ReCALL (2021), 33: 1, 34–48
doi:10.1017/S0958344020000129

ARTICLE

Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices

Muhammad M. M. Abdel Latif
Faculty of Graduate Studies of Education, Cairo University, Egypt (mmmabd@cu.edu.eg)

Abstract

With the increasing recognition of the pedagogical applications of corpus linguistics, there has been a growing interest in developing teachers' corpus literacy to popularize the use of corpora in language education. This longitudinal study investigated Arab Gulf EFL student teachers' immediate and long-term responses to corpus literacy instruction. After teaching a corpus literacy component to two classes of student teachers in a graduate computer-assisted language learning course they attended, the author collected focus group data about their views on this instruction and their own expected future uses of corpora in language learning, teaching and research. Two years later, a group of these student teachers ($n = 19$) responded to a follow-up questionnaire exploring their beliefs about corpus literacy integration and their multiple uses of corpora. The student teachers reported very positive immediate and long-term perceptions of corpus literacy instruction, but it was found that such instruction has not brought about all the desired changes in their long-term uses of online corpora as a linguistic and pedagogical resource, or their attitudes towards doing corpus-based TESOL research. However, it is expected that the popularization benefits gained from corpus literacy integration could lead to better future developments in using corpora for language education and research purposes in the target context.

Keywords: corpus literacy; corpora; corpus linguistics; data-driven learning; teacher education

1. Introduction

Technology has reshaped not only the content of language teacher education (LTE) programs but also their delivery (Farr, 2008). Nowadays, computer-assisted language learning (CALL) is no longer competing for a place in LTE programs (Hubbard & Levy, 2006) because it has already become an essential component in them to help trainee language teachers use technology effectively. As an area of important implications to CALL, corpus linguistics has gained increasing attention among language teacher educators. Due to the increasing recognition of the pedagogical applications of corpus linguistics, there has been a growing interest in integrating it into graduate LTE programs.

Corpus linguistics has revolutionized many aspects of language learning and teaching. There are two main pedagogical applications of corpus linguistics: the indirect application, which means making use of corpus linguistics analysis in writing dictionaries and language learning materials, and the direct one – known as data-driven learning (DDL) – which refers to involving learners in accessing corpora and working with concordances (Römer, 2011). Through this direct use of corpus concordances, learners – guided by their teachers – discover target language features in authentic materials and understand contextual language use and how it varies in different genres

Cite this article: 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

© European Association for Computer Assisted Language Learning 2020.

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

(Boulton & Tyne, 2014). Thus, DDL is used for developing learners' language awareness and knowledge through engaging them in linguistic pattern exploration tasks. It is noteworthy, however, that corpus integration into language instruction has not been largely popularized yet. Referring to some relevant survey research, Anthony, Chen and Flowerdew (2017) conclude that using corpora directly in language classes has been an uncommon practice for teachers at different educational stages.

Many researchers (e.g. Breyer, 2009; Callies, 2016; Farr & O'Keeffe, 2019) generally view that raising teachers' awareness of corpus-based instruction is a prerequisite for enabling learners to work with corpora. Although corpora have been increasingly available and the number of those supporting corpus-based language teaching steadily grows, we need to help teachers understand corpora and use them in their classes (Frankenberg-Garcia, 2012). Since language teachers have the ultimate decision to engage learners in working with corpora (Breyer, 2009), one main way for motivating them to do this is by incorporating corpus literacy into their education programs. Teachers' corpus literacy means their ability to use corpus linguistics tools to discover linguistic features and foster learners' language acquisition and development (Heather & Helt, 2012). According to Callies (2016), teachers' corpus literacy has a number of components:

> 1. Understanding basic concepts in corpus linguistics: What is a corpus and what types of corpora are available and how? What can you do – and cannot do – with a corpus? 2. Searching corpora and analysing corpus data by means of corpus software tools, e.g. concordancers: What is corpus software and how can it be used to search a corpus? How can corpus output be analysed? 3. Interpreting corpus data: How may general trends in language use/change be extrapolated from corpus data? 4. Using corpus output to generate teaching material and activities (p. 395)

Despite the above-mentioned importance of corpus literacy, the studies integrating it into LTE programs are generally few (e.g. Breyer, 2009; Çalışkan & Gönen, 2018; Naismith, 2017). The dearth of corpus literacy studies has also resulted in some contextual research gaps. In the Arab world, for example, corpora and corpus-based instruction have only started to gain the attention of a few researchers recently. On the one hand, there has been a recent research trend concerned with building Arabic corpora and integrating corpus-based teaching into Arabic classes (e.g. AbdelRaouf, Higgins, Pridmore & Khalil, 2010; Zaki, 2017). On the other hand, a few English corpus-based studies have been conducted so far in the Arab context, and most of these relied on corpus analyses to explore Arab students' use of particular lexical and grammatical features in their written or spoken English (e.g. Alangari, 2019; Alsharif, 2017). With regard to corpus literacy integration into the LTE programs in the Arab world, there does not seem to be any studies addressing this issue yet. Likewise, to the best of the author's knowledge, there have been no survey reports documenting the use of corpora in the Arab educational institutions. Arguably, the scarcity of corpus research in the Arab world may generally reflect the unpopular use of corpora and corpus-based teaching in the region. On the other hand, some issues have hardly been examined in the previous relevant research; for example, how corpus literacy may influence teachers' long-term use of corpora in their language learning and instruction. As Kavanagh (2019) notes, we lack knowledge about language teachers' use of corpora after receiving corpus instruction. To address these research gaps, the present study investigated Arab EFL student teachers' responses to corpus literacy instruction from a longitudinal perspective.

2. Previous studies

Not many studies have investigated corpus literacy integration into LTE. The relevant studies reported so far have addressed this issue in in-service and pre-service (or student teacher)

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

contexts, but the studies belonging to the latter are greater in number; this has perhaps resulted from researchers' reliance on convenience sampling – that is, having an easier and more convenient access to research participants. Each of these studies explored teachers' responses to the corpus literacy instruction/training they were exposed to.

In one of the earliest studies attempting to raise teachers' awareness of corpus tools, Mukherjee (2004) found that most trainee in-service teachers in Germany learned about corpora for the first time through training workshops, which led them to develop interest in introducing corpus tools in their classrooms. Recently, Anthony *et al.* (2017) introduced DDL activities to a group of English language teachers working in Hong Kong universities. Although the teachers in this study perceived the value of corpora as a source for academic writing and reported their enthusiasm for experimenting with corpus tools, they considered time a main obstacle to corpus-based teaching. Çalışkan and Gönen (2018) provided language teachers in Turkey with a four-week training in corpus-based vocabulary instruction, and found positive effects on their lexical awareness. However, the teachers reported concerns with regard to overcoming technical issues in corpus use and designing corpus-based materials. In the Taiwanese context, Lin (2019) trained one in-service teacher in using the DDL approach over four months, and found that the process of becoming a DDL teacher encompasses complex and transformative steps.

Different research approaches were adopted in the studies integrating corpus literacy training into pre-service or student teacher education. In addition to exploring pre-service/student teachers' beliefs about corpus literacy integration, some studies (Breyer, 2009; Heather & Helt, 2012; Leńko-Szymańska, 2017; Naismith, 2017) evaluated teachers' ability to design or use corpus-based teaching materials during the training. Breyer (2009) engaged student teachers in Germany in corpus learning experiences from a learner-teacher perspective. She helped her student teacher participants understand corpus use and analysis (the learner perspective), and engaged them in a group project for developing corpus-based teaching materials (the teacher perspective). Following a similar approach, Heather and Helt (2012) investigated corpus literacy instruction by evaluating six student teachers' corpus-based material design skills. In the Polish context, Leńko-Szymańska (2017) examined the effectiveness of a semester-long training course on student teachers' perceived benefits of corpora and their skills in implementing corpus-based instruction. Naismith (2017) also found that including corpus training in a CELTA program in Canada fostered trainee teachers' interest in and perceived benefits of using corpora as a linguistic resource, but the trainees made much less use of corpus tools in their teaching practice as compared to their lesson planning.

Another approach in the studies integrating corpus literacy in pre-service/student teacher education is tracing teachers' corpus-based instruction after training. This approach seems to have been followed in only one study. In an Irish university context, Farr (2008) provided corpus use training to 25 Master of Arts student teachers belonging to different nationalities and found positive effects on their perceived corpus use enjoyment and benefits, and on their willingness to incorporate corpora in future classroom instruction. The follow-up email survey she sent to them a few months after joining their workplaces showed that only two of them used corpus techniques in their instruction while three others reported trying to do so but their attempts were hindered by some technical and methodological difficulties. The other relevant pre-service LTE studies (Ebrahimi & Faghih, 2017; Leńko-Szymańska, 2014; Zareva, 2017) only explored trainee teachers' experience with corpus literacy instruction and expectations about using corpora in their teaching. These studies were conducted in the Iranian, Polish and US contexts, respectively.

Although these previous studies have enriched our understanding of how corpus literacy should be integrated into LTE, some gaps are yet to be addressed in this research area. As noted above, no study has yet looked at how Arab teachers perceive and use corpora after receiving corpus literacy instruction. Another under-explored research issue is teachers' long-term

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

responses to corpus literacy instruction. Farr's (2008) study is the only one that seems to have explored how teachers use corpora after receiving corpus literacy instruction, but her attempt is limited mainly because she traced the participants' corpus-based teaching only a few months after they joined their workplaces.

Taking these gaps and limitations into account, the present study explored Arab student teachers' immediate and long-term responses to corpus literacy instruction. Student teachers' immediate responses mean the views they provided immediately after receiving corpus literacy instruction with regard to their evaluation of this learning experience and their expected future uses of corpora, whereas their long-term responses refer to the views and uses of corpora they reported two years after receiving corpus literacy instruction. With these objectives, the study responds to the call made by Boulton and Tyne (2014) for investigating how teachers make use of corpora after their training.

3. The present study

In this study, the author integrated a corpus literacy component into a graduate CALL course he was teaching to a group of student teachers at a Saudi university, and explored their perceptions of corpora and corpus literacy instruction, and their expectations about using corpora in language learning, teaching and research. Two years after teaching the course, the author traced the teachers' long-term perceptions and uses of corpora in their own language learning, teaching, and research. Accordingly, the present study tried to answer the following three research questions:

1. How does a group of Arab EFL student teachers perceive corpus literacy instruction integrated into a graduate CALL course?
2. How do they view their future uses of corpora in language learning, teaching and research after receiving corpus literacy instruction?
3. To what extent are the immediate perceptions and expectations of these student teachers compared to their long-term views and multiple uses of corpora?

As noted, the novel dimension addressed by the present study is the longitudinal investigation of the student teachers' views and uses of corpora. The context of the study also adds to its originality.

4. Method

4.1 Participants

The participants who received the corpus literacy instruction in this study were student teachers attending a PhD program in TESOL/applied linguistics at a Saudi university. These student teachers had to complete a number of courses, one of which was in CALL, as a prerequisite for proceeding to the research stage in this program. The course was taught to two classes comprising 28 students; 19 of them were females and nine were males. All the student teachers were Saudis with the exception of two who were from Yemen. All the participants were working as teachers of English in different governmental educational institutions in Saudi Arabia before being admitted to the PhD program, and after completing the taught-course stage on a full-time basis they were supposed to return to their teaching jobs while doing their PhD research on a part-time basis. Most of the participants were university teachers of English, a few of them were high and middle school teachers, and two were teacher educators. The student teachers were of varied ages and teaching experiences, and all of them were awarded their Master of Arts degrees from different Saudi universities.

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

4.2 The corpus literacy component taught

The CALL course taught to the student teachers lasted for 15 academic weeks, three of which were allocated to the corpus literacy component. The CALL course edition taught to the students was the third one. Because a number of CALL issues were taught in the course, it was not possible to cover the corpus linguistics topics in more than three weeks. During the three weeks, the author, who was the teacher of the course, met each group of student teachers in three classes (three hours each; i.e. total time allocated to the corpus literacy component = nine hours). Given the level of the student teachers' academic study (i.e. the PhD degree), this part of the course aimed at raising their awareness of corpus tools and pedagogical applications and also of the relevant research issues. A number of corpus linguistics topics were covered in the three classes, as explained below:

1. *Class one*: The topics covered in this class included introducing corpus linguistics and discussing its importance, defining corpus types and the main concepts needed for understanding corpora and corpus analysis (e.g. token, key word in context, node, connotation, collocation, etc.), and working with concordancers and concordances. In this class, the students explored searching the following corpora: Corpus of Contemporary American English (COCA), Global Web-Based English (GloWbE), Corpus of American Soap Operas, and British National Corpus (BNC). Following this class, the URLs of the websites of these corpora were sent to the students for further corpus use practice.
2. *Class two*: This class dealt with the indirect and direct applications of corpus, and introduced the students to learner corpora (their characteristics, how they differ from the corpora explored in the first class – i.e. native speaker corpora – and what we can learn from them and use them for, and how they are built). The students were also introduced to the British Academic Written English Corpus (BAWE) as an exemplary learner corpus.
3. *Class three*: In this class, the students were provided with an overview of pedagogy-oriented corpus linguistics research. The students were particularly introduced to the vocabulary, grammar, writing, and DDL corpus research. After introducing each research type to the students, the teacher read with the students the titles, abstracts and (sometimes) research method parts of some reports of studies representing it. The students were then asked to find more similar studies through using Google and/or well-known CALL journals.

A learner-centered approach was used in teaching these corpus linguistics topics. Following the introduction of each topic in the class, the student teachers were engaged in discussing and reflecting upon this topic guided by the questions raised and/or tasks assigned by the teacher. The student-centered activities were assigned seven to eight times in each class, and each activity lasted for about 5–8 minutes. For further reading purposes, a list of well-cited journal papers on corpus-based applied linguistics and TESOL research was also given to the students who were to read about the above-mentioned topics as part of preparing for the final term exam. Completing the course also required the students to write a 2,000-word essay on one of five optional topics (two of them were about corpus-based vocabulary and grammar teaching). Eight students in the two classes chose to write about these two topics. All these reading activities and coursework assignments aimed at fostering the student teachers' corpus literacy and engaging them in using corpus tools, and getting them to discuss and reflect upon corpus pedagogical applications and related research issues.

4.3 Data sources

The study drew its data from two sources: (a) the focus group interviews conducted with the student teachers after teaching them the corpus linguistics part, and (b) the follow-up questionnaire sent to them two years after they completed the CALL course and returned to their workplaces. The focus group interviews were used to collect data about the student teachers'

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

immediate responses to and perceptions of corpus instruction and tools. During the course delivery stage, the focus group interviews were particularly preferred to the questionnaire because the purpose was to discuss with the student teachers their overall impressions and expectations following this learning experience, and to collect some qualitative insights through raising follow-up questions related to their answers. In the follow-up stage, the questionnaire was used to access the largest possible number of teachers who studied the corpus literacy component, given that it was expected many of them would not be available for the interviews. The questions of both the focus group interviews and the follow-up questionnaire were developed in light of the purpose and research questions of the study. The focus group interviews were guided by five questions that asked the student teachers about their views on the corpus literacy part taught to them and their attitudes towards using corpora in language learning, teaching and research purposes. Below are these guiding questions:

- Did you find the corpus linguistics part taught interesting or uninteresting? In what way?
- When looking for word meaning or use in future occasions, to what extent are you willing to use corpora versus online dictionaries? Please give your reasons.
- Do you think it is feasible to use corpora directly in language teaching? Please explain.
- Are you planning to introduce corpora to your students in your future teaching? If so, how?
- Are you planning to do corpus-based research? Please give your reasons.

As for the follow-up questionnaire, it started with a short bio-data section that included two questions asking the teachers about the educational stage in which they work and the courses they teach. Two drafts of the questionnaire were developed. The author edited some questions in the first draft in light of the feedback received from an expert language teaching researcher. The questionnaire's final draft is provided in the Appendix. As shown in the Appendix, the questionnaire included eight questions that asked the teachers about their views on including corpus literacy instruction in graduate TESOL programs, how frequently they had used online corpora since completing the course, and the extent to which they used them for language learning, teaching, and research purposes. As can be noted, some questions in the follow-up questionnaire were similar to those of the focus group interviews; this was intended to examine any changes or inconsistencies in the teachers' corpus-related immediate perceptions and expectations, and their long-term views and practices.

4.4 Data collection and analysis

The study started with assessing the students' previous knowledge and experiences with corpora during the first class in the course. The needs analysis assessment was conducted through oral discussion with the students in each group using the following questions: Have you ever heard about corpus or corpora? Have you ever used any corpora? And have you studied corpus linguistics in any previous undergraduate or graduate course? This initial needs analysis showed that the students received no corpus linguistics instruction in their bachelor's or master's programs. Although some students reported coming across the word "corpus" and its related terms while reading research papers, they were found to have no experience in using any corpora or knowledge of any corpus websites. The corpus literacy part was taught to the students for three academic weeks from the 10th academic week in the term to the 12th one. After teaching the corpus linguistics part to each class of students, the author conducted the focus interviews with them. Each focus group interview lasted for about 30–35 minutes. In each interview, the author used the guiding questions along with the follow-up ones, which were raised depending on the students' answers to each question. The focus group interviews were conducted in English (the medium of instruction of the course), and were recorded and transcribed at a later stage. Four academic terms – i.e. two years – after the end of the CALL course, the teachers were emailed

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

the follow-up questionnaire that focused on tracing their uses of corpora for learning, teaching and research purposes, and identifying any potential changes in their views about corpus linguistics and its teaching. At this time, the majority of the students who took the CALL course were still completing their PhD research and some of them were working at the author's workplace; that is why it was easy to access them via email. The follow-up questionnaire was sent to all the teachers who completed the course, but only 19 of them responded to it (14 females and five males). Of these 19 respondents, 15 were university teachers and four were teaching at high and middle schools. After completing the data collection stages, the author worked on sorting out the data and analyzing it. First, the student teachers' answers to the similar questions in the interviews and questionnaire were compared to identify any changes in their beliefs about corpus and using its tools. The student teachers' answers to the unique questions included in the interviews and questionnaire were analyzed to examine the other short-term and long-term dimensions of the teachers' corpus-related views and use practices. The data analysis process was guided by Lodico, Spaulding and Voegtle's (2006) guidelines that include preparing and organizing the data, exploring the data, categorizing it, providing rich descriptions of the participants' perspectives and experiences, and confirming the evidence emerging from the data. After organizing the data, the author read the participants' interview and questionnaire answers a number of times to identify the initial themes or categories, which were subjected to further cycles of analysis (Merriam, 1998).

5. Results

In the following four subsections, the author presents the results of the data analysis. In each subsection, the beliefs the student teachers had after studying the corpus literacy component are compared to the practices and beliefs they reported four terms – or two years – after learning about corpora. The four subsections cover the teachers' immediate and long-term perceptions and/or practices with regard to corpora and corpus literacy integration in LTE, using corpora as a learning resource, using corpora in languages instruction, and doing corpus-based research.

5.1 Perceptions of corpora and corpus literacy integration in LTE

The focus group interview and questionnaire data showed that the student teachers' immediate and long-term perceptions of corpora and corpus literacy integration in LTE programs are very positive. It was found that the corpus linguistics part taught to the student teachers made them aware of the importance and benefits of corpora as a learning, teaching and research resource. The focus group interviews revealed that all the student teachers found it quite interesting and beneficial to learn about corpus linguistics and its pedagogical applications and research. This is what the following comment summarizes:

> I liked the corpus and concordance classes indeed. I have always come across research papers that include the word "corpus", but I never knew what it means or how to use it and why we should do research on it. But these classes helped me to learn a lot about corpus and its multiple uses in teaching and research.

These positive responses were also confirmed in the final course evaluation form, but this data is not included in this study. Many of the students completing this form reported that the corpus linguistics was the part they liked most in the CALL course. Thus, all their views on the importance of integrating corpus linguistics in graduate CALL courses were unanimous.

The teachers' answers to the follow-up questionnaire also emphasized their long-term positive perceptions of corpora and corpus linguistics. Table 1 provides a summary of the participants' responses to the yes/no questions in the follow-up questionnaire. In answering the question about their views on integrating corpus linguistics into graduate TESOL programs, all the questionnaire

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

Table 1. The participants' responses to the yes/no questions in the follow-up questionnaire

Participants' responses to the question about:	Raw numbers		Percentages	
	Yes	No	Yes	No
Integrating corpora in graduate TESOL programs.	19	0	100	0
Using corpora more than other tools as a linguistic resource.	1	18	5.26	94.74
Introducing online corpora to their students.	9	10	47.37	52.63
Using corpus-based activities in their English classes.	6	13	31.58	68.42
Doing or planning to do corpus-based TESOL research.	5	14	26.32	73.68

respondents ($n = 19$) supported such integration due to the perceived importance of corpora as a language learning and research resource. As the following questionnaire answers indicate, the teachers viewed that teaching corpus linguistics in CALL courses creates new research and teaching dimensions for graduate students in their educational context:

> I strongly recommend integrating corpus linguistics in CALL courses because any graduate student in language teaching should be familiar with such area. Integrating corpora in CALL courses will reveal to graduate students new research topics and classroom activities. Many graduate students in our university community do not know about corpus linguistics and that's why teaching it is necessary.
>
> Of course, I think that teaching corpus linguistics should become the core curriculum for any CALL course because it familiarizes students with corpora, and makes them researchers and not just learners.

Despite these immediate and long-term perceived benefits of corpus linguistics and its integration in LTE programs, four respondents to the follow-up questionnaire viewed that the corpus literacy component taught to them should have included more practical activities in using corpus tools:

> I wish we have had a more practical part in using corpus. I suggest that at least six hours of the course should be assigned to training in corpus use in language teaching.
>
> I think the course should have included workshops in using corpus-based activities in the class.

As will be noted in two later sections, a few teachers also attributed their negative attitudes towards using corpora in language instruction and doing corpus-based research to the lack of extensive training.

5.2 Using corpora as a learning resource

One of the issues addressed in the two stages of the data collection is examining the student teachers' perceptions and uses of corpora as a learning resource as compared to online dictionaries and digital applications. The immediate responses of the student teachers generally showed that they preferred using online dictionaries and Google when looking for word meaning, and that they only used corpora in some exceptional cases such as looking for word collocations, and contextual and authentic use:

> I will use Google and online dictionaries more because I haven't heard about corpus before this semester. So, I always like to do it in the way I used to.

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

Table 2. Respondents' reported frequencies of using online corpora since receiving corpus literacy instruction in the follow-up questionnaire

Since completing the CALL course:	Raw numbers	Percentages
I have never used any online corpus.	4	21.06
I have used online corpora a few times.	8	42.10
I have frequently used online corpora.	7	36.84

> In fact, corpus is a completely different lexical resource. I will use it only when looking for how words are used in context. For example, how natives use a particular word, or which words collocate with a particular word.

The teachers' responses to the follow-up questionnaire revealed more details about the frequency of their use of corpora as a learning resource and the rationale for depending on it more or less. As Table 1 indicates, 18 questionnaire respondents (94.74%) reported their preferences to explore new word meanings and uses by depending on online dictionaries, technological applications and Google rather than corpora:

> I prefer to use Google and online dictionaries because I can find the word meanings and derivations I need easily and quickly without having to register in the website. They are also very well-known compared to corpus searches.

> I use some dictionary applications and online dictionaries to search for word meaning. It is an easy way and it will not take you much time to find the meanings of the words you are looking for.

These teachers' views of corpora as a linguistic resource were also confirmed by their reported long-term frequencies of consulting or using them. The follow-up questionnaire included a question about how often the teachers have been using online corpora since completing the course. Table 2 gives the participants' reported frequencies. As the table shows, four of the 19 follow-up questionnaire respondents reported they never used any online corpus since completing the course, eight said they used online corpora a few times, and seven reported using them regularly.

Some teachers also reported that the complex use of online corpora and the subscription needed for using them are the two main obstacles hindering their dependence on corpora as a learning resource:

> The use of corpora is a bit complicated and some of them require certain conditions such as being a corpus researcher or affiliated with a particular university.

> Corpus websites are highly designed and provide a lot of useful linguistic information, but I prefer user-friendly websites because I have to subscribe to some corpora to use them for a long time.

In their answers to the questionnaire, the teachers repeatedly mentioned COCA and BNC as two online corpora they have used. The main purpose mentioned for their corpus consultation was examining word contextual and authentic use:

> I have found corpus very useful. The use of a word in real life situations seems far-fetched to a learner or a user of English as a foreign language in my opinion. Hence, finding a bank of everyday English examples provides me with a better experience and a deeper understanding of the use of the word I'm looking for.

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

> I was intrigued by COCA when I first studied corpora and have been using it since then. It is a good reference especially in writing. It provides me with various contextual uses of this word and the genres in which these uses occur. This is really important for me as I need to make sure that my writing fits the academic standard.

As can be noted, the corpus literacy component taught has made the student teachers aware of corpora as an additional language resource. The teachers' responses to the interview and questionnaire indicate they have mainly depended on corpora as a resource for learning about lexical features of word use rather than the syntactic ones. Despite the maturity and development noted in their discussions of corpus-related issues, the student teachers made little use of online corpora due to the nature of accessing and searching them, and their language learning and use purposes.

5.3 Using corpora in language instruction

A third and important dimension the study has tackled is the influence of teaching the corpus literacy part on the teachers' beliefs and practices pertinent to the direct use of corpora and corpus concordances in their classes. After receiving corpus literacy instruction, the interviewed student teachers generally reported optimistic responses to the two related questions. Most of the answers emphasized their awareness of the great pedagogical benefits corpora offer; for example:

> Yes, this is expected. Corpus provides students with an authentic reference for language use. I think I'll be able to integrate it as a lexical source when teaching vocabulary and writing.

However, some student teachers reported concerns about the feasibility and practicality of corpus integration in their future classes. Specifically, the two reasons justifying these teachers' opinions are students' low language levels and the lack of adequate training in teaching using corpora:

> I don't think it will be easy to use corpora in classes because learners will not understand the content of concordances easily.

> In order to integrate corpora in my classes, I personally need a whole course for this purpose.

The follow-up questionnaire included two questions that asked the teachers about whether or not they have introduced corpora to their students and used them directly in their classes. The teachers' answers to the two questions were relatively different. With regard to the first question that asked the teachers whether or not they have ever tried to introduce online corpora to their students, 10 of the questionnaire respondents (52.63%) said no, while nine reported attempting to do so by drawing their students' attention to some corpus websites. A lower number of teachers ($n = 6$, 31.58%) responded positively to the second question concerned with the direct use of corpus in their language classes. Those teachers, however, reported just introducing some online corpora to their students in the class, and then getting them to do some course assignments on their own using them, but not getting students to explore concordances collaboratively during particular classroom lessons. As the following two questionnaire answers indicate, none of these teachers implemented DDL activities in their classes:

> I usually give my students a corpus-based assignment in some courses I teach. For example, in this semester I asked the students in one course to use corpus sites to identify the adjectives occurring with some words and compare their corpus search results with what they find in the dictionary search.

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

> Whenever I find corpus databases relevant to a particular course, I get my students to use them. I prefer to give the students assignments where they have to use these websites to complete their course-related projects.

In contrast, the teachers who reported not integrating corpora in their language instruction justified their practice by a number of factors, including the nature of the courses, the language level of the students they teach, and lack of accessible online corpora. As noted in the following answers given by two schoolteachers, these problems are more popular in the courses taught at the pre-university stages:

> I did not use corpus concordances in my classes because I'm teaching middle school students. So, I have different priorities in teaching these students who have many difficulties in learning to read and speak English.

> No, I didn't integrate corpus in my teaching. I don't have full access to online corpora in my school, and teaching corpus is beyond my students' levels as well.

On the other hand, the university teachers' responses to the follow-up questionnaire indicate that the type of courses played a major role in their initiatives to get students working with online corpora. In their responses to the bio-data section in the follow-up questionnaire, the six teachers who reported trying to engage students in working with corpora mentioned they taught vocabulary, CALL and reading courses. This means that these types of courses are rich environments for integrating corpus-based teaching. In addition to the type of course taught, the nature of students' language levels and learning needs seems to be another main factor hindering university teachers' integration of corpus in their instruction. The following questionnaire answer given by one teacher summarizes this:

> Unfortunately, the students in my classes do not motivate me to do so. With their current language levels, I believe that depending on the basic language learning materials is fair enough.

The above data parts and explanations of the teachers' immediate beliefs and long-term practices indicate that the optimistic expectations most of them reported about incorporating corpus in their instruction after attending the corpus linguistics classes were partially transformed into actual practices when they returned to their workplaces.

5.4 Doing corpus-based TESOL/applied linguistics research

A final issue the study addressed is related to the influence of corpus literacy instruction on the student teachers' attitudes towards doing corpus-based TESOL research. This issue was covered by one question in the interviews and another one in the follow-up questionnaire. The student teachers' immediate responses to the interview question about doing corpus-based research were generally negative. As the following answers indicate, it seems that the student teachers' short-term experiences with corpora after studying the taught part of the course led the majority of them to reject the idea of doing corpus-based TESOL/applied linguistics research:

> I think there are some easier alternatives than corpus research.

> I found corpus research very interesting, but I prefer to do research away from technology.

Since the corpus literacy part taught has only provided the student teachers with an overview of corpus linguistics research and exemplary studies of it, it is possible that the general negative

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

attitudes they showed towards doing corpus research has resulted from the lack of training in how to do such research, along with their short-term experiences with corpora.

The teachers' responses to the follow-up questionnaire indicate that some of them have developed a relatively more positive attitude towards doing corpus research in the future. Only one questionnaire respondent was completing her PhD research on using corpus-based teaching for improving students' use of lexical features in English writing; the other respondents were completing their PhD research in other TESOL areas. In addition to this respondent, four other teachers responded positively to the question about their possible plans to do future post-doctoral corpus research. The following two teachers explain why they are interested in doing future corpus research:

> Yes, I would like to do a future corpus study related to identifying the most frequent words in some coursebooks. I also suggested the corpus research area to one of my colleagues she wanted to apply for a PhD study programme.
>
> Yes, I'm interested in vocabulary teaching and learning in general and in any research approach that may improve students' vocabulary knowledge. Corpus-based research may yield generalizable results about which words students need to know more.

It seems the long-term experience in using online corpora and reading corpus research have led these four teachers to develop a more positive attitude towards doing corpus research. However, such positive attitudes towards doing corpus research should be fostered by training the teachers in how do it.

6. Discussion

From the data given above, the answers to the three research questions addressed by the study can now be summarized. Regarding the first research question that concerns the teachers' perceptions of corpus literacy instruction, the data shows that they unanimously had very positive immediate and long-term views of it and its benefits, and therefore they strongly supported corpus literacy integration into LTE. Despite this, a few teachers were not satisfied with the practical component in the corpus literacy part taught. As for the teachers' post-instruction expectations of their future uses of corpora in language learning, teaching and research (i.e. the second research question), their responses varied depending on each dimension. Despite recognizing the unique and authentic word collocation and use examples found in corpora, the teachers generally viewed that it is difficult to use corpora as an everyday linguistic resource instead of online dictionaries and other lexical tools. Many student teachers reported optimistic expectations about incorporating corpus-based instruction into their future classes; however, they all had no positive attitudes towards doing corpus-based research during the post-instruction stage, perhaps due to their short-term experiences with corpora at that time.

With regard to the third research question, the teachers' long-term actual uses of corpora differed relatively from the earlier expectations they reported at the post-instruction stage. First, the teachers' negative expectations of using corpora as a learning resource were consistent with their limited long-term uses of them. This could be attributed to the nature of corpus search and access, and to their language learning purposes. Second, in contrast to the optimistic expectations many teachers had about implementing corpus-based instruction, only some of them were able to introduce online corpora to their students after returning to their workplaces. The descriptions these teachers reported about integrating corpora into their language instruction imply that they did not implement real DDL activities. Rather, they only introduced corpora in their classes or gave their students assignments that involved them in working with online corpora on their own. Third, some of the teachers started to gradually develop a positive attitude towards doing corpus-based TESOL research, although they all did not expect doing such research when they

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

were first exposed to corpus literacy instruction. Thus, these few teachers' interest in such a research area could have been stimulated by their long-term experiences with using corpora and reading relevant published research.

Although corpus literacy instruction has not brought about the desired changes in the teachers' long-term practices, it has resulted in some positive gains. In line with Farr and O'Keeffe's (2019) view that corpus linguistics enhances the teacher's knowledge, pedagogy and research development, the corpus literacy instruction provided to the student teachers in this study has contributed to raising their awareness of corpora and their benefits. Such awareness has led some of them to use corpora for learning, teaching and research purposes, and to draw their students' attention to online corpora and/or engage them in using corpus tools.

The results of the present study concur with those of previous studies (e.g. Anthony *et al.*, 2017; Çalışkan & Gönen, 2018; Ebrahimi & Faghih, 2017; Farr, 2008; Leńko-Szymańska, 2017; Naismith, 2017) in emphasizing the importance of corpus literacy instruction to fostering teachers' awareness of corpus linguistics and its pedagogical applications. The results also support previous research findings indicating that corpus-based teaching may be hindered by technical and methodological challenges (Farr, 2008; Naismith, 2017), difficulties in designing corpus-based materials (Çalışkan & Gönen, 2018), and teachers' instructional priorities (Anthony *et al.*, 2017). The study indicates that teachers' inability to implement corpus-based instruction may be caused not only by these factors but also by the nature of the courses taught and students' language learning needs. These two obstacles were particularly obvious in the practices reported by the four high and middle school teachers in their responses to the follow-up questionnaire. Thus, the study supports Braun's (2007) conclusion that corpus-based teaching is very likely to be of limited use and success at the pre-university educational stages.

7. Conclusion

The present study compared Arab EFL student teachers' immediate responses to corpus literacy instruction with their long-term views and uses of corpora in language learning, teaching and research. Although the corpus literacy instruction has not greatly enhanced the participant teachers' practical uses of corpora, there are important popularization benefits gained from it. These benefits lie in raising the student teachers' awareness of corpora and their important uses, enabling some of them to help their students learn about corpora, and stimulating a few others' interest in doing corpus-based research. These positive gains could be regarded as a starting point for better near-future developments in popularizing the use of corpora and also in fostering pre-service and in-service language teachers' corpus literacy in the Saudi context.

The results of the present study along with previous research findings indicate that popularizing the direct use of corpora in language classes requires some support from language learning textbook designers, particularly those developing materials for university and high-level students. For example, designers of reading, vocabulary and writing textbooks could incorporate some corpus-based activities that could engage students and teachers alike in working with corpora in language classes. To accomplish this, collaboration between language textbook designers and online corpus developers will be necessary.

The issue of how teachers use corpora after receiving corpus literacy training is worth further research. As implied above, it seems that a very large portion of language teachers in the Arab context are unfamiliar with corpora and do know how to use corpus tools. That is why there is a need for more longitudinal research that introduces corpus literacy instruction to Arab in-service and student language teachers and examines how they make use of it later on. Given that a main limitation of the present study is the short period in which student teachers were taught the corpus literacy part (i.e. nine hours in three weeks), future studies may trace teachers' long-term responses to more extensive corpus literacy instruction. It is likely that

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

exposing teachers to longer or more intensive instruction and training in corpus linguistics may bring about better immediate and long-term responses in their uses of corpora for teaching and research purposes in particular.

Ethical statement. All respondents participated in the study voluntarily and provided informed consent prior to the commencement of the study. The confidentiality and anonymity of the research respondents was maintained throughout the study and the consequent analysis.

References

AbdelRaouf, A., Higgins, C. A., Pridmore, T. & Khalil, M. (2010) Building a multi-modal Arabic corpus (MMAC). *International Journal on Document Analysis and Recognition*, 13(4): 285–302. https://doi.org/10.1007/s10032-010-0128-2

Alangari, M. A. (2019) *A corpus-based study of verb-noun collocations and verb complementation clause structures in the writing of advanced Saudi learners of English*. University of Reading, PhD thesis.

Alsharif, M. (2017) The frequently used discourse markers by Saudi EFL learners. *Arab World English Journal*, 8(2): 384–397. https://doi.org/10.24093/awej/vol8no2.28

Anthony, L., Chen, M. & Flowerdew, J. (2017) Introducing in-service English language teachers to corpus-assisted academic writing pedagogy: A Hong Kong case. *Corpus Linguistics 2017 International Conference*. University of Birmingham, July 23–28.

Boulton, A. & Tyne, H. (2014) Corpus-based study of language and teacher education. In Bigelow, M. & Ennser-Kananen, J. (eds.), *The Routledge handbook of educational linguistics*. New York: Routledge, 301–312.

Braun, S. (2007) Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3): 307–328. https://doi.org/10.1017/S0958344007000535

Breyer, Y. (2009) Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2): 153–172. https://doi.org/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]*. Frankfurt: Peter Lang, 391–415.

Ebrahimi, A. & Faghih, E. (2017) Integrating corpus linguistics into online language teacher education programs. *ReCALL*, 29(1): 120–135. https://doi.org/10.1017/S0958344016000070

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 corpus approaches in English language teacher education. In Walsh, S. & Mann, S. (eds.), *The Routledge handbook of English language teacher education*. London: Routledge, 268–282. https://doi.org/10.4324/9781315659824-19

Frankenberg-Garcia, A. (2012) Raising teachers' awareness of corpora. *Language Teaching*, 45(4): 475–489. https://doi.org/10.1017/S0261444810000480

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.

Hubbard, P. & Levy, M. (2006) *Teacher education in CALL*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.14

Kavanagh, B. (2019) Using 'what already works' to 'bridge the gap' between corpus research and corpora in schools. *Learner Corpus Research Conference*. University of Warsaw, 12–14 September.

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. https://doi.org/10.1017/S095834401400010X

Leńko-Szymańska, A. (2017) Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3): 217–241. https://doi.org/10125/44628

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.

Lodico, M. G., Spaulding, D. T. & Voegtle, K. H. (2006) *Methods in educational research: From theory to practice*. San Francisco: Jossey-Bass.

Merriam, S. B. (1998) *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Mukherjee, J. (2004) Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In Connor, U. & Upton, T. A. (eds.), *Applied corpus linguistics: A multidimensional perspective*. Amsterdam: Rodopi, 239–250. https://doi.org/10.1163/9789004333772_014

Naismith, B. (2017) Integrating corpus tools on intensive CELTA courses. *ELT Journal*, 71(3): 273–283. https://doi.org/10.1093/elt/ccw076

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press

Römer, U. (2011) Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31: 205–225. https://doi.org/10.1017/S0267190511000055
Zaki, M. (2017) Corpus-based teaching in the Arabic classroom: Theoretical and practical perspectives. *International Journal of Applied Linguistics*, 27(2): 514–541. https://doi.org/10.1111/ijal.12159
Zareva, A. (2017) Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1): 69–79. https://doi.org/10.1093/elt/ccw045

Appendix

The follow-up corpus use questionnaire

Dear participant,
The purpose of this questionnaire is to explore your views on corpora and the extent to which you have used corpus tools for different purposes since completing the CALL course. I would be grateful if you would take a few minutes to complete the questionnaire. Please note that the questionnaire asks about how you have already used corpora rather than what you should do with them. Therefore, please try to provide as realistic answers as possible. Your answers will be used for research purposes only, and will be dealt with confidentially so that no one can identify who you are or where you work. Thank you for your cooperation.

Section 1: Please complete the following table:

The educational stage in which you teach
The courses you generally teach

Section 2: Please answer the following questions:

1. In case you have been asked about teaching corpus linguistics in graduate TESOL programs, will you recommend integrating it into them or not? Please give reasons for your answer.
2. Since you have been introduced to corpus linguistics in the CALL course, how often have been using online corpora?

I have never used any online corpus.
I have used online corpora a few times.
I have frequently used online corpora.

3. When searching for word meaning or use, do you usually use corpora more than other resources (such as Google and online dictionaries)? Please explain in detail.
4. If you have used any online corpora since completing the CALL course, how have you found this corpus search and use? Please explain in detail.
5. Have you ever tried to introduce online corpora to your students? Why or why not?
6. Have you ever tried to use corpus-based activities in your English classes? Why or why not?
7. Have you done corpus-based research in your PhD, or are you planning to do it in the future? Why or why not?
8. Do you like to add any more related information about your use of corpora for language teaching and learning purposes? If so, please explain.

About the author

Muhammad M. M. Abdel Latif is a lecturer of English language teaching at the Faculty of Graduate Studies of Education, Cairo University, Egypt. He has received international awards for his research and has published in a number of ranked and well-known international journals, including *Applied Linguistics*, *Assessing Writing*, *Canadian Modern Language Review*, *English Teaching: Practice & Critique*, *European Journal of Teacher Education*, *Journal of Multilingual & Multicultural Development*, *Journal of Research in Reading*, and *System*.

Author ORCiD. Muhammad M. M. Abdel Latif, https://orcid.org/0000-0003-4002-822X

https://doi.org/10.1017/S0958344020000129 Published online by Cambridge University Press