



Participatory Educational Research (PER)  
Vol.12(2), pp. 1-19, March 2025  
Available online at <http://www.perjournal.com>  
ISSN: 2148-6123  
<http://dx.doi.org/10.17275/per.25.16.12.2>

Id: 1548071

## Exploring Emerging Issues in Corpus Applications in ELT: A Systematic Review

Gamze EMIR

*Foreign Languages Education, Gazi University, Ankara, Türkiye*  
ORCID: 0000-0002-9705-9767

Nesli Çiğdem SARAL\*

*School of Foreign Languages, Gendarmerie and Coast Guard Academy, Ankara, Türkiye*  
ORCID: 0000-0001-5582-8495

---

### Article history

**Received:**  
11.09.2024

**Received in revised form:**  
13.11.2024

**Accepted:**  
03.12.2024

### Key words:

Corpora, corpus applications, data-driven learning, English language teaching and learning

The incorporation of corpora into language education has garnered a substantial amount of interest particularly within the context of second language and foreign language settings, with regard to language skills, components, and the teacher development. This systematic review explored emerging issues in corpus applications in English Language Teaching (ELT) by examining thirty-eight studies published in prominent journals from 2016 to 2023. In this review, the corpus studies were analysed with reference to the focused research topics, methodological design they employed, and the major findings that were presented. This systematic review provided valuable insights into the current state of corpus applications in ELT. The findings identified a predominance of quantitative research designs, highlighting a need for more qualitative and longitudinal studies to fully understand the dynamics of corpus use in language teaching and learning. Key findings revealed that while corpus applications as a Data-Driven Learning (DDL) tool enhanced learners' vocabulary, grammar, reading, writing, and pragmatic skills, there was limited focus on teachers' interaction with corpora. Increasing teachers' corpus literacy and fostering collaborative learning were found to be critical for teachers' effective integration of corpus tools into classroom practice. The review underscored the importance of comprehensive training and support for teachers to maximize the potential of corpus-based pedagogy.

---

## Introduction

With the advancement of computer technology, corpora have emerged as potent tools for language studies, increasingly finding their place in second/foreign language (SL/FL) classrooms (Ma et al., 2022). Particularly since Johns's (1990) seminal work on data-driven learning (DDL), where language learners employ corpus data for their language learning, there has been a notable increase in utilizing extensive language datasets for pedagogical purposes. This progression has not only reshaped our understanding of language but also transformed the methodologies of language acquisition (Tsai, 2019). Therefore, the integration of corpora has attracted significant attention, particularly within SL/FL settings.

---

\* Correspondency: [n.cigdemsaral@gmail.com](mailto:n.cigdemsaral@gmail.com)

Scholars now recognize the ‘multiple affordances’ of language corpora for DDL (Leńko-Szymańska & Boulton, 2015) across domains such as L2 acquisition and instruction, skills development, translation, materials development, and numerous other applications stemming from corpus-based pedagogical innovations. These advancements have proliferated as corpus software becomes more sophisticated and accessible, finding their way into language classrooms (Crosthwaite, 2017).

Corpora allows users to access vast, well-organized, and authentic language data (Pérez-Paredes, 2022). Corpus applications have been both indirect, which involves specialists utilizing information derived from corpora to enhance language descriptions, leading to the development of new dictionaries and other resources, and direct with language educators and learners independently accessing and utilizing corpora (Boulton & Vyatkina, 2021). Through the capabilities of computer software, learners and educators can explore extensive text collections relevant to their specific requirements, analysing frequencies, distributions, and multiple instances of target items in context (Boulton, 2017). This is where language corpora prove invaluable in what has become known as data-driven learning. The core principles of DDL emerged two decades after Johns’ initial publication on DDL, alongside such concepts as scaffolding, universal grammar, computer-mediated communication, and blended learning (Pérez-Paredes, 2022). DDL involves the hands-on utilization of authentic corpus data (concordances) by advanced foreign or second language learners in higher education for inductive, self-directed language learning of advanced usage (Boulton, 2011; Boulton & Vyatkina, 2021).

DDL is an exploratory learning methodology that empowers students to analyse lexical and grammatical patterns of keywords across authentic contexts, retrieved from a language corpus and organized into concordance lines (Hirata & Thompson, 2022). DDL enhances students’ linguistic capabilities and language awareness through their exploration and discovery of multiple related patterns from a wide array of examples (Chambers, 2010). Querying a corpus and comprehending concordance and collocation information offer an alternative to teacher-led, rule-based approaches to language pedagogy (Crosthwaite, 2017). Essentially, it is a constructivist, inductive approach wherein learners draw their own meaningful conclusions (Boulton, 2017; Pérez-Paredes, 2022). DDL is also conceptualized as fostering an independent and active learning environment that requires minimal teacher assistance; teachers primarily facilitate learning rather than engage in traditional instruction (Chambers, 2010). Therefore, in a DDL approach, students take responsibility for self-directed, learner-centred learning (Boulton & Cobb, 2017; Chambers, 2010; Hirata & Thompson, 2022). Flowerdew (2015) highlights how this approach promotes second language acquisition by encouraging students to consciously ‘notice’ gaps in their linguistic knowledge, as well as facilitating constructivist learning through student interaction with multiple data sources.

In terms of the benefits of corpora and data-driven learning, the relevant literature has demonstrated the efficacy of corpus in various aspects of SL/FL learning. Numerous studies have demonstrated the effectiveness of corpora in vocabulary acquisition (Karras, 2016; Lee et al., 2020; Lee & Lin, 2019; Li & Zhang, 2024; Otto, 2021; Yao, 2019), grammar comprehension (Crosthwaite & Steeples, 2022; Lin, 2021; Lin & Lee, 2015), writing proficiency (Chen et al., 2019; Crosthwaite, 2020; Larsen-Walker, 2017; Luo, 2016; Muftah, 2023; Sun & Hu, 2023), reading skills (Hadley & Hadley, 2021; Wilkinson, 2012), and pragmatic understanding (Li, 2019; Vyatkina & Belz, 2006; Zhang, 2021). Moreover, in addition to the effectiveness of corpora on language skills and components, several studies have shown that analysing learner corpora allows researchers to identify common errors



among learners (Lee et al., 2009; Rosen et al., 2014; Thewissen, 2013). Besides these studies with a pedagogical focus, recent meta-analyses (Boulton & Cobb, 2017) and critical assessments (Chen et al., 2019) have reviewed results and provided positive empirical support for the efficiency of corpus-aided approaches, leading to the prominence of corpus linguistics as a key strategy in language learning and research.

Recently, scholars have underscored the significance of corpora in fostering teacher development (Farr & O’Keeffe, 2019). They advocated for increasing teachers’ awareness of corpus-based instruction, which empowers learners to effectively engage with corpora (Callies, 2016; Farr & O’Keeffe, 2019). Given that language teachers hold the ultimate authority in introducing corpus work to learners (Breyer, 2009), scholars are directing their efforts toward incorporating corpus-based instruction into teacher education. In that regard, the relevant literature investigated pre-and in-service teachers’ attitudes toward corpus-aided language learning and teaching (Lin, 2016; Poole, 2020) and pre-and in-service teachers’ corpus literacy (Callies, 2019; Abdel Latif, 2021; Zareva, 2017). In the context of corpus literacy, these studies focused on equipping teachers with the skills to utilize corpus data as a learning tool (Çalışkan & Gönen, 2018), integrate corpus resources into language classrooms for effective teaching (Lin, 2019), and design effective teaching activities (Heather & Helt, 2012; Reppen, 2010). These studies concluded that pre-and in-service teachers should receive structured corpus training (Schmidt, 2023) so that “student teachers can experience the use of corpora for language learning from two perspectives: as learners and as teachers” (Breyer, 2009, p. 167).

The literature includes review studies regarding the significance of corpus applications in foreign/second language learning and teaching. In their review study, Boulton and Vyatkina (2021) presented that there were few studies based on theoretical foundations such as constructivism and sociocultural theory. They also indicated that more than half of the studies they analysed were conducted in similar research settings like the universities with language classes for several weeks or a semester. With regard to the methodology, the studies overwhelmingly employed quantitative design, especially experimental research. Likewise, the studies analysed by Pérez-Paredes (2022) showed that they were performed at universities with language students adopting qualitative research design mostly and lacked underlying theories. He also concluded that not being theory-led might keep DDL research focused on utility and language gains, preventing mainstream language teachers from comprehending DDL practice in SLA and language education contexts. In the same way, Dong, Zhao and Buckingham (2023) drew the conclusion from their bibliometric analysis that researchers could concentrate primarily on the integration of DDL into classroom instruction and teacher professional development in their future studies. These review studies also concluded that corpus applications have yet to reach mainstream status in second/foreign language teaching and learning (Chen et al., 2019; Pérez-Paredes, 2022). Despite nearly three decades of promising research, DDL remains relatively unknown and underutilized in everyday language learning classrooms (Hirata & Thompson, 2022). Scholars, in that regard, raise their concerns about a gap between advances in corpus studies and the practice of foreign/second language teaching (ELT), particularly in terms of a lack of teachers’ and material developers’ adaptation of corpus use (Perez-Parades, 2010, 2022; Vyatkina & Boulton, 2017).

As can be seen from the studies mentioned above, corpus application has yet to be fully implemented in the field of second/foreign language teaching. Thus, this article attempts to systematically review the corpus studies conducted in foreign language education and published in three prominent journals from 2016 to 2023. To this end, this review investigated

the corpus studies regarding their focused research topics, methodology designs, and results to reveal the emerging issues in corpus applications in foreign language education. Within this aim, this study addressed the following research questions:

- (1) What are the focal points of corpus studies conducted in English language education?
- (2) What are the methodological designs that the corpus studies in English language education adopted?
- (3) What are the major findings of the corpus studies conducted in English language education?

## **Methodology**

The present study examines the emerging issues in corpus research for English language education through a systematic review. In order to gather information for a systematic review, a specific research question must be addressed. To do this, all primary research that is pertinent to the review topic must be identified, the research must be critically evaluated, and the results criteria must be synthesized (Gough et al., 2012). Systematic reviews can integrate data from many research projects to provide a new integrated result or conclusion, or they might combine diverse types of evidence to analyse or clarify critical meaning (Pollock & Berge, 2018). As a result, findings from a systematic analysis can go beyond what a single study might provide (Hammersley, 2001).

The dataset for this systematic review was created by selecting articles for analysis from respected journals in the field of foreign language learning and teaching. Two distinct standards were set out to identify journals. The journals must, first and foremost, have a high indexing position in the Social Sciences Citation Index (SSCI). The journals were chosen according to their scope as well; in particular, they were required to center mostly on the use of technology in the education of foreign language teachers. Considering that corpus applications are a part of the larger field of computer-assisted language learning, this requirement was essential. As a result of their comprehensive examination of technology integration into foreign language education, three renowned journals—Computer-Assisted Language Learning (CALL), Language Learning and Technology (LLT), and ReCALL—were taken into account in the systematic review.

Particular inclusion and exclusion standards were set by this systematic review for research that was published in the designated journals. The search was initially limited to corpus studies that were carried out between 2016 and 2023, with an emphasis on the use of corpora in foreign language education. In the first search, the keywords "corpus," "corpora," "data-driven learning," and "DDL" were included. To make sure no items were missed, a manual search was also done. Following an extensive search, the review included only completed research papers; the publications like book reviews and systematic reviews were not included. After downloading the research studies, each one was carefully reviewed, paying particular attention mainly to the sections on the abstract, methodology, and findings—and, if needed, exploring other sections as well to address the research questions. Studies that used corpora as a compilation tool with a different purpose were ignored, and only articles that examined corpus applications in foreign language education were deemed eligible. To sum up, thirty-eight publications from three journals were examined in this study; Table 1 shows the distribution of these articles among the journals and publication years.



Table 1. The articles with their published journals and publication years

<i>Publication years/Journals</i>	<i>CALL</i>	<i>LLT</i>	<i>ReCALL</i>
2016	1	3	3
2017	2	7	1
2018	1	-	1
2019	1	-	1
2020	1	1	1
2021	1	1	1
2022	3	-	-
2023	6	-	2
Total	16	12	10

## Findings

The present study examined emerging issues in corpus research for foreign language education through a systematic review. The findings were categorized into three main areas: focal points of corpus applications, methodological designs of corpus studies, and major findings of these studies.

### *Focal points of corpus studies*

This review article categorizes the focal points of corpus studies in ESL/EFL contexts into two categories based on their research topics. These categories align with Römer's (2011) classification of corpus use in ESL/EFL contexts. The first category encompasses research on the indirect use of corpora, while the second category includes studies on the direct use of corpora. Table 2 provides a summary of the focal points of studies investigating corpus use in ESL/EFL settings.

Table 2. Focal points of corpus studies

<i>Focal points</i>	<i>f</i>
Indirect applications of corpus use	1
of Compiling corpus	1
of Compiling learner corpus	1
of Compiling academic corpus	1
Direct applications with corpus	9
with Learners' engagement with corpus	8
with Learning vocabulary	1
with Writing assistance	3
with Learning pragmatics	4
with Enhancing reading	3
with Learning grammar	3
with Concordance use	5
with Teachers' engagement with corpus	2
with Integrating corpus into language training	1
with Enhancing corpus literacy	1
with Detecting teachers' attitudes toward corpus applications	1

Table 2 illustrates the focal points of corpus studies conducted in ELT. The findings revealed two major issues: indirect applications of corpus use and direct applications of corpus use. Concerning indirect applications of corpus use, there was no direct engagement with the corpora in class to teach or learn English, but the studies compiled learner corpora to explore types of EFL learners' errors and reveal their interlanguage features. The studies also compiled academic corpora with the aim of assisting EFL learners in searching for appropriate collocations and phrases.

As for the direct application of corpus use, studies were delved into observing the aspects in which learners and teachers engaged with corpora. Regarding learners' engagement with corpora, the focus of studies was on enhancing vocabulary learning to assist learners in learning academic lexical bundles, understanding collocation patterns, discovering phraseology, improving retention of word knowledge, and inferring the meaning of vocabulary items. Studies also explored the efficacy of corpus use on learners' writing skills, which includes understanding types and structures of genre, encouraging error correction, and enhancing revision skills. Studies also investigated the role of corpora in grammar learning. The studies aimed to teach grammar patterns, increase grammar consciousness, and improve lexicogrammatical usage. Additionally, the scope of corpus studies extended to fostering pragmatics, which involves discovering routines for agreement, disagreement, and clarification. In the context of reading skills, studies focused on stimulating lexicogrammatical knowledge and reading speed, as well as raising awareness of discourse organizers. Lastly, studies explored concordance use among learners with the aim of revealing the construct of concordance and identifying learner preferences in its application.

Teacher engagement with corpora has also been a focus of studies. These research efforts delved into teacher training with the aim of effectively integrating corpora into foreign language teaching and developing teachers' competence in effective integration. Studies also investigated corpus literacy, in which they attempted to increase teacher knowledge to shed light on effective corpus integration. Additionally, studies explored teacher attitudes toward corpus-aided language learning and teaching.

### **Methodological designs of corpus studies**

As stated before, one of the parameters followed in the current review of 38 articles was their methodology. Through analyzing the methodology parts of pre-determined articles, the articles were classified according to their research designs.

Table 3. Research designs of corpus studies

<i>Research designs</i>	<i>f</i>
Quantitative designs	23
Qualitative designs	4
Mixed-method designs	11
Total	38

Table 3 indicates that the field is predominantly dominated by studies that utilized a quantitative research design, with a total of 23. Additionally, 11 studies employed a mixed-methods research design. Despite claims of a substantial increase in qualitative research related to corpus studies (Mueller & Jacobsen, 2016), the current investigation identified just 4 qualitative studies published from 2016 to 2023.

Considering quantitative studies, all articles focusing on the indirect use of corpus in ESL/ELF settings embraced the corpus-based approach. On the other side, most of the studies involving the direct use of corpus implemented an experimental design (Ackerley, 2017; Bardovi-Harlig et al., 2017; Larsen-Walker, 2017; Lee et al., 2017; Lee et al., 2018; Leńko-Szymańska, 2017; Li, 2017; Moon & Oh, 2017; Mueller & Jacobsen, 2016; Tsai, 2019; Wong & Lee, 2016; Vyatkina, 2016a; Vyatkina, 2016b). As for a mixed-methods design, it can be seen that there was a study in which an experimental design combined with observation and document analysis was employed (Hadley & Charles, 2017), a study in which an experimental study was followed by an interview (Kennedy & Miceli, 2016) and a study having a survey and interview combination (Yoon, 2016). All in all, the table showed that



quantitative research designs had outstanding popularity among corpus studies.

### **Major findings of corpus studies**

The findings of the studies were also reviewed in this study so as to provide certain invaluable insights into corpus use in English as a second or foreign language context. This review article analysed the major findings of the reviewed studies in two main categories: indirect applications and direct applications of corpus use in ELT. Table 4 presents the major findings related to the indirect applications of corpus use in ELT.

**Table 4. Findings related to indirect applications of corpus studies**

<i>Themes</i>	<i>Categories</i>	<i>Codes</i>
Compiling academic corpus	Easy search for target language item	Providing rich example sentences and collocations
	Easy search for three- to five-word sequence	Identifying more common combinations
	Searches using components of speech	Identifying parts of speech
Compiling learner corpora	Identifying learners' errors	Detecting mistakes
	Detecting learners' interlanguage features	Detecting two-, three-, four-, and five-word phrases

The studies revealed that scholars compiled corpora for academic and learner purposes. Academic corpora were compiled to facilitate learners in searching for target language items and three- to five-word sequences, as well as identifying parts of speech. Additionally, learner corpora were compiled to detect errors made by learners and identify common mistakes.

The systematic review also unveiled scholars' investigations into the effectiveness of directly applying corpora in ELT. Within this scope, they examined either learners' interaction with corpora or teachers' interaction with corpora. The major findings of studies focusing on learners' engagement with corpora are summarized in Table 5.

**Table 5. Learners' engagement with corpora**

<i>Themes</i>	<i>Categories</i>	<i>Codes</i>
Increased vocabulary understanding and usage	Increased awareness	Awareness to phraseology
	Increased definitional knowledge	Inferring meaning
	Acquiring complex elements	Improving collocational knowledge
	Applying new vocabulary understanding into practice	Utilizing more phrases Using academic bundles Using collocational patterns
Increased understanding of grammar	Inferring grammatical patterns	Identifying rules Identifying collocational patterns
	Applying grammatical rules	Improving correct grammar usage Improving collocational usage
	Enhancing reading	Being more conscious about discourse organizers Increased understanding of academic lectures
Enhancing reading	Increased understanding of discourse	Being more conscious about discourse organizers Increased understanding of academic lectures
	Increased understanding of authentic language	Improving comprehension of authentic academic lectures
	Increased lexical knowledge	Increasing lexicogrammatical knowledge Changes in lexical repetitions
Enhancing writing skills	Increased understanding of genre	Development of rhetorical, formal,

	norms	and procedural parts of genre knowledge
	Correcting errors	Correcting word choice and word form errors Correcting collocational errors Correcting phrasing errors
	Increased language accuracy	Improving the use of grammatical patterns Improving usage of lexical pattern
Increased understanding of pragmatics	Improved speech act clarity	Clarity of illocutionary force of turns
	Increased understanding of register	
Learners' concordancing preferences	Corpus type preferences Citation order preferences Task preferences	Small, focused corpora Ordered by the text citations Read to locate relevant examples Read for relationships between citations

Table 5 demonstrates the effectiveness of direct corpus use in ELT and reveals its impact on various aspects of language learning. The reviewed studies showed that corpus use led to improvements in learners' understanding and usage of vocabulary items, comprehension of grammar, enhancement of reading and writing skills, comprehension of pragmatics, and identification of learners' concordancing preferences. Regarding vocabulary learning, corpus use was shown to heighten learners' awareness of language items such as phraseology and expand their definitional knowledge by inferring meanings from concordance lines. Additionally, corpus use facilitated the acquisition of complex elements like collocations. It encouraged learners to apply their new vocabulary understanding by employing a wider range of phrases, utilizing academic lexical bundles, and integrating collocational patterns into their productions.

Corpus use was also found to enhance learners' understanding of grammar. In language class, it facilitated the inference of grammatical patterns by analyzing concordance lines to identify grammatical rules and patterns. It was also observed to aid learners in applying grammatical rules in new contexts; as such, learners were observed to increase their correct grammar and collocational usage. In addition to grammar, the reviewed studies indicated that learners' engagement with corpora enhanced their understanding of pragmatics. Particularly, learners were observed to improve their clarity in speech acts and develop a deeper comprehension of register nuances.

In terms of language skills, learners' interaction with corpora was found to enhance both their reading and writing skills. Regarding reading, the reviewed studies revealed that corpora facilitated an increase in learners' comprehension of discourse by heightening their awareness of discourse structures and organizers. Additionally, corpus use was shown to enhance learners' understanding of authentic language by exposing them to concordance lines and expanding their lexical knowledge. As for writing, corpus use among learners was found to enhance their understanding of genre norms by developing their rhetorical, formal, and procedural parts of genre knowledge. Learners were also observed consulting corpora to correct errors in their written productions, including issues with word choice and word form, collocations, and phrasing. Therefore, corpus use was shown to improve learners' language accuracy by encouraging the use of correct grammatical and lexical patterns in their written work.





The reviewed studies also unveiled learners’ preferences regarding concordancing. It was found that learners preferred utilizing small, focused corpora for easier searches, organized by text citations. Regarding task preferences, learners showed a preference for reading to locate relevant examples and to establish relationships between citations.

The studies also revealed various aspects of teachers’ engagement with corpora. Table 6 presents teachers’ engagement with corpora.

Table 6. Teachers’ engagement with corpora

<i>Themes</i>	<i>Categories</i>	<i>Codes</i>
Enhancing corpus literacy	Understanding	Understanding corpora Understanding concordance lines
	Searching	Understanding how to search with corpora Searching corpus data
	Analysis	Understanding how to analyze concordance lines Examining words before and after the keyword
	Understanding the benefits of corpora	Drawing conclusions about language use
Fostering corpus-based language instruction	Developing corpus activities	Fostering understanding of the function and textual structures of genre Fostering understanding of language features
	Fostering student-led corpus consultation	Self-detecting the usage of patterns
Teachers’ attitudes toward corpus applications	Affordances of corpora	Fostering autonomy Fostering authenticity Fostering discovery skills

The reviewed studies investigated corpus literacy, corpus-based language instruction, and teachers’ attitudes toward corpus applications. Studies aimed at enhancing teachers’ corpus literacy identified four sub-skills in corpus literacy. These include searching, which includes the capacity to interact with corpora and efficiently search corpus data, and comprehending, which entails knowledge of corpora and concordance lines. Analysis emerged as another dimension of corpus literacy, involving the ability to analyse concordance lines and examine words preceding and following the keyword searched. Additionally, benefits were highlighted as a facet of corpus literacy involving the ability to draw conclusions about language use.

The reviewed studies also aimed to promote corpus-based literacy instruction in language classrooms. Findings revealed that teachers developed corpus activities to enhance learners’ understanding of the functional and textual structures of genres, as well as to foster their comprehension of target language features. Additionally, teachers encouraged learners to utilize corpora in their learning process by leading them to identify usage patterns through contextualizing target items. Finally, the reviewed studies shed light on teachers’ attitudes toward corpus applications in language classrooms. The findings indicated that teachers perceived corpus applications as effective tools, in which they believed that corpus applications promote learners’ autonomy, authenticity in learning, and discovery skills.

## **Discussion**

This study systematically reviewed corpus applications in ELT to reveal emerging issues. By examining 38 articles published in prominent journals, it was found that corpus studies most frequently focus on students' direct engagement with corpora. Learners consult corpora to improve their reading and writing skills and enhance their understanding of vocabulary, grammar, and pragmatics. The use of corpora in language learning is well-documented in the literature (Chambers, 2005; Emir & Yangın-Ekşi, 2023; Gablasova et al., 2017; Huang, 2011). Boulton and Cobb (2017), in their systematic review, demonstrated that corpus use in language learning is effective both as a reference tool and a learning tool, particularly for vocabulary and lexicogrammar. Similarly, Lee and Lin (2019) found that learners' L2 proficiency and various features of corpus use (e.g., interaction types, corpus types, training, and duration) influence the effectiveness of corpus use in improving L2 vocabulary learning.

An emerging finding in this review was the integration of corpus use into teacher education. Recent studies have focused on corpus-based teacher training, which aims to enhance teachers' corpus literacy skills. Research indicated that developing corpus literacy is crucial for teachers to gain hands-on experience with corpus searching as a learning tool before teaching students how to use corpus tools (Breyer, 2009; Johns, 1990; Mukherjee, 2004). Teachers with higher levels of corpus literacy were found to be more likely to incorporate corpora into their classrooms (Ma et al., 2023). Given the importance of search skills, research recommended that more training activities be provided to help teachers become familiar with the various corpus search functions available on popular corpus websites. Additionally, improving their corpus search skills was observed to increase their confidence in working with corpora (Heather & Helt, 2012; Naismith, 2017).

Another emerging finding of this study is the need to foster corpus-based language pedagogy to increase corpus integration into language teaching. These studies primarily adopted the proficiency framework of technological pedagogical content knowledge (TPACK), as corpus applications in language teaching integrate subject knowledge (i.e., language), corpus technology, and pedagogical knowledge (Ma et al., 2022; Schmidt, 2023). Research indicates that corpus applications designed for teaching language, particularly through creating DDL-integrated lesson plans, enhance teachers' technological pedagogical knowledge (Crosthwaite et al., 2023). Corpus training adopting the TPACK framework allowed teachers to apply what they have learned to their teaching practices (Angeli et al., 2016; Schmidt, 2023). Studies also highlight the significance of guidance and scaffolding in ensuring teachers' continued engagement and enthusiasm during planning activities, especially when integrating DDL within the current curriculum and lesson-planning capacities (Meunier, 2019; Schmidt, 2023; Wicher, 2019). However, since teachers, and especially teacher trainees proficient in DDL practice, are still relatively rare, and there is a gap in using corpus resources for teaching purposes (Crosthwaite et al., 2023), having a corpus linguist available to guide trainees on DDL integration is essential for successful training interventions (Ädel, 2010).

Collaborative learning was found to be another crucial factor in encouraging teachers to apply DDL to language teaching. Studies revealed that teachers developed new perspectives and abilities in using corpora for teaching, particularly in transforming their knowledge into classroom practice (Ma et al., 2021). Therefore, teachers' engagement in collaborative learning in corpus-based language teaching is encouraged to energize and empower them, expand their teaching resources and ideas, and prompt them to review and refine their teaching techniques (Donnelly & Hume, 2015; Ma et al., 2021; Ma et al., 2022). Further,



orienting teachers to professional knowledge in corpus training has been shown to enhance their motivation to apply corpus technology in their teaching. Although the relevant literature indicates a lack of motivation among teachers to integrate corpus technology (Callies, 2019; Chambers, 2019; Farr, 2008; Leńko-Szymańska, 2017; Naismith, 2017; Zareva, 2017), recent studies revealed that helping teachers understand that corpus training enhances their professional development can motivate them to increase their use of corpus-based language teaching and incorporate more corpus technology into their practices (Abdel Latif, 2021; Ma et al., 2022).

Another emerging finding from the reviewed studies is that they generally adopted quantitative research designs (e.g., Ackerley, 2017; Balance, 2020; Lay & Yavuz, 2020; Li, 2023; Tsai, 2019). However, as Boulton and Cobb (2017) noted, there is a lack of longitudinal and delayed post-testing, which could provide more concrete evidence of DDL's long-term effects on learning, such as autonomy, learning to learn, consciousness-raising, and other forms of long-term change. Ma et al. (2021) also pointed out that studies on corpus applications in language learning and teaching often lack a focus on teachers' and learners' progress in working with corpus data or integrating corpus into classroom pedagogy. Considering the complexity of DDL, Schmidt (2023) highlighted the pressing need for long-term teacher support. In line with these findings, the current study suggests that more qualitative and longitudinal research designs should be adopted to investigate learners' and teachers' engagement with corpora and the long-term effects of DDL.

## **Conclusion**

This article reviewed thirty-eight studies on corpus use for teaching and learning foreign or second languages, published in three prominent journals from 2016 to 2023, to reveal emerging issues in corpus applications in ELT. The results of this review provided important insights into corpus use in the EFL/ESL context. Firstly, the articles reviewed primarily focus on corpus applications as a DDL tool in teaching or learning a second/foreign language, with most studies examining learners' interaction with corpus use. Learners' use of corpora as a DDL tool was found to foster a lexis-oriented view of the target language and was beneficial for the development of language skills (reading, writing), the improvement of language components such as grammar and vocabulary, and the enhancement of understanding pragmatics.

Studies focusing on teachers' interaction with corpora as a DDL tool appear to be quite limited. In the studies that do focus on teachers' engagement with corpora, the significance of increasing teachers' corpus literacy is highlighted. Enhanced corpus literacy enables teachers to work more effectively with corpora and integrate corpus applications into their teaching practices. Familiarity with corpora not only increases teachers' language awareness but also their ability to adopt corpus applications in the classroom. Additionally, the importance of teachers' collaborative learning emerged as another key finding from the review of studies.

Additionally, the studies were found to adopt quantitative, qualitative, and mixed-method designs, with the number of quantitative studies being much higher than the others. However, the lack of qualitative and longitudinal studies in investigating DDL has prevented researchers from fully understanding the nature of teachers' and learners' interactions with corpora and the long-term effects of DDL in language learning and teaching.

## 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. <https://doi.org/10.1017/s0958344020000129>
- Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology*, 21(3), 195–216. Retrieved from <http://llt.msu.edu/issues/october2017/ackerley.pdf>
- Ädel, A. (2010). Using corpora to teach academic writing: Challenges for the direct approach. In M. Campoy-Cubillo, B. Belles-Fortuno, & M. Gea-Valor (Eds.), *Corpus-based approaches to ELT* (pp. 39–55). Continuum.
- Alsehibany, R. A., & Abdelhalim, S. M. (2023). Overcoming academic vocabulary errors through online corpus consultation: the case of Saudi English majors. *Computer Assisted Language Learning*, 1–27. <https://doi.org/10.1080/09588221.2023.2249503>
- Angeli, C., Valanides, N. & Christodoulou, A. (2016). Theoretical considerations of technological pedagogical content knowledge. In Herring, M. C., Koehler, M. J. & Mishra, P. (Eds.), *Handbook of technological pedagogical content knowledge (TPACK) for educators* (pp. 11-32). Routledge.
- Ballance, O. J. (2016). Analysing concordancing: a simple or multifaceted construct? *Computer Assisted Language Learning*, 29(7), 1205–1219. <https://doi.org/10.1080/09588221.2016.1209527>
- Ballance, O. J. (2017). Pedagogical models of concordance use: correlations between concordance user preferences. *Computer Assisted Language Learning*, 30(3–4), 259–283. <https://doi.org/10.1080/09588221.2017.1307228>
- Ballance, O. J. (2020). Narrow reading, vocabulary load and collocations in context: Exploring lexical repetition in concordances from a pedagogical perspective. *ReCALL*, 33(1), 4–17. <https://doi.org/10.1017/s0958344020000117>
- Bardovi-Harlig, K., Mossman, S., & Su, Y. (2017). The effect of corpus-based instruction on pragmatic routines. *Language Learning & Technology*, 21(3), 76–103. Retrieved from <http://llt.msu.edu/issues/october2017/bardovi-harligmossmansu.pdf>
- Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Goźdz-Roszkowski (Ed.), *Explorations across languages and corpora* (pp. 563–580). Peter Lang.
- Boulton, A. (2017). Research timeline: Corpora in language teaching and learning. *Language Teaching*, 50(4), 483–506. <https://doi.org/10.1017/S0261444817000167>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language learning*, 67(2), 348-393. <https://doi.org/10.1111/lang.12224>
- Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning & Technology*, 25(3), 66-89. 10125/73450
- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–172. doi:10.1080/09588220902778328

- Çalışkan, G. & Gönen, S. İ. K. (2018). Training teachers on corpus-based language pedagogy: Perceptions on vocabulary instruction. *Journal of Language and Linguistic Studies*, 14(4): 190–210.
- Callies, M. (2016). Towards corpus literacy in foreign language teacher education: Using corpora to examine the variability of reporting verbs in English. In Kreyer, R., Schaub, S. & Güldenring, B. (Eds.), *Angewandte Linguistik in Schule und Hochschule* [Applied linguistics in secondary school and at university]. Peter Lang, 391–415.
- Callies, M. (2019). Integrating corpus literacy into language teacher education. In S. Götz & J. Mukherjee (Eds.), *Learner corpora and language teaching*. (pp. 245–263). John Benjamins.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning and Technology*, 9(2), 111-125. 10125/44022
- Chambers, A. (2010). What is data driven learning? In A. O’Keeffe & 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*, 52(4), 460–475. DOI:10.1017/S0261444819000089
- Chen, H. J. H., Lai, S. L., Lee, K. Y., & Yang, C. T. Y. (2023). Developing and evaluating an academic collocations and phrases search engine for academic writers. *Computer Assisted Language Learning*, 36(4), 641–668. <https://doi.org/10.1080/09588221.2021.1937229>
- Chen, M., Flowerdew, J., & Anthony, L. (2019). Introducing in-service English language teachers to data-driven learning for academic writing. *System*, 87, 102148. <https://doi.org/10.1016/j.system.2019.102148>
- Cotos, E., Link, S., & Huffman, S. (2017). Effects of technology on genre learning. *Language Learning & Technology*, 21(3), 104–130. Retrieved from <http://lt.msu.edu/issues/october2017/cotoslinkhuffman.pdf>
- Crosthwaite, P. (2017). Retesting the limits of data-driven learning: feedback and error correction. *Computer Assisted Language Learning*, 30(6), 447–473. <https://doi.org/10.1080/09588221.2017.1312462>
- Crosthwaite, P. (2020). Taking DDL online: Designing, implementing and evaluating a SPOC on data-driven learning for tertiary L2 writing. *Australian Review of Applied Linguistics*, 43(2), 169-195. DOI:10.1075/ara1.00031.cro
- Crosthwaite, P., & Steeples, B. (2022). Data-driven learning with younger learners: exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*, 1–32. <https://doi.org/10.1080/09588221.2022.2068615>
- Crosthwaite, P., Luciana, & Wijaya, D. (2023). Exploring language teachers’ lesson planning for corpus-based language teaching: a focus on developing TPACK for corpora and DDL. *Computer Assisted Language Learning*, 36(7), 1392–1420. <https://doi.org/10.1080/09588221.2021.1995001>
- Crosthwaite, P., Wong, L. L., & Cheung, J. (2019). Characterising postgraduate students’ corpus query and usage patterns for disciplinary data-driven learning. *ReCALL*, 31(3), 255–275. <https://doi.org/10.1017/s0958344019000077>

- Dong, J., Zhao, Y., & Buckingham, L. (2023). Charting the landscape of data-driven learning using a bibliometric analysis. *ReCALL*, 35(3), 339–355. doi:10.1017/S0958344022000222
- Donnelly, D. F., & Hume, A. (2015). Using collaborative technology to enhance pre-service teachers' pedagogical content knowledge in Science. *Research in Science & Technological Education*, 33(1), 61–87. <https://doi.org/10.1080/02635143.2014.977782>
- Emir, G., & Yangın-Ekşi, G. (2023). Corpus used as a data-driven learning tool in L2 academic writing: Evidence from Turkish contexts. *Teflin Journal*, 34(2), 209-225. <https://doi.org/10.15639/teflinjournal.v34i2/209-225>
- 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. <https://doi.org/10.2167/la414.0>
- Farr, F. & O’Keeffe, A. (2019). “Using corpora to analyse language”. In S. Walsh and S. Mann (Eds.), *Routledge handbook of English language teacher education* (pp. 268-2829). Routledge.
- Flowerdew, L. (2015). Data-driven learning and language learning theories: Wither the twain shall meet. In A. Lenko-Szymanska A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). John Benjamins.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language learning*, 67(S1), 155-179. <https://doi.org/10.1111/lang.12225>
- Gough, D., Oliver, S., & Thomas, J. (2012). Introducing systematic reviews. In D. Gough, S. Oliver, & J. Thomas (Ed.), *An introduction to systematic reviews* (pp. 1–16). Sage.
- Hadley, G., & Charles, M. (2017). Enhancing extensive reading with data-driven learning. *Language Learning & Technology*, 21(3), 131–152. Retrieved from <http://llt.msu.edu/issues/october2017/hadleycharles.pdf>
- Hadley, G., & Hadley, H. (2021). Exploring the impact of data-driven learning in extensive reading. In P. Pérez-Paredes, & G. Mark (Eds.), *Beyond Concordance Lines: Corpora in language education* (pp. 149–176). John Benjamins. <https://doi.org/10.1075/scl.102.07had>
- Hammersley, M. (2001). On ‘systematic’ reviews of research literatures: A ‘narrative’ response to Evans & Benefield. *British Educational Research Journal*, 27(5), 543–554. <https://doi.org/10.1080/01411920120095726>
- Harvey-Scholes, C. (2018). Computer-assisted detection of 90% of EFL student errors. *Computer Assisted Language Learning*, 31(1–2), 144–156. <https://doi.org/10.1080/09588221.2017.1392322>
- Heather, J., & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, 20(4), 415–440.
- Hirata, Y., & Thompson, P. (2022). Communicative data-driven learning: a two-year pilot study. *ELT Journal*, 76(3), 356-366. <https://doi.org/10.1093/elt/ccab066>
- Huang, L. S. (2011). Corpus-aided language learning. *ELT Journal*, 65(4), 481-484. <https://doi.org/10.1093/elt/ccr031>



- Johns, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning, *CALL*, 10,14–34.
- 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. doi:10.1017/S0958344015000154
- Kennedy, C., & Miceli, T. (2016). Cultivating effective corpus use by language learners. *Computer Assisted Language Learning*, 30(1–2), 91–114. <https://doi.org/10.1080/09588221.2016.1264427>
- Larsen-Walker, M. (2017). Can data driven learning address L2 writers' habitual errors with English linking adverbials?. *System*, 69, 26-37. <https://doi.org/10.1016/j.system.2017.08.005>
- Lay, K. J., & Yavuz, M. A. (2020). Data-driven learning of academic lexical bundles below the C1 level. *Language Learning & Technology*, 24(3), 176–193. <http://hdl.handle.net/10125/44741>
- Lee, H., Warschauer, M., & Lee, J. H. (2017). The effects of concordance-based electronic glosses on L2 vocabulary learning. *Language Learning & Technology*, 21(2), 32–51. Retrieved from <http://ilt.msu.edu/issues/june2017/leewarschauerlee.pdf>
- Lee, H., Warschauer, M., & Lee, J. H. (2018). Advancing CALL research via data-mining techniques: Unearthing hidden groups of learners in a corpus-based L2 vocabulary learning experiment. *ReCALL*, 31(2), 135–149. <https://doi.org/10.1017/s0958344018000162>
- Lee, H., Warschauer, M., & Lee, J. H. (2020). Toward the establishment of a data-driven learning model: Role of learner factors in corpus-based second language vocabulary learning. *The Modern Language Journal*, 104(2), 345–362. <https://doi.org/10.1111/modl.12634>
- Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, 81, 14-25. <https://doi.org/10.1016/j.system.2018.12.011>
- Lee, S. H., Jang, S. B., & Seo, S. K. (2009). Annotation of Korean learner corpora for particle error detection. *Calico Journal*, 26(3), 529-544. Retrieved from <http://www.jstor.org/stable/calicojournal.26.3.529>
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217–241. Retrieved from <http://ilt.msu.edu/issues/october2017/lenko-szymanska.pdf>
- Leńko-Szymańska, A., & Boulton, A. (2015). Introduction. In A. Leńko-Szymańska, A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 1–15). John Benjamins.
- Li, H., & Zhang, T. (2024). A Study on the Construction of College English Context Vocabulary Teaching Based on Hands-Off Data-Driven Learning in China. *International Education Studies*, 17(1), 1-7. DOI:10.5539/ies.v17n1p1
- Li, L. X. (2023). Promoting accuracy of collocation use in L2 writing: the role of data-driven learning in indirect corrective feedback. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2023.2292554>

- Li, Q. (2019). L2 Chinese learners' pragmatic developmental patterns in data-driven instruction and computer-mediated communication (CMC): A case of Chinese sentence final particle ne. *Applied Pragmatics*, 1(2), 154-183. <https://doi.org/10.1075/ap.19006.li>
- Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology*, 21(3), 153-171. Retrieved from <http://llt.msu.edu/issues/october2017/li.pdf>
- 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. <https://doi.org/10.1002/tesq.250>
- 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. (2021). Effects of data-driven learning on college students of different grammar proficiencies: A preliminary empirical assessment in EFL classes. *Sage Open*, 11(3), <https://doi.org/10.1177/21582440211029936>
- Lin, M. H., & Lee, J. Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, 69(3), 264-274. <https://doi.org/10.1093/elt/ccv010>
- Liu, T., & Gablasova, D. (2023). Data-driven learning of collocations by Chinese learners of English: a longitudinal perspective. *Computer Assisted Language Learning*, 1-26. <https://doi.org/10.1080/09588221.2023.2214605>
- Luo, Q. (2016). The effects of data-driven learning activities on EFL learners' writing development. *Springerplus*, 5(1). 1255. doi: 10.1186/s40064-016-2935-5
- 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. <https://doi.org/10.1017/s0958344022000180>
- 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. <https://doi.org/10.1080/09588221.2021.1895225>
- Ma, Q., Yuan, R. E., Cheung, L. M. E., & Yang, J. (2022). Teacher paths for developing corpus-based language pedagogy: a case study. *Computer Assisted Language Learning*, 1-32. <https://doi.org/10.1080/09588221.2022.2040537>
- Meunier, F. (2019). A case for constructive alignment in DDL: Rethinking outcomes, practices and assessment in (data-driven) language learning. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners* (pp.13-31). Routledge. doi:10.4324/9780429425899-2.
- Moon, S., & Oh, S. Y. (2017). Unlearning overgenerated be through data-driven learning in the secondary EFL classroom. *ReCALL*, 30(1), 48-67. <https://doi.org/10.1017/s0958344017000246>
- Mueller, C. M., & Jacobsen, N. D. (2016). A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills. *ReCALL*, 28(1), 3-21. <https://doi.org/10.1017/s0958344015000142>



- Muftah, M. (2023). Data-driven learning (DDL) activities: do they truly promote EFL students' writing skills development?. *Education and Information Technologies*, 28(10), 13179-13205. <https://doi.org/10.1007/s10639-023-11620-z>
- Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 239–250). Rodopi.
- Naismith, B. (2017). Integrating corpus tools on intensive CELTA courses. *ELT Journal*, 71(3): 273–283. <https://doi.org/10.1093/elt/ccw076>
- Otto, P. (2021). Choosing specialized vocabulary to teach with data-driven learning: An example from civil engineering. *English for Specific Purposes*, 61, 32-46. <https://doi.org/10.1016/j.esp.2020.08.003>
- 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. <https://doi.org/10.1080/09588221.2019.1667832>
- Pollock, A., & Berge, E. (2018). How to do a systematic review. *International Journal of Stroke*, 13(2), 138-156. DOI: 10.1177/1747493017743796
- Poole, R. (2020). “Corpus can be tricky”: revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641. <https://doi.org/10.1080/09588221.2020.1825095>
- Quan, Z., Grant, L., & Hocking, D. (2022). Comparing concordances of language patterns and words by ESL intermediate learners: a preliminary experiment with two mobile concordancers. *Computer Assisted Language Learning*, 1–27. <https://doi.org/10.1080/09588221.2022.2081581>
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.
- Rosen, A., Hana, J., Štindlová, B., & Feldman, A. (2014). Evaluating and automating the annotation of a learner corpus. *Language Resources and Evaluation*, 48, 65-92. <https://doi.org/10.1007/s10579-013-9226-3>
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual review of applied linguistics*, 31, 205-225. doi:10.1017/S0267190511000055
- Schmidt, N. (2023). Unpacking second language writing teacher knowledge through corpus-based pedagogy training. *ReCALL*, 35(1), 40–57. <https://doi.org/10.1017/s0958344022000106>
- Sun, X., & Hu, G. (2023). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research*, 27(3), 660-688. <https://doi.org/10.1177/1362168820954459>
- Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal*, 97(S1), 77-101. <https://doi.org/10.1111/j.1540-4781.2012.01422.x>
- 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. <https://doi.org/10.1080/09588221.2018.1527366>

- Vyatkina, N. (2016a). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(2), 207–226. <https://doi.org/10.1017/s0958344015000269>
- Vyatkina, N. (2016b). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, 20(3), 159–179. Retrieved from <http://llt.msu.edu/issues/october2016/vyatkina.pdf>
- Vyatkina, N., & Belz, J. A. (2006). A learner corpus-driven intervention for the development of L2 pragmatic competence. In K. Bardovi-Harlig, J. C. Félix-Brasdefer, & A. Omar (Eds.), *Pragmatics and language learning* (pp. 315-357). National Foreign Language Resource Center,
- Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1–8. 10125/44621
- Wicher, O. (2019). Data-driven learning in the secondary classroom: A critical evaluation from the perspective of foreign language didactics. In P. Crosthwaite (Ed.), *Data driven learning for the next generation: Corpora and DDL for pre-tertiary learners* (pp. 31–46). Routledge. doi:10.4324/9780429425899-3
- Wilkinson, D. (2012). A data-driven approach to increasing student motivation in the reading classroom. *Language Education in Asia*, 3(2), 252-262.
- Wong, T. S., & Lee, J. S. Y. (2016). Corpus-based learning of Cantonese for Mandarin speakers. *ReCALL*, 28(2), 187–206. <https://doi.org/10.1017/s0958344015000257>
- Wu, Y-j. A. (2021). Discovering collocations via data-driven learning in L2 writing. *Language Learning & Technology*, 25(2), 192–214. <http://hdl.handle.net/10125/73440>
- Yao, G. (2019). Vocabulary learning through data-driven learning in the context of Spanish as a foreign language. *Research in Corpus Linguistics*, 7, 18-46. <https://doi.org/10.32714/ricl.07.02>
- Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, 20(1), 209–229. Retrieved from <http://llt.msu.edu/issues/february2016/yoon.pdf>
- Zare, J., Karimpour, S., & Aqajani Delavar, K. (2023). Classroom concordancing and English academic lecture comprehension: an implication of data-driven learning. *Computer Assisted Language Learning*, 36(5–6), 885–905. <https://doi.org/10.1080/09588221.2021.1953081>
- Zareva, A. (2017) Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1). 69–79. <https://doi.org/10.1093/elt/ccw045>
- Zhang, Y. (2021). Combining computer-mediated communication with data-driven instruction: EFL learners' pragmatic development of compliment responses. *System*, 103, 102624. <https://doi.org/10.1016/j.system.2021.102624>

***Declarations:*** The authors have nothing to declare

***Funding:*** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

***Conflict of Interest:*** The authors declared no potential conflicts of interest concerning this article's research, authorship, and/or publication.

***Data availability:*** The data that support the findings of this study are available from the corresponding author upon reasonable request.