English as a foreign language (EFL) (Huang, 2018; Tribble, 2015). In addition to the creation of such dictionaries and grammar reference books by leveraging corpora to enhance language analysis across multiple dimensions, including word frequency and associations at the semantic level, thesaurus, and

concordancing, corpus-informed language learning and teaching materials have also been created for English language learning and teaching purposes (Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; McCarthy et al., 2005; O’Keeffe et al., 2007). This production has been regarded as one of the most outstanding contributions of corpus linguistics on language pedagogy in language teaching and learning contexts (McCarthy & O’Keeffe, 2010; Tsai, 2019).

Defined as the systematic, electronic collection of authentic language data derived from texts, the use of corpora in English language education has received a considerable amount of attention for it enables researchers, learners, and teachers to analyze the language on authentic data for pedagogical purposes with the progress in technology and ease of access to computers (Başal et al., 2024; Köse & Kuru Gönen, 2023; Römer, 2011; Schmidt, 2023; Tsai, 2019). Consulted for a vast number of purposes, any particular corpus comprises electronically compiled language in written form, derived from either written texts, transcribed spoken texts, or a combination of both (Çalışkan, 2020; Schmidt, 2023). Essentially, written texts collected in a corpus include written materials like books, articles, magazines, newspapers, and blogs, while spoken texts encompass recorded and transcribed radio and TV shows, movies, formal and informal conversation, and weather forecasts in digital forms (Farr & O’Keeffe, 2019; Poole, 2022). Consequently, corpora can vary significantly depending on the purpose and the inclusion and exclusion criteria of the specific corpus compiled. Starting with the 1st generation of major corpora in the 1970s, including the Brown Corpus, the Lancaster-Oslo/Bergen (LOB) Corpus, and the London-Lund Corpus of Spoken English (LLC), each modern corpus possesses different qualities and features, including their type (e.g., general language, specialized, native speaker, learner, etc.), vocabulary size (e.g., large with fourteen billion words, small with thirty thousand words), and the type of texts and genres compiled (Huang, 2018; Tsai, 2019).

Providing the actual, naturally occurring samples of the language, the majority of English language corpora used by researchers, teachers, and learners in EFL contexts are general/reference and specialized corpora that feature a wide array of texts and genres (e.g., Bunting, 2013; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Şimşek, 2020; Zareva, 2017). One of the free-access online corpora consulted in English language teaching and learning contexts is the British National Corpus (BNC), which contains 100 million words from magazines, fiction, academic texts, blogs, and movies (Godwin-Jones, 2017). In

addition to such large general/reference corpora with hundreds of millions of words as authentic language data, the British Academic Written English Corpus (BAWE) and the Michigan Corpus of Academic Spoken English (MICASE) can be given as samples of specialized corpora used for specific purposes in English language teaching contexts (Bouzekria et al., 2023; Wazni et al., 2023). While MICASE comprises more than 1.7 million words from nearly 200 hours of transcribed academic spoken English, BAWE contains 6.9 million words taken from the written academic work of students in the British context (Sanosi, 2022; Yiru & Le, 2024). Despite being smaller in scale, such corpora are tailored for specific language education or research purposes (Đurović & Bauk, 2022; Farr & O’Keeffe, 2019), like being used in English for Specific Purposes (ESP) or English for Academic (EAP) contexts and the demonstration of the direct use of corpora for *teaching* language learners *about* corpus linguistics and *teaching* them *to exploit* corpus data by making queries and working on concordance lines as proposed by Fligelstone (1993). Last but not least, one of the most popular and widely recognized free-access online corpora consulted in English language teaching and learning contexts, the Corpus of Contemporary American English (COCA) houses a vast one billion words taken from various academic texts, magazines, blogs, and movies (Godwin-Jones, 2017; Poole, 2022). Table 2.1. below outlines the features of the online corpora given above.

Table 2.1. *The list of available online corpora used in the EFL context*

Online Corpora	Features
BNC – British National Corpus	Focus on general language use 100 million words 1980-1990s Mix of written and spoken texts
BAWE - British Academic Written English Corpus	Focus on academic contexts 6.9 million words 2004-2007 Written academic English
COCA – Corpus of Contemporary American English	Focus on general language use The largest online corpora with over 1 billion words 1990-2020 Mix of written and spoken texts
MICASE – Michigan Corpus of Academic Spoken English	Focus on academic contexts 1.7 million words 1997-2001 Spoken academic English

As can be seen in Table 2.1, each online corpus given above excels in different areas due to their varying focus, types of texts, number of words included, and periods.

The choice of which corpus to consult depends on whether the focus is on British or American English, written or spoken language data, general usage, or academic contexts for linguistic analysis and practical applications in language teaching. As a versatile, up-to-date corpus with the most recent update in March 2020, COCA is often considered the most comprehensive due to its broad applications and used the most widely in language teaching and learning contexts (Bennett & Uí Dhonnchadha, 2023; Ma et al., 2023; Schmidt, 2023). Figure 2.1. illustrates the interface of COCA:

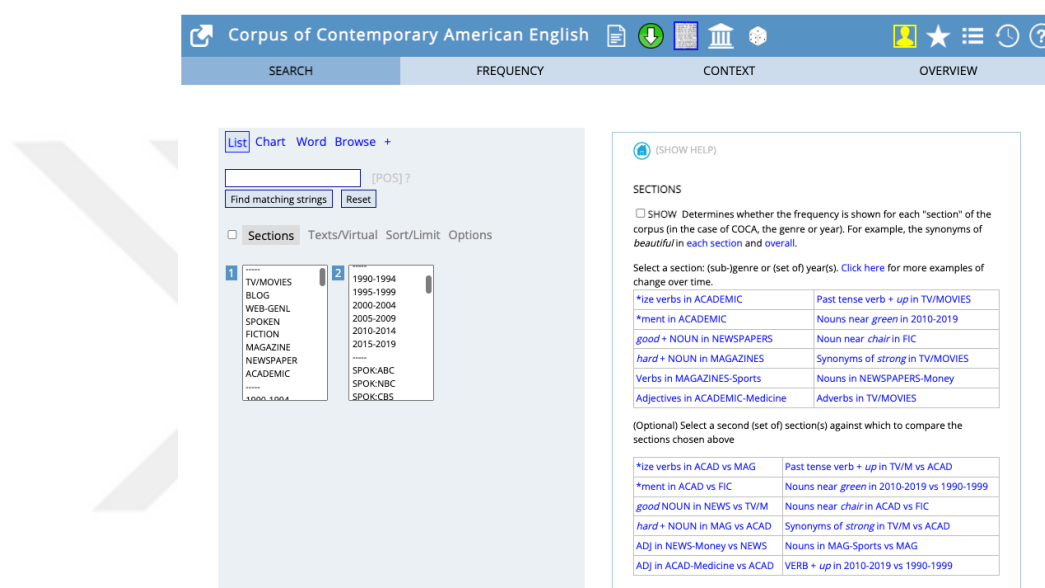


Figure 2.1. The interface of COCA

Alongside these online-accessible corpora, there exists a range of various computer software tools and applications (e.g., AntConc, LancsBox, Lextutor, Wordsmith tools, SketchEngine) for carrying out corpus analyses on the varieties in language use with pre-uploaded corpus data in text format (Özbay & Gürsoy, 2023; Şen Bartan, 2023). By utilizing these corpus software programs, researchers, language learners, and teachers can also create their own corpora for analysis rather than relying solely on ready-made online corpora (Godwin-Jones, 2017). In addition to their use in research endeavors for carrying out linguistic analysis through concordancing, such computer software has also been employed for language education purposes as corpus tools (Başal et al., 2024; Chen et al., 2019; Friginal, 2018; Kazaz, 2015; Köse & Kuru Gönen, 2023; Leńko-Szymańska, 2017; Poole, 2022). The common quality of both online corpora and corpus software is that each works as a concordancer for researchers, language learners, and teachers to

access each instance of language use across various contexts while facilitating the exploration and examination of language components within a corpus (Çalışkan, 2020; Schmidt, 2023; Şimşek, 2020).

Following the query of a word or a word group in a concordancer, users are provided with various corpus analysis features that unveil *authentic language evidence* (Bennett & Uí Dhonnchadha, 2023; Şimşek, 2020). By the process of searching for a specific word or a phrase in a corpus, concordancing, valuable insights into the query item can be obtained, including genre, collocations, phrasal verbs, idiomatic expressions, usage frequency, usage patterns, the use of modality and tenses, register and context (Farr & O’Keeffe, 2019; Ma et al., 2023). Typically, the result for a specific query is a display of concordance lines that show the distribution across texts in the form of Keyword in Context (KWIC). This form specifically presents the node, the searched target word or phrase, in the center of a number of lines where each occurrence of the target item in the desired genre, year, and text type is displayed with other words that surround it (Boulton & Cobb, 2017; Özbay & Gürsoy, 2023; Şimşek & Can, 2023). The effective analysis of concordance lines for making reliable deductions and inferences regarding the associations of a specific word or phrase with other words across different contexts involves employing particular techniques such as reading the concordance lines vertically and interpreting and identifying the usage patterns of the target item with other words on both sides (Bennett & Uí Dhonnchadha, 2023; Huang, 2018; Ma et al., 2023; McCarthy & O’Keeffe, 2010). Figure 2.2. below demonstrates some different features of COCA:

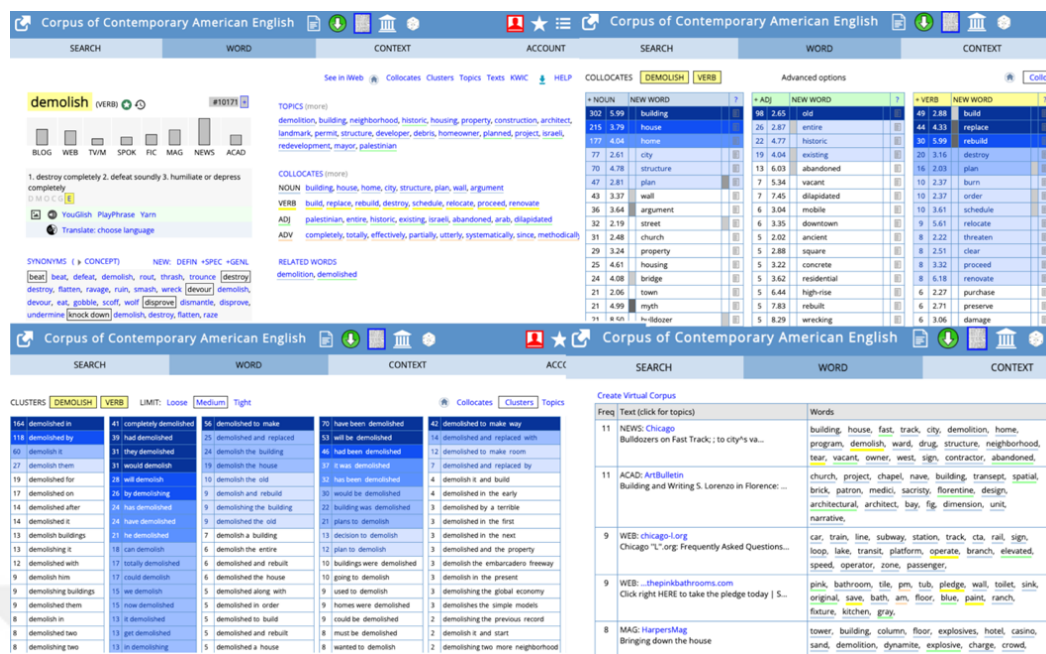


Figure 2.2. Some different features of COCA

With the emerging popularity of linguistic analysis of concordance lines and the pioneering work of Johns (1986) pointing out the benefits of the use of concordance lines with language learners, the consultation of corpora for English language teaching and learning purposes has become prominent (Chambers, 2019; O’Keeffe et al., 2007; Şimşek, 2020). Incorporating concordance lines into the English language pedagogy fostered the emergence of a new instructional approach by Johns (1991), the concept of data-driven language learning and teaching (DDL). Ensuring exposure of language learners to authentic language (Can, 2023; O’Keeffe et al., 2007) and discovery learning (Bernardini, 2004; Boulton & Cobb, 2017), this method involves learners examining concordance lines, recognizing and identifying language patterns and inferring grammar rules, and drawing insights autonomously at their own pace (Bennett & Uí Dhoonchadha, 2023; Boulton, 2010; Flowerdew, 2015; Johns, 1991). Data-driven language teaching activities and language learners’ engagement with discovery learning and linguistic analysis by working on concordance lines foster inductive learning (Boulton & Cobb, 2017; Michali & Patsala, 2020). As DDL activities are based on actual language data taken from concordance lines, they differ from ready-made learning activities that do not put learners at the center of language teaching and learning processes (Chambers, 2010; Flowerdew, 2015). With the integration of DDL activities into language pedagogy, language learners become more motivated to take charge of their own learning as it

requires them to analyze actual, naturally occurring language data themselves (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Chambers, 2005; Michali & Patsala, 2020; Şimşek, 2020). See Figure 2.3. below for sample concordance lines retrieved from COCA:

Corpus of Contemporary American English			
SEARCH	WORD	CONTEXT	ACCOUNT
WEBSITE		SORT	SORT
1 MAG:1999:TotalHealth	heart, liver and pancreas healthy. Add one to that	combination	and reduce an enlarged prostate. It has even been found
2 BLOG:2012:aperiodical.com	until I have a name. Some names are new	Combination	and Permutations was a name well before I ever had a show
3 ACAD:2013:AmjPubHealth	lives. Essentially, the roll-out of antiretroviral	combination	antiretroviral therapy in 1996 was a turning point: HIV
4 ACAD:1991:RoepersReview	regarding which identification methods and processes	combinations	are antiretroviral Related to this is the issue of ensuring
5 WEB:2012:housz.com	. I love how the windows in all their sizes and	combinations	are similarly shaped, pleasantly lengthening the house
6 NEWS:2003:Atlanta	economic and cultural issues, was supposed to be a new	combination	as he bought the Democratic presidential nomination. But his
7 ACAD:2002:MusicEduc	several rhythm patterns with which to compose as many new	combinations	as new When students have mastered fluent skills, they
8 FIC:2004:MassachRev	inside. I pumped my fist. I threw some new	combinations	at the air I picked up the phone and punched-up the
9 NEWS:2005:SanFranChron	infect birds. Out of this variety, a handful of	combinations	bearing names such has HSN1, HSN2, H7N7 and H7N3 regularly
10 NEWS:2005:NYTimes	apty for baseball -- before the Blue Jays broke up their	combination	by ending one of them to the minor leagues. # Twice
11 SPOK:2004:CBS_Morning	They call their collaboration the Miracle project , a	combination	coffee table book , CD and DVD collection. Ms-DION: It
12 NEWS:2014:Pittsburgh	clove. Depending on whether the chili and new peppercorn	combination	contains whole pulverized or ground ingredients, the heat can
13 MAG:1990:AmHeritage	a poor textbook, and frequently did. But that new project	combination	could exist independently of new math too. # The villain,
14 MAG:2010:TownCountry	like this one proliferate. On the surface, halau are	combination	dance troupes and hula schools, and the kumu hula is the
15 NEWS:2005:NYTimes	specific to this one man its effects are. # That	combination	drug resistance and rapid AIDS onset, the officials said, could
16 ACAD:1998:Raritan	a Chartist and also as an "individual." # That	combination	elicits new scrutiny from the stranger because it goes
17 MAG:2003:Fortune	You need extreme power and super-savvy course management . That	combination	enabled Jack Nicklaus to win six times-compared with five PGAs.
18 WEB:2012:colormatters.com	green create a harmonious design, regardless of whether this	combination	fits into technical formula for color harmony. #
19 NEWS:2011:NYTimes	to 17; free for 5 and under and members	Combination	food-and-admission tickets \$39; \$37 for 65+; \$32 for children
20 SPOK:2000:CNN_LiveSun	Warren Beatty getting the Thalberg Award, it is a new	combination	for Annette Bening , but she is not counting on anything.

Figure 2.3. Sample concordance lines retrieved from COCA

It is crucial to enable the proficiency and experience of language teachers in identifying and interpreting patterns in language data by corpus analysis on concordance lines, consulting corpora for enhancing their instruction, making corpus-based language pedagogy a part of their teaching practices by creating corpus-based language teaching materials (Bennett & Uí Dhonnchadha, 2023; Frankenberg-Garcia, 2012; Köse & Kuru Gönen, 2023; Leńko-Szymańska & Boulton, 2015; McCarthy et al., 2005; O’Keeffe et al., 2007). As Heather and Helt (2012) proposed, it is essential to enable the corpus literacy of language teachers to have the capacity to employ corpora in their language teaching to investigate the language and enrich their learners’ language development (Başal et al., 2024; Şimşek, 2020).

2.2. Consulting Corpora in Language Education

The natural integration of corpus use into English language education has profoundly influenced corpus-based language pedagogy and boosted the number of studies on utilizing corpora in language teaching and learning contexts (Bennett & Uí Dhonnchadha, 2023; Çalışkan, 2020; Godwin-Jones, 2017). Following the emergence of

DDL and the *corpus revolution* in English language practice, the integration of corpora into language education has increased significantly as the use of corpora has become a part of language classrooms with the extensive use of printed corpus-based textbooks, dictionaries, and grammar books (Chambers, 2019; Şimşek, 2020). Mukherjee (2006) further delineates the impact of such a revolution, focusing on the utilization of corpora both *for* and *in* English language teaching classrooms. Furthermore, Römer (2011) also provides an insightful exploration of corpus applications as corpora serve both *indirect* and *direct* means (Godwin-Jones, 2017; McEnery & Xiao, 2011; Sun & Hu, 2023). Figure 2.4. outlines the use of corpora in language learning and teaching by Römer (2011):

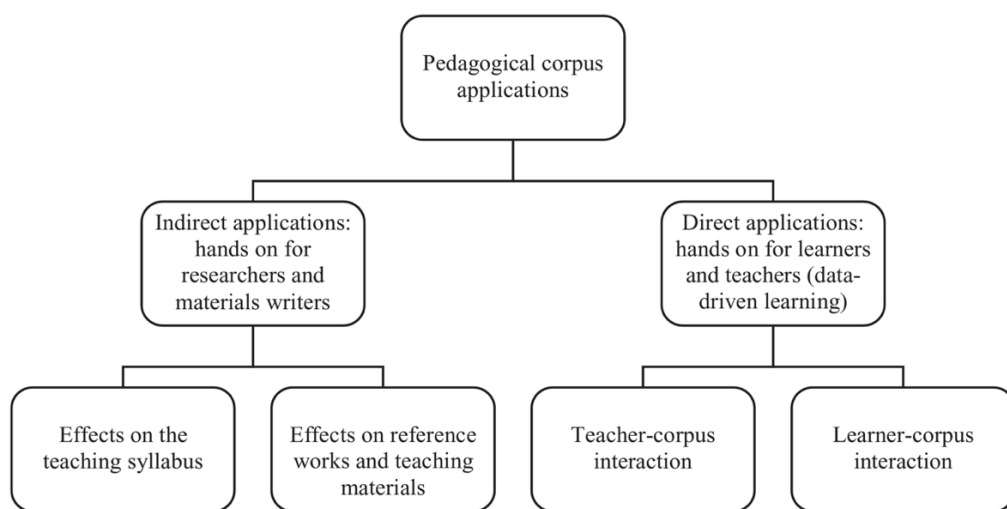


Figure 2.4. *The use of corpora in second language learning and teaching (Römer, 2011)*

As outlined by Römer (2011), the pedagogical applications of corpora in language learning and teaching contexts are delineated into two: the indirect applications of corpora by researchers and material writers for reference publishing, teaching and learning material development, syllabus design, and the direct applications of corpora by learners and teachers in language learning and teaching processes with concordance line analysis and DDL tasks (Çalışkan, 2020; Godwin-Jones, 2017; Sun & Hu, 2023).

2.2.1. Indirect corpus applications in language education

The indirect use of corpora in language education covers a vast number of applications, including corpus consultation and corpus-based language analysis, enabling

the creation and development of language teaching materials like textbooks and grammar books, reference publishing and dictionaries, syllabus design, and language test development and assessment (Kaya et al., 2022; McEnery & Xiao, 2011; Sun & Hu, 2023). In this respect, corpus-informed reference publishing and dictionaries are some of the most significant applications of corpora in language education (Cangır, 2022; Çalışkan, 2020; Godwin-Jones, 2017; Sinha, 2021; Şimşek, 2020). Following Sinclair's COBUILD monolingual learner's dictionary project in 1987, most of the dictionaries and grammar reference books published ever since have been corpus-driven (Breyer, 2009; Sinha, 2021). In addition to the COBUILD dictionary, many dictionaries, including the Cambridge Dictionary of American English (Landau, 2000), the Oxford Dictionary of English (Stevenson, 2010), the Longman Dictionary of Contemporary English (Summers & Gadsby, 1995) have been published with real-life, authentic language samples taken from corpora (Cangır, 2022; Godwin-Jones, 2017; Sinha, 2021). So much so that corpus-informed dictionaries are used by millions of people every day without their noticing that they are utilizing corpora indirectly (McEnery & Xiao, 2011; Sun & Hu, 2023).

Moreover, with the increasing number of published corpus-informed language education materials, many language learners and teachers use corpus-informed language teaching and learning materials, often without recognizing their corpus-based nature (Bennett & Uí Dhonnchadha, 2023; Godwin-Jones, 2017; Hunston, 2002; McEnery & Xiao, 2011; Sun & Hu, 2023). To ensure the authenticity of the materials and activities used, corpus-informed teaching materials highlight and represent the natural, descriptive aspect of the target language (Huang, 2018; Meunier & Reppen, 2015). As one of the most prominent and well-known corpus-based textbook series, the Touchstone textbook series by McCarthy et al. (2005) includes the usage of frequency of specific language items, as well as real-life language data and authentic language samples used in everyday language taken from the International Corpus of Cambridge (Cangır, 2022; Çalışkan, 2020; Godwin-Jones, 2017; Sun & Hu, 2023; Şimşek, 2020). See Figure 2.5. below for sample published corpus-informed dictionaries and textbooks as products of the indirect application of corpora for language teaching purposes:



Figure 2.5. Sample published corpus-informed dictionaries and textbooks

Numerous studies have emphasized the significance of the use of authentic and *use-based* language in language teaching and learning material development (Aşık, 2017; Bennett & Uí Dhonnchadha, 2023; Çalışkan & Kuru Gönen, 2018; Köse & Kuru Gönen, 2023; McCarthy, 2008; Römer, 2005), contrasting with conventional textbooks and language teaching syllabi, which have been often criticized for being influenced by subjective intuitions and judgments of material developers (Meunier & Reppen, 2015; Schmidt, 2023; Şimşek, 2020). Hunston (2002) also stressed that teaching syllabi for language education should also be informed by empirical evidence and usage frequency of the actual language with a lexical orientation rather than intuition and tradition (Bennett & Uí Dhonnchadha, 2023; Mindt, 1996). Given the growing recognition of corpus-based language teaching materials as authentic representations of the language and their role in fostering a research mindset for language educators, there arises a need for a shift in the role of language teachers in material development, adaptation, and evaluation to align with the evolving landscape of corpus-based language pedagogy (Başal et al., 2024; Çalışkan & Kuru Gönen, 2018; Meunier & Reppen, 2015; Şimşek, 2020). In addition to material development, reference publishing, and syllabus design, the exploitation of the indirect applications of corpora extends to language testing endeavors during the development and assessment phases of language tests (Cushing, 2017; Đurović & Bauk, 2022; Götz, 2022; Kaya et al., 2022; McEnery & Xiao, 2011; Park, 2014). The indirect application of corpora, as elucidated by Alderson (1996), plays a multifaceted

role in language testing as it covers activities such as test creation and development, content selection, and compilation (Kaya et al., 2022; McEnery & Xiao, 2011).

Despite the benefits of consulting corpora for creating language teaching materials for language instruction through the indirect applications of corpora, the number of published corpus-informed language teaching materials in the recently published textbooks remains limited, as well the lack of published or digital guidelines and handbooks to enable the *know-how* of corpus consultation for EFL teachers and learners (Đurović & Bauk, 2022; Farr & O’Keeffe, 2019; Gilmore, 2019; Götz, 2022). This calls for the need to create and enrich such corpus-informed language teaching materials for EFL contexts for language instruction and prepare booklets and handbooks for teachers to utilize corpora in their teaching through further research (Başal et al., 2024; Boulton & Cobb, 2017; Köse & Kuru Gönen, 2023).

2.2.2. Direct corpus applications in language education

The direct applications of corpora in language education refer to the hands-on use of corpora by language learners and teachers for practicing the naturally occurring, actual language data in concordance lines in language teaching and learning environments (Abdel Latif, 2021; Godwin-Jones, 2017; McEnery & Xiao, 2011; Römer, 2011; Sinha, 2021). The practice of the target language with corpus-based activities and the implementation of authentic language into actual language classrooms involve the hands-on use of concordancers either by teachers or learners, and the output taken from concordance lines as a result of corpus analysis in the form of language learning and teaching materials (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Boulton, 2010; Chambers, 2010; McEnery & Xiao, 2011). Ensuring the hands-on interaction between corpora and language learners and teachers, the direct applications of corpora in language teaching and learning pedagogy foster discovery learning for language learners and the facilitator role of language teachers (Bennett & Uí Dhonnchadha, 2023; Bernardini, 2004; Boulton & Cobb, 2017; Köse & Kuru Gönen, 2023; Şimşek, 2020).

Language learners’ hands-on practice with online-accessible corpora and corpus software themselves is regarded as the *hard version* of the direct application of corpora in language education as it requires them to autonomously access corpus data on concordancers for running a query for a specific word or phrase and analyze the keyword in context (KWIC) in the form of concordance lines (Alsahafi, 2022; Boulton, 2010;

Conza-Armijos & Celi-Celi, 2023; Gabrielatos, 2005; Timmis, 2015). As Gabrielatos put forward (2005), on the other hand, the *soft version* of the direct application of corpora in language education involves language learners' engagement with corpus-informed activities designed by language teachers with pre-selected concordance lines output to practice the language (Alsahafi, 2022; Chambers, 2010; Conza-Armijos & Celi-Celi, 2023; Hunston, 2002; McEnery & Xiao, 2011). Employing a learner-centered approach in language teaching, the use of such data-driven learning and teaching activities encourages the integration of authentic language samples into language teaching materials and the inductive learning and discovery process of patterns and rules in a language by learners (Bennett & Uí Dhonnchadha, 2023; Bernardini, 2004; Boulton & Cobb, 2017; Flowerdew, 2015; Köse & Kuru Gönen, 2023; Şimşek, 2020). Figure 2.6. illustrates a sample corpus-informed grammar teaching activity by Girgin (2011):

Concordance Lines (Passive Voice)	Leading Questions (Passive Voice)
1 bed is known as a St Ursula bed. The design <u>was inspired</u> by the Carpaccio picture of "The Dream 2. ausescu fell. He and a few of his colleagues <u>were inspired</u> by the early struggles of Solidarity in P 3. nducting a wedding ceremony, but nothing <u>was stolen</u>. Father Maire was said to have received 4. was involved in a fighting incident, a letter <u>was sent</u> by the principal to the parents describing 5. from which so few were to return. Kemper <u>was being shot</u> down; Armistead was falling as he 6 pt watching me coolly. I had the sense that I <u>was being watched</u> by two different people at 7 "DHSS should ensure that all general managers <u>are made</u> aware of all the possibilities for economy 8 into the shell lip with a hacksaw. The groove <u>is repaired</u> but the "scar" remains. It 9 with a "salvage" title from Georgia. The car <u>is sold</u> to SOM Auto Broker, which resells it to 10. Is of the mantle are damaged. Damaged shells <u>are repaired</u>, but the later- 11. Astronomer Carl Sagan may not believe we <u>are being visited</u> by aliens. But at the same time the 12. lable and unfilled whilst social support <u>is being paid</u> to an immobile work force. Wherever possible 13. you will see progress. By the roads that <u>are being built</u>, including one that will reach 14. old, and they just cannot maintain it. It <u>has been painted</u> by many great artists, including 15. porter waved a sign declaring: "Our votes <u>have been stolen</u>." # The Democrats have been using 16. in cash upon completion and the balance <u>will be paid</u> in the form of a loan note redeemable on 1st 17. YITZHAK SHAMIR: If such a decision <u>will be taken</u> by anybody, any foreign body, 18. item. If the growing lip is damaged it <u>will be repaired</u>, but the scar remains. As the animal grow 19. tr and producer met for lunch. The film <u>will be shot</u> in Blackpool, the director's 20. donow that ultimately the history of Russia <u>will be written</u> by Russians and the future of 21. the street? These are the questions which <u>must be answered</u> by American politicians, not to divert 22. when or how to wear the garment. This shirt <u>must be worn</u> as an overshirt,' reads t 23. Yeah. I think there are two things that <u>have to be said</u>. First of all, as people are saying, it's more 24. he cost of operating a business: Taxes <u>have to be paid</u>, and just as fuel and power bills are borne,	Task 1 : Analyze the concordance lines and answer the questions. 1. How do you form the passive in English? _____ _____ 2. By analyzing the concordance lines, write One example of the past passive: _____ One example of the present passive: _____ One example of the present perfect passive: _____ One example of the future passive: _____ One example of the present continuous passive: _____ One example of the past continuous passive: _____ 3. After examining the concordance lines for 'the passive', what can you say about the reason why the passive is used? _____ _____ 4. if you want to say who did the action, we use _____

Figure 2.6. A sample corpus-informed grammar teaching activity (Girgin, 2011)

Language learners play an active role in their own learning process while analyzing corpora output and concordance lines, in line with the noticing hypothesis (Schmidt, 1990), claiming that input turns into intake, meaning it is internalized and processed by the learner, and language patterns and forms are acquired by learners only when they are noticed (Boulton & Cobb, 2017; Flowerdew, 2015; Oveidi et al., 2022; Yılmaz & Koban Koç, 2020). This hypothesis suggests that mere exposure to language is not sufficient for language acquisition as learners must actively engage with and pay attention to language features for learning to occur (Boulton & Cobb, 2017; Oveidi et al., 2022; Schmidt, 1990). This active engagement is facilitated through the analysis of concordance lines by consulting corpora thanks to access to authentic examples of

language use in context and the identification of recurring patterns, grammatical structures, and vocabulary usage, thereby enhancing language awareness and understanding (Godwin-Jones, 2017; Timmis, 2015; Yılmaz & Koban Koç, 2020; Zareva, 2017).

Unlike more teacher-centered language teaching and learning methods, DDL is believed to enrich language education by enhancing learner autonomy, authenticity, and language awareness with a vast amount of naturally occurring language data taken from concordance lines for learners to observe the actual use of language in various natural contexts instead of intuitively created language learning materials (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Çalışkan, 2020; Farr, 2008; Gilquin & Granger, 2010; Zareva, 2017). As DDL utilizes a vast amount of naturally occurring language data drawn from concordance lines, the use of corpora in language classrooms allows learners to observe the actual use of language in various real-world contexts (Başal et al., 2024; Braun, 2005; Köse & Kuru Gönen, 2023). The active engagement with authentic language data through concordancing fosters the autonomy of the language learners as they can individually work on language data to identify and recognize language patterns (Bennett & Uí Dhonnchadha, 2023; Charles, 2022). By working on corpora output and corpus-informed materials for identifying language patterns and practicing the language, learners take an active role and responsibility as well as becoming more motivated for their own language-learning journey (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Çalışkan, 2020; Godwin-Jones, 2017; Lee, 2011; Timmis, 2015).

2.2.3. Using corpora for language skill development

DDL also stands out from traditional, teacher-centered methods in language teaching and learning contexts with its significant benefits in language instruction on language skills (Bennett & Uí Dhonnchadha, 2023; Boulton & Cobb, 2017; Çalışkan, 2020; Gilquin & Granger, 2010; Godwin-Jones, 2017; Lee et al., 2019; Mizumoto & Chujo, 2015; Sinha, 2021; Timmis, 2015; Xodabande & Nazari, 2022). In recent years, there has been a growing momentum in the number of studies on the incorporation of corpora and DDL into vocabulary instruction and learning as it provides language learners with the real-life usage of lexical items in various contexts through concordance lines (Alsaifi, 2022; Barabadi & Khajavi, 2017; Can, 2023; Chen, 2017; Daskalovska, 2015;

Eslek-Onur & Tosun, 2023; Karras, 2016; Koçak, 2020; Li, 2017; Lu & Dang, 2022; Saddhono et al., 2023; Sinha, 2021; Tsai, 2019; Uçar & Yükselir, 2015; Vyatkina, 2016; Zhang, 2022). Similarly, empirical studies put forward the efficiency of adopting a corpus-based method for vocabulary teaching as it enhances the acquisition of collocations, idiomatic expressions, and phrases in language classrooms (Aşık et al., 2016; Çalışkan & Kuru Gönen, 2018; Godwin-Jones, 2017; Greaves & Warren, 2010; Köse & Kuru Gönen, 2023; Li, 2017; Moon, 2010; Tsai, 2019; Uçar & Yükselir, 2015).

A research endeavor by Daskalovska (2015) found that utilizing BNC to employ corpus-based exercises resulted in the enhanced acquisition of verb-adverb collocations among university students and led to more significant gains compared to traditional textbook activities. In another study, Vyatkina (2016) explored the efficacy of DDL materials versus conventional teaching methods for teaching German verb-preposition word combinations to university students with low-level proficiency. The study revealed that corpus-based vocabulary instruction led to better learning outcomes even at lower proficiency levels. Barabadi and Khajavi (2017) also examined the impact of corpus-based vocabulary teaching activities on learning lexical items like synonyms and collocations in Iranian EFL settings. Their findings revealed that such DDL activities for vocabulary instruction outperformed traditional methods like consulting a dictionary or grammar books as the use of corpora offers real-life, authentic language data.

Tsai (2019) studied how the combination of corpora and dictionary use affects vocabulary learning in the Taiwanese context. Two classes of 50 students were formed, focusing on inductive vocabulary instruction by consulting corpora first and dictionaries second for word meaning in one class and deductive vocabulary instruction in the other by the reverse order of using dictionaries first and corpora later. It was revealed that the combination of corpora and dictionaries enhanced vocabulary learning for both approaches, while the inductive vocabulary instruction class students showed better learning results in collocational knowledge. In the Chinese context, Zhang (2022) investigated the efficiency of the use of corpora through corpus-informed vocabulary teaching activities via COCA and BNC for developing morphological knowledge with online corpora. Conducted as an action research with 31 undergraduate students majoring in English in China, this study revealed that the participants' vocabulary knowledge, morphological knowledge, and self-regulated learning motivation increased through the use of corpora for vocabulary teaching. The findings pointed out that a learning

community was established while the participants were learning vocabulary through corpus-based vocabulary teaching activities and increasing their corpus literacy. In another study conducted by Sinha (2021) in Bangladesh, the perceptions and attitudes of 32 EFL learners toward the integration of corpora into vocabulary instruction by corpus-based vocabulary teaching activities were explored by administering a questionnaire. It was revealed that the participants favored the use of COCA and corpus-informed vocabulary instruction as they found the in-class application of corpus tools positive for vocabulary learning. The participants stressed that there is a need for guidance from the instructors for EFL learners with a lack of familiarity and experience with COCA.

In the Turkish EFL context, Uçar and Yükselir (2015) explored the impact of integrating COCA-based activities on students' acquisition of verb-noun collections, and their findings indicated superior outcomes in vocabulary learning and collocation recognition compared to conventional methods, revealing substantial differences between the experimental and control groups. In another study by Kayaoğlu (2013) involving preparatory program students with intermediate-level language proficiency, it was concluded that employing corpora enhances the learning of synonyms and the word choice of language learners. Similarly, Aşık, Vural, and Akpınar (2016) also advocated for the integration of corpora into language instruction on vocabulary, highlighting that it enhances learners' awareness and recognition of collocations and synonyms. Çalışkan and Kuru Gönen (2018) conducted a study to explore the perspectives of three instructors on using corpora in vocabulary teaching. Following the four-week training on the consultation of corpora in the EFL context, along with the incorporation of corpus-informed language teaching materials and activities into EFL classrooms, it was revealed that the training increased the awareness of instructors on corpora and its use in language instruction. Despite the technical issues and time constraints experienced, the instructors perceived corpus consultation as an effective way to teach vocabulary to high-level proficiency learners and teach prepositions and collocations.

In another study by Koçak (2020), the perceptions of EFL students at a state university in Türkiye were investigated regarding the integration of COCA into vocabulary instruction. While the participants spent three hours with the use of COCA to develop their vocabulary knowledge and completed several corpus-informed vocabulary exercises for four weeks, their perceptions were obtained by the administration of a questionnaire by the author. The findings revealed that the participants held positive

perceptions regarding the implementation of COCA and corpus-based vocabulary teaching activities. However, the need for training language learners and teachers to effectively use corpus tools and online corpora to develop their vocabulary knowledge was called for. Despite their prior lack of familiarity and experience with corpora, they found the use of corpora valuable for lexical development. Similarly, Can (2023) investigated the efficiency of the use of corpus-informed teaching activities and questioned if the use of corpora develops the vocabulary knowledge of advanced EFL learners more than traditional vocabulary teaching with conventional materials following the principles of action research at a state university in Türkiye. While corpus-informed vocabulary teaching materials were used in one of the C1 proficiency level EFL classes, the students in the other C1 class received vocabulary instruction through conventional materials. The findings revealed that there was not a statistical difference regarding their vocabulary development between the two groups. It was revealed that despite the positive attitudes of the participants and the perceived benefits of corpora, such as access to authentic language usage and developing the language awareness of language learners, the use of corpora was not practical according to the participants as they found using corpora time-consuming and the high number of language data hard to analyze.

According to the relevant research on the integration of corpora and DDL activities into vocabulary instruction and learning, it was revealed the use of corpus-informed language teaching materials for vocabulary instruction outperformed traditional methods with using conventional materials, consulting a dictionary or grammar books as the use of corpora offers real-life, authentic language data and enhances vocabulary development through better retention and lexical knowledge. Despite the benefits of the integration of corpora into vocabulary instruction, it was pointed out that corpora consultation for vocabulary teaching might be time-consuming due to difficulties in selecting suitable language data from a corpus and technical difficulties.

In addition to the integration of corpora into vocabulary instruction, studies examining its use for grammar teaching also highlight benefits that stem from the ability of corpora and DDL to offer authentic examples of grammar usage across different contexts via concordance lines and corpus-informed grammar teaching activities for learners (Girgin, 2011, 2019; Emir & Yangın-Ekşi, 2023; Hughes, 2010; Khan et al., 2021; Lin & Lee, 2019; Lakew et al., 2021; Lin, 2019; Özer & Özbay, 2022; Rodríguez-Fuentes & Swatek, 2022; Römer, 2004; Smart, 2014; Vannestal & Lindquist, 2007; Zhong

& Wakat, 2023). The potential benefits of corpora use for teaching particular grammar structures, including conjunctions, modal verbs, prepositions, gerunds, and infinitives, were emphasized in the relevant research (Girgin, 2011; Lin & Lee, 2019; Lin, 2019; Smart, 2014). Grammar instruction can be enhanced by utilizing corpus-derived insights such as usage frequency of grammar structures while creating grammar teaching materials for the teaching of language structures in language classrooms for learners to identify language patterns and infer grammar rules (Girgin, 2019; Khan et al., 2021; Rodríguez-Fuentes & Swatek, 2022; Römer, 2004).

In a study, Lin (2019) conducted a case study investigating an EFL teacher's use of corpus-based grammar activities in teaching patterns of *used to* despite challenges like the time-consuming nature of material creation by using COCA, the teacher responded positively to the use of corpora for teaching grammar, noting increased classroom interaction and learner-centeredness. In the Taiwanese context, Lin and Lee (2019) examined a new teaching approach of employing three grammar groups of one traditional deductive and two blends of deductive and DDL approaches with university students by utilizing corpus-informed grammar teaching activities for teaching relative clauses, passive voice and purpose, cause and result phrases. While learners developed positive perceptions towards using corpus-informed activities and using DDL in a learner-centered manner for grammar learning, the results revealed no significant learning difference among groups. In another study in the Ethiopian context, Lakew et al. (2021) investigated the perceptions of 31 EFL learners and 1 instructor regarding the use of corpus-informed grammar instruction for speaking skills lessons. The study aimed to explore the effectiveness of corpora integration for the grammar development of EFL learners over the six-week intervention by weekly consecutive tests. It was revealed that the participants showed significant achievements in terms of grammar development, and they favored the use of corpus-informed grammar activities throughout the intervention. The need for more training and practice was called for by the participants. In order to compare the effectiveness of corpora integration into grammar instruction via corpus-based homework materials and traditional, conventional homework activities, Rodríguez-Fuentes and Swatek (2022) assigned two different types of homework materials randomly to a group of EFL learners in the Colombian context. Pre-, post-, and delayed post-tests revealed that the group of EFL learners with the corpus-informed homework materials outperformed the group of learners with the conventional homework materials. It was

pointed out that the use of corpus-based teaching materials and activities enhances the grammar performance and development of EFL learners as supplementary teaching materials.

In the Turkish context, Girgin (2011) investigated the impact of corpus-based grammar teaching activities on grammar learning of Turkish EFL university students with low-level proficiency in English. While learners' attitudes and perceptions towards the use of corpora were neutral, no significant improvement was observed in terms of learning grammar structures like question tags, relative clauses, and indirect questions compared to conventional textbook methods. In another study by Girgin (2019), the views and attitudes of 55 EFL learners regarding the use of corpora for learning phrasal-prepositional verbs were investigated by the administration of a questionnaire and focus-group interviews. It was revealed the participants held positive attitudes toward the use of corpora and favored corpora consultation for learning phrasal-prepositional verbs. Similarly, Özer and Özbay (2022) aimed to explore the effectiveness of DDL-based corpora use on grammar instruction in the EAP context at a preparatory school in Türkiye. A total of 75 EFL students were taught grammar by the use of a specific corpus compiled for the purposes of the study as well as DDL materials over 7 weeks in two consecutive semesters. The findings showed that the participants showed increased grammar achievements through the use of DDL and corpus-based teaching materials. It was also emphasized that the use of DDL was autonomy-enhancing for the EFL students.

Despite claims regarding the efficiency of corpora integration into grammar instruction as it promotes inductive and autonomous learning of grammar structures, its adoption in language classrooms remains limited due to factors related to teacher preferences, learning styles of language learners, and concerns about curriculum constraints (Emir & Yangın-Ekşi, 2023; Khan et al., 2021; Lin & Lee, 2019; Lakew et al., 2021; Lin, 2019; Rodríguez-Fuentes & Swatek, 2022; Zhong & Wakat, 2023). The scarcity of research on corpora integration into grammar teaching underscores the need for more studies to explore the effect of corpus use and understand the influence of challenges and limitations in language classrooms on grammar instruction. Drawing insights from the research on corpora integration into grammar teaching, it is also essential to further investigate EFL learners' and teachers' attitudes and perspectives toward corpus-based grammar teaching.

Empirical research on the integration of corpora into language education has also highlighted the value of utilizing corpora for error correction in writing instruction (Anthony, 2016; Chang, 2014; Emir, 2019; Gaskell & Cobb, 2004; Gilmore, 2008; Luo & Liao, 2015; O'Sullivan & Chambers, 2006; Quinn, 2015; Yoon & Hirvela, 2004; Yoon, 2008). The research studies in the literature focused on the impact of the use of corpora for the feedback process of error correction in the teaching of writing as teachers utilized the use of corpora by enabling learners to work on their writing by consulting corpora by themselves (Gaskell & Cobb, 2004). As language learners identify and rectify their mistakes through corpus analysis of authentic language samples, the use of corpora for writing and error correction enhances learners' autonomy and active participation and lightens the workload of language teachers (Gilmore, 2008; O'Sullivan & Chambers, 2006).

Yoon and Hirvela (2004) conducted a study incorporating corpora into academic writing courses for intermediate and advanced language learners, and the results revealed improved writing skills, increased confidence, and positive perceptions towards the use of corpora as it was found helpful in accessing word usage and collocations by the learners. In another study by Luo and Liao (2015) in the Chinese EFL context, 30 low-level proficiency university students highlighted the significant impact of corpora on improving their writing skills. The participant students utilized corpora to make corrections and reduce language mistakes and wrong word usage. Anthony (2016) also explored the integration of corpora and DDL in technical writing classes and its potential for evaluating written work by emphasizing its value as a tool for teachers to offer students tailored corpus data. Studies investigating learners' views on the integration of corpora for error correction and writing revealed positive perceptions and results for helping learners correct their language. In the Turkish EFL context, Emir (2019) set out to explore the impact of corpus consultation for writing and the perceptions of learners regarding the integration of corpora during essay writing. The results indicated enhanced writing proficiency among those who used corpora, effectively addressing their vocabulary and grammar errors. Bal-Gezegin et al. (2022) conducted research involving pre-service EFL teachers, and the participants noted that utilizing corpora while providing corrective feedback on students' writing fostered their confidence and enthusiasm as they benefitted from corpus consultation.

However, several studies also highlighted challenges in using corpora due to the vast amount of data to be analyzed in corpora, the language proficiency of learners, learners' lack of experience and skills related to computer and corpus use, and technical problems related to concordancers (Breyer, 2009; Çalışkan & Kuru Gönen, 2018; Chujo et al., 2015; Farr, 2008; Flowerdew, 2009; Lee, 2011; Luo & Liao, 2015; Varley, 2009; Yoon & Hirvela, 2004). Especially in error correction and writing, Chang (2014) inferred that learners with low-level proficiency in the target language were less inclined to employ corpus tools for error correction. Similarly, Quinn (2015) noted that some students encountered challenges while consulting corpora due to the enormous amount of data that is hard to navigate despite the efficacy of corpora for error correction as a reference tool. In order to overcome such challenges and concerns, language teachers need systematic and comprehensive training programs on theoretical and practical aspects of corpus software and online corpora on how to effectively analyze and interpret concordance lines and data (Frankenberg-Garcia, 2012; Lee, 2011), as well as finding out and practicing how to benefit from corpora in language teaching with hands-on corpora use for creating data-driven, authentic language teaching materials and actual teaching practice with the use of DDL materials (Çalışkan, 2020; Leńko-Szymańska, 2017).

2.3. Corpora in Language Teacher Education

Another key indirect application of corpora in language education contexts is the effort to equip foreign language teachers with the ability to utilize corpora for language teaching purposes by training them through pre-service and in-service courses and training as teachers play a crucial role in imparting corpus practice and access to authentic language to language learners (Breyer, 2009; Römer, 2009). It is of utmost importance to make language teachers familiar with what a corpus is, how to benefit from corpora for designing teaching materials, enhancing language skill development, and using corpus tools for error correction as well as test creation and development (Başal et al., 2024; Breyer, 2009; McEnery & Xiao, 2011; Park, 2014; Römer, 2009).

In his discussion of corpus linguistics, Fligelstone (1993) identified three primary uses of corpora in language teaching contexts as a scaffolder for teachers and language learners: *teaching* learners and teachers *about corpora* and corpus linguistics, *teaching* learners *to exploit corpora*, and *exploiting corpora to teach* with corpus-informed teaching resources. In the context of language teacher education, it is essential to teach

language teachers about corpora and train them on how to make use of corpora and work on language data for corpus analysis through hands-on experience to be able to equip them with pedagogical and technical skills of corpora use with a degree of *corpus literacy* to utilize corpus in their teaching practices later on (Breyer, 2009; Callies, 2019; Gavioli, 2005; Heather & Helt, 2012; Hunston, 2002; Mukherjee, 2006; Tribble, 2000). In addition to the multiliteracies in today's language education world, such as technology skills and digital competency for teachers to raise the language awareness of their learners and enhance self-confidence (Barnes & Murray, 1999; Breyer, 2009; Farr, 2008; Tammelin, 2001; Warschauer, 2000), corpus literacy is another type of literacy required for language teachers as the ability to utilize corpora in language teaching to explore the language and promote their learners' language development (Heather & Helt, 2012). In order to gain this literacy, teachers must be aware of what a corpus provides and the *know-how* of corpora by gaining experience and knowledge in language analysis by concordancing (Johns, 1986; Mukherjee, 2004). Additionally, research studies in relevant literature emphasize the importance of training language teachers on how to consult corpora for creating authentic corpus-informed materials by the analysis of corpus data and integrating them into language classrooms meaningfully to meet the needs of their students as a vital aspect of their teacher development (Başal et al., 2024; Breyer, 2009; Farr, 2010; Farr & O'Keeffe, 2019; Gilmore, 2019).

Despite being one of the most prominent issues in the relevant research with an emerging interest, the use of various corpus tools in EFL contexts as well as the corpus literacy and experience of EFL teachers on corpus-based language teaching pedagogy remains limited (Başal et al., 2024; Boulton, 2009; Callies, 2019; Chambers, 2019; Ebrahimi & Faghih, 2017; Leńko-Szymańska & Boulton, 2015; Ma et al., 2022; Mukherjee, 2004; Römer, 2009). It is stressed by researchers that language teachers' awareness of corpora and necessary knowledge on how to utilize corpora is one of the most significant reasons for such little use of corpora in language classrooms (Breyer, 2009; Frankenberg-Garcia, 2012; Gilquin & Granger, 2010; Naismith, 2017; Tribble, 2015). According to Gilquin and Granger (2010), the extensive adoption and integration of corpora into language instruction remain uncommon in everyday language teaching practices as this lack of awareness and knowledge of language teachers on available corpus software and online corpora and how to apply them in their language instruction by either teaching learners to exploit corpora or exploiting corpora to teach with corpus-

based teaching materials (Bennett & Uí Dhonnchadha, 2023; Boulton, 2009; Breyer, 2009; Fligelstone, 1993; Poole, 2022).

Bunting (2013) surveyed EFL teachers to assess their familiarity with corpus tools, and it was revealed that most of the 116 EFL teachers lacked knowledge of how to integrate corpora into their teaching despite their familiarity with corpus linguistics from their university courses. Participants expressed interest in learning to use corpora and corpus tools if provided with training opportunities. Aşık (2015) surveyed 42 EFL instructors in the Turkish EFL context and discovered that nearly half the instructors lacked familiarity with corpora, with only a few having prior experience in using corpora. A recent investigation by Pérez-Paredes (2022) revealed that teachers still do not utilize corpora in their classrooms. These results underscored the importance of training language teachers in the pedagogical applications of corpora by devoting more training and courses on the *know-how* of corpus consultation (Bennett & Uí Dhonnchadha, 2023; Mukherjee, 2004; Römer, 2011). Based on the results of relevant literature, it is evident that language teachers must be exposed to corpora and need instruction on integrating its use into their teaching practices through concordancing (Leńko-Szymańska, 2014; Lin, 2019). Several researchers argue for the inclusion of corpora into both initial teacher training for prospective pre-service EFL teachers and the professional development of in-service language teachers (Chambers et al., 2011; Chen et al., 2019; Çalışkan & Kuru Gönen, 2018; Frankenberg-Garcia, 2012; Lin, 2019; McCarthy, 2008; Mukherjee, 2004; O’Keeffe & Farr, 2003; Römer, 2011; Şimşek & Can, 2023). Despite the emphasis on the significant importance of educating teachers about corpora use in language education and exploring their attitudes and perceptions toward corpus utilization in the literature (Breyer, 2009; Chen et al., 2019; Lin, 2019; Zareva, 2017), the number of empirical studies on the topic remains limited.

In the context of in-service EFL teacher training on corpora use, Bunting (2013) explored the corpus use of three EAP writing course instructors at a U.S. university by individualized training to suit the needs of each instructor and lesson planning was done collaboratively to integrate corpora into their teaching practices. Qualitative data from interviews and stimulated sessions revealed diverse teacher perspectives influenced by their prior knowledge and experiences with corpus consultation and corpus exposure throughout the study. The findings stressed the necessity for in-service training to explore corpora consultation of the instructors further. Similarly, Chen et al. (2019) organized a

three-hour in-service teacher training workshop for instructors working in Hong Kong on the integration of corpora into academic writing classes. After surveying 54 in-service teachers, the findings showed that instructors favored consulting corpora and the training. The study identified four factors that affect the perceptions of teachers toward corpora use: prior experience with corpora, willingness, and enthusiasm for teaching, motivation for professional development, and teaching experience. In another study by Bennett and Uí Dhonnchadha (2023), a model for training in-service EFL teachers was presented to develop the corpus literacy of five EFL instructors over five weeks. While the instructors expressed a notably positive experience in using corpora, they also stressed that corpus use required less time for lesson and activity prep time than conventional methods and materials.

Investigating the attitudes and views of 6 instructors regarding corpus use in language instruction in the Turkish EFL context, Özbay and Kayaoğlu's study (2015) revealed that the corpus training effectively enhanced instructors' awareness of available corpora and their use for English teaching purposes. Following the eight-week training on corpus consultation in language teaching via several corpus tools, In addition to the potential of corpora for verifying language usage accurately and accessing authentic language data, the instructors highlighted corpora's suitability for advanced-level language learners, especially for writing development. In another study in the Turkish EFL context, Çalışkan (2020) examined the familiarity of 10 EFL university instructors with corpora and their views on integrating corpora in language instruction over a seven-week training program. While the initial data from 89 instructors revealed low familiarity with corpora, following the training, the instructors enhanced their knowledge, awareness, and skills related to corpora consultation in the EFL context. In spite of the perceived benefits of consulting available corpora for language instruction, such as improved error correction and vocabulary instruction, the participants stressed several challenges, such as difficulties in designing materials and using corpora with low-level proficiency students. In a more recent study with 29 in-service English teachers, Başal et al. (2024) investigated the effects of a three-day corpus literacy course and the perceptions of the participant language teachers toward corpora. The study's findings revealed that even a three-day training on corpus use can positively influence the abilities of language teachers and enhance their awareness of corpus consultation.

In addition to such training for in-service EFL teachers, researchers have stressed the need for offering courses and training for pre-service EFL teachers to familiarize them with corpora, features of corpus data and tools, and the use of corpora for concordance analysis and designing language teaching materials before their graduation as such training is essential and *best-situated* in initial language teacher education (Breyer, 2009; McCarthy, 2008; Mukherjee, 2006; O’Keeffe & Farr, 2003). Along with such training, the investigation of pre-service language teachers’ perceptions and views on the integration of corpora into language instruction is highlighted (Chen et al., 2019; Lin, 2019). In 2009, Breyer conducted a case study with pre-service English teachers in the form of an eleven-week course as a part of their language teacher education. During the training on corpus consultation for language teaching, the 18 student-teachers progressed to both theoretical and practical applications of corpora while creating corpus-informed teaching materials via hands-on experience. The positive impact of corpora on their awareness of corpora’s potential in language instruction was highlighted. The study was predominantly theoretical in nature, as the pre-service teachers were not given the opportunity to implement the activities they designed in a real teaching environment. In another study with 25 student-teachers, Farr (2008) explored the views and perspectives of English teacher candidates in relation to the utilization of corpus tools for language teaching. Integrated into an ELT MA program, the training included exploring the language system through corpus-informed tasks and writing a research paper based on corpus analysis and investigation. Despite challenges like time constraints and technological issues experienced, participant pre-service EFL teachers viewed corpora positively with access to authentic and naturally occurring language samples in various contexts, enhancing their language awareness. Overall, they expressed enthusiasm for incorporating corpora into their future teaching and further training on corpus use.

In the Polish context, Leńko-Szymańska (2014) surveyed the students of an MA course on the basics of corpora and corpus tools, corpus integration into language instruction, and compiling a corpus and its analysis. Participants of the study showed positive attitudes towards corpus use, emphasizing its benefits for teaching lexical items. The study underscores the importance of acquiring knowledge on corpus consultation for pre-service EFL teachers. In another study by Leńko-Szymańska (2017), the impact of a course on corpus use was assessed. The course engaged participants with various activities on corpus use, and the pre-service EFL teachers created a small ESP corpus as

the course's final project. Analysis of the 53 projects over five years revealed that the course provided basic skills in corpus analysis. However, more exposure to corpus applications in language pedagogy was recommended to promote the implementation of corpus-informed teaching materials and activities in the EFL context. Naismith (2017) focused on the investigation regarding the impact of the integration of corpus tools into the Certificate in Teaching English to Speakers of Other Languages (CELTA) courses for initial teacher training and the perceptions of trainees and trainers on corpora use for language teaching revealed that the integration of corpus tools into CELTA courses would be beneficial for enhancing language awareness and lesson planning. However, it was noted that extending the use of corpus-based teaching materials in language instruction was unrealistic as the participant trainees utilized corpus tools significantly less in their practice compared to their lesson planning.

Abdel Latif (2021) studies Arab EFL student teachers' perceptions and usage of corpus tools. Despite the recognition of the benefits of corpora in English language teaching, only a few participants incorporated corpus use into their teaching due to limited practice time during the three-week corpus training and technical issues related to the interface of corpus software. Similarly, Ebrahimi and Faghieh (2017) investigated the reflections and perceptions of MA-level pre-service EFL teachers regarding the seven-week online course on corpora use in language teacher training. Aiming to introduce basic information regarding corpora and several corpus software in the EFL context, the results indicated that 32 Iranian pre-service EFL teachers needed more emphasis and exposure to the use of corpora for pedagogical purposes, also stressing that they had varying attitudes and perceptions regarding corpora use and corpus software as they found their use beneficial for language teachers and researcher but challenging for language learners due to technical issues and the time needed for being familiar with corpora. Zareva (2017) examined 21 Teaching English to Speakers of Other Languages (TESOL) trainee teachers' perceptions of integrating corpora into a grammar course. After introducing corpus software and mini corpus-informed tasks to enhance their corpus literacy, they concluded with a corpus-based project on a grammar topic. Despite low initial familiarity with corpora, participant trainee teachers appreciated the course design and the use of corpora in English teaching and for research. While the participants noted a need for more hands-on practice, they stressed several technical challenges, including navigating the interface of corpus tools and running a corpus analysis.

Şimşek (2020) designed a course at a state university in Türkiye on the corpus literacy of pre-service EFL teachers and investigated the possible impact of corpora on language awareness, teacher roles in material development, and reaching authentic language use. While stressing a lack of familiarity and basic corpus knowledge of Turkish pre-service EFL teachers, the study's findings indicated that the course on corpus literacy was found to enhance language awareness and positively affect the participants' perceptions of their teacher roles in material development. The pre-service teachers found corpus use beneficial thanks to access to authentic language, raising awareness, and the integration of technology and innovative content into language teaching. The study also emphasized the need for more courses targeting pre-service teachers and expanding their corpus literacy and consultation of corpora for language instruction. A recent study by Köse & Kuru Gönen (2023) explored how Turkish English teacher candidates perceived the integration of corpora into language instruction following a seminar on introduction to corpus use for two weeks. The seminar sessions focused on the introduction to corpus linguistics and basic practical applications of corpora for creating teaching materials. Results of the study showed that the pre-service teachers held positive views towards consulting corpora in vocabulary and grammar instruction. Despite some challenges and drawbacks, including query limit per day for several different concordancers and difficulties with finding appropriate language items through corpora consultation, the participants were willing to incorporate corpora into their future teaching practices. They stated their interest in attending further training on corpus use for language teaching.

All in all, the aforementioned research in the literature underscores the importance of providing teaching with training corpus and its pedagogical applications in language classrooms. Additionally, empirical studies emphasize the necessity of systematic, comprehensive training programs to equip language teachers with practical skills for the effective integration of corpus-based language teaching materials into actual teaching in real classrooms (Çalışkan, 2020; Chen et al., 2019; Lin, 2019). As already indicated in recent studies, gaining hands-on practice and experience through systematic training is essential for encouraging EFL teachers to integrate corpora and corpus-informed activities, thereby enriching their teaching and students' learning process under their guidance (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Ma et al., 2022; Şimşek & Can, 2023).

Considering the initial low familiarity with corpora and lack of research and training on corpora in the Turkish EFL context, it is crucial to investigate Turkish pre-service EFL teachers' corpus knowledge, their views on corpora and their experience in its practical applications in language instruction to gain deeper insights and develop tailored systematic, comprehensive training programs on the integration of corpora into language teaching environments. The literature also suggests the necessity for pre-service training to familiarize EFL teachers with corpus use and corpus-based language teaching pedagogy, equip them with practical skills for consulting corpora, demonstrate how corpora can be integrated into various language domains, and provide opportunities for gaining hands-on experience for implementing corpus-informed teaching activities into their EFL teaching practices (Başal et al., 2024; Köse & Kuru Gönen, 2023).

In order to meet this need and to assist prospective EFL teachers in the effective integration of corpora, this study aimed at designing a systematic and comprehensive training program on the integration of corpora into EFL teaching environments and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms. The CBLT Training was tailored for the specific context of the study and implemented into the practicum processes of senior pre-service EFL teachers as an attempt to equip them with the theoretical knowledge and practical skills of consulting and utilizing corpora, demonstrating how corpus use can be integrated across various language domains and skills, and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms.

3. METHODOLOGY

The primary aim of this chapter is to offer a comprehensive overview of the methodological details of the present study. It begins by introducing the research design, the research setting, and the participants, followed by a thorough explanation of the CBLT training procedure and the tools used during the training. This chapter also presents the data collection instruments, outlines the data collection procedure, and describes the data analysis process of the study.

3.1. Research Design

The present study aimed to reveal the pre-service EFL teachers' knowledge and awareness of corpora, as well as their experiences with corpora use for language teaching before the CBLT Training. Furthermore, during and after the training, the perceptions of the English teacher candidates regarding the use of corpora for language teaching and the CBLT Training program were investigated. This study employed an embedded mixed-methods research design with a priority on the qualitative data within the overall design (Creswell, 2012). This type of research design involves the collection and analysis of both qualitative and quantitative data during the study simultaneously or sequentially, and while one form of data is the primary source of data, the other form of data is given the secondary status to and the supportive role of the quantitative data to the other form of data (Creswell & Plano Clark, 2011). Figure 3.1. below outlines the research design of the present study.

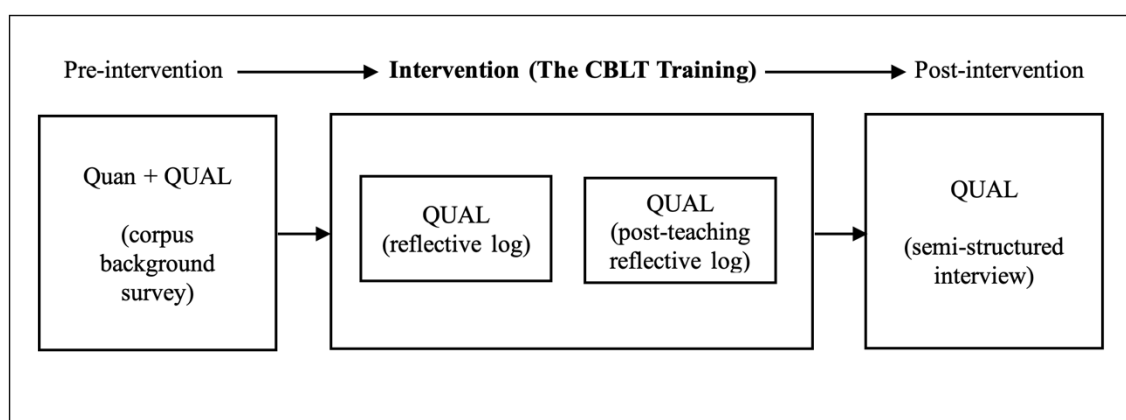


Figure 3.1. Outline of the present study's research design

As can be seen in Figure 3.1., the present study includes both quantitative and qualitative data collection methods, prioritizing qualitative data to address each research question. The quantitative components of the Corpus Background Survey administered before the training, alongside its qualitative components, were embedded into a qualitative research design framework that included qualitative data obtained from reflective logs, post-teaching reflective logs, and semi-structured interviews during and after the training.

3.2. Setting and Participants

3.2.1. Setting

The present study was conducted in the Department of Foreign Language Education at the Middle East Technical University (METU) in the Fall Semester of the 2023-2024 Academic Year. METU is an EMI (English-medium instruction) university and is recognized as an esteemed research institution.

The Department of Foreign Language Education (FLE) was established as a part of the Faculty of Education in 1982. In addition to the centralized, nationwide exams for being admitted to the FLE program, prospective FLE students and pre-service teachers must take a passing grade of 70 out of 100 on the English Proficiency Exam (EPE) or provide evidence of achieving a satisfactory score on the IELTS Academic or TOEFL IBT to certify their proficiency in English. Offering B.A., M.A., and Ph.D. programs, this department provides core courses categorized into the subject matter, pedagogical formation, and world knowledge, as well as compulsory and elective courses related to teaching pedagogy and English Language Teaching (ELT) methodology, linguistics, and literature. Moreover, FLE programs cover a range of other subjects, including language acquisition, technology in education, practicum, translation, testing, and evaluation (METU, 2024).

In the last two semesters of the METU FLE program, senior pre-service English teachers are to complete the requirements of the ELT Practicum I and II courses. According to the curriculum of the Higher Education Council (HEC) of Türkiye, the main requirement of the practicum course is observing the lessons of the cooperating mentor teacher at one of the Ministry of National Education (MoNE) state and private schools in Ankara for 12 weeks and 72 lesson hours in total. The observations of the student-teachers

include the teaching techniques, classroom management strategies, and lesson planning of the mentor teacher. Moreover, senior pre-service English teachers of the FLE program must teach two forty-minute lessons each semester under the supervision of the mentor teacher and the course instructors at FLE. In addition to the two assessed teachings, they engage in more unassessed teachings, various activities, and tasks to enhance their teaching skills and classroom management techniques throughout the practicum courses, including reflection and observation tasks, lesson planning assignments, and designing materials. It should be noted that no compulsory or elective courses or training are offered on corpus use in the FLE program. However, it's also worth mentioning that corpus was touched upon in a few courses due to its recent significance. Despite having a little theoretical familiarity with corpus linguistics, there has been no specific course on the integration of corpus into EFL teaching in actual teaching contexts for the prospective English teachers of the FLE program.

3.2.2. Participants

The present study was conducted in two phases: the first phase with the administration of the Corpus Background Survey, and the second phase that included the 9-week Corpus-based Language Teaching (CBLT) Training and the following data collection procedures. While a total of 102 senior Turkish pre-service EFL teachers took part in the first phase of the study, there were 20 pre-service EFL teachers who attended the CBLT training in the second phase of the study.

In line with the purposes of the study, criterion sampling was used while choosing the participants for the first phase of the study as it enables the researcher to identify all participants that meet predetermined criteria and characteristics (Patton, 2002). The criterion for participating in the Corpus Background Survey was taking the Practicum course as the study set out to explore the views and perspectives of Turkish English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts. In the initial phase of the study, there were a total of 102 senior pre-service EFL teachers who met the criterion of all 110 senior pre-service EFL teachers in the department. The main motive behind choosing the senior pre-service EFL teachers taking the practicum course was to work with teacher candidates who will have real teaching contexts where they will experience the need to create teaching materials for their own teaching purposes

according to the topic, theme, grade, and level of the class they will be teaching. The researcher is also a research and teaching assistant working at the Department of Foreign Language Education, actively participating in the practicum processes.

The Corpus Background Survey was administered to a total of 102 senior pre-service EFL teachers taking the FLE431 ELT Practicum I course in the Fall Term of the 2023-2024 Academic Year via Google Forms. Of all 102 pre-service EFL teachers who opted to participate voluntarily in the initial phase of the study, including the administration of the Corpus Background Survey, the majority attended their practicum processes at various high schools in Ankara at the time of the study as a part of their undergraduate education. Their ages ranged between 18 to 27. Table 3.1. demonstrates detailed demographic information of the participants who attended the initial phase of the study.

Table. 3.1. *Demographic information of the participants*

Gender	<i>n</i>	Age	<i>n</i>	Observation Schools	<i>N</i>
Female	62	18-21	70	Primary school	12
Male	39	22-25	18	Secondary school	12
Non-binary	1	25+	14	High school	78

Apart from administering the background survey to provide a comprehensive overview of participants' background knowledge, familiarity, and perceptions regarding corpus and corpus-based language teaching in EFL contexts with both quantitative and qualitative items, the researcher also utilized the survey to identify the pre-service EFL teachers who were willing to attend the Corpus-based Language Teaching Training and take part in the second phase of the study. The researcher informed the participants regarding the content, framework, and objectives of the CBLT Training and reminded them of the *Corpus Use and Corpus-Based Materials Seminar* conducted by the researcher the year before (Köse & Kuru Gönen, 2023). Of the 102 participants of the present study, 57 attended the two one-hour seminar sessions mentioned above on corpus use before. This seminar mainly covered background information on corpus and included theoretical details on the use of corpus in language teaching. Almost all participants of the aforementioned seminar stated they would be willing to attend a more comprehensive training on corpus use to learn more about corpora and its practical use in EFL contexts

as they had not received any course or professional development training or any other exposure to the use of corpora or concordancers for teaching purposes aside from the seminar. Further details regarding the seminar conducted by the researcher are outlined in the Corpus-Based Language Teaching Training section. 36 potential participants informed the researcher of their interest in attending the 9-week systematic and extensive CBLT training. However, 16 potential participants who expressed interest in attending the training sessions were unable to do so due to their busy schedules. Despite their willingness to participate, they could not accommodate the training sessions into their schedules.

In the second phase of the study, 20 pre-service EFL teachers were selected for the training as they were available, voluntary, and willing to attend all training sessions in line with the principles of convenience sampling (Creswell, 2012). The rest of the data collection procedures of the study were conducted with the 20 participants of the CBLT Training. The detailed demographic information of 20 pre-service EFL teachers in the training sessions is outlined in Table 3.2.

Table. 3.2. *Demographic information of the participants in the training sessions*

Gender	<i>n</i>	Age	<i>n</i>	Observation Schools	<i>N</i>
Female	14	18-21	5	Primary school	4
Male	6	22-25	14	Secondary school	3
		25+	1	High school	13

3.3. Corpus-based Language Teaching Training

The primary purpose of the study was to investigate the views and perspectives of Turkish English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts before and after a systematic and comprehensive training program on the theoretical and practical aspects of corpus software and online corpora. Additionally, the study aimed to assess the hands-on experience of creating, evaluating, and adapting corpus-informed teaching materials, as well as the use of these materials in real EFL classrooms as a part of the practicum process of the participants. Taking into account the scope and purposes of the study, a need for devising such training that embodies all the aspects regarding corpus use in EFL contexts given above emerged.

As Nunan (2003) suggests, a small-scale pilot study was conducted by the researcher in order to explore the needs, interests, and characteristics of pre-service EFL teacher candidates as well as their familiarity with and perceptions regarding corpus and the integration of corpora into EFL classrooms. Carried out in the Spring Term of the 2021-2022 Academic Year, the pilot study was conducted in the same context as the present study with 63 pre-service EFL teachers to create corpus awareness. The piloting revealed that they were either unfamiliar with consulting corpora or had *little* familiarity before the two one-hour seminar sessions conducted by the researcher. Upon receiving the seminar, they were willing to attend a more comprehensive training on corpus use to learn more about corpora and its practical use in EFL contexts, as it was their first introduction to corpora (Köse & Kuru Gönen, 2023). It was also found that there was a need to integrate corpora into language teaching contexts to consult corpora to decide what would be selected and taught from ‘the total corpus of the language’ (Richard, 2017) and create authentic teaching materials to be used in actual teaching contexts. The aim of the seminars in piloting was to develop an awareness of corpora and its potential use for language teaching purposes and to identify any need for corpus integration. 57 participants enrolled in the seminars participated in the recent study. Before the current study, they were asked whether they used any of the ideas in the piloting via a Corpus Background Survey, and it was revealed that only a few used corpus software or online corpora for teaching purposes after the seminar as they required more hands-on experience to improve their corpus literacy. It highlighted the need for systematic and comprehensive training and exposure to the practical applications of corpora. Thus, the present study was designed based on the need that arose from the piloting.

Following the experiential learning cycle of Kolb (1984), the Corpus-based Language Teaching Training was designed to engage its participants with actual hands-on corpus consultation, analysis, and building corpus practices to create ‘concrete experiences’ which was followed by reflecting on the process and their experiences of using corpora for their own corpus literacy as in the role of a learner. Training sessions were designed in line with the experiential learning cycle, which prioritizes the cyclical process of observing, reflecting, conceptualizing, and experimenting with the practical application of corpus consultation and the actual integration of corpus-informed language teaching materials into their actual teaching contexts. The overview of the experiential learning cycle by Kolb (1984) is demonstrated in Figure 3.2.

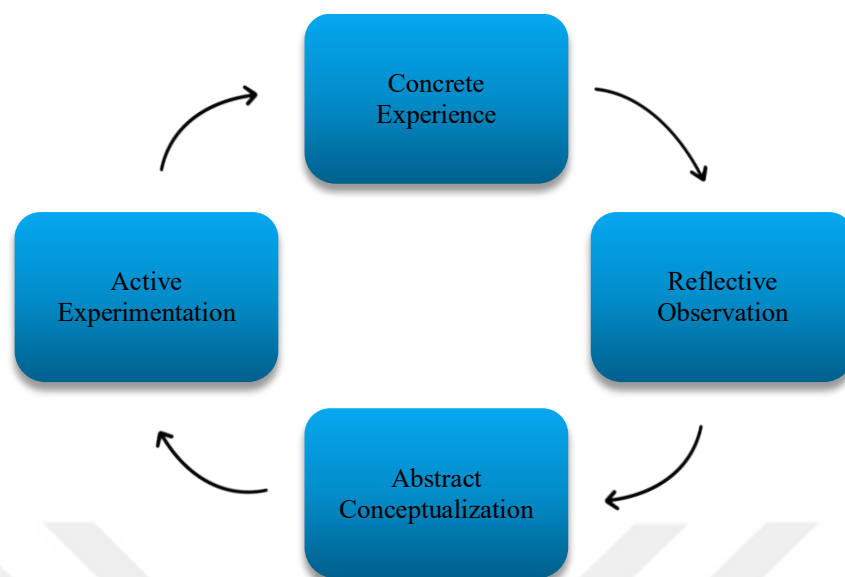


Figure 3.2. *Overview of the experiential learning cycle by Kolb (1984)*

The participants were provided with various tasks and assignments in order to experience corpus use to develop their skills and fully grasp the concept of corpus consultation. While the majority of the tasks and assignments of the training focused on the principles of indirect corpus applications (60%) with an emphasis on the design and utilization of corpus-informed language teaching materials for language teaching, the rest focused on the direct application of corpora (40%) by the participants to have hands-on corpora consultation. Through the conceptualization process of the experiential learning cycle, they learned from their own practices, and, eventually, during the active experience stage, they applied their concept of corpora and the products they created by using corpus for creating materials both at micro and macro levels. The pre-service EFL teachers were required to teach three micro-teachings during the training sessions and at least one macro-teaching as a part of their practicum process by utilizing corpus-based EFL teaching materials they created.

The content of the training included presentations on the theoretical aspects and practical applications of corpora, in-class and follow-up tasks and assignments, assigned chapters on each session's objectives, podcasts and YouTube videos on the integration of corpora into language teaching, recommended journals for accessing the corpus-based language teaching pedagogy research, and corpus-informed language teaching resources to be used in EFL classrooms. Such corpus-based informed language teaching resources included online corpora for vocabulary and grammar teaching and ready-made lesson

plans utilizing corpus-based language teaching materials. The content of the training was compiled and adapted from the relevant studies in the corpus-based research literature (Çalışkan, 2020; Girgin, 2011; Leńko-Szymańska, 2015; Şimşek, 2020) and free online corpus-based teaching pedagogy resources and tools. Two experts in the field of corpus-based language teaching pedagogy were consulted about the scope and content of the training for their expert opinions.

Accordingly, the content was reviewed and finalized. In addition to the content of the training with the aforementioned tasks, assignments, and materials above, the researcher introduced multiple corpus software and online-accessible corpora and demonstrated their hands-on use to the English teacher candidates during the training sessions in order to familiarize them with navigating corpus software and online corpora and explore their direct and in-direct applications for language teaching purposes before their own hands-on practice with corpora for creating corpus-based language teaching materials. The corpus tools and software and online-accessible corpora in Table 3.3. were demonstrated and practiced during the training sessions in line with the objectives of the training and the relevant literature, as the participants were guided on the hands-on practice of the corpora below for maximum access to authentic language data with their easy-to-access interfaces.

Table. 3.3. *The list of corpus software and online corpora demonstrated in the training*

Corpus Software & Resources	Online Corpora
AntConc (Building a corpus)	BNC – British National Corpus
AWL Highlighter (Vocabulary teaching)	COCA – Corpus of Contemporary American English
Compleat Lexical Tutor (Corpus tools)	MICASE – Michigan Corpus of Academic Spoken English
CorpusMate (Online corpus platform)	NOW – News on the Web
Just the Word (Online corpus platform)	WebCorp Live
LancsBox (Corpus analysis software)	
SketchEngine (Building a corpus)	
Skylight (Online corpus tool)	
Tim Johns' Kibbitzers (Teaching materials)	

3.3.1. Procedure of the Corpus-based Language Teaching Training

The researcher designed the Corpus-based Language Teaching Training to cover the theoretical background of corpus and corpus-based language pedagogy, hands-on practical use of corpus software, and online corpora to create and adapt corpus-based teaching materials for EFL contexts. The training lasted for nine weeks in the form of

two-hour sessions each week. The syllabus and schedule of the training sessions are outlined in Table 3.4.

Table. 3.4. *The syllabus and schedule of the Corpus-based Language Teaching Training*

Session	Topic
Session 1	<i>Introduction to the Training</i> • Training content and objectives, Corpus Linguistics, Defining a corpus, Types of Corpora
Session 2	<i>Introduction to Corpus Use in EFL</i> • Quiz on corpus literacy, Consulting corpora in EFL, Concordancers, COCA Practice
Session 3	<i>Building a Corpus</i> • Introducing AntConc and SketchEngine, How to build a corpus, Building our own corpus
Session 4	<i>Consulting Corpora for Teaching Vocabulary (Assignment 1)</i> • Corpus-based language teaching pedagogy, Creating vocabulary teaching materials
Session 5	<i>Consulting Corpora for Teaching Grammar (Assignment 2)</i> • Micro-teachings & presentations, Creating grammar teaching materials
Session 6	<i>Consulting Corpora for Error Correction & Skill Integration (Assignment 3)</i> • Micro-teachings & presentations, Creating materials for error correction and skill integration
Session 7	<i>Material Evaluation and Adaptation Practice (Assignment 4)</i> • Micro-teachings, Evaluation and adaptation of MoNE Textbooks & Corpus-informed Textbooks
Session 8	<i>Material Evaluation and Adaptation Practice (Assignment 5)</i> • Evaluating materials created by the participants of the seminar conducted the year before
Session 9	<i>Overall Reflection</i> • Discussion on creating and using corpus-informed teaching materials in EFL classrooms

As presented above, each training session focused on discussions and both in-class and homework assignments on corpus use in EFL contexts. The training participants were given five assignments (e.g., creating corpus-based EFL teaching materials for teaching grammar, vocabulary, error correction, skill integration, and material evaluation and adaptation) throughout the training. The researcher briefed the participants on the ethical guidelines of the study with a consent form (see Appendix A), ensuring data confidentiality and participant anonymity before the training started officially. Participants were also informed that they were free to drop out at any stage of the training and the study.

Session 1 was the only online session to inform the participants about the research, the researcher, and the training. The following eight training sessions were conducted face-to-face to ensure participant interaction and collaboration. Details regarding the training schedule, objectives and training requirements, assignments, and data collection

procedure were outlined in the session's first hour. The second hour was about corpus linguistics, the concept, definition, and types of corpora. The researcher explained and demonstrated what a corpus provides to its users, as well as different types, features, and versions of online corpora, by using one presentation with several tasks to raise the participants' awareness of what we can do with what a corpus provides. The researcher designed a specific presentation, multiple tasks, and assignments for each session by benefitting from several published materials such as dissertations, books, and online resources and adapting them to the study's context to draw the attention of the participants to the features of corpora and their use in EFL contexts. The first session ended with assigning several videos on corpus and the 'Touchstone: From Corpus to Course Book chapter by McCarthy (2004) with four discussion questions.

Session 2 and all successive training sessions were held in the same room, at the same hour, and on the same day for eight weeks. The researcher surprised the participants with a pop quiz in order to attract their attention. The participants were expected to write down their answers to the ten multiple-answer questions about the previous session's content. Following the quiz, discussions were shaped around the chapter by McCarthy (2004), as well as the why and how to consult corpora. The researcher defined and demonstrated what a concordancer is in detail. As the session focused on introducing corpus use in EFL contexts and classrooms, corpora in language pedagogy with both indirect and direct applications were demonstrated and exemplified to the participants (Römer, 2011). The researcher also clarified how to exploit corpora and corpus-based language teaching materials by a brief lecture and short discussions. In the second hour, the researcher demonstrated the features and interface of the Corpus of Contemporary American English (COCA) before its hands-on practice by the participants. COCA, as the most 'widely-used' online corpus with over 1 billion words derived from about 500,000 texts between 1990-2019, provides authentic and naturally occurring language data from multiple genres, including academic texts, newspaper articles, fiction, magazines, and so on (Çalışkan, 2020). COCA offers various features to its users, facilitating word or phrase searches in keyword-in-context usage as in concordancing to analyze their collocations, frequencies, and part of speech positioning. Additionally, COCA also stands out as one of the online corpora commonly utilized in recent research (Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Poole, 2022; Şimşek & Can, 2023; Xodabande & Nazari, 2022). After the researcher demonstrated the features

of COCA and how to run a search on COCA with its specific search syntax and query options, the participants were given ten tasks to practice searching for particular words and phrases, finding their collocations, and analyzing keyword-in-context/concordance lines of the target search in groups of two or three. The researcher uploaded ReadMe files and manuals for practicing COCA to the course page and provided assistance to the groups while they were completing the assigned tasks.

The aim of **Session 3** was to familiarize the pre-service EFL teachers with the concept of text analysis on corpus data and guide them to build their own corpus. The session started with a discussion run by the researcher on why to analyze a corpus and how to do it. Afterward, the researcher introduced two major corpus software for building and analyzing a corpus, AntConc and SketchEngine (Leńko-Szymańska, 2017), with their features and applications. These two corpus software are used for the investigation of language either by building an original corpus with text files of written or spoken language or simply running a corpus analysis on ready-made corpus data (Godwin-Jones, 2017). The researcher first demonstrated how to download the AntConc software and then compile and upload texts to create a corpus on the software. All features of AntConc, including running word or phrase search with query syntax, n-grams clusters, collocations, keyword-in-context/concordancing, and word clouds. Due to time constraints during the session, the participants were encouraged to practice SketchEngine independently after the session. In the second hour of the session, the researcher provided the participants with both user manuals for the software and multiple available text files so they could have hands-on practice with building their own corpus. In addition to providing them with the ZIP files of such available texts from different genres such as the NLTK Movie Review Corpus, Gothic Fiction and Abraham Lincoln Speeches and Writings, etc., the researcher also encouraged them to create their own corpus with a group of texts of their liking. They were also assigned an in-class assignment to run an analysis of text corpora with a handout devised by Prof. Wendy Anderson (2018) of the University of Glasgow for the *Methods for Text Mining and Analysis of Text Corpora Workshop* conducted in 2018. Having submitted their assignments, they received feedback from the researcher for this task and the COCA Practice task they submitted the previous week. A corpus search result demonstrated for searching the clusters in the NLTK Movie Review Corpus using AntConc in Session 3 is illustrated below.

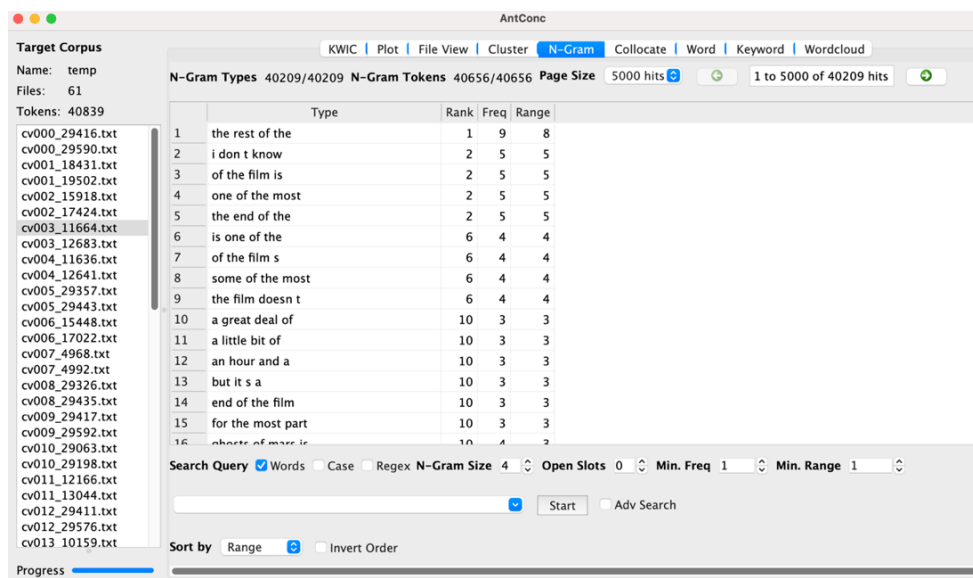


Figure 3.3. A screenshot of a sample corpus search on AntConc used in session 3

Having revised what corpora can offer in language teaching and learning with the assigned chapter by McEnery and Xiao (2011), the participants were provided with ready-made sample lesson plans for teaching language skills to different age groups and proficiency levels and further resources to be used in EFL contexts before **Session 4**. The researcher first introduced corpus-based language pedagogy and the design principles for corpus corpus-based learning and teaching materials by Ma et al. (2021), including testing the knowledge of students and then having them practice hands-on corpus queries for the inductive exploration of the language use or vocabulary item by the students before the output practice in the form of exercises. With a discussion on how to use corpora for vocabulary teaching, the researcher drew the attention of the participants to strategies and types of exercises to be used and challenges experienced while teaching lexical items through questions. By discussing which lexical items are the most problematic for the students in their practicum contexts, the focus was set on how to benefit from corpus data to teach lexical items according to the needs in our contexts.

The researcher provided the participants with various corpus-informed language teaching materials for EFL classrooms, and discussions were held on developing and adapting them to their contexts. The sample corpus-informed EFL teaching materials were adapted from the relevant literature (Çalışkan, 2020; Girgin, 2011; Şimşek, 2020). At the end of the session, the participants were given an assignment on designing at least two corpus-based vocabulary teaching materials for the lexical items at any proficiency

level they wanted. The researcher reminded them of the corpus-based teaching materials they are required to create for their macro-level assessed teachings in their practicum schools as a part of the study and the practicum processes and encouraged them to create their teaching materials according to their contexts. Considering and specifying the elements such as language proficiency level, target lexical items, genre, activity type, and activity format, they were allowed to work individually or in groups of two. The researcher provided them with a guideline on these elements, and they were encouraged to either micro-teach or present their teaching material in the following session. Guiding questions for the in-class discussions were also provided beforehand. Each individual/pair made their submissions to ODTÜClass before the session and prepared a presentation for their corpus-informed teaching materials and activities. Further information regarding ODTÜClass and the other tools will be elaborated in the subsequent section.

The procedures for sessions four, five, and six on consulting corpora for teaching vocabulary, grammar, error correction, and skill integration were similar. **Session 5** started with the micro-teachings and presentations of the participants. Discussions were held on the materials created and their appropriateness for the specified elements of their contexts. In the second hour, the same procedure was followed as in session four. Following the discussions on consulting corpora and using corpus-based teaching materials for teaching grammar, the researcher provided the participants with various corpus-informed grammar teaching materials. At the end of the session, the participants were given an assignment on designing at least two corpus-based grammar teaching materials for the grammar structures for the proficiency level of their liking for session six, mainly for their practicum contexts. In **Session 6**, a similar procedure was followed. After the micro-teachings and presentation of the pre-service EFL teachers, the researcher shifted the session's focus to creating teaching materials for error correction and skill integration. Following the discussion on revealing the views of the pre-service EFL teachers on the use of corpus for error correction, which common errors they encounter in their practicum contexts, and how they give corrective feedback, the researcher provided the participants with various corpus-informed teaching materials for error correction and skill integration. The participants were assigned to design at least two error correction and skill integration materials, targeting any proficiency levels.

Session 7 started with micro-teachings and presentations on corpus-informed teaching materials on error correction and skill integration. Following the presentations,

the researcher clarified the session's focus as material evaluation and adaptation of textbooks. Having provided the participants with one MoNE textbook and one corpus-informed textbook, the researcher guided the participants in evaluating both textbooks and comparing them in groups of two or three. While the MoNE textbook was *Spice Up* for 11th-grade English students, the corpus-informed textbook was the 2nd edition of *Touchstone* by McCarthy et al. (2014). As the participants were also taking FLE405 Materials Adaptation and Development course at the time of the study, they were already familiar with textbook evaluation and adaptation. The participants were encouraged to choose either one of the coursebook evaluation checklists by Demir & Ertaş (2014) and Razavi (2014) uploaded to the course page. The researcher provided assistance to the groups while they were evaluating the textbooks if needed. Each group made their submissions to ODTÜClass after the session and received feedback on their evaluations. During the in-class assignment, discussions were held on different aspects of the two textbooks.

Session 8 focused on the evaluation and adaptation of the corpus-based EFL teaching materials created by the participants of the seminar conducted the year before. As a part of the in-class assignment of the session, participants were provided with a total of twenty teaching materials. Working in groups of two or three, participants were encouraged to work on at least three teaching materials of their liking and adapt them to their contexts. Discussions were held on each material created earlier, including which parts of these corpus-based teaching materials could have been adapted and how. The researcher provided assistance to the groups and held whole-class discussions while they were adapting the materials. Each group made their submissions to ODTÜClass after the session and received feedback on their adaptations. Participants were then informed about the last session of the training.

In **Session 9**, the researcher and the participants engaged in a comprehensive discussion and reflection on using corpus-based teaching materials in EFL contexts. The pre-service EFL teachers reflected on their experience of creating and utilizing corpus-informed teaching materials and integrating corpora into their EFL teaching context during the practicum process. Participants took the floor to express their experiences, gains, challenges, and final remarks on the CBLT training.

3.3.2. Tools Used in the Corpus-based Language Teaching Training

A range of specific tools was utilized to enable ease of access to the training materials, including assigned chapters and readings, podcasts, YouTube videos, recommended journals, and resources to be used in EFL classrooms, and facilitate efficient communication throughout the training and scheduling for the training sessions and the semi-structured interviews. By leveraging such specific tools, the CBLT Training was able to optimize communication, collaboration, and organization, ultimately enhancing the learning experience for the pre-service EFL teachers.

3.3.2.1. ODTÜClass

As the official learning management system (LMS) of METU, ODTÜClass functions as an online platform aimed to support METU students in their academic pursuits, both within and beyond the classroom environment. Within this virtual learning environment, students benefit from the seamless delivery of course materials as it provides a structured framework for course management. This LMS has been actively used by the academic staff and students of METU since the Fall Term of 2014-2015 Academic Year. ODTÜClass was utilized by the researcher so that participants could access relevant resources and activities anytime and anywhere on a structured platform. The overview of the CBLT Training on ODTÜClass is demonstrated in Figure 3.4.

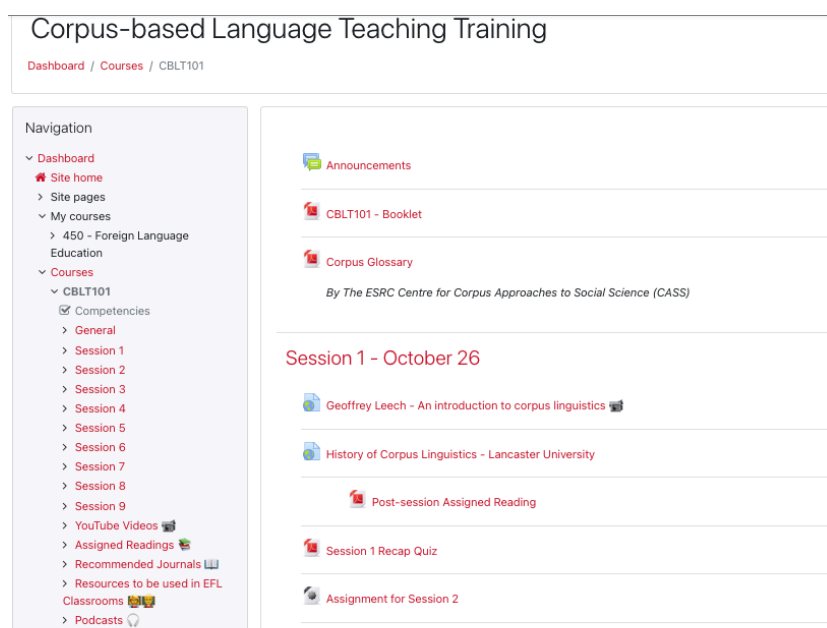


Figure 3.4. Overview of CBLT Training course page on ODTÜClass

All CBLT training-related materials were made available to the participants on ODTÜClass under different weeks designed for each session, enabling easy access to the assigned chapters and readings. Direct URL links were given to the specific corpus resources to be used during the training sessions for in-class tasks and after the session for follow-up tasks and assignments. Participants were provided various YouTube videos and podcasts on the ODTÜClass platform regarding corpus and corpus use in EFL contexts. See Figure 3.5. for the resources and podcasts posted on the ODTÜClass course page.

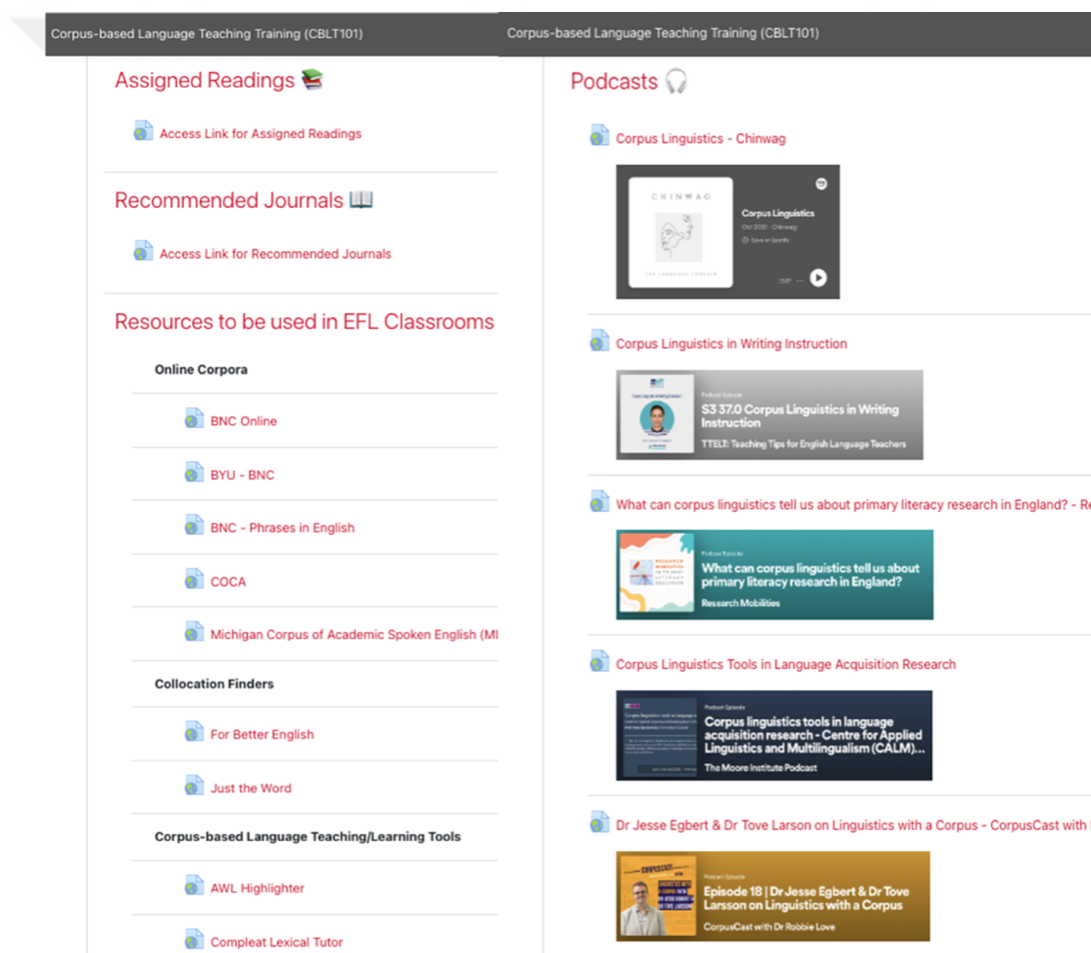


Figure 3.5. Screenshots of the resources to be used in EFL Classrooms and the Podcasts posts

3.3.2.2. Doodle

The Doodle app is a widely used scheduling and calendar tool for booking 1v1 or group meetings with multiple guests in a large group. Doodle was pivotal in facilitating

coordination and scheduling for the training sessions and the semi-structured interviews. The researcher created polls to schedule the training sessions and the 1v1 interviews with potential day and time options and sent the polls to the CBLT Training WhatsApp Group to arrange the sessions and interview meetings. Figure 3.6. offers a glimpse of the Doodle poll interface used for scheduling the post-training interviews.

Doodle

Berkcan Köse is organizing

CBLT101 - Post-training interviews

Zoom

Please choose only ONE time slot that works for you.

10 of 20 invitees responded

Yes (1 click)

If need be (2 clicks)

No

Pending (yet to vote)

Select your preferred times

We'll let you know when the organizer picks the best time

Turkey, Istanbul (GMT+3) 20 options

Order by Date | Most popular

	TUE 16 JAN 10:00 AM - 10:30 AM	TUE 16 JAN 10:30 AM - 11:00 AM	TUE 16 JAN 11:00 AM - 11:30 AM	TUE 16 JAN 11:30 AM - 12:00 PM	TUE 16 JAN 12:00 PM - 2:00 PM
You					
CE C*** E***	X	X	X	X	X
OA O*** A***	X	✓	X	X	X
BD B*** D***	✓	X	X	X	X

Figure 3.6. The interface of the Doodle poll for semi-structured interviews

3.3.2.3. WhatsApp

In addition to the e-mails and announcements posted on the course page on ODTÜclass, WhatsApp was chosen to establish a class group with the participants of the CBLT Training. Having created a WhatsApp group with all 20 participants, the researcher provided the training team with a convenient and accessible platform to foster interaction and facilitate communication and collaboration by posting mini-polls, sending reminders and announcements, and creating a 7/24 Q&A and discussion atmosphere for any questions that emerged during the 9-week systematic and extensive CBLT training. Participants actively used the class group to seek support for their follow-up assignments, where the researcher promptly provided assistance when needed and discussed assigned chapters with other participants.

3.4. Data Collection Instruments

The primary purposes of the present study were to investigate the views and perspectives of Turkish English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts before and after receiving the systematic and comprehensive Corpus-based Language Teaching (CBLT) Training tailored for the specific purposes of the study. Additionally, it was aimed to reveal the participants' views and perceptions of the CBLT Training program. In light of these objectives, a combination of four data collection instruments was employed. Quantitative data were collected through the Corpus Background Survey, incorporating qualitative and quantitative data collection. Qualitative data were retrieved from Reflective Logs, Post-teaching Reflective Logs, and Semi-structured Interviews.

Experts' opinions were obtained to ensure the appropriateness and applicability of all four data collection instruments to the study's objectives. In accordance with this, six experts in corpus linguistics and corpus-based language teaching were informed about the purpose of the current study and presented with each data collection instrument for their expert opinions. Incorporating their feedback and recommendations, essential adjustments were implemented for each data collection instrument. Research questions of the study and relevant data collection instruments employed before, during, and after the CBLT Training are outlined in Table 3.5. below.

Table. 3.5. *Research Questions and Data Collection Instruments*

Research Questions	Data Collection Instruments
RQ1: What do Turkish pre-service EFL teachers think about corpora and the integration of corpora into language teaching contexts?	Corpus Background Survey (n=102)
RQ2: What are the perceptions of Turkish pre-service EFL teachers regarding the	Reflective Logs (n=60)
a. integration of corpora into actual teaching contexts?	Post-teaching Reflective Logs (N=20)
b. systematic Corpus-Based Language Teaching Training?	Semi-structured Interviews (n=20)

3.4.1. Corpus Background Survey

Both quantitative and qualitative data were gathered through a background survey on corpus use to explore the Turkish pre-service EFL teachers' background knowledge, familiarity, and awareness of corpora, as well as their experiences with using corpus tools and online corpora for corpus-based language teaching. The survey was adapted for the specific purposes and context of the present study by the adaptation of existing surveys in the relevant literature (Bunting, 2013; Çalışkan, 2020; Geluso & Yamaguchi, 2014) to investigate the characteristics, opinions, and attitudes of the participants (Nunan, 1992). Six experts in the field of English language teaching were consulted in order to ensure the reliability of the survey, while two EFL instructors were also consulted to proofread the survey items to eliminate any potential ambiguities before its administration to the pre-service EFL teachers. With the feedback from the experts and the EFL instructors, adjustments were made to certain items, and several expressions and the order of items underwent revision. Comprising both open-ended and five-point Likert scale items, the corpus background survey was meticulously designed to provide a comprehensive overview of participants' background knowledge, familiarity, and perceptions of corpus and its utilization in ELT contexts, facilitating a robust analysis of the data collected. The revised and finalized background survey (see Appendix B) was structured into three sections: Demographic Information, Corpus Knowledge and Literacy, and Using Corpus Resources in Teaching.

As the participants of the study were senior pre-service EFL teachers doing their practicum at the Ministry of National Education (MoNE) schools, the demographic information section included four items related to gender, age, and grade/level the pre-service teachers were teaching at their practicum schools. Focusing on the corpus knowledge and literacy of the participants, the subsequent section was composed of five open-ended and two five-point Likert scale items. Pre-service EFL teachers were tasked with assessing their familiarity with corpus software and online corpora and their level of knowledge of using them for various educational and academic purposes. Participants selected options ranging from *no knowledge* to *expert knowledge*. In the third section on using corpus resources in teaching, there were three sub-sections for the participants of the seminar conducted by the researcher earlier. The pre-service EFL teachers who indicated they attended the seminar were additionally instructed to proceed to the following sub-sections, where they were asked to provide additional explanations to gain

further insight into how they became familiar with corpus software and online corpora, whether they had taken a course or professional development training or had any other exposure to the use of corpora or concordancers aside from the seminar, and their experiences in using corpora for language learning and teaching. The survey was administered in English due to the participants' proficiency level, competence, and confidence in using English for academic and professional purposes, and language support was provided when needed while completing the Corpus Background Survey. The survey was administered via Google Forms.

3.4.2. Reflective Logs

In order to gain insight into the evolving perceptions and experiences of the participants throughout the Corpus-based Language Teaching (CBLT) Training, three reflective logs (see Appendix C) were employed at the end of the second, fifth, and eighth training sessions via Google Forms. Tailored specifically to the objectives of the study, these logs posed guiding questions designed to delve into the participants' reflections on the impact of training sessions on their understanding and skills related to corpus use, specific aspects of corpus use explored, potential real-world applications, research, challenges encountered and strategies to overcome them in future sessions. The questions were strategically crafted to explore the participants' evolving attitudes toward corpus-based language teaching and their perceptions of the effectiveness of the training program over nine weeks, as the same guiding questions were presented for each reflective log. By requiring the participants of the CBLT Training to write three reflective logs, this study aimed to provide a perspective on their journey and development in integrating corpus-based approaches into their teaching practices. While the guiding questions were provided in English, participants had the option to respond in either English or their native language, Turkish. All participants wrote their responses in English, citing proficiency and confidence in expressing their thoughts and experiences effectively.

3.4.3. Post-teaching Reflective Logs

Participants were further tasked with completing a post-teaching reflective log (see Appendix D) after conducting a 40-minute lesson at their observation schools as part of their practicum requirement. This reflective log was specifically designed to capture

participants' reflections on their use of corpus-based teaching materials they created as a part of the CBLT Training according to the topic, theme, grade, and class level they taught. Guided by four structured questions, participants reflected on their utilization of corpus tools and online corpora in preparing teaching materials, their experiences integrating corpus-based teaching materials into their teachings during practicum, including any perceived benefits or drawbacks, and the impact of these materials on student participation in the classroom. Through these post-teaching reflective logs following their integration of corpus-based teaching materials into their own teaching, the study aimed to gain comprehensive insights into participants' experiences and perceptions of integrating corpus-based teaching pedagogy into their teaching practices, thereby providing a holistic understanding of the effectiveness and challenges of corpus-based language teaching. All participants filled in one post-teaching reflective log via Google Forms after teaching English at their practicum school. Even though the post-teaching reflective logs were provided both in Turkish and English, all participants wrote their reflections in English, similar to the reflective logs.

3.4.4. Semi-structured Interviews

Semi-structured interviews (see Appendix E) were conducted as a crucial component of this study to delve deeper into the perspectives and experiences of participants who attended the CBLT Training for nine weeks as semi-structured interviews enable each participant to share their individual experiences (Merriam & Tisdell, 2015). A one-to-one semi-structured interview protocol was followed due to its flexibility in gathering a thorough understanding of the qualitative data (Creswell, 2013; Nunan, 1992). Therefore, once the training sessions ended, ten participants of the CBLT training volunteered to attend the interviews. These interviews provided an opportunity for participants to articulate their interpretations of the CBLT Training, discuss their experiences in integrating corpus-based approaches into their teaching practices, and express their views on the effectiveness and challenges of corpus-based language teaching. The five questions tailored for the semi-structured interviews were revised based on the feedback of six experts in the field. Each one-to-one semi-structured interview was conducted in English as participants stated they were confident in expressing themselves in English at the beginning of the interviews. All interviews took place on Zoom, initially audio-recorded and later transcribed.

3.5. Data Collection Procedures

The data for the present study were gathered using multiple data collection instruments in the Fall Term of the 2023-2024 Academic Year to explore the views and perspectives of Turkish English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts before and after receiving the CBLT Training. Before initiating the data collection procedures, all necessary permissions and ethical approval were obtained from the Ethical Committee of Anadolu University (see Appendix F) and the Human Research Ethics Committee of Middle East Technical University (see Appendix G). Upon receiving the permissions and ethical approval from both universities, the data collection process that lasted 14 weeks was initiated.

As the first step, the four instructors of the Practicum course at the Middle East Technical University Foreign Language Education Department were informed about the study and the upcoming data collection procedures. The researcher initiated the data collection procedure in the second week of the semester by introducing Corpus-based Language Teaching Training, its aims and schedule, and the data collection procedure to all 4th-year practicum students (pre-service EFL teachers) with the CBLT Training Booklet (see Appendix H) to introduce the materials and activities to be used throughout the training and administering the Corpus Background Survey. As well as obtaining the consent of all practicum students with the Corpus Background Survey, the researcher stated that those who are interested in attending the CBLT Training could inform the researcher regarding their interest, availability, and contact information with the mini-survey at the end of the Corpus Background Survey. Having analyzed the availability of all interested participants, the researcher came up with a day and time that worked for most of the willing pre-service EFL teachers. Twenty pre-service EFL teachers who were both interested in attending the CBLT Training and available for all training sessions were informed about the training and data collection process in more detail. The researcher briefed the participants on the training's requirements, assignments, and data collection procedures. They were provided with the ethical guidelines of the study with a consent form, ensuring data confidentiality and participant anonymity before the first training session.

The second set of data was collected by three reflective logs at the end of the second, fifth, and eighth training sessions of the training. Having provided all training materials, including assigned chapters and readings, podcasts, YouTube videos, recommended journals, and resources to be used in EFL classrooms on the course page of the CBLT Training on ODTÜClass, the researcher conducted the first Introduction to the Training Session online to inform the participants regarding the training content, framework, and objectives. The following eight training sessions were conducted face-to-face to ensure participant interaction and collaboration. The following eight training sessions were conducted face-to-face to ensure participant interaction and collaboration. Three reflective logs were employed at the end of the second, fifth, and eighth training sessions via Google Forms to explore their evolving attitudes and perceptions toward corpus-based language teaching and the effectiveness of the training program over nine weeks.

The participants of the training created corpus-based EFL teaching materials to be used in their teaching contexts at their practicum observation schools. All participants filled in a post-teaching reflective log right after conducting a 40-minute lesson at their observation schools to share their reflections on their use of corpus-based teaching materials they created as a part of the CBLT Training according to the topic, theme, grade, and level of the class they taught. While each participant submitted their post-teaching reflective logs in different weeks, the due date for the submissions was the end of the eighth training session, as the teaching schedules were already over by the week of the eighth training.

Having completed all training sessions, participants received their certificate of attendance to the training at the end of the ninth week following the overall reflection about the CBLT training, training content, and assignments, the effectiveness of consulting corpora for the practicum processes in relation to the implementation corpus-based language teaching activities in EFL classrooms in the last session. The researcher also informed the English teacher candidates about the semi-structured interviews to be conducted to delve deeper into their perspectives and experiences. One-to-one online semi-structured interviews were conducted with ten voluntary participants of the CBLT Training. All 1v1 interviews were conducted on Zoom, initially audio-recorded and later transcribed. The data collection procedure is summarized in Table 3.6.

Table. 3.6. Data Collection Procedure

Week 1	Administration of the Corpus Background Survey to 4 th year pre-service EFL teachers
Week 2	Selection of the participants of the Corpus-Based Language Teaching Training
Week 3	1 st Training Session
Week 4	2 nd Training Session – Reflective Log 1
Week 5	3 rd Training Session
Week 6	4 th Training Session
Week 7	5 th Training Session – Reflective Log 2
Week 8	6 th Training Session
Week 9	7 th Training Session
Week 10	8 th Training Session – Reflective Log 3 & Post-teaching Reflective Logs
Week 11	9 th Training Session
Week 12	Scheduling semi-structured interviews with voluntary participants
Week 13	Semi-structured interviews
Week 14	Semi-structured interviews

3.6. Data Analysis

The primary aim of the current study is to delve into the perceptions held by Turkish English teacher candidates concerning the consultation and implementation of corpora within the framework of corpus-based language teaching pedagogy. As well as investigating how these perceptions evolve and shift before and after participating in the Corpus-based Language Teaching (CBLT) Training, a significant focus of the research lies in exploring the practical implementation of corpus-based teaching materials within EFL contexts, examining the challenges, benefits, and effectiveness perceived by participants. Adopting a comprehensive approach by utilizing a variety of data collection instruments, the study also seeks to shed light on participants' perspectives regarding the training program itself, aiming to discern its effectiveness and relevance in addressing the needs and expectations of the participants. The Corpus Background Survey was administered to find answers to RQ1 to find out the views and perceptions of the English teacher candidates on corpus consultation and implementation into EFL contexts. Analysis of the quantitative data gathered from the two numeric items in the Corpus Knowledge and Literacy section of the Corpus Background Survey involved employing descriptive statistical techniques, including frequencies and percentages. Descriptive statistics were performed using IBM SPSS 29.0 software. The open-ended items of the background survey were simply analyzed by identifying the frequencies of participants' brief responses to the open-ended questions to explore their perceptions of and familiarity with corpora.

The data obtained through reflective logs, post-teaching reflective logs, and semi-structured interviews were qualitatively analyzed in accordance with the principles of the Constant Comparative Method (CCM). The CCM, rooted in grounded theory by Glaser and Strauss (1967), allows researchers to delve into participants' views, ideas, and perceptions inductively by revealing codes and themes that may not initially be evident rather than following any predefined categories or themes (Boeije, 2002; Corbin & Strauss, 2015). Following a cyclical process wherein qualitative data are broken down into smaller units (codes), this method involves systematically and constantly comparing each unit/code to another unit/code so as to identify recurring themes, thereby forming categories grounded in these recurring themes to understand the relationship among the data (Charmaz, 2006; Leong et al., 2010). In line with the principles of the CCM, the following steps were followed in the qualitative analysis: (1) the initial coding and listing of each word or phrase related to the ideas, perceptions, views, and experiences of the pre-service EFL teachers regarding corpus-based language teaching and its practical implementation in English language teaching contexts, and the CBLT Training program, (Allan, 2003), (2) making constant comparisons among the codes while revising the codes that occurred in the data, (3) combining and grouping similar codes while constantly comparing them with one another to represent specific subcategories, and then (4) delineating main categories by the constant comparison of the subcategories with one another, given in the concepts maps taken from MAXQDA 24.0 Software (Appendix I). The CCM was employed using MAXQDA 24.0 software, a robust qualitative data analysis software that enables researchers to analyze qualitative data with its powerful query and retrieval functions and facilitates exploring relationships within the data (Guetterman & James, 2023). Following the transcription of reflective logs and semi-structured interviews, the qualitative data were imported to MAXQDA for coding and analysis. MAXQDA's user-friendly interface initially allowed for the identification of codes and then the delineation of main and subcategories as a result of constant comparison of each code with one another by using the software.

To ensure the validity and reliability of the qualitative data, two independent raters, including the researcher and a co-rater with an MA degree and expertise in ELT and prior experience with qualitative data analysis, were involved in the qualitative data analysis process. Given the large dataset, the co-rater analyzed 30% of the qualitative data collected, which is regarded as an adequate portion of data for measuring inter-rater

reliability (Lombard et al., 2004; O'Connor & Joffe, 2020). In line with Huberman and Miles's (2002) percentage agreement formula $[\text{number of agreements} \times 100 / \text{number of agreements} + \text{number of disagreements}]$, the inter-rater reliability between the two raters was found to be .90, which indicates a high inter-rater reliability. Any disagreements and mismatches between both raters were negotiated and resolved during the negotiation session conducted to reach a consensus. The data analysis procedure is outlined in Table 3.7 in accordance with the research questions and the data collection instruments.

Table. 3.7. *Data Analysis Procedure*

Research Questions	Instrument	Data Analysis Procedure
RQ1: What do Turkish pre-service EFL teachers think about corpora and the integration of corpora into language teaching contexts?	Corpus Background Survey (n=102)	Descriptive statistics A simple analysis of the frequencies of participants' brief responses to the open-ended questions for exploring their perceptions of and familiarity with corpora
RQ2: What are the perceptions of Turkish pre-service EFL teachers regarding the	Reflective Logs (n=60)	The Constant Comparative Method
a. integration of corpora into actual teaching contexts?	Post-teaching Reflective Logs (N=20)	
b. systematic Corpus-Based Language Teaching Training?	Semi-structured Interviews (n=20)	

4. FINDINGS

4.1. Overview of the Study

The primary purpose of the current study was to investigate the perceptions and reflections of English teacher candidates on corpora, corpus-based language teaching and the practical implementation of corpus-informed teaching activities and materials in EFL contexts before and after receiving the 9-week systematic and comprehensive Corpus-Based Language Teaching (CBLT) Training tailored for the specific context of the current study. Based on the relevant need for providing training opportunities for pre-service FL teacher education, this study aimed at designing a corpus training program with a specific focus on hands-on practices of corpus integration into actual teaching contexts aligned with the needs of the learners and FL teaching environments along with the necessary theoretical information on corpus-based language teaching. Moreover, this study explores the views and perceptions of English teacher candidates regarding the training program, which aimed to engage the participants with corpus consultation for designing corpus-informed language teaching materials and integrating these materials into their actual teaching experiences in language classrooms in their practicum schools. Aligned with the scope and the purposes mentioned above, the current study seeks to answer the research questions below:

- 1) What do Turkish pre-service EFL teachers think about corpora and the integration of corpora into language teaching contexts?
- 2) What are the perceptions of Turkish pre-service EFL teachers regarding the
 - a. integration of corpora into actual teaching contexts?
 - b. systematic Corpus-Based Language Teaching Training?

The first set of data was collected from 102 senior pre-service EFL teachers through the Corpus Background Survey to explore Turkish pre-service EFL teachers' background knowledge, familiarity, and views regarding corpora and their experiences with consulting available corpus software and online-accessible corpora in EFL contexts. While the quantitative data gathered from the Corpus Background Survey were analyzed by employing descriptive statistics (i.e., frequencies and percentages), the qualitative data received from the participants' brief responses to the open-ended questions were analyzed by identifying and listing their utterances in a descriptive manner. The second phase of the data collection procedure took place during and after the CBLT Training. In order to explore the perceptions and views of the 20 senior English teacher candidates on the

CBLT training and the implementation of corpus consultation into EFL contexts, the second set of data was collected by reflective logs at the end of the second, fifth, and eighth training sessions and follow-up interviews after the training. Moreover, as the training participants created corpus-based EFL teaching materials to be used in their teaching contexts at their practicum observation schools, they were assigned a post-teaching reflective log to share their reflections on the use of corpus-based teaching materials they designed and used for their language instruction. The data received from the post-session reflective logs, post-teaching reflective logs, and semi-structured interviews were qualitatively analyzed in accordance with the principles of the Constant Comparative Method (CCM) to answer the research questions. Following the analysis of the data collected over fourteen weeks, the findings presented in the subsequent sections related to the familiarity, views, and perspectives of the English teacher candidates regarding the consultation of corpora for English language teaching and their views on the CBLT training program were reached.

4.2. RQ1: Corpus Familiarity, Knowledge, and Experiences of Pre-service EFL Teachers in EFL Contexts

In order to explore the background knowledge, familiarity, and prior experience of pre-service EFL teachers on corpora and its use in EFL contexts, the Corpus Background Survey was administered to 102 senior pre-service EFL teachers doing their practicum at the Ministry of National Education (MoNE) schools. The survey was adapted for the specific purposes and context of the present study by the adaptation of existing surveys in the relevant literature (Bunting, 2013; Çalışkan, 2020; Geluso & Yamaguchi, 2014). The responses to the first section on the participants' demographic information related to their gender, age, and the level/grade they were teaching at their practicum schools were already provided in the methodology chapter section. Focusing on the corpus knowledge and literacy of the participants and experience in using corpus resources in teaching, the two subsequent sections were composed of both open-ended questions and five-point Likert scale items. To assess the participants' familiarity with corpus software and online corpora and their knowledge of using them for various educational and academic purposes, they were expected to evaluate their knowledge by selecting one of the options ranging from *no knowledge* to *expert knowledge*. The participants were also asked to provide explanations to gain further insight into whether

they were familiar with corpus software and online corpora and if they had taken a course or professional development training or had any other exposure to the use of corpora or concordancers.

The quantitative and qualitative data obtained from the Corpus Background Survey were analyzed meticulously to answer the first research question related to the corpus familiarity, knowledge, and experiences of English teacher candidates in EFL contexts. The familiarity and knowledge of the pre-service EFL teachers of available corpora was one of the key foci of the Corpus Background Survey. The two quantitative items of the survey aimed to investigate their **knowledge about online-accessible corpora** (e.g., BNC, COCA, MICASE) and **corpus software** (e.g., AntConc, LancsBox, SKELL). The descriptive analysis of the two items related to the degree of knowledge of the participants is outlined in Table 4.1.

Table 4.1. *The pre-service EFL teachers' knowledge of online corpora and corpus software*

The degree of knowledge	Online-accessible Corpora (e.g., BNC, COCA, MICASE)		Corpus software (e.g., AntConc, LancsBox, SKELL)	
	<i>n</i>	%	<i>n</i>	%
No knowledge	67	66	78	77
A little knowledge	32	31	21	20
Some knowledge	3	3	3	3
Extensive knowledge	-	-	-	-
Expert knowledge	-	-	-	-
Total	102	100	102	100

Table 4.1 shows that only 3% of the 102 Turkish pre-service EFL teachers had *some* knowledge about corpora available online according to their own evaluation of their knowledge. While 66% of the participants stressed they had *no knowledge* about online-accessible corpora, 31% reported that they had *little* knowledge about these online concordancers. None of the pre-service teachers claimed their degree of knowledge about online corpora *extensive* or *expert*. In simpler terms, most participants were unfamiliar with or had little knowledge about online-accessible corpora (e.g., BNC, COCA, MICASE). As can also be found in Table 4.1., %77 of the participants were unaware of available corpus tools and software, and only 20% of the pre-service EFL teachers had *little* knowledge about such concordancing tools. That is, more than three-quarters of the

102 participants were unfamiliar with corpus software, and only 3% of them had *some* knowledge. Just like online corpora, having *extensive* or *expert* knowledge about corpus software was not reported. Overall, the analysis of the two items regarding the knowledge of the pre-service EFL teachers about available corpora showed that most participants were neither familiar with online corpora nor corpus software. Moreover, it was also revealed that the participants were more familiar with online-accessible corpora than concordancing software. In relation to the participants' knowledge of different corpus tools and online corpora, open-ended questions were asked to gain further insights regarding their familiarity and knowledge of the types of corpora. According to the brief responses obtained from the pre-service EFL teachers, 32 participants had prior experience in using COCA, and only 3 participants experienced the use of AntConc.

In order to delve into their experiences and views regarding corpora and its use in EFL contexts, as well as how they became familiar with corpora, four open-ended items were asked to find out whether they had taken a course or professional development training or had any other exposure to the use of corpora or concordancers, and their experiences in using corpora for language learning and teaching. The analysis of the brief responses of the participants to specify how they were informed about corpora revealed that 51 pre-service EFL teachers had been informed about corpora during their teacher training education. Table 4.2. below demonstrates how the participants were informed about corpora.

Table. 4.2. *The results related to the ways of being informed about corpora in EFL*

Sources of knowledge on corpora	<i>n</i>
• Previous seminar	40
• BA education	7
• Academic articles	3
• Workshops	1

As outlined in Table 4.2., 40 participants had been informed about corpora as they had attended the *Corpus Use and Corpus-Based Materials Seminar* conducted by the researcher the year before, which mainly covered background information on corpora and included theoretical details on the use of corpus in language teaching. Concerning how the participants had been informed about corpora, 7 participants reported that they had

been informed about corpora in their linguistics courses during their BA education. Three participants stated that they became familiar with corpora through academic articles, whereas only 1 participant was introduced to corpora in the EFL context via a TÜBİTAK workshop. According to the participants' responses to the open-ended questions, it was revealed that half of the pre-service EFL teachers had been informed about corpora, while the rest had not heard about corpora as a part of their undergraduate-level education or any specific training or seminar.

To further investigate their familiarity and knowledge regarding corpora, the data related to the knowledge of the pre-service EFL teachers about using corpora for various pedagogical purposes (i.e., using corpora for creating corpus-based teaching materials, using corpora to build their own language knowledge, using corpora for designing testing items, using corpora for assessing language learners' language use and using corpora directly with students) in the EFL context were gathered through five quantitative items. The first item (i.e., using corpora to create teaching activities and materials) emphasized the *indirect* consultation of available corpora by consulting online corpora or corpus software to create authentic corpus-based language teaching materials for language instruction. Concerning using corpora for language knowledge and awareness, the second item aimed at investigating the knowledge of the pre-service EFL teachers about using corpora for building their language knowledge and awareness. Moreover, the third item was related to the knowledge and awareness of the participants in the context of using corpora for creating language test items. Item 4 referred to using corpora as a reference tool to check language learners' language use and provide error correction. The last item, using corpora directly with students, emphasized the knowledge of English teacher candidates in relation to the *direct application* of available corpora by hands-on practice by language learners. Overall, all five items focused on the knowledge of the English teacher candidates regarding either the *indirect* or *direct* use of corpora in the EFL contexts. Table 4.3. below shows the pre-service EFL teachers' knowledge of using corpora for different pedagogical purposes.

Table. 4.3. *The pre-service EFL teachers' knowledge of using corpora for different pedagogical purposes*

	Creating corpus-based materials		Building language knowledge		Creating language test items		Checking students' language use		Directly with students	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
<i>Degree of knowledge</i>										

Table. 4.3. (Continued) *The pre-service EFL teachers' knowledge of using corpora for different pedagogical purposes*

<i>Degree of knowledge</i>	Creating corpus-based materials		Building language knowledge		Creating language test items		Checking students' language use		Directly with students	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
No knowledge	72	70,6	64	62,7	98	96,1	76	74,5	93	91,2
A little knowledge	26	25,5	26	25,5	4	3,9	22	21,6	8	7,8
Some knowledge	4	3,9	12	11,8	-	-	4	3,9	1	1,0
Extensive knowledge	-	-	-	-	-	-	-	-	-	-
Expert knowledge	-	-	-	-	-	-	-	-	-	-
Total	102	100	102	100	102	100	102	100	102	100

As seen in Table 4.3., 70,4% of English teacher candidates had *no* idea how to use corpora to create corpus-based materials. The percentage of the participants that evaluated their knowledge of consulting corpora in material preparation as *little* was 25,5%. Only four pre-service EFL teachers reported *some* knowledge about creating corpus-informed language teaching materials, while none declared *extensive* or *expert* knowledge.

Concerning the knowledge about using corpora for building their language knowledge, 62,7% of the participants did not know anything. 25,5% of the pre-service EFL teachers had *little* knowledge about using corpora to improve their language knowledge and awareness. The percentage of participants with *some* knowledge about benefitting from corpora for building language knowledge was 11,8%. Similar to the results of the other four items, none of the participants had *extensive* or *expert* knowledge about using corpora to improve their English language knowledge.

On using corpora for creating language test items, 96,1% of the pre-service EFL teachers had *no* knowledge at all. While none of the participants reported *some*, *extensive* or *expert* knowledge, the percentage of pre-service EFL teachers who had *little* knowledge was 3,9%. With regard to checking students' language use by using corpora, 74,5% of the participants reported that they had *no* knowledge about how to use corpora for this purpose. The percentage of the pre-service EFL teachers with *little* knowledge was 21,6%, while only four pre-service EFL teachers reported *some* knowledge about

using corpora to check students' language use. None of the 102 participants declared *extensive* or *expert* knowledge.

In relation to the use of corpora directly with students, the results of the fifth item revealed that 91,2% of the English teacher candidates were unaware of the ways of consulting available corpora with learners. While 7,8% of the pre-service EFL teachers evaluated their degree of knowledge about using corpora directly with students as *little*, only 1 participant reported *some* knowledge. None of the teacher candidates evaluated their knowledge of corpora consultation directly with language learners as *extensive* or *expert*.

The analysis of the five quantitative items investigating the familiarity and knowledge of the pre-service EFL teachers revealed that most participants had *little* or *no* knowledge about using corpora for various pedagogical purposes in EFL contexts. While *none* of the participants declared extensive or expert knowledge on the use of corpora for such purposes, the percentage of participants who had *no* knowledge about using corpora for creating test items (96,1%) and directly with students (91,2%) was relatively higher than the other different usage of corpora in EFL context. Similarly, the percentage of pre-service EFL teachers who had *no* idea about using corpora for checking students' language use (74,5%) was higher than those who reported *no* knowledge about consulting corpora for creating language teaching materials (70,6%). It was also revealed that the participants had relatively more knowledge about using corpora for building their language knowledge and awareness as prospective EFL teachers.

The pre-service EFL teachers were also asked if they had used corpora for pedagogical purposes to gain insights into their experiences and views on consulting corpora in language teaching contexts. While some participants (20%) reported that they only experienced the use of corpora for basic word queries during the seminar conducted by the researcher the year before, according to their brief responses, only 18 of the 102 pre-service EFL teachers had hands-on experience in using corpora for different purposes other than their practice during the seminar. Regarding using corpora as a reference tool, 9 participants declared that they used COCA and SketchEngine for their writing assignments to check the usage frequency of specific words and phrases. Their responses to the open-ended items also showed that six pre-service EFL teachers consulted corpora to search for collocations and synonyms using online-accessible corpora (i.e., BNC, COCA). One participant indicated that they consulted COCA to analyze the English

language and find politeness strategies by checking the frequency of language use for research purposes. In relation to the consultation of available corpora for EFL teaching, one pre-service EFL reported the use of COCA for developing teaching materials and worksheets for tutoring purposes. Furthermore, it was also revealed that another participant consulted corpora while tutoring a student to see the differences in the language use of specific words and phrases and check the accurate usage of some others.

Concerning the views and perspectives of the English teacher candidates on consulting corpora and corpus-based materials (e.g., corpus-informed dictionaries, teaching materials, worksheets, etc.) in EFL classrooms before the CBLT Training, their brief responses to two open-ended questions were obtained to reveal their familiarity and prior experiences with corpora use for teaching purposes. 35% of the participants reported that they were unfamiliar with corpora and had no idea about its use for teaching purposes. The excerpts below illustrate their lack of familiarity and knowledge about corpora use for teaching:

PST5: *"I have no idea what corpus is or how it can help our teaching experience."*

PST29: *"Unfortunately, I do not have an idea about corpus use for language teaching."*

Similar to the participants above, other participants stated that they had *no idea about* the consultation of corpora and corpus-informed language teaching activities for language teaching purposes. In addition to those unaware of what a corpus is, 22% of the participants indicated that they were aware of it; however, they were not very familiar with its use for teaching purposes. They indicated their interest and openness in attending training to learn more about what a corpus is and how it is consulted for language teaching purposes. The following quotes summarize the interest of the participants in learning more about corpora:

PST1: *"I am not well-informed about corpora, but I am open to learning about it."*

PST27: *"I'm not sure about what it means, but I would be open to using it for teaching."*

All in all, while 57% of the participants reported that they had no idea about how to use corpora for teaching purposes or they were interested in learning about it despite their lack of awareness on how to consult corpora, the rest of the participants indicated their views on the use of corpora for teaching purposes. Their perceptions and views before the CBLT Training included perceived practicality, effectiveness, and challenges of consulting corpora for teaching purposes. Some participant teacher candidates emphasized that consulting corpora provides authentic language across various texts. The excerpts below illustrate the perceived benefits of using corpora for teaching purposes:

PST7: *"I believe that they would provide an authentic source of information as the used language and its frequency would come from daily life."*

PST63: *"It can be beneficial to teach English to students in real authentic contexts."*

In addition to the benefits of using corpora by providing access to authentic language, the brief responses of the English language teachers also revealed the practicality of corpora consultation for creating teacher materials despite their lack of experience and knowledge, as explained above. The quotes below also present their perceptions and views regarding the practicality of using corpora for language teaching:

PST3: *"It can be helpful to prepare different materials for students as most of the English books that students are using have repetitive exercises."*

PST16: *"It could be beneficial to use them in EFL classrooms as with the materials, students can analyze different texts that are for specific needs."*

PST35: *"It is a time-saving and efficient method for teachers while preparing materials."*

Despite the perceived benefits of using corpora by providing access to authentic language and practicality for preparing corpus-based teaching materials, the brief responses of the participants also revealed some challenges of using corpora in language teaching contexts. The excerpts below, derived from the brief responses of the participants, emphasize these challenges:

PST75: *"I honestly had a hard time grasping the use of corpora. I tried to use some of them, but I kept struggling."*

PST94: *"I think it is difficult, and I feel like I am doing fine without it because it is time-consuming."*

In sum, the brief responses of the pre-service EFL teachers to the two open-ended questions revealed that most were unfamiliar with corpora or using corpus tools for language teaching. Some reported they were somewhat familiar with corpora and open to learning more about it. The rest of the participants indicated the perceived benefits and challenges of using corpora for language teaching. They emphasized the practicality of using corpora for creating teaching materials, the benefits of corpora for enabling access to authentic language samples, and the challenges of using corpora in language teaching contexts. The participants' responses revealed that some found corpora complicated and time-consuming for teachers, even with little experience in consulting corpora for teaching purposes.

All in all, the analysis of the quantitative and qualitative data obtained from the Corpus Background Survey revealed that most of the senior, prospective pre-service EFL teachers had *no* or *little* knowledge about corpus software and online-accessible corpora,

as well as consulting such corpora for teaching purposes. The findings revealed that most English teacher candidates were not aware of or knowledgeable about ways of consulting corpora in EFL contexts, such as using corpora for creating corpus-based teaching materials, using corpora to build their own language knowledge, using corpora for creating language test items, using corpora for checking and assessing language learners' language use and using corpora directly with students. Furthermore, the analysis of the brief responses of the participants to specify how they were informed about corpora revealed that 51 pre-service EFL teachers had been informed about corpora during their teacher training education. One of the most prominent findings is that while 40 participants had been informed about corpora as they had attended the seminar conducted by the researcher the year before, only seven pre-service EFL teachers had been informed about corpora as a part of their BA education through other courses. What is more, the students who attended the previous seminars acknowledged a need for a more comprehensive training with hands-on experience opportunities on the integration of corpora into EFL contexts. Thus, these findings regarding the background knowledge and familiarity with corpora underscores the need for systematic and comprehensive training for pre-service EFL teachers on using corpora with hands-on practice and integrating corpus-based language teaching materials into language classrooms.

4.3. RQ2: Perceptions of Pre-service EFL Teachers after the Training

The second research question was addressed to explore the perceptions and views of the pre-service EFL teachers after the systematic, comprehensive CBLT Training that was developed to equip prospective English teachers with the theoretical knowledge and skills related to the practical, hands-on experience of using corpora for actual language teaching. Regarding the perceptions of the English teacher candidates, the second research question encompassed two distinct inquiries:

- a. The integration of corpora into actual teaching contexts
- b. The systematic Corpus-Based Language Teaching Training

In order to answer the second research question with its two sub-questions, qualitative data were obtained through multiple qualitative data collection tools. Regarding the evolving perceptions of the participants throughout the CBLT Training, reflective logs were kept at the end of the second, fifth, and eighth training sessions. As the participants of the training created corpus-informed English language teaching

materials to be used for teaching at their practicum observation schools according to the topic, theme, grade, and level of the class they taught, they were assigned a post-teaching reflective log to share their reflections and perceptions on the use of corpus-based teaching materials in actual teaching. After the training, semi-structured interviews were conducted with 10 voluntary participants. The data obtained through reflective logs, semi-structured interviews, and post-teaching reflective logs were qualitatively analyzed in accordance with the principles of the Constant Comparative Method (CCM) to answer the second research question. The qualitative data findings regarding each sub-question of the second research question are given below for each sub-question.

4.3.1. RQ2.1: Perceptions on the integration of corpora into actual teaching contexts

The qualitative data that addressed the perceptions of the participants on the corpora integration were obtained via reflective logs, post-teaching reflective logs, and semi-structured interviews. The analysis of the qualitative data revealed a total of 929 codes regarding the integration of corpora into actual teaching contexts. Two main categories emerged following the identification of these codes, namely, the benefits of corpora use for language teaching (688 codes) and the *drawbacks of corpora use for language teaching* (241 codes). The main categories are outlined in Table 4.4.

Table 4.4. *The main categories related to the integration of corpora into actual language teaching*

Main categories	N*
Benefits of corpora use for language teaching	688
Drawbacks of corpora use for language teaching	241
Total	929

N*: Number of codes

4.3.1.1. *Benefits of corpora use for language teaching*

The first main category that emerged from the perceptions views of the English teacher candidates on the implementation and consultation of corpora for EFL teaching contexts was the benefits of corpora use for language teaching with seven sub-categories. It was revealed that the prospective EFL teachers found the use of corpora beneficial for *creating language teaching materials* (186 codes), *enhancing vocabulary instruction*

(132 codes), *enhancing grammar teaching* (117 codes), *improving error correction* (86 codes), *accessing authentic language samples* (72 codes) and *accessing contextualized language* (57 codes), and *accessing usage frequency of language items* (38 codes). The sub-categories are displayed in Table 4.5.

Table 4.5. *The sub-categories related to the benefits of corpora use for language teaching*

Sub-categories	N*
Creating language teaching materials	186
Enhancing vocabulary instruction	132
Enhancing grammar teaching	117
Improving error correction	86
Accessing authentic language samples	72
Accessing contextualized language	57
Accessing usage frequency of language items	38
Total	688

N*: Number of codes

As can be seen in Table 4.5, concerning the perspectives and views of the English teacher candidates on the benefits of consulting corpora in and for EFL classrooms, the most prominent and remarkable benefit of corpora use for language teaching was *creating language teaching materials* (186 codes). As the participants experienced corpora practice while designing and adapting EFL teaching materials both for their micro-teachings for the assignments of the training and for their macro-teachings, the actual teaching at their practicum schools, they found the authenticity of corpus-based language teaching materials significant. The excerpts below illustrate how beneficial using corpora was for preparing authentic materials:

PST2: “*The authenticity of the material is very critical, especially while you're preparing your own teaching material from scratch, so using corpora helped me a lot during the practicum.*”

PST7: “*It is nice to design and polish an activity using authentic language from a corpus. Especially while using a corpus to create activities or games to teach the target topic. Students find it lovely and funny.*”

PST8: *“As teachers, we can prepare worksheets using corpus data to create authentic teaching materials, which is significant for enhancing students’ learning and making learning relevant to real-world language use.”*

As reflected in these excerpts, the use of corpora was found beneficial for creating and adapting language teaching materials since consulting corpora ensures access to authentic language samples. As the participants consulted various corpus tools and online corpora for preparing corpus-informed teaching materials for training assignments and their teachings during practicum, they stated that the use of corpora was beneficial for tailoring teaching materials for their specific teaching contexts. PST13 also stressed that the use of corpora provided opportunities for creating teaching activities they had not even considered before.

The majority of the participants also acknowledged that they benefitted from corpora while creating teaching materials, as using corpora provided contextualized language. Since consulting corpora shows the relationship between words and how these words change in different contexts, the participants emphasized that one of the positive sides of using corpora was accessing various examples from different modes, genres, and contexts. The quotes below are related to the benefit of using corpora for creating teaching materials, thanks to access to a huge context:

PST3: *“Using corpora for creating contextual teaching materials for practicum was very beneficial because it was like asking another instructor about how to modify my lesson plans. It was just like that, but even better and much easier to communicate with.”*

PST11: *“I think using corpora has been very helpful and practical for me since I hadn’t spent too much time preparing material and coming up with sentences myself before. I was sure of the authentic use and huge context of the sentences I chose as modal sentences. I was actually amazed by how short a time it took me to come up with what I had thought about.”*

PST17: *“Considering that the meanings of words can change within specific contexts and that words can frequently appear in collocations in the target language, I realized how beneficial corpus tools are in developing materials for students.”*

As can be understood from these excerpts, the participants favored consulting corpora as they accessed language samples across various contexts with their specific collocations and word usage while creating corpus-based language teaching activities. PST11 stressed that it was time-saving to use corpora instead of trying to find a context for teaching a particular language item. In addition to the benefits of using corpora for designing authentic and contextual materials, the participants also emphasized that they

benefitted from consulting corpus tools for designing corpus-based language teaching materials with a focus on teaching language skills, especially vocabulary and grammar teaching. It is worth mentioning that they benefitted from the use of corpora for teaching various vocabulary items (e.g., parts of speech, commonly confused words, collocations) and grammar structures (i.e., conjunctions, modal verbs, prepositions, gerund/infinitive). The excerpts below summarize how the pre-service EFL teachers favored using corpora for creating language teaching material for vocabulary and grammar instruction:

PST8: *"Instead of just getting some fantasy-world vocabulary items and also not authentic materials from the coursebooks, I think we can create our own materials for teaching vocabulary items taken from various contexts in their authentic usage."*

PST10: *"I have learned how to use corpora from scratch. It made me feel like I was doing magic while using COCA and searching for the words for creating materials. I find it really helpful while preparing vocabulary and grammar materials as well as playing word games."*

PST17: *"I gained insights into how corpora can be used for creating materials for grammar teaching, which was a relief since I couldn't come up with various activities on the spot."*

PST19: *"I am grateful for learning about corpora as I am doing my practicum. Using corpora has made material development for vocabulary teaching so much easier and way more accurate than I could achieve on my own."*

These excerpts showed that the pre-service EFL teachers found the use of corpora useful for creating corpus-informed vocabulary and grammar teaching materials. They pointed out that consulting corpus tools was relieving for them while preparing for teaching vocabulary items and grammar structures as they had easy access to how words are used across various contexts. The participants also emphasized that they found using corpora beneficial for creating language teaching materials and activities for teaching specific lexical items and grammar structures as will be elaborated below.

While the majority of the participants expressed the perceived benefits of using corpora for creating materials, two pre-service EFL teachers emphasized that they would not even think about using corpora for developing teaching materials before the training as they had felt corpora would only be used for research purposes. After they received training, their perspective toward using corpora for language teaching shifted. The quotes below illustrate this finding:

PST5: *"I wouldn't think that corpora would be used so broadly for language teaching purposes like material design because I thought corpus is just kind of a research tool that"*

we can only use for research-related purposes. It is so useful now I always use corpora for preparing materials for my practicum class.”

PST18: *“I previously thought that corpora can only be used for linguistics research purposes only. However, learning that it can also be in EFL classrooms and for material design broadened my ideas on the use of corpora.”*

As shown in the quotes above, the use of corpora has been beneficial for creating authentic, contextualized language teaching materials for teaching language skills or creating gamified activities according to the participants. They also highlighted that the training has been beneficial in providing sufficient practice for creating language teaching materials for different language skills. Due to the lack of ready-made corpus-based, authentic, contextualized language teaching materials for EFL teachers’ disposal, equipping prospective EFL teachers with the skills required for using corpora for creating corpus-based language teaching materials is one of the contributions of the CBLT Training.

In addition to the benefits of using corpora for designing language teaching materials, the majority of the participants also reported that using corpora was significant for *enhancing vocabulary instruction* (132 codes). According to the participants of the training, using corpora can be helpful while teaching new vocabulary or designing activities for the practice of previously introduced lexical items, as consulting corpora enables access to authentic language samples. They stressed that using corpora was helpful for teaching lexical items, usage of propositions and collocations, commonly confused words, and in which contexts a lexical item is used. The excerpts below show the participants’ perceptions regarding the use of corpora for vocabulary instruction:

PST2: *“I think corpus tools are the most useful when it comes to vocabulary teaching. This is mainly due to their ability to present countless authentic examples for any given item.”*

PST5: *“I think using corpora is good for finding some striking, authentic sentences as it really buries the vocabulary in the student's mind.”*

PST11: *“When it comes to education purposes, we, as EFL teachers, can see how authentic and contemporary a vocabulary item is so that we can teach our students the real, authentic usage of up-to-date vocabulary items rather than archaic ones.”*

PST16: *“I used concordance lines in COCA for the vocabulary part of my teaching so that I could come up with authentic, better sentences. I was inspired by them but tweaked it for my own class.”*

As can be understood from the excerpts above, the participants found the use of corpora beneficial for vocabulary teaching since consulting corpora provided access to

the authentic, real usage of vocabulary items of the target language. PST5 pointed out that using corpora for teaching vocabulary is highly useful as it enhances vocabulary learning and retaining. While the majority of the participants emphasized how beneficial using corpus is for accessing authentic usage of lexical items, they also found using corpus beneficial in accessing language samples for vocabulary teaching as it shows how a lexical item is used across various contexts. PST3 stated that consulting corpora fostered teaching commonly confused words since the learners accessed the usage of words across their natural contexts. It was emphasized that teaching vocabulary items by using corpora was the most useful feature of corpora consultation. The quotes below show the English teacher candidates' views on the consultation of corpus tools and online corpora for vocabulary instruction with the benefit of accessing lexical items in context:

PST9: *"After our session on the use of corpora for vocabulary teaching, I compared non-corpus-informed vocabulary materials with my corpus-informed vocabulary materials and saw the difference clearly. It was given in a meaningful context."*

PST16: *"In a traditional way of teaching while you're teaching ten vocabulary items, having a basic, traditional organization is just too boring. So, by using corpora, I actually organized my example sentences very well and showed different instances of how these words are used in which context. So, instead of just giving the definition, using corpora allowed me to show some kind of differences for each vocabulary item."*

PST20: *"I think using corpus-based vocabulary materials got the attention of my students on topic and helped them understand the meanings of the words in different contexts. Although they were not familiar with some of the items at first, they had no difficulties later in the other tasks."*

As seen by the statement of PST16, the use of corpora for vocabulary instruction provides access to authentic language samples and lexical items in different contexts with multiple instances of its usage. This enabled both EFL teachers and learners with the opportunity to access the real usage of language across various meaningful contexts. Moreover, the participants also pointed out that using a corpus is beneficial for teaching several particular lexical items and word groups, such as parts of speech, collocations, and confusing words. The excerpts below also emphasize that consulting corpora provides access to how a word or phrase is used with other words:

PST3: *"To show the difference between literal and nonliteral meanings, I have used the 'ADJ vision' feature of COCA to see which meaning of 'vision' is used mostly with adjective combinations. It helped me a lot, and my students loved that activity."*

PST8: *"While I was teaching at practicum, I showed multiple vocabulary items like collocations and confusing words in one context, and I couldn't do it in the traditional*

way, so I found different sentences that are connected to each other. I combined the language usage of these items in paragraphs and asked students to see how the same word is used differently in different contexts with different combinations and collocations. The key feature of corpora is organization."

PST15: *"I made use of corpora in my vocabulary teaching to show students which words come together and how often; I think it was super helpful, especially with collocations and phrasal verbs."*

As can be seen from the excerpts above, the English teacher candidates benefitted from consulting corpora for vocabulary teaching and found consulting corpora useful as this enables access to authentic, contextualized language usage across various contexts for teaching multiple lexical items and word groups like collocations, word combinations, commonly confused words and parts of speech. Therefore, consulting corpora for teaching vocabulary items has been found more beneficial when given in meaningful, authentic contexts rather than given in isolation or in made-up contexts. According to the participants, using corpora for lexical items might be beneficial for their future teaching practices.

In addition to vocabulary instruction, another sub-category related to the benefits of using corpora in language teaching is *enhancing grammar teaching* (117 codes). The participants emphasized that using corpus-informed teaching materials was beneficial for grammar instruction as concordance lines provided authentic, naturally occurring samples of the language to enhance teaching several grammar structures (i.e., *conjunctions, modal verbs, prepositions, gerund/infinitive*). The quotes below are related to the benefit of using corpora for grammar teaching:

PST3: *"Corpora is useful in determining which structure is used, how, and which words are used most frequently so I can avoid certain words that may convey some meaning I did not foresee."*

PST6: *"Compared to vocabulary teaching, I believe grammar materials are more eligible for integration of the corpora as corpora perfectly show certain forms and structures in similar ways in various sentences across different contexts."*

PST7: *"Now I know how beneficial it is to use a corpus to teach grammar structures like conjunctions and modal verbs by looking at concordance lines and using them in class."*

PST9: *"I think by using authentic language samples from corpora, teachers can help students train their weakest muscles in languages. For instance, if a student is experiencing trouble learning the difference in the use of "as" and "because" or prepositions, the teacher can produce a corpus-based authentic material to target the weak grammar points students have."*

As summarized by the participants' statements above, consulting corpora was promising for teaching particular grammar structures such as conjunctions, modal verbs, prepositions, and gerunds/infinitives. PST9 pointed out that using corpus-informed teaching material and activities was beneficial as the grammar instruction is enhanced thanks to authentic language usage by corpora consultation. PST17 further emphasized that consulting concordance lines was helpful for preparing grammar activities that were problematic for their students. Furthermore, the pre-service EFL teachers also stressed that corpora can be used to raise students' language awareness through corpus-based grammar teaching materials as well as their awareness of grammar structures of the English language. The excerpts below illustrate the benefit of using corpora for teaching grammar and raising students' grammar and language awareness:

PST1: *"I believe using corpora can help students gain more command of certain grammar structures, and it will raise their awareness of the target language."*

PST8: *"Structures that will be taught can be provided based on the frequency of occurrence in the language, and materials can be gradually developed to improve and follow the students' grammar proficiency and enhance their language awareness."*

PST15: *"Implementing corpus on grammar teaching materials is really important. But I think corpus tools themselves can also be materials for teaching structures. I think I would use corpora with my students to study grammar and improve their language awareness."*

As clarified in the statements of the participants, teaching grammar can be enhanced by consulting corpora in the form of creating corpus-informed grammar teaching materials and activities and using corpora directly with students to raise their language and grammar awareness while teaching particular grammar structures. Overall, the prospective EFL teachers found corpora consultation highly beneficial as the use of corpora enhances the grammar proficiency of EFL learners thanks to access to authentic, contextual language samples across various genres and contexts.

According to the views and perspectives of the English teacher candidates on consulting corpora for language teaching, consulting a corpus was also found beneficial for *improving error correction* (86 codes). The prospective EFL teachers pointed out that using corpus tools or online-accessible corpora along with corpus-informed activities was highly beneficial for correcting the errors of their students. The participants highlighted that using corpus-based exercises was highly effective, as simply correcting or giving written feedback on their learners' errors did not provide permanent learning. They

stressed that after being engaged with corpus-informed activities, their learners also noticed and corrected their own errors. The excerpts taken from the participants' statements show their perceptions regarding the use of corpora for error correction:

PST1: *"The application of corpora that stood out the most for me was error correction. I find using corpora for correcting my students' errors very significant since providing real language samples for helping my students correct their errors resonates with my role as a teacher and a facilitator."*

PST7: *"As I started to understand how to use corpora, it turned out to be a very fun and useful tool for me now that I can use it for checking the language of my students and correcting their errors."*

PST9: *"Most of the corpus-based activities my peers designed and did micro-teaching on had very specific purposes in that they were focused on their actual students' problems. This proved to me that corpus-based activities are most efficient when it comes to error correction and fixing problems. I think using corpus is great for error correction, especially for my students at my practicum school."*

In addition to the use of corpora by EFL teachers for correcting the errors of their students and common mistakes in the classroom, two participants (i.e., PST9 and PST13) also pointed out the benefits of encouraging the direct use of corpora by language learners to detect and correct their own errors as well as recognize and identify language patterns to infer grammar rules and word usage and draw insights from concordance lines autonomously. The quotes below outline the benefits of the use of corpora by language learners according to these participants:

PST9: *"I believe that encouraging students to recognize their own mistakes in sentences and correct their errors by consulting corpora by themselves can be effective in teaching target vocabulary items and grammar usage, providing a semantic perspective and contextual understanding. Of course, teachers must supervise because it might be hard."*

PST13: *"I didn't know that we could use corpora for error correction and I found it very effective in my own practicum experience. I think I can also use it by showing my practicum students how to use corpora to correct their errors."*

Based on their experiences with using corpora, the pre-service EFL teachers emphasized that it is the most beneficial for designing language teaching materials and vocabulary and grammar instruction. In addition to these, another sub-category emerged related to the benefits of using corpora in language teaching, namely, *accessing authentic language samples* (72 codes). The prospective EFL teachers pointed out that consulting corpora through online-accessible corpora and corpus software can be beneficial for accessing authentic language samples to show the actual usage of the language by its

speakers. According to PST3, their students saw the real usage of English in the sentences provided while teaching vocabulary items, as they stated that a corpus is all about language in use. Since each modern corpus possesses different qualities and features, including type (e.g., general language, specialized, native speaker, learner, etc.), vocabulary size (e.g., large with fourteen billion words, small with thirty thousand words), and the type of texts and genres compiled, they all provide different authentic samples of the target language. The participants also pointed out that corpora can be used to show their students the real usage of language, especially the specific usage of vocabulary items in authentic texts. The following excerpts show how beneficial using online-accessible corpora and different corpus software is for accessing authentic language samples from the perspectives of the pre-service EFL teachers:

PST2: *"I think using corpora was like using Control + F (find command in computers) to find whatever I wanted to find in the whole English language. It was actually even more than that because we had access to authentic language samples."*

PST4: *"One of the most beneficial aspects of corpus tools is their authenticity. They directly showcase the usage of words, along with higher-order structures in a language."*

PST8: *"I am so amazed by the authenticity of corpora and variety of language usage, for example, for the phrases, nouns before or after a verb, frequencies, etc."*

PST11: *"Seeing all the possible usage of words and authentic language data must be a chest of resources for teachers, and we should know how to invest them."*

Moreover, they also pointed out the benefit of using corpus-based language teaching materials and textbooks after evaluating and comparing one corpus-based and one non-corpus-based (MoNE) textbook based on two checklists provided by the researcher, which were already familiar to the participants from another undergraduate course based on material adaptation and development as mentioned in the methodology chapter. They evaluated and compared both textbooks regarding their general impressions, appropriacy of language items and reading materials, authenticity and contextuality of the language content, and language usage through activities for language skills, sub-skills, and phonology. The findings revealed that the participants found corpus-based textbooks more authentic and practical to be used in the EFL context. The quotes below illustrate the participants' perspectives on the use of corpus-based textbooks over non-corpus-based textbooks:

PST3: *"After comparing a corpus-informed textbook with a non-corpus-based one, I think corpus-informed teaching materials and textbooks are more authentic and practical to be used in language classrooms with students."*

PST13: *"It was observable to me that corpus-based textbooks were more about communication, using more widely used words with authentic language samples, compared to MoNE books with no authentic material and lots of passive vocabulary."*

PST17: *"I am confident that corpus-informed textbooks seem to be more beneficial, especially in terms of more focus on authentic language usage and communicative competence for language learners."*

As reflected in the statements of the pre-service EFL teachers, they found using corpora beneficial for language teaching in actual classrooms because consulting corpora provides access to authentic language samples for language students and teachers. They also added that using corpus-based language teaching materials and textbooks provides authenticity to language instruction and practicality for EFL teachers.

Another sub-category focusing on the benefits of consulting corpora in EFL contexts was *accessing contextualized language* (57 codes). The pre-service EFL teachers reported that the use of corpora was useful for language instruction as the accessed language was full of contextualized language data. It was stated that using corpora shows the relationship among words and how their usage change across different contexts, and the pre-service EFL teachers emphasized that consulting corpora was useful for accessing various examples from different genres and contexts. According to PST9, taking language samples from a corpus and adapting them for teaching purposes for their own context was valuable as consulting corpora provided a better context and more authentic language usage than a chatbot may create. They stressed that showing different usage of nouns, verbs, idiomatic expressions, and phrasal verbs was beneficial for language learners. The excerpts taken from the statements of the participants outline their perspectives on the benefits of using corpora for contextualized language samples:

PST10: *"I'm so thankful that context and genre selection exist in COCA as I had a lot of times when I was trying to teach formal and informal language, and the choice of genre was hard to teach. But using corpora for that made that a breeze for me as a teacher."*

PST12: *"Seeing different usage of a verb or phrasal verb across various contexts and mediums is a beneficial thing for any learner of general English."*

PST19: *"I loved using corpora to provide contextual language as a major part of teaching vocabulary is to provide a context for language learners. One of the activities I designed included eliciting an item from the context provided with different sentences taken from concordance lines. They loved it."*

PST20: *"I could use concordance lines to show students some real-life sentences taken from a corpus and make them guess which words are used in an idiomatic way, etc."*

Additionally, two pre-service EFL teachers (i.e., PST14 and PST18) stressed that using corpus for language instruction is also beneficial as having access to contextualized language also enhances the learning of different meanings of words in various contexts. The excerpts below focus on the benefits of accessing contextualized language samples through corpora in terms of teaching different meanings of words:

PST14: *"Using corpus can be very beneficial in terms of seeing how the words can be used in different contexts with different meanings. For instance, checking how words' usage differs (e.g., literal, metaphorical meaning) in every sentence in different contexts can be observed through corpus use."*

PST18: *"Oftentimes, when a student comes to a teacher with a question like 'Can we use this word here?' or 'How do we use this word?' the teacher is unable to answer. With the help of a corpus, the teacher would be able to show the different meanings and usage of words in real life in different contexts."*

In addition to the benefit of accessing contextualized language, another sub-category of different benefits of corpora use in language teaching that emerged was *accessing usage frequency of language items* (38 codes), as shown in Table 4.5. The prospective teachers of EFL pointed out that using corpora was beneficial for language teaching as the frequency information of each specific word or phrase provides insights and benefits into the selection of vocabulary items and grammar structures by EFL teachers. The following excerpts show how they found using corpora beneficial thanks to accessing the usage frequency of specific language items:

PST4: *"I'm impressed by getting mathematical frequency information of different words and collocations and getting different results for each specific search. Using corpora is beneficial for choosing what to teach."*

PST12: *"Using corpora can provide comprehensive information about the word use frequency in spoken or written language, and this information can give a clue to educators about what they should teach in accordance with real language usage."*

PST18: *"I really enjoyed learning how corpus could be used as a frequency indicator, helping the decision of the use of idioms or real meanings or expressions."*

As shown in the statements of the participants above, it was revealed that prospective EFL teachers found using corpora useful for accessing the frequency information of authentic language items and their usage across various contexts. All in all, the findings related to the benefits of consulting corpora in EFL contexts showed that using corpora in language classrooms is highly beneficial, creating corpus-based language teaching materials for teaching lexical items and grammar structures and improving error correction. The perceptions of the prospective EFL teachers revealed that

integrating corpora into language classrooms would benefit both EFL teachers and learners. The participants reported that consulting corpora for language teaching equips language teachers with a great skill of accessing the actual usage of the language as well as reaching authentic and contextualized language items to be used in language classrooms.

4.3.1.2. Drawbacks of corpora use for language teaching

The second main category that emerged from the perspectives and views of the English teacher candidates on the consultation and use of corpus tools in actual language teaching was the drawbacks of corpora use for language teaching with four sub-categories. The four sub-categories showed that the prospective EFL teachers found some drawbacks of using corpora for language teaching, namely, *difficulty in finding appropriate language items* (75 codes), *limitations for low proficiency level students* (68 codes), *technical problems with different concordancers* (56 codes) and *difficulty in using search functions* (42 codes). The sub-categories are outlined in Table 4.6.

Table 4.6. *The sub-categories related to the drawbacks of corpora use for language teaching*

Sub-categories	N*
Difficulty in finding appropriate language items	75
Limitations for low proficiency level students	68
Technical problems with different concordancers	56
Difficulty in using search functions	42
Total	241

N*: Number of codes

As seen in Table 4.6, the perceptions of English teacher candidates on the challenges and drawbacks of corpora consultation in EFL classrooms revealed that the most notable drawback of corpora use for language teaching was the *difficulty in finding appropriate language items* (75 codes). According to the experiences of the participants with various corpus tools and online-accessible corpora during the training, they encountered difficulties related to finding suitable language items and search results while

using corpora. The excerpts below illustrate the difficulties they experienced while using corpora to find appropriate language items through concordance lines:

PST3: *"It was hard for me sometimes to find the right samples of the language I was looking for through concordance lines while searching on COCA, yet I got used to it by practice."*

PST6: *"The difficulty of using corpora for me was that as most of the corpora I used are not annotated, you need to pick the adverbial usages, for instance, on your own. This took a long time for me."*

PST9: *"As I was doing the tasks and designing materials for the assignment after the session, I had a hard time finding what I needed in the concordance lines."*

As can be seen from the statements of the pre-service EFL teachers above, they had trouble finding the suitable language items they were looking for through concordance lines as there were too many sentences and samples of the language to be chosen from. One participant stressed that practicing more helped in finding ways to overcome this difficulty, while another reported that this difficulty made using corpora time-consuming for particular instances. Some participants also pointed out that there might be some *incorrect* usage samples of language items as the language used is authentic and taken from the actual written and spoken corpus of the whole English language. The quotes below are related to the difficulty of finding appropriate language samples due to incorrect language items in some of the concordance lines:

PST1: *"There are some questions about corpora, especially COCA, in my mind. For example, in COCA, we see a large data set, but we do not know which of these usages of language is "correct" because they need to be corrected or adapted to be used in my practicum class."*

PST6: *"One should be careful while using corpora since it may sometimes be "too" authentic that some "incorrect" usages can be seen."*

PST17: *"It takes time to manually eliminate the bad strings, but it is worth it because it helps me generate authentic and contextual materials for my practicum students."*

In addition to the participants who reported the difficulties they experienced while searching for appropriate language items due to incorrect and raw language samples, some participants also stressed that there is a need for adaptation for some concordance lines, as the language samples are in the raw form, and they might not be appropriate for classroom use. The excerpts below summarize how the prospective EFL teachers pinpointed the difficulties they experienced and the need for adaptation for some language items accordingly:

PST4: *“Using corpora is perfect for extracting natural, real, authentic sentences however, there is a need for intense adaptation for some concordance lines, so the selection of language items is very important.”*

PST9: *“It’s not like we are taking a sentence directly from corpora and presenting it to class, but we also need to adapt as we do with every other material.”*

PST17: *“I think corpus is a great tool because it gives thousands of examples, although it needs to be manually checked for each concordance line.”*

PST19: *“Sometimes, I felt some trust issues with some corpus tools, and while I was preparing the material for teaching, I couldn't get all the examples from corpora, so I had to double-check every time.”*

As shown in the quotes above, the participants experienced some difficulties while deciding which language items to take from concordance lines as they found some raw sentences that needed adaptation for their context and the needs and levels of their students. Despite the difficulties they encountered, the pre-service EFL teachers reported that more practice in using corpora and manually checking the concordance lines helped them to overcome these difficulties.

Another sub-category related to the drawbacks of consulting corpora and corpus-based teaching materials in EFL contexts was *limitations for low proficiency level students* (68 codes). The participants emphasized that corpus-informed teaching materials and activities can be too hard for low proficiency level learners, and the time and effort required to create appropriate activities for these learners is impractical. Despite their positive perceptions related to the use of corpora for creating language teaching materials for their classrooms and experiences of using such materials with low proficiency level students in their practicum context, the pre-service EFL teachers highlighted that the required language awareness and knowledge for making the best use of corpus materials cannot be met by many low proficiency level learners and it takes time and effort to implement corpus data into such activities for these learners. The following excerpts are related to the perceptions of the participants regarding the limitations of using corpus-based activities with low proficiency level students:

PST3: *“I have a difficult time finding authentic examples for lower proficiency levels because the authentic use of English in COCA is not always appropriate for my students at practicum.”*

PST6: *“I can say that using corpora is not manageable for low-level learners because it would be hard to implement materials into their lessons.”*

PST11: *“While preparing a material, we cannot get stuck in one corpus tool like COCA because sometimes they include very high-level sentences for my low-level students.”*

PST20: *"My biggest challenge was finding the correct, right, and most suitable sentences for my low proficiency level practicum students."*

While some participants emphasized the limitations of using corpus-based activities with low proficiency level students, others pointed out how they overcame these limitations. It was revealed that using different corpus tools and online-accessible corpora for different proficiency-level students helped the participants overcome the limitations they experienced while selecting suitable language items for their students with low English proficiency. The quotes below illustrate how three pre-service EFL teachers came up with some ideas for overcoming such limitations:

PST6: *"Thanks to the variety of corpus tools and online corpora we practiced, I didn't have trouble with my low proficiency level students at practicum as I know that we must adapt the language usage according to our students' needs and proficiency levels as well as consulting different corpus tools."*

PST9: *"I have figured out that I can use different corpus tools for my different proficiency level students at practicum. While SKELL is more practical for students with lower levels of proficiency as the sentences are shorter and simpler, COCA and MICASE are more appropriate for my more advanced level learners."*

PST17: *"Considering the level of my low proficiency level practicum students, I had to actually adapt the concordance lines I found from different corpora."*

Another sub-category of the drawbacks of using corpora for language teaching was *technical problems with different concordancers* (56 codes), according to the pre-service EFL teachers. While most participants enjoyed using COCA both for the in-class tasks and assignments of the CBLT training and for their own studies outside the session hours, they reported multiple difficulties and problems they encountered while consulting COCA. As specified in the methodology section, the participants of the study were introduced to multiple corpus software (i.e., AntConc, AWL Highlighter, Compleat Lexical Tutor, CorpusMate, Just the Word, LancsBox, SKELL, Skylight, and Tim John's Kibbitzers) and online-accessible corpora (i.e., BNC, COCA, MICASE, NOW, and WebCorp Live). Despite the number of different corpus tools demonstrated and practiced, the participants favored using COCA more than the others. Below are the quotes taken from the statements of the English teacher candidates regarding the technical problems they encountered while using COCA:

PST2: *"I think one of the most disturbing challenges of using corpora for me was that COCA kept requiring a premium account after 20 queries a day and showed ads all the time."*

PST6: *"Last week, I tried to design an activity for my students with COCA, and I exceeded the limit right half of the activity, and I could not complete it. I was enjoying it anyway, but I wish I could have finished it."*

PST9: *"In COCA, I had to reset my search before my next search, and each time that there was an error, I felt like, 'Oh my god, what did I do wrong?'"*

PST11: *"COCA was easy to surf through despite some difficulties regarding subscription, search limit, and ads."*

As can be seen from the excerpts above, the participants experienced problems while navigating COCA due to technical aspects of the specific corpus. The participants not only had technical problems with using COCA but also with other corpus tools as they found the interface of most corpora primitive and *user-unfriendly*. The excerpts below illustrate the views of the prospective EFL teachers about consulting the unfriendly and primitive interfaces of such corpus tools and online-accessible corpora:

PST4: *"I think most corpora are probably developed to be user-unfriendly as much as possible."*

PST9: *"I encountered difficulties as most corpus software and online corpora are not user-friendly due to their interface and primitive nature as a website."*

PST16: *"I find using some corpus tools hard, mostly due to their interfaces being sort of a mess."*

Despite the technical problems they experienced while using corpus tools and online-accessible corpora due to the query limit or their primitive, user-unfriendly interfaces, the participants stated that they found using corpora enjoyable and fun as they overcame these problems by practicing more and getting used to their interface. One participant pointed out that having the skill to navigate as many corpus tools and online corpora as possible is a must for EFL teachers to overcome such problems, as given in the excerpt below:

PST3: *"Since the COCA demands premium subscription and sticking out to one corpus tool is not the best idea for teachers, learning about more tools will be quite helpful in the future while creating corpus-based language teaching materials."*

In addition to the technical problems they encountered while using corpus software and online corpora, the participants' perceptions regarding the drawbacks of using corpora for language teaching were shaped by another sub-category, namely, *difficulty in using search functions* (42 codes). As the last sub-category of the perceived drawbacks of using corpora, the pre-service EFL teachers reported that they experienced difficulties while making queries on various corpus tools and online corpora. The excerpts

taken from the statements of the participants below outline the difficulties they encountered while consulting corpora:

PST4: *“It was a bit challenging to use corpora until I got used to the searching methods and specific wording for search, but then I got very excited to use it in my teachings as I learned how to do specific search, especially on COCA.”*

PST6: *“We as a group had some difficulties while working on some corpus tools as it was a bit confusing to learn the search functions while searching for what we were looking for.”*

PST9: *“I had difficulties remembering and using which function was going to show me which result. Even leaving a blank would change the results.”*

PST19: *“The only thing that I had a hard time with was determining which corpus tool I should use for every specific query and how to do each search as each tool serves different purposes with different search syntax even though they still have so many features in common.”*

As the statements of the participants point out, they experienced challenges and difficulties with corpora use related to the nature of the corpora they consulted for and the context they were teaching at their practicum observation schools. The findings revealed that the majority of the codes regarding the drawbacks of corpora use for language teaching were obtained via the 3rd reflective log, post-teaching reflective logs, and semi-structured interviews as the participants were assigned the design and implementation of corpus-based language teaching materials for their assessed practicum teachings. The distribution of the codes regarding the drawbacks of corpora use for language teaching is outlined in Table 4.7.

Table 4.7. *The distribution of the codes regarding the drawbacks of corpora use for language teaching*

Data Collection Instruments	N*
Reflective Log 1	9
Reflective Log 2	22
Reflective Log 3	86
Post-teaching Reflective Log	57
Semi-structured Interviews	67
Total	241

N*: Number of codes

As can be seen in Table 4.7, it was found that their perceptions regarding the drawbacks of the use of corpora evolved throughout the training, and they expressed the highest number of drawbacks in the 3rd reflective log. It is revealed that the participants experienced most of the challenges and drawbacks of corpora consultation towards the end of the training as they were assigned the highest number of assignments for the last three sessions of the training. Considering the individuals who expressed the highest number of drawbacks of corpora use, the findings showed that PST3, 6, and 9 were affected by the challenges of corpora use in terms of difficulties with corpora and different concordancers and limitations while designing teaching materials for their low proficiency level learners. It can be summarized that these three prospective EFL teachers experienced more difficulties than the other participants towards the end of the training due to the number of tasks and assignments they were assigned as well as their practicum teaching experiences with the integration of corpora.

All in all, the findings related to the perspectives and views of the English teacher candidates revealed multiple benefits and drawbacks of consulting corpora in EFL teaching. While the majority of the perceptions stated that the use of corpora is highly beneficial for creating corpus-based language teaching materials and activities for teaching vocabulary and grammar and correcting their students' errors, they also stressed that consulting corpora for language teaching equips language teachers with the skill of navigating different available corpora to access authentic and contextualized language items to be used in language classrooms. Despite the perceived benefits of using corpora for language teaching, the pre-service EFL teachers also experienced some drawbacks based on the technical issues they encountered while searching for specific language items due to the nature of different corpus tools and online-accessible corpora. Moreover, they also experienced the limitations of using corpora with low proficiency level students. However, for each sub-category of the drawbacks they experienced, the participants also reported how they overcame the problems they encountered.

Overall, most prospective EFL teachers stated that they enjoyed consulting various corpus tools and online-accessible corpora in EFL contexts and that they are also willing to consult corpora and implement corpus-based teaching activities in their actual teaching classrooms as well as the classrooms they taught English by utilizing corpora in their practicum schools. Despite the drawbacks they encountered, the perceptions of the majority of the participants were positive toward using corpora in language teaching.

4.3.2. RQ2.2: Perceptions regarding the CBLT training

The second sub-question of the second research question was addressed to explore the perceptions and views of the English teacher candidates regarding the CBLT training. Regarding the evolving perceptions of the participants about the training program, reflective logs were kept at the end of the second, fifth, and eighth training sessions, and semi-structured interviews were held with 10 participants. The qualitative data were obtained via reflective logs and semi-structured interviews to delve into their experiences throughout the CBLT training. The analysis of the data on the perceptions and views of the participants on the systematic CBLT Training revealed 308 codes in total. Two main categories emerged following the identification of these codes, namely, the *strengths of the CBLT training* (234 codes) and the *weaknesses of the CBLT training* (74 codes). The main categories are outlined in Table 4.8.

Table 4.8. *The main categories related to the systematic corpus-based language teaching training*

Main categories	N*
Strengths of the CBLT training	234
Weaknesses of the CBLT training	74
Total	308

N*: Number of codes

4.3.2.1. *Strengths of the CBLT training*

The first main category based on the perspectives and views of the English teacher candidates about the systematic Corpus-Based Language Teaching (CBLT) Training was the strengths of the training with five sub-categories. The five sub-categories showed that the prospective EFL teachers found the systematic CBLT training significant for *equipping EFL teachers with the required skills for consulting corpora* (79 codes), *providing opportunities for professional development* (48 codes), *building their own corpus* (43 codes), *raising EFL teachers' language awareness* (38 codes) and *creating a reflective and collaborative environment* (26 codes). The sub-categories related to the perceived strengths of the CBLT according to the pre-service EFL teachers of the training are displayed in Table 4.9.

Table 4.9. *The sub-categories related to the strengths of the systematic corpus-based language teaching training*

Sub-categories	N*
Equipping EFL teachers with the required skills for consulting corpora	79
Providing opportunities for professional development	48
Building their own corpus	43
Raising EFL teachers' language awareness	38
Creating a reflective and collaborative environment	26
Total	234

N*: Number of codes

As can be seen in Table 4.9 above, concerning the perspectives and views of the English teacher candidates regarding the strengths of the CBLT training, the most prominent strength of the training the pre-service EFL teachers benefitted from was *equipping EFL teachers with the required skills for consulting corpora* (79 codes). The participants expressed that they gained the skills needed for consulting corpora for language teaching purposes as they attended the CBLT training and completed the in-class tasks and assignments for micro-teachings during the session hours and macro-teaching experiences they had at their practicum schools for teaching EFL learners. The CBLT training was found effective in teaching the *know-how* of consulting corpora to the prospective EFL teachers. The excerpts below illustrate their perceptions related to the benefits they gained by attending the training in terms of being equipped with the required skills for navigating different corpus tools for creating materials and using corpora directly with their students:

PST3: *"Nowadays, teachers have to be pretty eclectic, and it's great to have such a wonderful tool at our disposal so that we can decide on our own while creating teaching materials."*

PST5: *"I believe every teacher must learn how to use corpora as I have gotten help from corpora in both of my teachings in practicum, and I believe with this skill I have now, I will always have one tab with a corpus tool open in my browser because I believe it is 'the' real source of a language."*

PST9: *"I personally believe that EFL teachers must learn about corpora for their teaching as corpus tools cannot be replaced with any other tool (like AI, etc.) in terms of scanning and accessing authentic language samples and texts (like in the AntConc)"*

and searching a word with specific filters. This training helped me get the skill of how to do that.”

PST14: *“I feel more competent about using corpora for language teaching and more confident as a language teacher now that I have gained a different perspective on the English language and skill for using such a great helper for teachers.”*

As can be seen through the excerpts taken from the statements of the participants, they have gained the skill for navigating different corpora for language teaching purposes. PST19 further pointed out that they became more proficient corpora users throughout the training. It was also added that they kept looking forward to the next sessions of the CBLT training to practice corpora consultation more. Furthermore, the pre-service EFL teachers also reported that the CBLT training was beneficial in *providing opportunities for professional development (48 codes)* for the participants. Expressing their gratefulness, they indicated that they hadn't been informed about corpora as a part of their undergraduate education. Participants stated that they felt privileged to be a part of the training in their undergraduate education during their practicum process as they were provided with the opportunity to work with corpus tools. The excerpts in relation to this finding are given below:

PST1: *“Honestly, I am grateful for this training as it's one of those things that will separate me from the other applicants when we're looking for a job because having this type of training is not exactly the most common thing in our department and a lot of workplaces will fancy that.”*

PST3: *“This was a genuinely wonderful opportunity to learn about corpora during my practicum. I wish I'd known about corpora before, as nobody told me in my whole four-year BA education. I fell in love with using corpora.”*

PST8: *“I think it was very valuable for us to attend such training during my undergraduate years. Because I know that most teacher candidates are not informed about corpora as well as we are currently, as I know from my friends in our department.”*

PST11: *“I feel really privileged to learn corpus in my undergrad years as a future EFL teacher. I'm really grateful for the opportunity as this training was perfect for my career and professional development.”*

PST17: *“I'm genuinely really happy that I got to learn about it because it's not something that you could just Google; one could read articles or watch YouTube videos, but it would not be like this training because we had the ability to ask someone who knows about corpora, how to or how not to do something. I really appreciate this opportunity as a future English teacher.”*

The participants stressed that the CBLT training was an irreplaceable part of their practicum process as they benefitted a lot from the use of corpora for teaching English to

their students in the practicum contexts. The prospective EFL teachers stated that they were grateful to be equipped with the skill of using corpora and to have the chance to build their own corpus. Another sub-category related to the strengths of the CBLT training was *building their own corpus* (43 codes). Most participants acknowledged that they benefitted from consulting online-accessible corpora, while others enjoyed and benefitted from creating their own corpus more. The participants pointed out that having hands-on experience with building their own corpus was enlightening as they experienced the use of such software for creating language teaching materials from scratch for the first time. The excerpts below illustrate the perceptions of the participants regarding the benefit of the training for building their own corpus:

PST6: *"I think the most interesting part of corpora for me was creating my own corpus with AntConc as it was very useful for it allows us to create a corpus by uploading any text, script, or book to detect and analyze concordance lines for real-life usage of a certain word or phrase for language teaching and research purposes."*

PST8: *"I enjoyed creating my own corpus and having the desire to at least go and do something for myself out of nothing, out of scratch as a language teacher. While using online corpora, we are provided with already derived compiled corpora, but by using corpus software, you can create your own corpus, and this is amazing."*

PST14: *"I loved working on AntConc so we could create our own corpus with different texts, and we could work on it to analyze the language. This was the best part of the training for me."*

PST20: *"The AntConc was so fun to work with, and it looked simple in design, which I liked more than Coca. It was more fun learning using AntConc for me."*

As can be seen through the excerpts taken from the statements of the participants, they favored using AntConc over COCA since building their own corpus was interesting and fun, as well as being practical and useful for some participants. One participant also stressed that building a corpus was the best part of the CBLT training. The pre-service EFL teachers also indicated another strength of the CBLT training, which was *raising EFL teachers' language awareness* (38 codes) for the participants. The participants emphasized that using corpora for in-class tasks and assignments and creating language teaching materials were eye-opening for them. It was revealed that consulting corpora opened new doors for them in the English language as they learned more about how the language works. The quotes below highlight this finding:

PST6: *"Using corpora during the training actually helped me understand the multidimensionality of English. I feel like I learned a lot about the language. "*

PST7: *"It was eye-opening for me to learn that much about the language while learning how to use corpora for teaching language vocabulary. I see teaching with a different eye now as I am more knowledgeable about the language I am teaching."*

PST13: *"What stuck out to me the most about corpora use is how easy it is to find answers to questions I couldn't come up with myself while teaching English as an EFL teacher. This training has definitely raised my awareness and knowledge about the language I thought I was teaching the best."*

As can be seen from the excerpts above, the participants stressed that working on the language by consulting corpora taught them *about* English as they felt more knowledgeable about the multidimensionality of the language. While some pre-service EFL teachers found the CBLT training eye-opening about the English language, others reported that the CBLT training was beneficial in *creating a reflective and collaborative environment* (26 codes). The majority of the in-class tasks and assignments required the participants to work in groups of two or three to experience the use of corpora in a collaborative atmosphere so that peers would learn from each others' struggles and accomplishments while learning how to navigate a corpus and how to consult corpora for creating language teaching materials. The excerpts below show the last perceived benefit of the CBLT training according to the pre-service EFL teachers:

PST2: *"Sometimes, I struggled while preparing materials, but by working with my groupmates to learn about corpora, I overcame such instances."*

PST5: *"It was a collaborative training as we worked together for various tasks and assignments. I think learning starts here."*

PST11: *"During the training, you gave us a space to learn more about the tools and corpora in a reflective and collaborative environment. I learned a lot from my peers."*

As can be seen above, the participants stated that the collaborative and reflective environment throughout the training enhanced the overall success of the training in equipping them with the required skills for corpora consultation. All in all, the perspectives and views of the English teacher candidates about the CBLT training revealed multiple strengths and benefits. The participants reported that the training was successful in equipping them with the required skills for consulting corpora, providing opportunities for their professional development, building their own corpus, raising their language awareness, and creating a reflective and collaborative learning environment. The participants stressed that the CBLT training was beneficial in equipping them with the skill of consulting corpora for language teaching, and they were grateful for the opportunity to attend this training during their undergraduate education.

4.3.2.2. Weaknesses of the CBLT training

The second main category based on the perspectives and views of the English teacher candidates about the systematic Corpus-Based Language Teaching (CBLT) Training was the weaknesses of the training under two sub-categories: *too much focus on production* (47 codes) and the *need for more practice* (27 codes). The sub-categories related to the perceived weaknesses of the CBLT according to the pre-service EFL teachers of the training are outlined in Table 4.10.

Table 4.10. *The sub-categories related to the weaknesses of the systematic corpus-based language teaching training*

Sub-categories	N*
Too much focus on production	47
Need for more practice	27
Total	74

N*: Number of codes

As reflected in Table 4.11, the perspectives and views of the English teacher candidates on the weaknesses of the CBLT Training uncovered that the participants experienced difficulty due to too much focus on production by using corpora. At the same time, there was a need for more emphasis on practice with corpus tools. The most remarkable weakness of the training was *too much focus on production* (47 codes). According to the experiences of the prospective EFL teachers with available corpora throughout the training, they encountered difficulties while creating language teaching materials since they believed the training was too loaded with tasks and assignments. The excerpts below are taken from the statements of the participants regarding the course load of the training:

PST2: *"We focused too much on production, and the training was too loaded with assignments."*

PST10: *"I believe there could be fewer assignments because, at the end of the training, it became harder for me to create more materials."*

As emphasized by the pre-service EFL teachers above, they had trouble with the number of assignments and tasks they had to complete throughout the training, and they

found the content of the training too loaded with compulsory assignments. Moreover, they also stressed the *need for more practice* (27 codes) as they found it hard to keep up with the number of materials they were to design. The participants noted the need for more intense training sessions focusing on using different corpus tools. The quotes below show the perceptions of the participants regarding the need for more practice throughout the training:

PST4: *"I would like to use a different corpus tool in each session instead of creating teaching materials and other assignments. We could have practiced the tools together more."*

PST12: *"I would love to have more intense training sessions on practicing corpora and hands-on use."*

PST16: *"I wish I could have gotten more practice with you, not creating teaching materials and presenting them to my peers. Maybe fewer people could present, and we could have focused on more practice with corpus tools."*

PST20: *"If we had more time, we could have been introduced to more tools in detail, and we could have used them right after to produce materials."*

As seen above, the participants reported a need for assigning particular corpus tools and online corpora for particular assignments despite the number of corpus tools and online-accessible corpora that were introduced, demonstrated, and required to be used for creating language teaching materials.

All in all, the findings related to the perspectives and views of the English teacher candidates about the systematic, comprehensive Corpus-Based Language Teaching Training revealed multiple strengths and weaknesses of the training. The majority of the perceptions of the participants showed that they found the CBLT training highly beneficial for equipping them with the required skills for consulting corpora, providing opportunities for their professional development, building their own corpus, and raising their language awareness, as well as creating a reflective and collaborative environment. In addition to such perceived strengths of the training, some pre-service EFL teachers also stressed that the training was too loaded and there was too much focus on production with the number of tasks and assignments. They also indicated a need for more practice with corpus tools rather than having that many assignments during the training. Overall, the perceived benefits of using corpora for language teaching and the perceived strengths of the CBLT training showed that most pre-service EFL teachers benefitted from the CBLT training and the consultation of corpora in EFL contexts for actual teaching practices at their practicum schools.

5. DISCUSSION

The current study aimed to investigate the knowledge, experiences, and perceptions of English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts before providing them with the systematic and comprehensive Corpus-Based Language Teaching (CBLT) Training tailored for the specific context of the current study. Furthermore, during and after the implementation of the CBLT Training, the aim of the present study was to investigate the views and perceptions of the English teacher candidates on the consultation of corpora for EFL teaching and the CBLT Training. The findings obtained at each stage of the study are discussed in this section based on relevant literature as well as the observations and experiences obtained during the study process.

Based on the analysis of the data obtained in the initial phase of the study from 102 pre-service EFL teachers regarding their perceptions, knowledge, and experiences with corpora prior to the training through the Corpus Background Survey, the overall familiarity of the participants with available corpora was very low. The findings revealed that most prospective EFL teachers were not familiar with available corpus tools and resources (e.g., AntConc, LancsBox, SKELL) or online-accessible corpora (e.g., COCA, MICASE, BNC). Almost two-thirds of the pre-service EFL teachers evaluated their knowledge about such corpora as *none*, while the rest of them reported having *little* knowledge, and only a handful of them evaluated their knowledge of available corpora as *some*. Despite the low familiarity and knowledge about available corpora due to their lack of awareness regarding the use of corpora for language instruction, it was revealed that the pre-service EFL teachers were more familiar with online-accessible (i.e., COCA and BNC) corpora than concordancing software. These findings, obtained through the Corpus Background Survey, aligned with the empirical studies that employed the surveys adapted for the present study's specific context (Bunting, 2013; Çalışkan, 2020; Geluso & Yamaguchi, 2014) as they concluded that only a handful of participants had little knowledge and familiarity about corpora. These findings may stem from the lack of focus on the integration of corpora into language teaching in the teacher training curriculum since there is not a specific course in the study context that introduces corpora to prospective EFL teachers.

The investigation regarding the familiarity and awareness of the pre-service EFL teachers about using corpora for various pedagogical purposes (i.e., using corpora to

create corpus-based teaching materials, build their own language knowledge, design testing items, assess and check language learners' language usage and to utilize corpora directly with students) in the EFL context uncovered that the prospective EFL teachers were not familiar with or knowledgeable about how to consult corpora for corpus-based language teaching pedagogy. The results showed that only one-third of the participants had *little* or *some* knowledge about using corpora for building their own language knowledge and awareness, while the awareness and knowledge of the majority of the group regarding the use of corpora for creating language test items and consulting corpora directly with language learners was almost non-existent. In line with the relevant literature, the overall familiarity and level of knowledge of English teacher candidates of consulting corpora for different pedagogical purposes was extremely low, despite the minor changes regarding the knowledge of the participants with various applications of corpora for language pedagogy (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Chen et al., 2019; Köse & Kuru Gönen, 2023; Ma et al., 2022; Poole, 2022). This finding may be due to the fact that the participants were not introduced to available corpora in the forms of corpus software and online-accessible corpora until their last year as a pre-service EFL teacher before the CBLT training. While these findings may stem from how language teaching and learning is perceived in the specific context of a study, it is obvious that the integration of corpora into FL teacher training programs in Türkiye is still not a common practice (Aşık, 2015; Başal et al., 2024; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Şimşek 2020). Although it is a well-known fact that making prospective EFL teachers familiar with computer technologies and tools that enhance and assist language instruction is a must for teacher training programs in light of recent developments, the current programs and courses offered fail to familiarize them with tools such as corpora, AI-based generative language models and many others.

These findings in the present study also complied with the results of the surveys employed by Bunting (2013) and Çalışkan's (2020) studies, as their participants also showed low familiarity and knowledge about such various pedagogical applications of corpora despite minor changes related to the relatively higher knowledge about the pedagogical application of corpora for building their own language knowledge. Besides, this study also revealed that most pre-service EFL teachers were unaware of the practical applications of corpora *with* and *for* language students. Rather, the knowledge and awareness of the participants were more related to the use of corpora as a reference tool

by and for themselves (Bunting, 2013; Çalışkan, 2020; Mukherjee, 2004; Römer, 2009; Tribble, 2015). In this regard, these findings are in line with prior research in revealing teachers' lack of awareness and knowledge about available corpora (Bennett & Uí Dhonnchadha, 2023; Chen et al., 2019; Ebrahimi & Faghih, 2017; Köse & Kuru Gönen, 2023; Leńko-Szymańska & Boulton, 2015; Ma et al., 2022; Naismith, 2017) as EFL teacher candidates do not receive a sufficient amount of interaction with computer technologies like corpus software and online corpora during their teacher education experiences. Despite the recent developments in technology and the emerging role of computer technologies in language education, the number of courses offered and instructors who are willing or capable of offering such courses are limited.

Concerning the investigation regarding the pre-service EFL teachers' experience with consulting corpora in language teaching contexts, it was highlighted that only a handful of the participants had limited hands-on experience in using available corpora as a part of the seminar conducted by the researcher the year before. The findings in relation to the investigation of whether the pre-service EFL teachers had been informed about corpora during their undergraduate teacher training education before the CBLT Training revealed that only a few of them had been informed about corpora as a part of some of their courses during undergraduate teacher training education. It was revealed that less than half of the pre-service EFL teachers had heard about corpora, while the rest had not been informed about corpora as a part of their undergraduate-level education or any specific training or seminar. These findings of the current study complied with the results of the empirical studies in the relevant literature in that consulting corpora is still not a common practice in teacher training curricula (Aşık, 2015; Başal et al., 2024; Chen et al., 2019; Çalışkan, 2020; Ebrahimi & Faghih, 2017; Leńko-Szymańska & Boulton, 2015; Köse & Kuru Gönen, 2023; Şimşek 2020; Tribble, 2015). Stemming from the lack of focus on the use of corpora in teacher training programs, it was revealed that the specific context of the study fails to familiarize teacher candidates with the use of corpora and equip them with the skills required for the consultation of corpus tools and online corpora for language teaching pedagogy. This finding also underlines the need for the integration of systematic and comprehensive training into initial teacher training for prospective EFL teachers on the use of corpora with hands-on practice and the integration of corpus-based language teaching materials into actual language teaching practices, as highlighted in the relevant literature (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Çalışkan, 2020;

Chen et al., 2019; Köse & Kuru Gönen, 2023; Lin, 2019; Ma et al., 2022; Şimşek & Can, 2023). As discussed above, there is a lack of focus on the integration of corpora into language teaching in EFL teacher training programs. The problems pre-service EFL teachers experience during their teaching practices at their practicum schools in terms of not knowing how to create authentic language teaching materials or adapt the ready-made ones and troubles they encounter while correcting the errors of their students may be due to not knowing about corpora, or how to consult a corpus. The participants pointed out during the study that using corpora made it easier for them to create, critically evaluate and adapt, and integrate teaching materials as well as offer error correction exercises for their students at their practicum schools.

In short, these findings revealed the need for integrating systematic and comprehensive training into initial teacher training for prospective EFL teachers on using corpora with hands-on practice and integrating corpus-based language teaching materials into actual language teaching practices. In order to meet this need and to assist prospective EFL teachers in the effective integration of corpora, this study aimed at designing a systematic and comprehensive training program on the integration of corpora into FL language teaching environments and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms. The CBLT Training was tailored for the specific context of the study and implemented into the practicum processes of senior pre-service EFL teachers. The main objectives of the training were to equip them with the theoretical knowledge and practical skills of consulting and utilizing corpora, demonstrate how corpus use can be integrated across various language domains and skills, and provide opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms. The primary purpose of the second phase of the study was to investigate the perspectives and views of the English teacher candidates during and after the training on the implementation of corpora into actual teaching practices in EFL contexts and the CBLT training program.

The analysis of the qualitative data based on the investigation of the perspectives and views of the English language teachers on the use of available corpora in actual teaching practices in EFL contexts revealed two main themes, namely, the benefits and the drawbacks of corpora use for language teaching. The overall findings shed light on

the pre-service EFL teachers' experiences in using online corpora and software for basic queries, building their own corpus, searching for specific language usage, creating corpus-informed language teaching materials, and comparing a corpus-based textbook with a non-corpus-based textbook for evaluation and adaptation. Complying with the relevant literature, it was revealed that the use of corpora raised the awareness of each participant on language and corpus-based language pedagogy despite their low familiarity and knowledge of available corpora and different pedagogical applications of using corpora in language teaching (Breyer, 2009; Çalışkan, 2020; Ebrahimi & Faghih, 2017; Farr, 2008; Özbay & Kayaoğlu, 2015; Römer, 2011; Zareva, 2017). The striking aspect of this finding is that despite their initial unfamiliarity, the EFL teacher candidates significantly enhanced their awareness and capabilities in using corpora through targeted training. This underscores the importance of incorporating such training on corpora use into teacher education programs to equip future EFL teachers with effective teaching tools and methodologies.

According to the prospective English teacher candidates, the corpora consultation for language teaching was found beneficial as it provided ease of access to the usage frequency of authentic and contextualized language samples. The findings revealed that the prospective EFL teachers found consulting corpora through online-accessible corpora and corpus software beneficial for accessing authentic language samples to show the actual usage of the language by its speakers. With different qualities and features each corpus possesses, including type (e.g., general language, specialized, native speaker, learner, etc.), vocabulary size (e.g., large with fourteen billion words, small with thirty thousand words), and the type of texts and genres compiled, users access to different authentic samples of the target language across various genres. The findings uncovered the perceptions and views of the English language teachers on the benefits of consulting corpus tools and online-accessible corpora for language instruction, with access to language samples full of contextualized language data showing different usage of nouns, verbs, idiomatic expressions, and phrasal verbs (Bennett & Uí Dhonnchadha, 2023; Can, 2023; Cangır, 2022; Ma et al., 2023; Sinha, 2021; Sun & Hu, 2023; Şimşek, 2020). The participants also pointed out that using corpora was beneficial thanks to the frequency of usage information for each specific word or phrase, providing insights and benefits into the selection of vocabulary items and grammar structures. This finding emphasizes the value of integrating corpus tools in language instruction, as it enables teachers to present

real-world language use across various genres, thereby enriching the learning experience of their students. Moreover, the use of corpora may contribute significantly to their professional identity development as teachers, fostering greater language awareness and confidence in their ability to create effective, data-driven teaching materials and activities. This, in turn, enhances their overall teaching efficacy and prepares them to implement more informed and authentic language instruction in their future teaching practices.

Regarding the perceived benefits of corpora use for language teaching, creating language teaching materials for language instruction was found highly beneficial as corpus-informed activities and materials involve naturally occurring samples of the real language usage retrieved from concordance lines. In this regard, the design and implementation of language teaching materials enable access to authentic language samples across various contexts. The participants benefited from using corpora for creating such corpus-based teaching materials for both their micro-teachings for the training assignments and their macro-teachings for the actual teaching at their practicum schools. (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Çalışkan & Kuru Gönen, 2018; Kaya et al., 2022; Köse & Kuru Gönen, 2023; Sun & Hu, 2023; Şimşek, 2020). Utilizing corpora for creating teacher materials not only enhances the authenticity and relevance of language instruction but also fosters teachers' ability to design effective, contextually rich activities, thereby improving their overall teaching competence and confidence. This integration of real-world language data into their teaching practices contributes significantly to their professional development and effectiveness as future language educators.

Most participants acknowledged that the authenticity of corpus-based language teaching materials is significant as using corpora provided contextualized language, and it was time-saving to use corpora instead of trying to find a context for teaching a particular lexical item or grammar structure. In line with the relevant literature, the prospective EFL teachers emphasized the effectiveness of creating and implementing corpus-informed language teaching materials as the language used in teaching materials should be informed by empirical evidence and usage frequency of the actual language with a lexical orientation rather than intuition, tradition and invented language samples (Aşık, 2017; Başal et al., 2024; Çalışkan, 2020; Hunston, 2002; Meunier & Reppen, 2015; Mindt, 1996; Mukherjee, 2004; Özbay & Kayaoğlu, 2015). This finding may stem from

that they recognized the importance of authentic, contextualized, and meaningful language usage through corpora. Regarding their future teaching practices, the prospective EFL teachers pointed out that they will consult corpora for EFL teaching as this consultation leads to more effective and empirically grounded teaching resources and teaching materials become more relevant to language learners because they reflect actual language use, increasing engagement and comprehension. As these prospective EFL teacher candidates are highly likely to work in K12 or Higher Education contexts in Türkiye upon their graduation, equipping them with the skills of using corpora for teaching purposes may contribute to enhancing the overall language teaching quality with access to authentic and contextualized language samples.

Another finding related to the benefits of corpora use for language teaching was that the use of corpora was found beneficial and valuable for vocabulary teaching, as corpora consultation offers access to the authentic, naturally occurring usage of lexical items across various contexts. According to the participants of the training, using corpora can be helpful while teaching new vocabulary or designing activities for the practice of previously introduced lexical items, as emphasized in the relevant literature (Alsahafi, 2022; Barabadi & Khajavi, 2017; Can, 2023; Chen, 2017; Daskalovska, 2015; Eslek-Onur & Tosun, 2023; Karras, 2016; Koçak, 2020; Li, 2017; Lu & Dang, 2022; Saddhono et al., 2023; Sinha, 2021; Tsai, 2019; Uçar & Yükselir, 2015; Vyatkina, 2016; Zhang, 2022). The findings revealed that using corpora for vocabulary teaching was highly beneficial, especially for particular vocabulary items, including collocations, idiomatic expressions, commonly confused words, and the parts of speech of lexical items. This finding was in line with the empirical studies in the relevant literature, as the high potential of the use of corpora for teaching such particular vocabulary items was emphasized frequently (Aşık et al., 2016; Çalışkan & Kuru Gönen, 2018; Godwin-Jones, 2017; Köse & Kuru Gönen, 2023; Li, 2017; Tsai, 2019; Uçar & Yükselir, 2015). According to the EFL teacher candidates, using corpora for lexical items might benefit their future teaching practices. The consultation of corpora for teaching vocabulary items has been found more beneficial when given in meaningful, authentic contexts rather than in isolation or made-up contexts. Since the focus of instruction on teaching vocabulary is limited during their teacher education, these teachers lack the required skills and awareness on how to teach vocabulary, which lexical items to include in their teaching, and how to access authentic language data. While the majority of language teachers in

our K12 and Higher Education contexts in Türkiye teach lexical items given in the textbooks provided or refer to their prior experiences, through this study and the training they received, the awareness of the prospective EFL teachers raised in that consulting corpora provides access to authentic, real-world language usage (Alsahafi, 2022; Barabadi & Khajavi, 2017; Can, 2023; Eslek-Onur & Tosun, 2023; Koçak, 2020; Lu & Dang, 2022; Saddhono et al., 2023; Sinha, 2021; Zhang, 2022).

In addition to the benefits of corpora to vocabulary instruction, the findings uncovered the perceived benefits of using corpora for grammar instruction, according to the pre-service EFL teachers. Corpus-based language teaching materials were found beneficial for grammar instruction with access to authentic, naturally occurring language samples to enhance teaching grammar structures via concordance lines. Furthermore, the participants also pointed out that using corpora can raise language learners' language awareness through corpus-based grammar teaching materials and their awareness of the grammar structures of the target language. The findings of the study revealed that using corpora for grammar instruction posed significant benefits for teaching particular grammar structures, including conjunctions, modal verbs, prepositions, and gerund/infinitive. These findings are consistent with other studies in the relevant literature, as the potential benefits of corpora use were emphasized for teaching particular grammar structures along with the positive perceptions of the EFL teachers toward consulting corpora for grammar teaching (Girgin, 2011, 2019; Emir & Yangın-Ekşi, 2023; Khan et al., 2021; Lin & Lee, 2019; Lakew et al., 2021; Lin, 2019; Özer & Özbay, 2022; Rodríguez-Fuentes & Swatek, 2022; Zhong & Wakat, 2023). The findings revealed that teaching grammar can be enhanced by consulting corpora in the form of creating corpus-informed grammar teaching materials and activities and using corpora directly with students to raise their language and grammar awareness while teaching particular grammar structures. While it is a well-known fact that grammar teaching taking place in K12 education in the Turkish context relies highly on the explicit, direct instruction of grammar rules, the use of corpora allows for a contextualized, meaningful grammar instruction that enhances the engagement of learners with the grammar of the target language, how the language system works. As pointed out by the participants, a teacher can tailor their grammar instruction to focus on areas of grammar that are particularly challenging for Turkish learners since this targeted approach helps address the specific needs of learners and improves their overall language skills. Despite the particular focus

on grammar instruction in the EFL teacher training programs in Türkiye through courses that emphasize the significance of contextual grammar, these courses are far from providing prospective EFL teachers with ways of accessing authentic, contextualized language to be utilized for grammar teaching.

Another perceived benefit of corpora use for language teaching was enhanced error correction. The pre-service EFL teachers reported that using available corpora (i.e., corpus tools or online-accessible corpora) along with corpus-based activities was highly beneficial for correcting the errors of EFL students. In line with the findings of relevant studies, the findings revealed that the use of corpora was believed to contribute to an effective error correction procedure with the paper-based error correction activities created by the English teacher candidates and the consultation of corpora for writing instruction was suggested (Anthony, 2016; Crosthwaite, 2017; Çalışkan, 2020; Emir, 2019; Luo & Liao, 2015; Quinn, 2015; Tono et al., 2014). In addition to the use of corpora by EFL teachers for correcting their students' errors and common mistakes in the classroom, the benefits of encouraging the direct use of corpora by language learners to detect and correct their own errors. It was pointed out that by doing this, they recognize and identify language patterns to infer grammar rules and word usage and draw insights from concordance lines autonomously. While utilizing available corpora for error correction enables teachers to gain access to vast databases of authentic language that serve as a reference point for identifying and correcting errors made by EFL learners, there is a lack of emphasis on error correction in teacher training programs. There is a need to address this lack of focus on error correction as teacher candidates are not introduced to ways of providing error correction and how and what type of exercises or materials are to be designed and used in their future teaching practices. By recognizing error correction as a fundamental aspect of language teaching, providing opportunities for teacher candidates to consult corpora for error correction revealed that they were not equipped with adequate knowledge and guidance.

Despite the perceived benefits of corpus consultation for language teaching, several drawbacks and challenges of using corpora were stressed by the English teacher candidates. Based on the findings of the study, one of the drawbacks of using corpora for language teaching was the difficulty in finding appropriate language items. As the raw data accessed through concordance lines needed adaptation due to incorrect usage or undesired results, the pre-service EFL teachers revealed that they experienced difficulties.

It was revealed that the participants had trouble finding the suitable language items they were looking for through concordance lines, as there were too many sentences and samples of the language to be chosen from. This selection of appropriate concordance lines and language items was time-consuming for the pre-service EFL teachers. This finding is consistent with the findings of several other empirical studies in the relevant literature, as some EFL teachers experienced similar difficulties and found the selection of appropriate language items by consulting corpora time-consuming and impractical due to the tight schedules and workload of EFL teachers (Breyer, 2009; Çalışkan, 2020; Lin, 2019; Tribble, 2015). As Tribble (2015) highlighted, the participants encountered problems while consulting corpora for selecting appropriate language samples as they found spending too much time for it to be one of the most striking obstacles in language teaching contexts. In another study by Çalışkan (2020), the participants reported that scanning and selecting suitable concordance lines took a lot longer than they anticipated. Providing further guidance and instruction on eliminating undesired language samples and selecting appropriate concordance lines effectively and ready-made corpus-informed language teaching materials and activities might make corpus queries through concordancers more efficient. This finding was also in line with another drawback revealed in the present study related to the limitations of corpora consultation for language learners with low proficiency levels. It was pointed out that corpus-based activities can be too hard for low proficiency level learners, and the time and effort required to create suitable activities for such learners is impractical for EFL teachers. Despite their positive experiences and perceptions related to the use of corpora for creating language teaching materials, the participants reported that it takes time and effort to implement corpus data into teaching activities for low proficiency level learners. One participant pointed out that using SKELL was more practical for low proficiency level students due to shorter and simpler language samples and sentences, and using COCA and MICASE was more convenient for more advanced-level language learners. Similarly, as relevant literature also suggested, the use of a variety of different corpus tools and online-accessible corpora for varying proficiency levels of students can help overcome these limitations (Breyer, 2009; Chen, 2017; Çalışkan, 2020; Ebrahimi & Faghih, 2017; Luo & Liao, 2015; Özbay & Kayaoğlu, 2015; Tono et al., 2014). As each classroom is unique with learners from different backgrounds and experiences, tailoring the language teaching materials and activities to be utilized in the classroom is crucial for EFL teachers.

The last drawback of the use of corpora for language teaching was related to the technical problems faced during corpora consultation. The findings highlighted that the pre-service EFL experienced technical problems while using corpus tools and online-accessible corpora due to the nature of different concordancers. Even though most participants favored using COCA over other corpus software and online corpora during the training for in-class tasks and assignments, they reported multiple difficulties and problems they encountered while navigating COCA due to query limits or the primitive, user-unfriendly interface, as well as difficulties they experienced while using search functions of some corpus tools and online corpora. In line with the relevant literature, technical difficulties encountered by the pre-service EFL teachers were believed to negatively affect the consultation of corpora and the creation of corpus-based language teaching materials (Breyer, 2009; Çalışkan & Kuru Gönen, 2018; Ebrahimi & Faghih, 2017; Farr, 2008; Heather & Helt, 2012; Zareva, 2017). Many online-accessible corpora and corpus software are not without technical difficulties and shortcomings due to their interfaces and need for funding. In the era of AI, the use of corpora can also be enhanced with the possible benefits of AI consultation, and such technical difficulties can be overcome in the future with the combined usage of corpus tools and AI-based generative language models.

In addition to the perceptions and views of the English teacher candidates on the consultation and implementation of corpora for actual teaching practices in EFL contexts, the investigation of their perceptions of the CBLT training program revealed two main categories, including strengths and weaknesses of the training. According to the prospective EFL teachers, the CBLT training was found successful in multiple different aspects. One of the most prominent strengths of the training was equipping EFL teachers with the required skills for consulting corpora, as the participants expressed that they gained the essential skills for corpora use for language teaching purposes throughout the training. It was also revealed that while the specific context of the study fails to equip these skills required for the consultation of corpus tools and online corpora for language teaching pedagogy, the overall findings based on the perceptions and experiences of the pre-service EFL teachers' experiences revealed that the training was successful in familiarizing them with corpus linguistics and training them on corpora consultation, corpus practice and building their own corpus, searching for specific language usage for completing tasks during sessions and their actual teaching practices at their practicum

schools. In line with the relevant literature, it was revealed that raising the awareness of the prospective EFL teachers on language and corpus-based language pedagogy was one of the strengths of the nine-week CBLT training program as well as equipping them with corpus literacy (Ebrahimi & Faghih, 2017; Farr, 2008; Özbay & Kayaoğlu, 2015; Römer, 2011; Şimşek, 2020; Zareva, 2017). The participants emphasized that using corpora for in-class tasks and assignments and creating language teaching materials were eye-opening. Raising the awareness of EFL teachers on the language and corpora may contribute to overall language teaching as teachers also need to adapt their language instruction according to the needs of their learners, the unique nature of each topic that takes place in the classroom, and lexical items and grammar structures taught. Equipping prospective EFL teachers with corpus literacy is one of the most striking gains of the training as these future language teachers were introduced to corpora consultation with hands-on practice during their practicum processes, right before their graduation, and starting to work as EFL teachers.

It can also be inferred from the study findings that the CBLT training successfully provided opportunities for the pre-service EFL teachers' professional development. The findings showed that the prospective EFL teachers were grateful for being a part of the training, and they indicated that they hadn't been informed about corpora as a part of their undergraduate education as prospective teaching. According to the qualitative findings, it was pointed out that the participants found the CBLT training to be an irreplaceable part of their practicum process as they consulted corpora to teach English to their students with access to authentic, contextualized language usage. The prospective EFL teachers stated that they were grateful to be equipped with corpus literacy and to have the chance to build their own corpus. Calling it better late than never, these future EFL teachers were introduced to the use of corpora for language teaching for the first time in the last year of their teacher education program. Even though they grew up with the most recent advancements in technology, they were not familiar with corpora as there are no courses, training, certificate programs, or workshops available on the *know-how* and actual practice of corpora consultation for language teaching. Despite the poor familiarity of a few, they were not instructed on how to actually consult corpora as future teacher candidates and language learners in the past. These findings in the present study also complied with the results of previous research in that most participants acknowledged that the use of corpus tools and software was highly beneficial for creating their own

corpus for their actual teaching practices and being equipped with the essential skills for corpora use was highly beneficial for their teacher capabilities (Farr, 2008; Şimşek, 2020; Zareva, 2017). They emphasized that creating a corpus based on the needs of their students would also enhance language instruction as each classroom is unique with the learners inside, considering their needs, experiences, and motivations.

It was also revealed that creating a reflective and collaborative learning environment was one of the strengths of the CBLT training. The findings related to the prospective EFL teachers' perceptions regarding the CBLT training and the feedback mechanisms employed throughout the training revealed that creating a learning environment enriched with guidance and constant support from the trainer and collaboration among participants encouraged their practice with available corpora. The majority of the tasks and assignments of the training encouraged the pre-service EFL teachers to work in groups of two or three to experience the corpora consultation in a collaborative atmosphere so that they would learn from each others' struggles and accomplishments while learning how to navigate a corpus and how to consult corpora for creating language teaching materials. Similarly, Çalışkan (2020) pointed out that the participants of her study benefitted from the interaction and collaboration in group discussions and workshops, as well as the material pool created for the EFL instructors' disposal as a part of the training. It can be concluded that ensuring a reflective and collaborative atmosphere for prospective EFL teachers fostered their experience during the training. This finding may stem from the need for cultivating communities and cooperation of practice for prospective EFL teachers as such initiatives are also almost non-existent in our education system and educational infrastructure of faculties of education. Providing more opportunities for language teachers to work together and learn from each other in communities of practice where mutual support and collaboration are emphasized and practiced.

According to the experiences and perceptions of English teacher candidates with the CBLT training, it was uncovered that the prospective EFL teachers also had some challenges due to too much focus on production throughout the training. They experienced difficulties that stemmed from the number of tasks and assignments for creating language teaching materials since they believed the training was too loaded. They pointed out the need for more emphasis on practice with several corpus tools and online corpora. In line with the findings of previous research, it was revealed that the pre-

service EFL teachers found it hard to keep up with the number of materials they were to design for the compulsory assignments of the training (Farr, 2008; Şimşek, 2020; Zareva, 2017). Despite their highly positive remarks and experiences with corpora use for creating corpus-informed language teaching materials for both their micro-teachings during the training and their actual teaching practices at their practicum schools, they found the training content too loaded. The participants stressed that the CBLT training was too loaded for a nine-week training, independent of the teacher training education curriculum, with no grades or credits. It was suggested that the CBLT training could have focused on more hands-on practice with corpus software and online-accessible corpora, and one way of achieving this could be assigning particular corpus tools and online corpora for particular assignments according to the pre-service EFL teachers. This finding may stem from the fact that the use of corpora for language education is not a common practice in our education system for language learners or teachers (Aşık, 2015; Başal et al., 2024; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Şimşek 2020). Since the participants of the study were not introduced to corpora either as EFL learners in the past or EFL teacher candidates throughout their teacher education, the training was found to be too loaded with the number of assignments and tasks. Turning this individual effort of equipping prospective EFL teachers with the required skills of consulting corpora for EFL teaching purposes into a longitudinal aspect of the education system where learners and teachers are introduced to the use of corpora and provided with opportunities to experience using corpus tools for language teaching and learning purposes would enhance the overall language education in Türkiye. Then, the training would not have been found to be too loaded; not only the participants of the training but also each language learner and teacher in our society would have learned about corpora before and utilized corpus tools and online-accessible corpora for accessing more authentic and contextualized language usage.

All in all, the findings of the current study pointed out the importance of the integration of corpora into pre-service EFL teaching education for equipping teachers with the required practical skills for using corpora by providing opportunities for prospective EFL teachers to gain hands-on experience with using corpora and integrate corpus-based language teaching materials into their actual language practices in EFL contexts offers various potential benefits for English language instruction.

6. CONCLUSION

This chapter aims to present the conclusions of the current study by initially providing a brief summary of the study and the findings. Subsequently, conclusions drawn from the findings of the current study are presented along with the pedagogical implications of the present study and suggestions for further research endeavors based on the findings of the study.

6.1. Summary of the Findings

The initial purpose of the current study was to explore the knowledge, experiences, and perceptions of English teacher candidates on corpus-based language teaching and the practical implementation of corpus-informed EFL teaching activities and materials in EFL contexts before providing them with the systematic and comprehensive Corpus-Based Language Teaching (CBLT) Training tailored for the specific context of the current study. In line with the relevant need emphasized in the literature for providing training opportunities on and for the use of corpora for language teaching for pre-service EFL teacher education, this study aimed at designing a corpus training program with a specific focus on involving prospective EFL teachers with hands-on practice of corpus integration into actual teaching practices. Furthermore, during and after the implementation of the CBLT Training, the aim of the current study was to investigate the views and perceptions of the English teacher candidates on the consultation of corpora for EFL teaching and the CBLT Training.

The data collection procedure of the present study took a total of fourteen weeks, and the first set of data was collected from 102 prospective EFL teachers via the Corpus Background Survey to investigate English teacher candidates' background knowledge, familiarity, and views regarding corpora and their experiences with consulting available corpora in EFL contexts for language instruction in the first week of the study. The first session of the CBLT Training was conducted in the third week of the study following the selection of the 20 voluntary pre-service EFL teachers, and the training was held for nine weeks. The second set of data was collected by reflective logs at the end of the second, fifth, and eighth training sessions, and follow-up interviews after the training to explore the perceptions of the 20 English teacher candidates on the consultation and implementation of corpora into EFL teaching contexts, and the CBLT Training itself. In addition to reflective logs and semi-structured interviews, qualitative data were obtained

by a post-teaching reflective log written by the participants to share their reflections on using corpus-based teaching materials they designed to be used in their teaching contexts at their practicum observation schools. While the data obtained through the survey were analyzed by descriptive statistics, the open-ended items of the background survey were simply analyzed by identifying the frequencies of participants' brief responses to the open-ended questions to explore their perceptions of and familiarity with corpora. The qualitative data obtained via the reflective logs, semi-structured interviews, and post-teaching reflective logs were analyzed qualitatively by following the principles of the Constant Comparative Method (CCM).

Based on the analysis of the quantitative data obtained from the Corpus Background Survey, it was revealed that most pre-service EFL teachers had no or little knowledge about or experience in using corpora for pedagogical purposes. Only 20% of the participants reported that they experienced the use of corpora during the seminar conducted by the researcher the year before, and the rest of the English teacher candidates in the context of the study did not have knowledge about or hands-on experience in using corpus tools and online corpora. The findings in relation to the investigation of whether they were informed about corpora before the CBLT Training revealed that less than half of the participants had been informed about corpora during the seminar conducted by the researcher the year before, and only a handful of participants reported that they had been informed about corpora in their linguistics courses during their undergraduate teacher training. In short, the analysis of the first set of data underscored the need for integrating systematic and comprehensive training into initial teacher training for prospective EFL teachers on using corpora with hands-on practice and integrating corpus-based language teaching materials into actual language teaching practices. Therefore, the CBLT Training was tailored for the specific context of the study and implemented into the practicum processes of senior pre-service EFL teachers as an attempt to equip them with the theoretical knowledge and practical skills of consulting and utilizing corpora, demonstrating how corpus use can be integrated across various language domains and skills, and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms.

The analysis of the qualitative data based on the investigation of the perceptions and views of the English teacher candidates on the consultation and implementation of

corpora into actual teaching practices in EFL contexts revealed the perceptions of the participants as the benefits of corpora use for language teaching and drawbacks of corpora use for language teaching. According to the pre-service EFL teachers, using corpora for language teaching was highly beneficial for creating corpus-informed language teaching materials for teaching vocabulary items (e.g., parts of speech, commonly confused words, collocations) and grammar structures (e.g., conjunctions, modal verbs, prepositions, gerund/infinitive), and enhancing error correction. Furthermore, the perceptions of the prospective EFL teachers uncovered that integrating corpora into language classrooms would benefit both EFL teachers and learners as corpus consultation was found useful for accessing the frequency information of authentic language items and their usage across various contexts. It was also stressed that using corpora provides access to contextualized language samples. In contrast to such benefits, the participants reported several drawbacks of using corpora for language teaching, including difficulties in finding appropriate language items (e.g., raw data needs adaptation due to incorrect usage) and using search functions of several corpus tools (e.g., difficulties while making queries on corpus tools). Moreover, they also reported technical problems (e.g., query limit, primitive and user-unfriendly interfaces of different concordancers) and limitations of corpora use for low proficiency level students (e.g., impractical for beginner and elementary level learners).

The perceptions of the prospective EFL teachers on the CBLT training uncovered the strengths of the CBLT training and its weaknesses. According to the pre-service EFL teachers, the CBLT training successfully equipped them with the required skills for using corpora for language teaching, provided opportunities for their professional development, and raised their language awareness in a reflective and collaborative learning environment. The participants emphasized that one of the most prominent strengths of the CBLT training was providing them the opportunity to build their own corpus, and they were grateful for attending this training during their undergraduate education as prospective EFL teachers. Despite such prominent strengths of the training, the pre-service EFL teachers also stressed that the training focused too much on production with too many tasks and assignments. They also pointed out a need for more practice with corpus tools and online-accessible corpora that were introduced, demonstrated, and encouraged to be used for hands-on practice and creating language teaching materials.

6.2. Conclusions

The findings related to the initial purpose of the study to investigate the perceptions, knowledge, and experiences of English teacher candidates on the consultation of corpora and the practical implementation of corpus-based language teaching materials into English language teaching before the CBLT Training revealed that the majority of the prospective EFL teachers were not familiar with the use of corpora for language pedagogy as well as available corpus tools and online-accessible corpora. It can be highlighted that while the participants were more familiar with online-accessible corpora than corpus tools and software, the overall familiarity of the participants with available corpora was very low. It was revealed that the level of knowledge of the participants regarding the use of corpora for different pedagogical purposes was extremely low (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Ma et al., 2022; Poole, 2022). This shows that prospective EFL teachers were not informed about corpora to be used as a professional technological tool for language teaching throughout their teacher training education.

The findings revealed that two-thirds of the English teacher candidates were not aware of using corpora for building their language knowledge and awareness. It was highlighted that the knowledge and experience of the majority of the group regarding the use of corpora for creating testing items, checking and assessing language learners' language usage, creating corpus-based teaching activities and materials, or consulting corpora directly with language learners was almost non-existent. These findings underscore the necessity to revise and update teacher training programs to include comprehensive instruction on the use of corpora, as well as provide ongoing professional development opportunities for current and future EFL teachers. The integration of corpus-based teaching pedagogy methods into the curriculum can equip prospective EFL teachers with valuable skills and knowledge, ensuring they are well-prepared to utilize corpora consultation effectively in their classroom practices.

The investigation regarding the participants' experience with consulting corpora in language teaching contexts highlighted that only a handful of the participants had hands-on experience in using corpus tools and that they had been informed about corpora through some of their linguistics courses and academic articles. It was revealed that the specific context of the study fails to familiarize and equip teacher candidates with knowledge about corpora and the skill of using corpus tools and online corpora for

language teaching pedagogy. As highlighted by many researchers (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Ma et al., 2022; Şimşek & Can, 2023), this finding underlines the need for the integration of systematic and comprehensive training into initial teacher training for prospective EFL teachers on the use of corpora with hands-on practice and the integration of corpus-based language teaching materials into actual language teaching practices. In order to meet this need and to assist prospective EFL teachers in the effective integration of corpora, this study aimed at designing a systematic and comprehensive training program on the integration of corpora into EFL teaching environments and providing opportunities to gain hands-on experience with corpora for designing and integrate corpus-based language teaching materials into their actual language classrooms. Moreover, in the emerging era of AI-based generative language model use in education that brought about critical concerns regarding the reliability and credibility of AI-generated content, the integration of corpora into language teaching environments for accessing reliable and authentic language data is essential.

The qualitative findings regarding the perceptions and views of the English teacher candidates after the training revealed several perceived benefits and challenges of consulting corpora for language teaching as well as strengths and weaknesses of the CBLT Training. As the findings related to the use of corpora put forward, it can be concluded that consulting corpus tools and online-accessible corpora can be highly beneficial for creating corpus-informed language teaching materials, teaching particular vocabulary items (e.g., parts of speech, commonly confused words, collocations) and grammar structures (e.g., conjunctions, modal verbs, prepositions, gerund/infinitive), as well as enhancing error correction in language classrooms. It can be highlighted that corpus consultation might also be useful for accessing the frequency information of authentic language items and their usage across various genres and contexts (Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Özbay & Gürsoy, 2023; Poole, 2022; Şen Bartan, 2023). Another benefit of corpora for language teaching is accessing contextualized language items and samples. By employing corpus-informed language teaching materials and activities for effective practice of corpus-based language pedagogy, the use of corpora for and with language students might foster language learning and instruction.

Although they emphasized the perceived benefits of corpus consultation for language teaching, the participants reported several drawbacks of using corpora, which might disrupt the benefits and effectiveness of corpus-based language pedagogy. One drawback of using corpora for language teaching might be related to difficulties in finding appropriate language items, as the raw data accessed through concordance lines might need adaptation due to incorrect usage or undesired results. Providing further guidance and instruction on eliminating undesired language samples and selecting appropriate concordance lines effectively might make corpus queries through concordancers more efficient (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Köse & Kuru Gönen, 2023; Ma et al., 2023; Michali & Patsala, 2020; Şimşek & Can, 2023). As stressed by the participants, technical problems related to query limits, ads, user-unfriendly interfaces of different concordancers, and problems related to query functions of different corpus tools might hinder the benefits of corpus consultation for language teaching. The introduction and demonstration of various other corpus tools and online corpora, along with more practice, might be one way to overcome such problems. As stressed by the pre-service EFL teachers, another drawback of using corpora for language teaching might be that using corpora with low proficiency level learners is not practical due to the lack of required language awareness and proficiency. Consulting a variety of corpora and corpus tools for creating language teaching materials for students with different proficiency levels might be effective in overcoming this shortcoming, as proposed and employed by the participants. Overall, it can be concluded that using corpora for language teaching might be highly beneficial for language teachers and learners despite several potential drawbacks.

The qualitative findings related to the perceptions and views of the prospective English teacher candidates about the training uncovered that the participants found the CBLT training comprehensive and successful in equipping them with the skills of navigating and utilizing corpus tools and online-accessible corpora for various pedagogical purposes. Consequently, the design process, content, and structure of the CBLT Training program might enlighten the design processes of future training on corpus literacy and practice or its adaptation and application to other teacher training contexts with the different types of corpus tools and resources, practice tasks, and assignments employed and offered via a booklet. In addition to the direct and indirect applications of corpora throughout the training, the CBLT training was found to be significant in

providing the pre-service EFL teachers with professional development opportunities as their awareness of the English language and the concept of corpora was formed, shaped, and raised by the hands-on practice with corpora for corpus-based language pedagogy. The participants also reported the success of the training in creating a reflective, collaborative, and interactive learning environment that enhanced the instruction and the practice. This finding might encourage future training programs to enable such a collaborative learning environment for prospective EFL teachers. The participants also stressed that one of the most prominent strengths of the CBLT training was the opportunity to build their own corpus, and they were grateful for being a part of the CBLT training and during their undergraduate education as pre-service EFL teachers. It can be implied that paying attention to the processes of building context-specific corpus activities and assignments might enhance language instruction and the corpus literacy of the participants.

Despite such prominent strengths and benefits of the training, the pre-service EFL teachers also reported a few weaknesses as it focused too much on production, with the number of tasks and assignments based on creating corpus-informed language teaching materials. They also indicated the need for more practice with corpus tools and online-accessible corpora despite the number of corpus tools and online corpora introduced for hands-on practice and creating language teaching materials throughout the training. In this regard, this might shed light on the need for the inclusion of more practice with different corpus tools and resources to enable the participants to be more confident in the consultation of corpora.

All in all, based on the findings regarding the perceptions of the participants, the integration of corpora into initial teaching training for equipping teachers with practical skills for consulting and utilizing corpora and providing opportunities for pre-service EFL teachers to gain hands-on experience with consulting corpora and integrate corpus-informed EFL teaching activities and materials into actual language practices in EFL contexts offers various potential benefits for English language instruction. This study sheds light on the need to make the use of corpora a part of EFL teachers' teaching practices by training them on the use of corpora before their graduation, as such training is essential and best situated in initial language teacher education. Thus, it is evident that offering training for pre-service EFL teachers is crucial for integrating corpora into actual teaching practices and classrooms.

6.3. Implications

According to the findings of the present study, some pedagogical implications can be proposed for curriculum designers, material writers and publishers, researchers, teacher education programs, and pre-service EFL teachers in terms of educational uses of corpora and its integration into L2 contexts.

As the findings of the study revealed, prospective EFL teachers lacked the knowledge and awareness on the consultation and implementation of corpora into EFL teaching instruction. The implementation of corpus-based language pedagogy in language teaching into initial teacher education courses might help equip future EFL teachers with the skills required for consulting corpora and integrating corpus-based teaching materials into their future practices (Başal et al., 2024; Bennett & Uí Dhonnchadha, 2023; Çalışkan, 2020; Köse & Kuru Gönen, 2023; Şimşek & Can, 2023). The findings of this study can contribute to the teacher training program curricula in the Faculty of Education of universities in the Turkish EFL context for training English teacher candidates. It can be suggested that in addition to the need for such training programs and workshops like the CBLT training, the interaction of prospective EFL teachers with the use of corpora as a part of a specific course on corpus-based language pedagogy or the material evaluation and adaptation and methodology courses since the first days of their teacher training education and practicum processes might raise their language awareness and prepare them for their future teaching practices with more opportunities to design and adapt corpus-based language teaching materials and use such materials for their teaching practices even before their professional teaching career. Since consulting corpora can contribute to accessing the usage frequency information of authentic language items and their usage across various contexts, integrating corpora into such courses can equip prospective EFL teachers with the required skills and knowledge for exposing language learners to authentic language through corpora. The benefits of corpora consultation for creating corpus-informed language teaching activities for teaching particular vocabulary items (e.g., parts of speech, commonly confused words, collocations) and grammar structures (e.g., conjunctions, modal verbs, prepositions, gerund/infinitive), as well as correcting the errors of language learners can be highlighted in such interaction in the initial teacher training.

In addition to the integration of corpora into initial teacher training through such courses, training programs, and workshops, comprehensive in-service teacher training is essential for effectively integrating corpora into language teaching practices. This can be achieved by organizing dedicated in-service training sessions, workshops, and online training opportunities through webinars with the initiatives and efforts of the Ministry of National Education to support and extend widespread adoption. Implementing such training would enhance EFL teachers' professional development and competencies, enabling them to effectively utilize corpora in their teaching practices, while future programs can also be developed for both pre-service and in-service EFL teachers to improve their skills and stay updated with the latest advancements in language teaching methodologies. These measures would not only foster better language teaching practices but also contribute to the overall quality of education and professional development of EFL teachers. Moreover, it can be noted that the findings regarding integrating corpora into language classrooms would also contribute to the teacher training curricula for the teacher education of teachers of other languages. According to the findings of the study, it can be noted that integrating corpora into language classrooms would benefit both the teachers and learners of other languages as well as English.

As the findings of the study suggest, the pre-service EFL teachers benefitted from the use of corpus-based language teaching materials for their micro-teachings throughout the training and the macro-teachings as a part of their practicum process. Consequently, such instruction on corpus-based language teaching pedagogy can equip the teacher candidates with the required theoretical conceptualization and knowledge, skills to navigate and utilize corpora, and experience in the hands-on experience to consult corpora for language instruction to be applied in their future language teaching practices. The findings of the study regarding the drawbacks of the CBLT training revealed that further training should focus on more practice with the hands-on, guided practice of various corpus tools and resources. Considering the perceived benefits of the use of corpora for language teaching according to the pre-service EFL teachers, it was revealed that the hands-on consultation of various corpus tools and online-accessible corpora poses significant benefits for designing teaching materials as the participants were both provided with sample corpus-informed activities and materials and tasked with creating their own teaching materials to be implemented into their actual teaching practices at the practicum schools. It can be inferred that there is a need for more corpus-based activities

and textbooks for teacher candidates and in-service EFL teachers to interact with and encourage them to create their own materials based on their context-specific needs, as the number of published corpus-based activities and materials for teachers' disposal is limited.

Furthermore, the findings revealed that the lack of focus on EFL teachers' corpus literacy is related to the lack of published or digital guidelines and handbooks to enable the *know-how* of corpus use for EFL teachers. This finding calls for the need to create and enrich such booklets and handbooks for teachers to utilize corpora in their teaching and the need for published corpus-informed teaching materials to foster the indirect application of corpora by further research. Considering such a need, another implication of this study is that textbook writers and ELT publishing companies can develop additional platforms, handbooks, and materials accessible to language teachers and learners.

As the findings related to the prospective EFL teachers' perceptions regarding the CBLT training and the feedback mechanisms employed throughout the training revealed, creating a learning environment enriched with guidance and constant support from the trainer and collaboration among participants encouraged their practice with corpus tools. It can be concluded that future training programs on corpus literacy must enable a reflective and collaborative atmosphere for prospective EFL teachers and various feedback mechanisms to get in-depth insights regarding the training and the implementation of corpus-based language teaching materials into actual language classrooms. This underscores the importance of fostering communities of practice where support, cooperation, and discussion groups can provide assistance for gaining new ideas and insights in consulting corpora for language teaching purposes. Such communities become even more effective when they are supported by faculties, decision-makers in administrative positions, and textbook publishers.

Regarding the findings related to the weaknesses and drawbacks of the use of corpora in language teaching, it was revealed that the use of corpora for creating teaching materials, especially for low proficiency level students, can be challenging for EFL teachers due to their workload and the time and efforts required for corpus consultation and designing activities. Another implication is that more simplified, level-appropriate language teaching materials can be developed for low-proficiency learners by textbook and material writers, while high-proficiency learners benefit from more advanced and

interactive corpus-informed materials. Moreover, it was also revealed that some participants experienced several technical problems with some concordancers (i.e., query limit, primitive and user-unfriendly interfaces of different concordancers, using search functions). Demonstrating and encouraging the practice of a vast number of varying corpus tools and resources and online-accessible corpora can be highly beneficial for future training programs in order to provide different corpora alternatives to overcome such challenges. Regarding technological requirements at schools, it is implied that there is a need for decision-makers to provide financial support to overcome the lack of technological infrastructure in schools and develop infrastructural projects to support the use of corpus tools on smartboards. Figure 6.1. outlines the implications of the study.

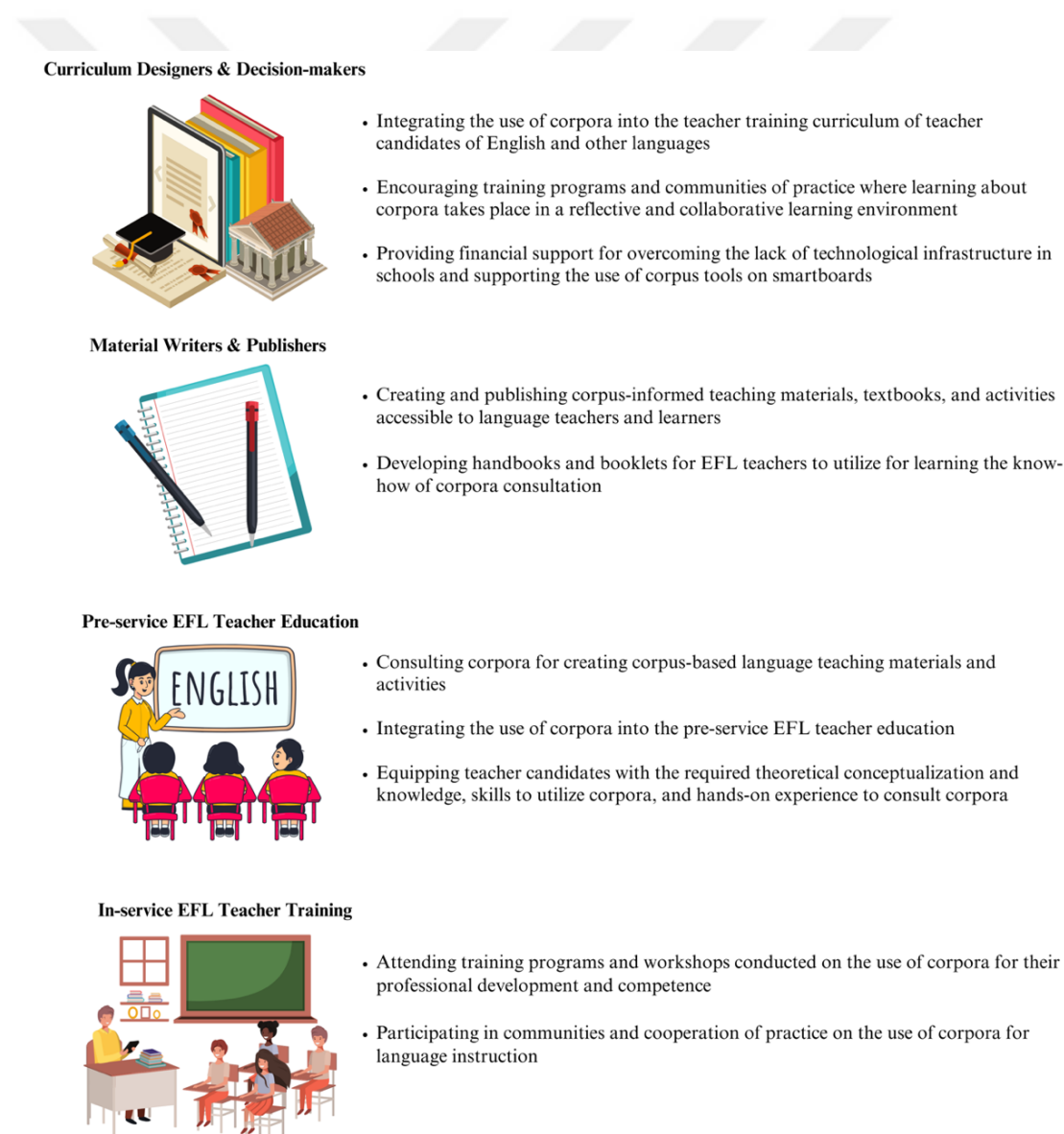


Figure 6.1. *The outline of the implications of the study*

Overall, the current study suggests pedagogical implications regarding integrating the use of corpora for language teacher training programs to equip EFL teacher candidates with corpus literacy and the integration of corpora into actual teaching practices of EFL teachers. These pedagogical and practical implications might shed light on future endeavors of effective corpus consultation and integration into L2 educational settings.

6.4. Suggestions for Further Research

The findings of this study offer significant insights into English teacher candidates' perceptions and views on the implementation of corpora consultation into their actual teaching practices and the CBLT Training program. Concerning the findings related to the prior low or non-existent familiarity and experience of the participants with using corpora for language teaching, it can be indicated that there is a need for conducting similar studies with larger sample sizes across various contexts to further explore prospective EFL teachers' perceptions regarding corpus consultation for language teaching. It is essential to provide the interaction of pre-service EFL teachers with the consultation of corpora in initial teacher training during their undergraduate education in more longitudinal studies to equip EFL teacher candidates with corpus literacy in future studies. This could offer deeper insights into the effects of integrating corpora into language teacher training programs and whether integrating corpora into initial teacher training gains traction in the teaching practices of EFL teachers.

Further studies might also be designed with a more extended implementation of the CBLT training program and a more comprehensive scope to focus on the use of corpora for other pedagogical purposes of corpus consultation, including the creation of language test items and guiding teacher candidates on the direct use of corpora with language learners. Future studies might also pay attention to the processes of building context-specific corpus and corpus-informed activities as these might enhance language instruction and their corpus literacy. Due to the nature of the practicum, the participants of the study could only teach two lessons throughout the semester and the training. Implementing longitudinal training programs with more opportunities for pre-service EFL teachers to integrate corpora in their actual language teaching practices can shed light on the long-term effects of such training and the effectiveness of using corpus-informed teaching materials in language classrooms. Future studies might be conducted

to investigate the influence of affective factors such as individual differences, learning styles and strategies, motivation, and emotional states of language practitioners on their use of corpora and teaching preferences in their language teaching practices.

As the participants were assigned to various schools and grades for their practicum process, the specific classroom, students, and the topic they were assigned to teach might have an influence on their perceptions and views regarding the implementation of corpora into language teaching. Future empirical studies might investigate the effectiveness of incorporating corpus-based language pedagogy and corpus-informed teaching materials and activities into different grade levels and contexts to bring light to the perceptions of EFL teachers and their students regarding the actual application of corpus in language pedagogy. Future studies might also be conducted to investigate whether EFL teachers can effectively integrate such corpus-based activities and apply the skills they were equipped with into their actual teaching practice. This might provide in-depth insights into the design and integration processes of further training on integrating corpora for language teaching with the feedback received from the participants of earlier training.

Finally, future studies might be conducted to explore the impact of using AI-based generative language models and such emerging technologies hand in hand with the integration of corpora into language teaching environments. By examining the interaction between corpora and AI, the benefits of the combined use of such technologies in enhancing effective corpora consultation, identifying language patterns, and generating insights to further enrich teaching practices might be uncovered. This investigation has the potential to shed light on how emerging technologies can work together to enhance language education. Additionally, future research can explore how corpora consultation and integration can be applied to teaching different languages, broadening the scope beyond English.

REFERENCES

- Abdel Latif, M. M. M. (2021). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices. *ReCALL*, 33(1), 34–48.
- AlAfnan, M. A., Samira Dishari, Marina Jovic, & Koba Lomidze. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3(2), 60–68.
- Alderson, C. (1996). Do corpora have a role in language assessment? In J. Thomas & M. Short (Eds.), *Using corpora for language research* (pp. 248–259). Longman.
- Allan, G. (2003). A critique of using grounded theory as a research method. *Electronic Journal of Business Research Methods*, 2(1), 1-10.
- Alsahafi, M. (2022). Corpora and Concordancing in second language vocabulary learning. *Theory and Practice in Language Studies*, 12(12), 2651-2657.
- Anderson, W. (2018). *Methods for text mining and analysis of text corpora basic analysis handout* [Class Handout]. School of Critical Studies, University of Glasgow. <https://www.gla.ac.uk/schools/critical/aboutus/resources/stella>
- Anthony, L. (2016). Introducing corpora and corpus tools into the technical writing classroom through Data-Driven Learning (DDL). In J. Flowerdew & T. Costley (Eds.), *Discipline-specific writing: Theory into practice* (pp. 162-180). Routledge.
- 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.
- Aşık, A., Vural, A. S., & Akpınar, K. D. (2016). Lexical awareness and development through data driven learning: Attitudes and beliefs of EFL learners. *Journal of Education and Training Studies*, 4(3), 87–96.
- Bal-Gezegin, B., Akbaş, E., & Başal, A. (2022). “Corpus made my job easier”: Preservice language teachers' corrective feedback practices in writing with

- corpus consultation. In G. Yangın-Ekşi, S. Akayoğlu, & L. Anyango (Eds.), *New directions in technology for writing instruction* (pp. 279–295). Springer Cham.
- 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.
- Barnes, A. & Murray, L. (1999). Developing the pedagogical ICT competence of modern foreign languages teacher trainees. Situation: All change and plus ça change. *Journal of Information Technology in Teacher Education*, 8(2), 165-180.
- Başal, A., Akbaş, E., & Bal-Gezegin, B. (2024). Corpus literacy training for in-service English language teachers. *Profile: Issues in Teachers' Professional Development*, 26(1), 65–80.
- Bennett, C., & Uí Dhonnchadha, E. (2023). Becoming corpus literate: An in-service EFL teacher education framework for integrating corpora into EFL teaching. *Applied Corpus Linguistics*, 3(1), 100048.
- Bennett, G. R. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. University of Michigan Press.
- 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). John Benjamins.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews. *Quality & Quantity*, 36, 391-409.
- Boulton, A. (2009). Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 35(1), 81-106.
- Boulton, A. (2010). Data-driven learning: taking the computer out of the equation. *Language Learning*, 60(3), 534-572.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393.
- Bouzekria, H., Mashaqba, B., Al Khalaf, E., & Huneety, A. (2023). Production of pragmatic routines by Algerian EFL learners: The effect of corpus-based instruction. *Ampersand*, 10, 100122.

- Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, 17(1), 47-64.
- Breyer, Y. (2009). Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153-172.
- Bunting, J. D. (2013). *An investigation of language teachers' explorations of the use of corpus tools in the English for Academic Purposes (EAP) class*. [Doctoral Dissertation, Georgia State University]. Georgia State University Research Depository. https://scholarworks.gsu.edu/alesl_diss/26/
- Can, H. (2023). Using corpora in teaching vocabulary to advanced EFL learners in a higher education context. *Language Learning in Higher Education*, 13(2), 461-477.
- Callies, M. (2019). Integrating corpus literacy into language teacher education. *Learner Corpora and Language Teaching*, 92, 245.
- Cangır, H. (2022). Using corpus tools for academic writing in EFL settings: A data-driven learning approach. In G. Yangın-Ekşi, S. Akayoğlu, & L. Anyango (Eds.), *New directions in technology for writing instruction* (pp. 171–191). Springer Cham.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, 9(2), 11-125.
- Chambers, A. (2010). What is data-driven learning? In A. O’Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 345- 358). Routledge.
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research-practice gap. *Language Teaching*, 1-16.
- Chambers, A., Farr, F., & O’Riordan, S. (2011). Language teachers with corpora in mind: From starting steps to walking tall. *Language Learning Journal*, 39(1), 85–104.
- Chang, (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, 26(2), 243-259.
- Charles, M. (2022). Corpora and autonomous language learning. In R. Jablonkai and E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 406–419). Routledge.
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Sage Publications.

- Chen, L. (2017). Corpus-aided business English collocation pedagogy: An empirical study in Chinese EFL learners. *English Language Teaching*, 10(9), 181-197.
- Chen, M., Flowerdew, J., & Anthony, L. (2019). Introducing in-service English language teachers to data-driven learning for academic writing. *System*, 87, 1-12.
- Chujo, K., Oghigian, K., & Akasegawa, S. (2015). A corpus and grammatical browsing system for remedial EFL learners. In A. Leńko-Szymańska, & A. Boulton (Eds.). *Multiple affordances of language corpora for data-driven learning* (pp. 109–128). John Benjamins.
- Conrad, S. (2000). Will corpus revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34(3), 548-560.
- Conza-Armijos, H. I., & Celi-Celi, L. F. (2023). Data-driven Learning in an EFL class: A study of Ecuadorian learners' perceptions. *INNOVA Research Journal*, 8(3), 37-50.
- Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Sage Publications.
- Creswell, J. W., and V. L. Plano Clark. 2011. *Designing and Conducting Mixed Methods Research* (2nd ed.). Sage Publications.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.
- Creswell, J. W. (2013). *Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
- Crosthwaite, P. (2017). Retesting the limits of data-driven learning: feedback and error correction. *Computer Assisted Language Learning*, 30(6), 1-26.
- Cushing, S. T. (2017). Corpus linguistics in language testing research. *Language Testing*, 34(4), 441-449.
- Çalışkan, G. (2020). *An investigation into EFL university instructors' perspectives on the use of corpora in language teaching* [Unpublished master's thesis]. Anadolu University.
- 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), 130-144.

- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 190-210.
- Demir, Y., & Ertaş, A. (2014). A suggested eclectic checklist for ELT coursebook evaluation. *The Reading Matrix*, 14(2), 243-252.
- Đurović, Z., & Bauk, S. (2022). Testing corpus linguistics methods in ESP vocabulary teaching. *Journal of Teaching English for Specific and Academic Purposes*, 10(2), 319-337.
- Ebrahimi, A. & Faghih, E. (2017). Integrating Corpus Linguistics Into Online Language Teacher Education Programs, *ReCALL*, 29(1), 120-135.
- Emir, G. (2019). *An examination of the effectiveness of corpus use as a data-driven learning tool in process writing* [Unpublished master's thesis]. Gazi University.
- Emir, G., & Yangın Ekşi, G. (2023). Corpus used as a data-driven learning tool in 12 academic writing: Evidence from Turkish contexts. *TEFLIN Journal - A publication on the teaching and learning of English*, 34(2), 209-225.
- 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*, 6(2), 110-123.
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43.
- Farr, F. (2010). How can corpora be used in teacher education? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 620- 632). Routledge.
- Farr, F., & O'Keeffe, A. (2019). Using corpus approaches in English language teacher education. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 268–282). Routledge.
- Farr, F., & Leńko-Szymańska, A. (2023). Corpora in English language teacher education: Research, integration, and resources, *TESOL Quarterly*, 58.
- Fligelstone, S. (1993). Some reflections on the question of teaching from a corpus linguistics perspective. *ICAME Journal*, 17, 97-109.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393-417.

- Flowerdew, L. (2015). Data-driven learning and language learning theories. In A. Leńko-Szymańska, & A. Boulton (Eds.). *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). John Benjamins.
- Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, 45(4), 475- 489.
- Friginal, E. (2018). *Corpus Linguistics for English Teachers: Tools, Online Resources, and Classroom Activities* (1st ed.). Routledge.
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells? *TESLEJ*, 8(4), 1-37.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32(3), 301–319.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. John Benjamins.
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(2), 225–242.
- Gilmore, A. (2008). Using online corpora to develop students' writing skills. *ELT Journal*, 63(4), 363-372.
- Gilmore, A. (2019). Materials and authenticity in language teaching. In S. Mann & S. Walsh (Eds.), *The Routledge handbook of English language teacher education* (pp. 299–318). Routledge.
- Gilquin, G & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359- 370). Routledge.
- Girgin, U. (2011). *Corpus-based activities at lower levels of EFL proficiency: The effectiveness of using concordance lines on grammar learning* [Unpublished master's thesis]. Bilkent University.
- Girgin, U. (2019). Perceptions of Turkish EFL student teachers towards learning phrasal-prepositional verbs through corpus-based materials. *Language Teaching and Educational Research (LATER)*, 2(1), 1-19.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.
- Godwin-Jones, R. (2017). Data-informed language learning. *Language Learning & Technology*, 21(3), 9–27.

- Götz, S. (2022). Learner corpora to inform testing and assessment. In R. Jablonkai and E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 311–325). Routledge.
- Greaves, C & Warren, M. (2010). What can a corpus tell us about multi-word units? In A. O’Keeffe and M. McCarthy (eds), *The Routledge handbook of corpus linguistics* (pp. 212- 226). Routledge.
- Guetterman, T. C., & James, T. G. (2023). A software feature for mixed methods analysis: The MAXQDA Interactive Quote Matrix. *Methods in Psychology*, 8, 100116.
- 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.
- Huang, L. S. (2011). Corpus-aided language learning. *ELT Journal*, 65, 481– 484.
- Huang, L. S. (2017). Has corpus-based instruction reached a tipping point?: practical applications and pointers for English-as-an-additional-language teachers. *TESOL Journal*, 8(2), 295-313.
- Huang, L. S. (2018). Taking stock of corpus-based instruction in teaching English as an international language. *RELC Journal*, 49(3), 381-401.
- Huberman, M., & Miles, M. B. (2002). *The qualitative researcher's companion*. Sage Publications.
- Hughes, R. (2010). What a corpus tell us about grammar teaching materials? In A. O’Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 401- 413). Routledge.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Journal of Mixed Methods Research*, 1(3), 271-294.
- Johns, T. (1986). Micro-concord: A language learner's research tool. *System*, 4(2), 151-162.
- Johns, T. (1991), Should you be persuaded?: Two examples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.
- Karras, J.N. (2016). The effects of data-driven learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL*, 28(2), 166–186.

- Kaya, Ömer F., Uzun, K., & Cangır, H. (2022). Using corpora for language teaching and assessment in L2 writing: A narrative review. *Focus on ELT Journal*, 4(3), 46–62.
- Kayaoğlu, M. N. (2013). The use of corpus for close synonyms. *The Journal of Language and Linguistic Studies*, 9(1), 128-144.
- Kazaz, İ. (2015). *Corpus-aided language pedagogy: the use of concordance lines in vocabulary instruction* [Unpublished master's thesis]. Bilkent University.
- Khan, M. Y., Rasheed, N., & Rasheed, S. (2021). Teaching grammar through literature to EFL learners: A corpus approach. *Journal of English Language, Literature and Education*, 2(03), 19-52.
- Koçak, A. (2020). EFL students' perceptions of using COCA to develop their vocabulary. *Journal of Language Research*, 4(1), 12-26.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.
- Köse, B., & Kuru Gönen, S. İ. (2023). Training on Using Corpora in EFL Classrooms: Perceptions of Pre-Service Teachers. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi*, 21(4), 259-278.
- Kučera, H., Francis, W. N., Twaddell, W. F., Marckworth, M. L., Bell, L. M., & Carroll, J. B. (1967). *Computational analysis of present-day American English*. Brown University Press.
- Landau, S. I. (Ed.). (2000). *Cambridge Dictionary of American English* (Klett Ed.). Ernst Klett Sprachen.
- Lakew, A. K., Teshome, S., & Negede, D. (2021). Introducing corpus-informed spoken grammar instruction in EFL classrooms: The Ethiopian experience. *Corpus Pragmatics*, 5(4), 487-510.
- Lee, S. (2011). Challenges of using corpora in language teaching and learning. *Linguistic Research*, 28(1), 159–178.
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721–753.
- Leńko-Szymańska, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL*, 26(2), 260–278.

- Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. In A. Turula, B. Mikolajewska & D. Stanulewicz (Eds.). *Insights into technology enhanced language pedagogy* (pp. 129–144). Peter Lang GmbH, Internationaler Verlag der Wissenschaften.
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217–241.
- Leńko-Szymańska, A., & Boulton, A. (2015). Introduction: data-driven learning in language pedagogy. In A. Leńko-Szymańska, & A. Boulton (Eds.). *Multiple affordances of language corpora for data-driven learning* (pp. 1–14). John Benjamins.
- Leong, P., S. R. H. Joseph, & R. Boulay. (2010). Applying Constant Comparative and Discourse Analyses to Virtual Worlds Research. *Journal of Virtual Worlds Research* 3(1), 3–26.
- Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology*, 21(3), 153–171.
- Lin, M.H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*, 50(4), 871–893.
- Lin, M. H. (2019). Becoming a DDL teacher in English grammar classes: A pilot study. *The Journal of Language Teaching and Learning*, 9(1), 70-82.
- Lin, M. H., & Lee, J. (2019). Pedagogical suitability of data-driven learning in EFL grammar classes: A case study of Taiwanese students. *Language Teaching Research*, 23(5), 541-561.
- Liu, D. (2013). Using corpora to help teach difficult-to-distinguish English words. *English Teaching*, 68(3), 27-50.
- Lombard, M., Snyder-Duch, J. & Bracken, C. (2004). Practical resources for assessing and reporting intercoder reliability in content analysis research projects. *Human Communication Research*, 28(4), 587-604.
- Lu, C., & Dang, T. N. (2022). Vocabulary in EAP learning materials: What can we learn from teachers, learners, and corpora? *System*, 106, 102791.
- Luo, Q., & Liao, Y. (2015). Using corpora for error correction in EFL learners' writing. *Journal of Language Teaching and Research*, 6(6), 1333-1342.

- Ma, Q., Tang, J. & Lin, S. (2021) The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, 35(9), 2731-2760.
- Ma, Q., Yuan, R., Cheung, L. M. E. & Yang, J. (2022) Teacher paths for developing corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*, 37(3), 1-32.
- Ma, Q., Chiu, M. M., Lin, S., & Mendoza, N. B. (2023). Teachers' perceived corpus literacy and their intention to integrate corpora into classroom teaching: A survey study. *ReCALL*, 35(1), 19–39.
- Mayoh, J., Bond, C. S., & Todres, L. (2012). An innovative mixed methods approach to studying the online health information-seeking experiences of adults with chronic health conditions. *Journal of Mixed Methods Research*, 6(1), 21-33.
- McCarthy, M. (2004). *Touchstone: From corpus to classroom*. Cambridge: Cambridge University Press.
- McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, 41(4), 563–574.
- McCarthy, M., McCarten, J., & Sandiford, H. (2005). *Touchstone series*. Cambridge University Press.
- McCarthy, M., McCarten, J., & Sandiford, H. (2014). *Touchstone series* (2nd Ed). Cambridge University Press.
- McCarthy, M. & O'Keeffe, A. (2010). Historical perspective: what are corpora and how have they evolved? In A. O'Keeffe and M. McCarthy (eds), *The Routledge Handbook of Corpus Linguistics* (pp. 3- 13). Routledge.
- McEnery, & Xiao. (2011). What corpora can offer in language teaching and learning? In Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (1st Ed.). Routledge.
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Meunier, F., & Reppen, R. (2015). Corpus versus non-corpus-informed pedagogical materials: Grammar as the focus. *The Cambridge Handbook of English Corpus Linguistics*, 498–514.

- Michali, M., & Patsala, P. (2020). *Learning experiences in secondary education: A study exploring Greek EFL learners' attitudes towards corpus-based teaching*. In *14th International Technology, Education and Development Conference*.
- Mindt, D. (1996). English corpus linguistics and the foreign language teaching syllabus. *Using Corpora for Language Research*, 232-247.
- Mizumoto, A., & Chujo, K. (2015). A meta-analysis of data-driven learning approach in the Japanese EFL classroom. *English Corpus Studies*, 22, 1–18.
- Moon, R. (2010). What can a corpus tell us about language? In A. O’Keeffe and M. McCarthy (eds), *The Routledge handbook of corpus linguistics* (pp. 197- 211). Routledge.
- Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor and T. Upton (eds) *Applied corpus linguistics: A multi-dimensional perspective* (pp.239-250). Rodopi.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: the state of the art and beyond. In S. Braun, K. Khon & J. Mukherjee (eds.) *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5–24). Frankfurt am Main: Peter Lang.
- Naismith, B. (2017). Integrating corpus tools on intensive CELTA courses. *ELT Journal*, 71(3), 273–283.
- Ninaus, M., & Sailer, M. (2022). Closing the loop – The human role in artificial intelligence for education. *Frontiers in Psychology*, 13.
- Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.
- Nunan, D. (2003) *Practical English Language Teaching*. McGraw-Hill.
- O’Connor, C. & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guideline. *International Journal of Qualitative Methods*, 19, 1–13.
- O’Keeffe, A., & Farr, F. (2003). Using language corpora in initial teacher education: Pedagogic issues and practical applications. *TESOL Quarterly*, 37(3), 389–418.
- O’Keeffe, A., McCarthy, M. & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge.

- O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing, 15*, 49-68.
- Oveidi, M., Sepehri, M., & Shafiee, S. (2022). Differential effects of input/output tasks on learning English collocations by Iranian EFL learners through the corpus-based instruction. *Journal of Modern Research in English Language Studies, 9*(4), 143-168.
- Özbay, A. Ş. (2017). Turkish EFL English majors' perceptions towards the implementation of "Introduction To Corpus Linguistics" course in an English department in Turkey. *Electronic Turkish Studies, 12*(34).
- Özbay, A. Ş., & Gürsoy, Z. (2023). Computerized corpus as a tool for educational technology and learning in the analysis of four-word recurrent expressions. *Journal of Educational Technology & Online Learning, 6*(1), 249-272.
- Özbay, A. Ş. & Kayaoğlu, M. C. (2015). EFL teachers' reflections towards the use of computerized corpora as a teaching tool in their classrooms. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 44*(1), 85-104.
- Özer, M., & Şükrü Özbay, A. (2022). Exploring the data-driven approach to grammar instruction in the ELT context of Turkey. *Australian Journal of Applied Linguistics, 5*(2), 35-63.
- Park, K. (2014). Corpora and language assessment: The state of the art. *Language Assessment Quarterly, 11*(1), 27-44.
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative Social Work, 1*(3), 261-283.
- Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning, 35*(1–2), 36–61.
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice, 20*(2).
- Pokrivcakova, S. (2023). Pre-service teachers' attitudes towards artificial intelligence and its integration into EFL teaching and learning. *Journal of Language and Cultural Education, 11*(3), 100-114.

- Poole, R. (2022). "Corpus can be tricky": revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641.
- Rodríguez-Fuentes, R. A., & Swatek, A. M. (2022). Exploring the effect of corpus-informed and conventional homework materials on fostering EFL students' grammatical construction learning. *System*, 104, 102676.
- Quinn, C. (2015). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, 69(2), 165–177.
- Razavi, A. R. (2014). ESL/EFL Coursebook Evaluation Checklist. https://www.academia.edu/7561620/ESL_EFL_course_book_evaluation_checklist
- Richard, J. C. (2017). *Curriculum development in language teaching* (2nd Ed). Cambridge University Press.
- Römer, U. (2004). A corpus-driven approach to modal auxiliaries and their didactics. In J. McH. Sinclair (Ed.) *How to use corpora in language teaching* (pp.185-199). John Benjamins.
- Römer, U. (2005). *Progressives, patterns, pedagogy: A corpus-driven approach to English progressive forms, functions, contexts and didactics*. John Benjamins.
- Römer, U. (2009). Corpus research and practice. In K. Aijmer (Ed.) *Corpora and language teaching* (pp.83-98). John Benjamins.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 37, 205–225.
- Saddhono, K., Rohmadi, M., Setiawan, B., Suhita, R., Rakhmawati, A., Hastuti, S., & Islahuddin, I. (2023). Corpus linguistics use in vocabulary teaching principle and technique application: A study of Indonesian language for foreign speakers. *International Journal of Society, Culture & Language*, 11(1), 231-245.
- Sanosi, A. B. (2022). The use and development of lexical bundles in Arab EFL writing: A corpus-driven study. *Journal of Language and Education*, 8(2), 106-121.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–58.
- Schmidt, N. (2023). Unpacking second language writing teacher knowledge through corpus-based pedagogy training. *ReCALL*, 35(1), 40–57.

- Sinclair, J. (1987). *Collins COBUILD English language dictionary*. Harper Collins Publishers.
- Sinclair, J. McH. (2004). Introduction. In J. McH. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 1-10). John Benjamins.
- 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), 106-119.
- Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26, 184–201.
- Stevenson, A. (Ed.). (2010). *Oxford Dictionary of English*. Oxford University Press.
- Summers, D., & Gadsby, A. (Eds.). (1995). *Longman Dictionary of Contemporary English*. Longman.
- Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research*, 27(3), 660-688.
- Şen Bartan, Ö. (2023). Using corpus tools for stylistic analysis in translator training. *Istanbul University Journal of Translation Studies*, 18, 19–40.
- Şimşek, T. (2016). Turkish EFL learners' reflections on corpus-based language teaching. *Global Journal of Foreign Language Teaching*, 6(1), 21-27.
- Şimşek, T. (2020). *Corpora in foreign language teacher education: Introducing a corpus literacy course to ELT pre-service teachers* [Unpublished doctoral dissertation]. Çukurova University.
- Şimşek, T., & Can, C. (2023). Integration of corpus literacy consultation into language teacher education: Course design, implementation, and evaluation. *Porta Linguarum Revista Interuniversitaria De Didáctica De Las Lenguas Extranjeras*, 39, 193-211.
- Tammelin, M. (2001). Empowering the Language Teacher through ICT Training and Media Education. Case HSEBA. In *ESADE, CALL in the 21st Century*.
- Tekin, B., & Soruç, A. (2016). Using Corpus-assisted learning activities to assist vocabulary development in English. *TOJET: The Turkish Online Journal of Educational Technology, Special Issue for INTE 2016*, 1270–1283.
- Timmis, I. (2015). *Corpus linguistics for ELT*. Taylor & Francis.

- Tognini Bonelli, E. (2010). Theoretical overview of the evolution of corpus linguistics. In A. O’Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 14- 27). Routledge.
- Tono, Y., Satake, Y., Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, 26(2), 147–162.
- Tribble, C. (2000) Genres, keywords, teaching: Towards a pedagogic account of the language of project proposals. In L. Burnard and T. McEnery (Eds.), *Rethinking Language Pedagogy from a Corpus Perspective. Papers from the Third International Conference on Teaching and Language Corpora* (pp. 75–90). Peter Lang.
- Tribble, C. (2015). Teaching and language corpora: Perspectives from a personal journey. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 37–62). John Benjamins.
- Tsai, K. J. (2019). Corpora and dictionaries as learning aids: inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning*, 32(8), 805-826.
- Uçar, S., & Yükselir, C. (2015). The effect of corpus-based activities on verb-noun collocations in EFL classes. *The Turkish Online Journal of Educational Technology*, 14(2), 195–205.
- Vannestel, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 19(3), 329-350.
- Varley, S. (2009). I’ll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, 22(2), 133-152.
- Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(2), 207–226.
- Vyatkina, N., (2020). Corpora as open educational resources for language teaching. *Foreign Language Annals*. 53(2), 359–370.
- Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1–8.
- Warschauer, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly*, 34(3), 511-535.

- Wazni, A. K., Mansouri, S., & Sattar Boroujeni, S. (2023). A contrastive study of boosters in a corpus of academic spoken English. *BELT - Brazilian English Language Teaching Journal*, 14(1), e45028.
- Xodabande, I. & Nazari, M. (2022). Impacts of a Corpus Linguistics Course on in-Service EFL Teachers' Corpus Literacy. *CALL-EJ*, 23, 318-346.
- Yılmaz, N., & Koban Koç, D. (2020). Developing pragmatic comprehension and production: Corpus-based teaching of formulaic sequences in an EFL setting. *Journal of Language and Linguistic Studies*, 16(1), 474-488.
- Yiru, Z., & Le, Z. (2024). A corpus-driven perspective on the verb valency structure of Chinese EFL learners: A case study of "Suggest" and "Advise". *Sino-US English Teaching*, 21(1), 14-18.
- Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology*, 12(2), 31–48.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.
- Zareva, A. (2017). Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1), 69–79.
- Zhang, R. (2022). Developing morphological knowledge with online corpora in an ESL vocabulary classroom. *Frontiers in Psychology*, 13.
- Zhong, X., & Wakat, G. (2023). Enhancing grammar proficiency of EFL learners through corpus-integrated lessons. *Ampersand*, 11, 100139.

Appendix-A. Sample Consent Form

Gönüllü Katılım Formu

Bu çalışma, “Derlem ve Veri Yönetimli Öğrenmenin İkinci Dil Olarak İngilizce Öğretimi Bağlamında Kullanımı Üzerine Bir Çalışma: İngilizce Öğretmeni Adaylarının Algıları” başlıklı bir araştırma çalışması olup İngilizce öğretmeni adaylarının derlem araçlarının ikinci dil olarak İngilizce öğretiminde kullanımı üzerine düşüncelerini belirleme amacını taşımaktadır. Çalışma, Berkcan Köse tarafından yürütülmekte ve sonuçları ile İngilizce öğretiminde derlem araçlarının kullanımına dair gelişime ışık tutulacaktır.

- Bu çalışmaya katılımınız gönüllülük esasına dayanmaktadır.
- Çalışmanın amacı doğrultusunda, bir anket kullanılarak sizden veriler toplanacaktır.
- İsminizi yazmak ya da kimliğinizi açığa çıkaracak bir bilgi vermek zorunda değilsiniz/araştırmada katılımcıların isimleri gizli tutulacaktır.
- Araştırma kapsamında toplanan veriler, sadece bilimsel amaçlar doğrultusunda kullanılacak, araştırmanın amacı dışında ya da bir başka araştırmada kullanılmayacak ve gerekmesi halinde, sizin (yazılı) izniniz olmadan başkalarıyla paylaşılmayacaktır.
- İstemeniz halinde sizden toplanan verileri inceleme hakkınız bulunmaktadır.
- Sizden toplanan veriler korunacak ve araştırma bitiminde arşivlenecek veya imha edilecektir.
- Veri toplama sürecinde/süreçlerinde size rahatsızlık verebilecek herhangi bir soru/talep olmayacaktır. Yine de katılımınız sırasında herhangi bir sebepten rahatsızlık hissederseniz çalışmadan istediğiniz zamanda ayrılabilirsiniz. Çalışmadan ayrılmanız durumunda sizden toplanan veriler çalışmadan çıkarılacak ve imha edilecektir.

Gönüllü katılım formunu okumak ve değerlendirmek üzere ayırdığınız zaman için teşekkür ederim. Çalışma hakkındaki sorularınızı Orta Doğu Teknik Üniversitesi İngilizce Öğretmenliği bölümünden Berkcan Köse’ye yöneltebilirsiniz.

Araştırmacı Adı : Berkcan Köse

Adres :

İş Tel :

Cep Tel :

E-posta adresi :

Bu çalışmaya tamamen kendi rızamla, istediğim takdirde çalışmadan ayrılabileceğimi bilerek verdiğim bilgilerin bilimsel amaçlarla kullanılmasını kabul ediyorum.

(Lütfen bu formu doldurup imzaladıktan sonra veri toplayan kişiye veriniz.)

Katılımcı Ad ve Soyadı:

İmza:

Tarih:

Appendix-B. Background Survey



Corpus-based Language Teaching Training CBLT101 - Background Survey

Dear Participant,

This survey is designed to gather information about your ideas on corpus use and your experience on consulting a corpus. The study is being conducted by Berkcan KÖSE, a researcher pursuing his MA at the department of English Language Teaching at Anadolu University.

Your replies to this survey will be kept anonymous. Participation is completely voluntary, and your answers will be confidential and used for research purposes only. The data will not be shared with third parties.

By clicking "Yes, I agree to take part in this study," you acknowledge that your participation in the study is entirely voluntary and that you may withdraw at any moment for any reason. Your time and responses are highly valued.

If you have any questions about the study, feel free to contact the researcher.

Berkcan KÖSE

Research Assistant

Middle East Technical University

Do you agree to take part in this study? *

☐ Yes, I agree to take part in this study.

Demographic Information

1. Gender *

☐ Female

☐ Male

☐ Diğer:

2. Age: *

- ☐ 18-21
- ☐ 22-25
- ☐ 25+

3. Which grade/level are you teaching at the practicum school this term?

- ☐ Primary School - Grades 1-4
- ☐ Secondary School - Grades 5-8
- ☐ High School - Grades 9-12
- ☐ Diğer: _____

4. Please state the name of your current practicum school. *

Yanıtınız _____

Corpus Knowledge and Literacy

1. Have you ever heard the terms 'corpus' and/or 'corpora'? If **yes**, please explain *
where you heard them from, e.g. in a course, from a friend, etc.

Yanıtınız _____

2. Have you ever been informed on the use of corpora as a part of your *
undergraduate education? If **yes**, please indicate where and how.

Yanıtınız _____

3. Do you know any corpus tools and/or online corpora? If **yes**, please state *
briefly.

Yanıtınız _____

4. Have you ever used any corpus tools and/or online corpora? If **yes**, please explain how and for what purposes briefly. *

Yanıtınız

5. What is your level of knowledge about the following corpus tools and online corpora? *

	no knowledge	a little knowledge	some knowledge	extensive knowledge	expert knowledge
Corpus Tools (e.g., AntConc, Sketch Engine)	☐	☐	☐	☐	☐
Online Corpora (e.g., BNC, COCA)	☐	☐	☐	☐	☐

6. How would you evaluate your knowledge level related to the following topics? *

	no knowledge	a little knowledge	some knowledge	extensive knowledge	expert knowledge
Using corpus tools directly with students (e.g. analyzing concordance lines)	☐	☐	☐	☐	☐
Using corpus tools to create corpus-informed classroom materials	☐	☐	☐	☐	☐
Using corpus tools to create language test items	☐	☐	☐	☐	☐
Using corpus tools to check students' language use	☐	☐	☐	☐	☐
Using corpus tools to build my own language knowledge	☐	☐	☐	☐	☐

*

- ☐
- No

CLT101 Corpus Literacy Training

Please indicate your interest to the training by completing one final item.

*

- ☐ **Diğer:**

If no, just leave a dot in the box.

If you are keen on participating in the training, please indicate your availability by selecting the time slots that are convenient for you.

[illegible]

If you have additional notes regarding your availability, please indicate.

Yanıtınız

Thank you for your participation.

Thank you for your genuine answers. If you have any questions about the study, feel free to contact the researcher.

Res. Asst. Berkcan KÖSE

If you would like to add anything further, please state in this box. Thank you.

Yanıtınız

Appendix-C. Reflective Logs



Corpus-based Language Teaching Training CBLT101 - Reflective Log

Dear Participant,

This online reflective journal is designed to gather your feelings and ideas regarding the Corpus Training session.. Your time and responses are highly valued.

If you have any questions about the study, you are always welcome to contact me.

If you have any questions about the study, feel free to contact the researcher.
Berkcan KÖSE
Research Assistant
Middle East Technical University

[Hesap değiştir](#)

* Zorunlu soruyu belirtir

Please provide your initials (e.g. BK) *

Yanıtınız

Reflective Log

You are expected to write a reflective log following the training session of each week you attended by keeping the questions below in mind. Think about your experience in the specific training and reflect on your positive and negative ideas and feelings.

- Reflect on your experience in today's training session. How did it impact your understanding and skills related to corpus use?
- In today's session, what specific aspects of corpus use did you choose to delve into, and why did you find them significant?
- Consider the corpus use techniques you explored today. Can you reflect on how these techniques might be applied in real-world scenarios, and what advantages they could offer in research or language analysis?
- Were there any instances during today's session where you encountered difficulties while consulting corpora? If so, can you provide a deeper analysis of these challenges and potential strategies to overcome them in future practice-based sessions? (only for practice-based sessions)

Please share your thoughts, insights, or observations regarding the following questions. *

Yanıtınız

Which three adjectives would you choose to define today's training session? *

Yanıtınız

Thank you for your participation.

Thank you for your genuine answers. If you have any questions about the study, you are always welcome to contact me.

Berkcan KÖSE

Appendix-D. Post-teaching Reflective Logs



Corpus-based Language Teaching Training CBLT101 - Post-Assessed Teaching Reflective Journal

Dear Participant,

This online reflective journal is designed to gather your feelings and ideas regarding your use of corpus-based teaching materials in your assessed teaching. Your time and responses are highly valued.

If you have any questions about the study, you are always welcome to contact me.

If you have any questions about the study, feel free to contact the researcher.

Berkcan KÖSE

Research Assistant

Middle East Technical University

Post-Assessed Teaching Reflective Journal

You are expected to write a reflective journal regarding your use of corpus-based teaching materials in your assessed teaching.

1. Which corpus tools and/or online corpora did you use while preparing teaching materials for your teaching in the practicum school? *

Yanıtınız

2. How did you use corpus tools and/or online corpora for preparing teaching materials? *

Yanıtınız

3. How did you feel while integrating corpus-based teaching materials into your teaching practicum? Write down any benefits or drawbacks. *

Yanıtınız

4. Did using corpus-based teaching materials contribute to student participation in the classroom? If yes, how? *

Yanıtınız

Thank you for your participation.

Thank you for your genuine answers. If you have any questions about the study, you are always welcome to contact me.

Berkcan KÖSE
Research Assistant
Middle East Technical University



Appendix-E. Semi-structured Interview Questions

- 1. What do you think about the corpus training you have received?**
- 2. Did you use any corpus tools and/or online corpora while preparing teaching materials for your teaching practicum?**
- 3. How did you feel while using corpus-based teaching materials during your teaching practicum?**
- 4. What do you think about the integration of corpus-based materials into EFL classrooms?**
- 5. How do you feel about using corpus-based materials in your future teaching experiences?**

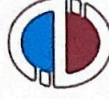


Appendix-F. The Ethical Committee Approval of Anadolu University

Evrak Kayıt Tarihi: 15.05.2023

Protokol No: 526481

Tarih: 30.05.2023



ANADOLU ÜNİVERSİTESİ
SOSYAL VE BEŞERÎ BİLİMLER BİLİMSEL ARAŞTIRMA VE YAYIN ETİĞİ KURULU
KARAR BELGESİ

ÇALIŞMANIN TÜRÜ:	Yüksek Lisans Tez Çalışması
KONU:	Eğitim Bilimleri
BAŞLIK:	Derlem ve Veri Yönetimli Öğrenmenin İkinci Dil Olarak İngilizce Öğretimi Bağlamında Kullanımı Üzerine Bir Çalışma: İngilizce Öğretmeni Adaylarının Algıları
PROJE/TEZ YÜRÜTÜCÜSÜ:	Doç. Dr. S. İpek KURU GÖNEN
TEZ YAZARI:	Berkcan KÖSE
ALT KOMİSYON GÖRÜŞÜ:	-
KARAR:	Olumlu
Prof. Dr. Saime ÖNCE (Başkan-İkt. ve İdari Bil. Fak.)	
Prof. Dr. M. Erkan ÜYÜMEZ (Başkan Yardımcısı -İkt. ve İdari Bil. Fak.)	Prof. Dr. Fatime GÜNEŞ (Edebiyat Fak.)
Prof. Dr. Yıldız UZUNER (Eğitim Fak.)	Prof. Dr. Erkan YÜKSEL (İletişim Bil. Fak.)
Prof. Dr. Handan DEVE (Eğitim Fak.)	Prof. Dr. Kamil ÇEKEROL (Açıköğretim Fak.)

Appendix-G. The Human Research Ethics Committee Approval of Middle East Technical University

Berkcan Köse - Tez - İnsan Araştırmaları Etik Kurulu (İAEK) Değerlendirme Süreci Başvurusu



İnsan Araştırmaları Etik Merkezi

To: Berkcan Köse

Cc: İpek

Sayın Safiye İpek Kuru Gönen

Danışmanlığını yaptığınız Berkcan KÖSE'nin "Ölçme Derlem ve Veri Yönetimli Öğrenmenin İkinci Dil Olarak İngilizce Öğretimi Bağlamında Kullanımı Üzerine Bir Çalışma: İngilizce Öğretmeni Adaylarının Algıları" başlıklı araştırması Anadolu Üniversitesinden aldığınız Etik Onay doğrultusunda üniversitemizde araştırma yapmanız uygun bulunmuştur.

Saygılarımızla bilgilerinize sunarız.

Aysel KÖSEGİL

UEAM ve İAEK Koordinasyon



Appendix-H. CBLT Training Booklet



23/27 October – 18/22 December 2023
Corpus Literacy Training
Training Booklet

CLT 101 Training Booklet

A Journey for Being
Corpus Literate

Prepared by
Res. Asst. Berkcan Köse

Table of Content

- 1 Introduction to the Training**
- 2 Training Schedule**
- 3 Training Content**
- 4 Extra Materials**
- 5 Conclusion**

We are delighted to embark on this educational journey with you as we delve into the world of corpus linguistics and its invaluable applications in the field of language education. This booklet serves as your guide to the comprehensive training program that awaits you.

Here, we will provide you with essential information about the training schedule, content, requirements, and the intricate details that will shape your experience over the next nine weeks.

Introduction

Welcome to the "Corpus Literacy Training"

The Corpus Literacy Training is set to commence on the week of 23-27 October and is designed to run for a total of nine weeks. During this training, we will explore corpus linguistics in depth, ensuring a well-rounded understanding and hands-on experience. The aim is to create a diverse and engaging learning environment for all, fostering collaboration, knowledge sharing, and the opportunity to build lasting professional connections.

To cater to the varied preferences of our participants, we have incorporated both online and face-to-face sessions. While the majority of our training sessions will be conducted in person, we have strategically included a few online sessions to maximize accessibility and convenience. The training is designed to be off-the-record and voluntary, yet it comes with a certification. There will be no formal evaluation or grading. Your main criterion for success in this training is full participation throughout the eight-week program.

The Corpus Literacy Training is designed to empower you as educators and provide you with the skills and knowledge needed to utilize corpora effectively in your teaching practices. Throughout the training, you will engage in various activities, including full participation in discussions and micro-teaching sessions, evaluating, adapting, and creating corpus-informed teaching materials, and even crafting your very own corpus.

This program also places a strong emphasis on applying corpus-informed teaching materials to your practicum teachings, ensuring that you leave with a practical and valuable skillset that can be immediately applied to your professional endeavors.

Training Schedule

Session 01 Module 1 - Introduction Introduction to the Training Discussion on creating teaching materials and the former corpus training - Introducing training content and framework - exploring learner needs and experiences - training objectives - defining a corpus - types of corpora - what can we do with a corpus?	Session 02 Module 1 - Introduction Introduction to Corpus Use in EFL Introducing basic concepts and corpus tools & online corpora, corpus use in EFL - What is a corpus? - What is a concordancer? - What is corpus linguistics? - How big should a corpus be? - What are corpora used for? - Different types of corpora - Key terms (keywords, lemma, lexical density, span token, type, word lists, concordances, frequency data, registers, genres.
Session 03 Module 2 - Exploitation Building a Corpus - Introduction to AntConc and SketchEngine - How to build a corpus? - Introducing building a corpus and its features.	

Training Schedule

Session 07

Module 3 - Evaluation

Micro-teachings on
Concordancer Practice & Tasks
Error Correction & Skill Integration

Material Evaluation and Adaptation Practice

MONE Textbooks
&
Corpus-informed Textbooks

Session 08

Module 3 - Evaluation

Evaluating materials
created in the previous
corpus training by the
participant pre-service
EFL teachers

Training Schedule

Session 04

Module 2 - Exploitation

Consulting Corpora for Teaching Vocabulary

Task 1:
Creating teaching materials
for vocabulary teaching

Session 05

Module 2 - Exploitation

Micro-teachings on Concordancer
Practice & Tasks
Teaching Vocabulary

Consulting Corpora for Teaching Grammar

Task 2:
Creating teaching materials for
grammar teaching

Session 09

Module 4 - Reflection

Discussion on creating corpus-
informed teaching materials
through concordances

How to use them in language
classrooms & overall reflection

Session 06

Module 2 - Exploitation

Micro-teachings on
Concordancer Practice & Tasks
Teaching Grammar

Consulting Corpora for Error Correction & Skill Integration

Task 3:
Creating teaching materials for
error correction & skill integration

Training Content

Online session

Session 01

Module 1 - Introduction Introduction to the Training

1. PowerPoint presentation about the training content and framework, training objectives, and introduction to the training booklet.
2. Discussion on creating teaching materials and the former corpus training; exploring learner needs and experiences; defining a corpus, types of corpora, and what we can do with a corpus.

Discussion context

3 An innovative corpus-based language pedagogy (CBLP)

As pointed out by Mukherjee (2006), corpus-based data-driven learning requires high learner autonomy and can only be successful for learners with basic corpus literacy (CL), which involves understanding what a corpus is, knowing what can or cannot be done with a corpus, knowing how to analyse corpus data and knowing how to draw conclusions about language use based on corpus data. Ma, Teng, and Lin (2021) argue that CL is about using corpora as a learning tool but may not automatically lead to pedagogical skills to apply corpora to classroom teaching. They further point out that teachers need both corpus-based language pedagogy (CBLP) in order to integrate corpora into classroom teaching. CBLP is built upon CL and is defined as: "the ability to integrate corpus linguistics technology into classroom language pedagogy to facilitate language teaching" (Ma et al., 2021, p. 2).

2. Corpus literacy

Römer (2009) carried out a teachers' needs analysis to gain information about the situation of teachers of English, their problems and wishes. Römer's questionnaire-based survey examined topics such as the quality of existing teaching materials, authenticity in language teaching, and the teachers' language competence and exam marking. The participants' answers resulted in the following wish list:

- "a wider range of better teaching materials with more interesting, longer and genuine texts, and with more exercises
- textbooks that reflect actual language use and cover variation
- textbooks that keep track of language development
- tools which enable an ad-hoc creation of suitable materials, e.g. work-sheets on particular lexical items or structures
- help with exam marking
- more reliable and more quickly accessible resources or reference tools that help them solve the question "Can you say that in English?", perhaps "an online service or database", as one teacher wrote
- an always available native-speaker consultant" (Römer 2009: 89).

Assignment

Read the article Touchstone from Corpus to Course Book by Michael McCarthy and answer the following questions for the discussion in the next session.

1. What do you use as a reference (dictionaries, textbooks, etc.) for teaching materials? In what ways, are corpora similar to / different from your references?
2. Look at Figure 4 in the booklet by McCarthy. Which function of the target structure is generally taught in EFL classrooms in Turkey? Why?
3. Do you think that the textbook that is used in your practicum context includes corpus evidence for the choice of words or grammar structures?
4. What might be the advantages of using corpus-based materials in EFL classrooms?

Training Content

Online session

Session 02

Module 1 - Introduction Introduction to Corpus Use in EFL

1. PowerPoint presentation about corpus, types of corpora, different concordancing tools and online corpora, and corpus use in EFL.
2. Reading assignment for discussion about using corpora in language teaching
3. Multiple tasks will be used to show how to be informed about how to consult corpus tools and online corpora, the frequency of words in English using corpora, make participants familiar with corpus-based language teaching, and provide further information about the use of corpora in language classrooms.
4. COCA practice.

What is a corpus?

A corpus (plural corpora) is a collection of electronic and manually recording spoken and written texts that are stored electronically.

What is collected in a corpus?

Spoken
Written

Representative texts (e.g. newspaper, magazine, website, etc.)
Representative texts (e.g. newspaper, magazine, website, etc.)

Task 1

What are the 10 most frequent words?

Task 2

do you think are the top three nouns
it calculate with the verb "make"?

Task 3

How many of the examples use "not" as an adverb?
How many of the examples use the word "not" in the plural context?

Task 4

What does a corpus provide?

A corpus

- is permanently accessible to anyone
- stores and recalls all the information that is stored in it.
- provides a vast number of contextualized, reliable, and authentic samples of the language
- gives accurate statistics regarding what is common and typical.
- reflects even the recent changes in the language if it is a constantly updated corpus
- can be used to design a syllabus, develop materials and classroom activities (grammar and vocabulary)

Figure 4: Touchstone from English grammar in use (grammar and word lists)

How many of the examples use "not" as an adverb?
How many of the examples use the word "not" in the plural context?

Concordance - Key word in context

A concordance is a screen display of a word or phrase in its context.

Training Content

Session 06

Module 2 - Exploitation
Consulting Corpora for Error Correction & Skill Integration

Assignment

Design two corpus-based teaching materials: one for error correction and one for skill integration at any proficiency level that you want. Take notes based on the following elements. Prepare your material accordingly, and please bring it to the next session for presentation.

Language proficiency level _____
Aim of the activity (presenting new grammar items, or practice) _____
Target grammar items _____
Genre (If any, e.g. spoken, academic, newspapers, etc.) _____
Activity type (e.g. multiple-choice, matching, gap-filling, etc.) _____
Activity format (printout materials vs. computer-based version) _____

Before the following session, participants will be assigned a task in which they are expected to prepare a corpus-based activity for their upcoming assessed teaching in light of the guidelines provided in the task in order to be involved in the material preparation process with firsthand experiences. The participants will prepare their teaching material based on the proficiency level of their teaching group and the expected theme/objective by the mentor teacher at the observation school.

Training Content

Session 07

Module 3 - Evaluation
Material Evaluation and Adaptation Practice
MONE Textbooks & Corpus-informed Textbooks

The 7th Session will start with participant presentations & micro-teachings on the teaching material they created and they will be given reflection questions after each micro-teaching (vocabulary, grammar error correction & skill integration) and macro-teachings (assessed teachings during practicum).

1. Discussion on the place and use of corpus tools and online corpora for material evaluation and adaptation practice.

2. Comparison and evaluation of MONE Textbooks used in the practicum observation schools and Corpus-informed textbooks (Touchstone series (McCarthy, McCarten, and Sandiford, 2005)

Dear Students,

Using the "Textbook Evaluation Checklist" provided, compare one of the books written by the Turkish Ministry of Education and one of the books from the Touchstone series written based on Cambridge International Corpus data.

Do not forget to provide specific notes from the books to prove your comments.

Group Members:

- 1.
- 2.
- 3.



Based on the assessed teaching material task, participants will show the first draft of their teaching material before the actual teaching. Based on the reflection from the researcher and their peers, they will adapt the material for the second draft and implement that material into their teaching.

They will be given reflection questions after the macro-teaching (assessed teaching during practicum).

Training Content

Session 08

Module 3 - Evaluation
Evaluating materials created in the previous corpus training
by the participant pre-service EFL teachers

The participants of the previous, small-scale, training conducted by the researcher created 19 teaching materials by consulting COCA. The participants of the current study will evaluate and, if needed, adapt, these materials by working in pairs. Those materials will be provided to the whole group and they will choose one or two materials of their liking.

Session 09

Module 4 - Reflection

Online session

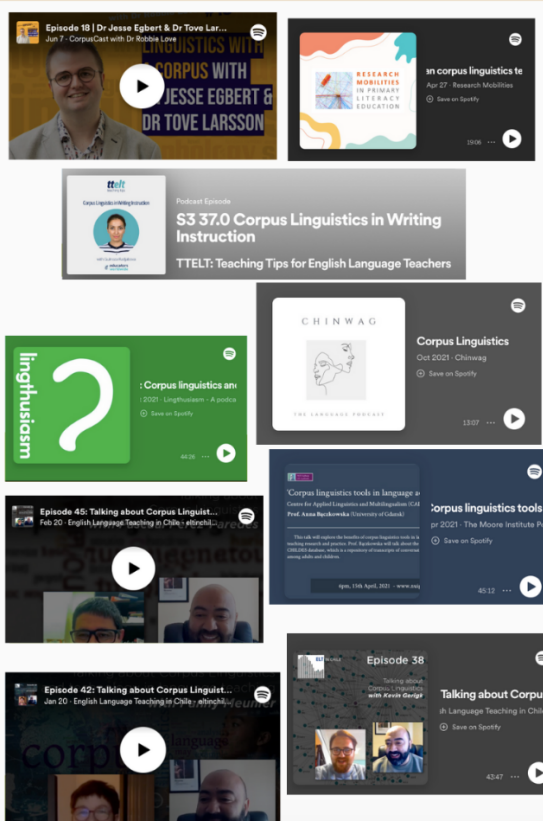
Discussion on creating corpus-informed teaching materials through concordances and how to use them in language classrooms & overall reflection

Semi-structured interviews will be scheduled for the following two weeks.

Extra Materials

- 01 Podcasts
- 02 Youtube Content & Videos
- 03 Assigned Readings
- 04 Recommended Journals
- 05 Resources to be used in EFL Classrooms

01 corpus podcasts



02 youtube content & videos

[Corpus Linguistics: The Basics](#)

[Corpus Linguistics for Beginners](#)

[What is corpus linguistics?](#)

[MOOC - Corpus linguistics: method, analysis, interpretation](#)

[Corpus linguistics: how can it help with English language teaching and learning? With Niall Curry](#)

[Key concepts in corpus linguistics](#)

[\(Very\) Practical Applications of Corpus Linguistics by Daniel Zuchowski](#)

[Corpus Linguistics, Learner Corpora & SLA. Prof Tony McEnery, 5 October 2022](#)

[THE INTERSECTION OF CORPORA AND LANGUAGE TEACHING: Practical Tips and Suggestions](#)

[Corpus Linguistics in the classroom - Teacher training video](#)

[Using Corpora in the Language Classroom | The New School](#)

[Corpus linguistics for language lovers. - Ondřej Matuška | PG 2019](#)

[Using a language corpus to design classroom activities](#)

[Using Corpus Linguistics in Language Teaching](#)

05 Resources to be used in EFL Classrooms

Online Corpora

[BNC Online](#)

[BYU - BNC](#)

[COCA](#)

[BNC - Phrases in English](#)

[Michigan Corpus of Academic Spoken English \(MICASE\)](#)

Collocation finders

[Just the word - collocation finder \(based on the BNC\)](#)

[For Better English - another collocation finder \(powered by Sketch Engine\)](#)

[Word and Phrase \(based on COCA\)](#)

Dictionaries

[Longman Dictionary of Contemporary English](#)

[Cambridge Dictionaries Online](#)

[Oxford Dictionaries](#)

[Merriam-Webster Dictionary and Thesaurus](#)

[Dictionary.com \(dictionary and thesaurus\)](#)

Corpus-based language teaching/learning tools and materials

[Tim Johns' Kibbitzers](#)

[MICASE ESL/EAP Teaching Materials](#)

[MICASE/MICUSP Kibbitzers Originals](#)

[MICUSP Simple - The Michigan Corpus of Upper-Level Student Papers](#)

[Michigan Corpus Linguistics - Archived Materials](#)

[Compleat Lexical Tutor](#)

[AWL Highlighter](#)

Web as Corpus

[Web as Corpus](#)

[WebCorp Live](#)

03 Assigned Readings

Research Papers & Dissertations

Grouped under each session
for classroom discussion

04 Recommended Journals for Corpus-related Research

URL link for access



Conclusion

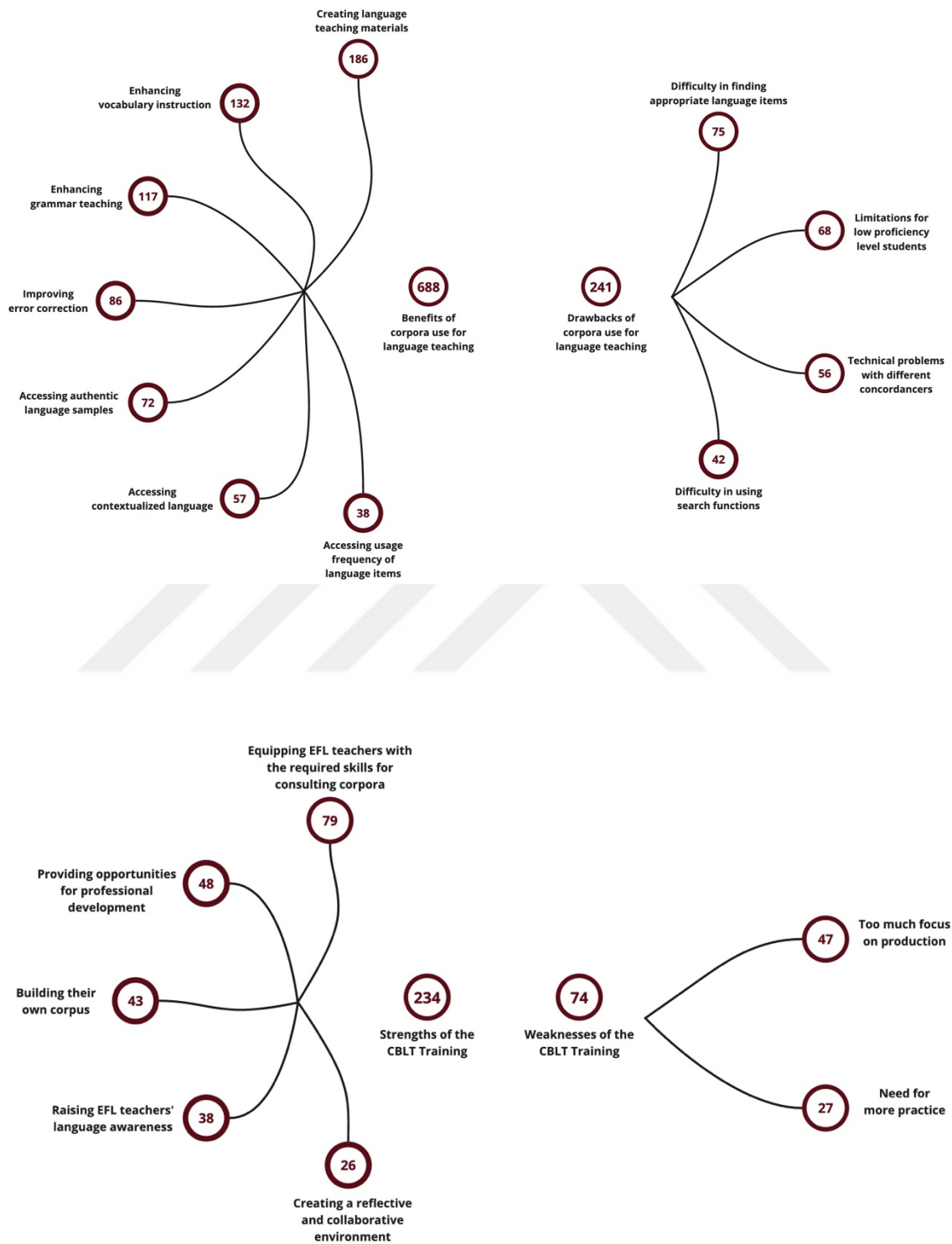
The "Corpus Literacy Training" offers a unique and transformative opportunity for English Language Teachers.

As you embark on this journey, you will not only enhance your pedagogical skills but also contribute to the future of language education.

We look forward to guiding you through this exciting and enriching experience.

Let's get started on this remarkable learning adventure! 🇪🇺🌟

Appendix-I. The Concept Maps of the Qualitative Analysis Taken from MAXQDA



CURRICULUM VITAE

Name, Surname : Berkcan KÖSE
Foreign Languages : English
Place and Year of Birth: :
E-mail :

Educational Background

B.A. 2021-Anadolu University, English Language Teaching Program,
Eskişehir

Work Experience

2022-Present Research Assistant, Middle East Technical University, Faculty of
Education, Department of Foreign Language Education, Ankara

Publications

Köse, B., & Kuru Gönen, S. İ. (2023). Training on using corpora in EFL classrooms: perceptions of pre-service teachers. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi*, 21(4), 259-278.