ReCALL **29**(1): 120–135. 2016 © European Association for Computer Assisted Language Learning
doi:10.1017/S0958344016000070
First published online 24 June 2016

Integrating corpus linguistics into online language teacher education programs

ALICE EBRAHIMI

AlzahraUniversity, Iran
(email: aebrahimi09@gmail.com)

ESMAIL FAGHIH

Islamic Azad University, Iran
(email: esfaghih@alzahra.ac.ir)

Abstract

This study reports on a qualitative study which explored 32 pre-service teachers' evaluations of two online seven-week introductory courses in corpus linguistics (CL). Data were gathered through questionnaires, participants' written journals, post-course semi-structured email interviews, and discussion forum entries. The qualitative analysis of data revealed that for successful integration of CL into online language teacher education programs, several procedures should be employed. These include: providing the necessary technological infrastructure in educational settings; incorporating CL instruction in initial stages of language teacher education degree programs and extending it throughout the whole curriculum; focusing more on the practical aspects of CL with much emphasis placed on the necessary pedagogical knowledge and skills for successful exploitation of CL; introducing user-friendly tools and encouraging indirect use of corpora in the absence of necessary technological facilities; providing adequate and effective instructional materials (text-based reading materials, screen capture videos, hands-on activities, etc.) along with sufficient instructor support; and encouraging the participants to reflect on the approach critically. The findings may promise implications for language teacher educators to effectively introduce CL to student teachers in virtual learning environments.

Keywords: CALL, teacher education, corpus linguistics, corpus-based tools

1 Introduction

In recent decades, educating language teachers in computer-assisted language learning (CALL) has grown exponentially. Nowadays there is a general consensus that CALL teacher education should become an integral part of the teacher preparation courses (Jones & Youngs, 2006; Peters, 2006). According to Chapelle (2006: vii), teacher education in CALL "is one of the many areas that compete for a place in the curriculum" of language teacher education programs.

Despite such recognition, a common concern among language teacher educators is to decide on the content of such CALL training programs. Though educating teachers in the

use of some technologies such as emails, wikis, blogs, podcasts, web quests, etc. has received a great deal of attention, training teachers in other areas, e.g. the use of corpora – collections of written and spoken texts in electronic form – has been widely neglected (Boulton, 2010; Römer, 2010).

As regards the potential value of corpora for language teachers, several researchers assert that corpora can contribute to the personal and professional growth of language teachers and student teachers (STs) in various ways. For example, instead of relying on their intuition to make sense of linguistic patterns, teachers can use corpora to increase their language awareness (Farr, 2008), as well as develop teaching materials and resources (Hewings, 2012; Oakey, 2011); moreover, they can compile corpora from learners' artifacts, textbooks or the internet (Granger, 2013), or exploit corpora in language classes to encourage data-driven learning (DDL) (Boulton, 2011b; Breyer, 2011).

Despite the affordances of corpora, the reality is that they are still rarely used by language teachers. One reason may be that, so far, corpora have been primarily used by corpus linguists and researchers rather than teachers or students (Hable, 2010). Therefore, many researchers in the field now argue strongly for the inclusion of corpora and corpus linguistics (CL) in the syllabus of language teacher education programs (Boulton, 2011a; Flowerdew, 2012; Leńko-Szymańska, 2014).

Several researchers to date have examined the efficacy of offering CL training courses for pre-service or in-service language teachers. For example, Farr (2008) reports on integrating corpora activities over a two-year master of arts (MA) program in English language teaching (ELT). The responses of 25 STs to a questionnaire on the perceptions and attitudes to corpus-based study indicated that there was a positive predisposition among the students toward the use of corpora. The program was successful in promoting students' spirit of inquiry and research; a large majority of the STs chose to use corpora in their MA dissertation research.

In another attempt, Boulton (2011a) offered a course in CL entirely via distance as part of an MA in English for students of literature, cultural studies, and linguistics. The instructor adopted a problem-solving, hands-on approach, i.e. he enforced the students to explore the tools themselves to discover their applications, as the participants were required to be relatively autonomous and were used to managing their work on their own. Evaluation was based on students' final reports on a research project in an area of their choice. Boulton reports on 30 papers and highlights the diversity of approaches and topics taken by the students who used the internet, ready-made reference corpora, and small home-built corpora. He concludes that students were able to grasp the concept of corpora and their applications even with minimal training.

Recently Heather and Helt (2012) examined six pre-service teachers' development of corpus literacy during a grammar course offered to 52 English as a second language students. The course included CL activities, and data gathered through surveys, interviews, and narratives were analyzed for five components of corpus literacy. Based on the positive findings indicative of the effectiveness of the training, the researchers encouraged the continuation of this type of training for pre-service teachers.

More recently Leńko-Szymańska (2014) carried out another study during which an elective course in CL was offered to thirteen MA students. The findings indicated that they reacted positively to the course, yet expressed dissatisfaction with too big a workload and the amount of feedback they received on the activities completed. Moreover, the students

found the instruction insufficient. The author suggests that in order for the STs to gain full command of the CL resources and software, extensive exposure to corpora and more guidance on pedagogical application of CL are needed.

Clearly, research on the integration of CL in language teacher education programs is scarce. In particular, there are very few studies reporting on such courses offered fully online. Considering the increased popularity of online education, it is imperative to conduct more studies in this regard. To fill this gap in the literature, the present study was conducted to explore the optimal way of offering a fully online CL course to STs majoring in teaching English as a foreign language (TEFL) at the MA level in an Iranian educational context. Given that CL is largely neglected in TEFL in Iran, the aim of the course was to help STs explore the potential pedagogical, personal, and research applications of CL techniques and tools. This study is distinctive from previous ones in that it examines the participants' reflective experiences of taking a CL course delivered fully online to suggest effective ways of designing such courses in similar contexts. Worth mentioning is that the study presented here is part of a larger research project and solely data reflecting the STs' evaluations of the CL course are discussed. The following research question guided the study:

> How do the STs evaluate the quality and effectiveness of the online CL instruction and in what direction should the training be developed in order to bring the maximum benefits to the STs?

2 The study

2.1 Context and participants

A total of 32 Iranian graduate students of TEFL at the MA level from three universities participated in the study. The research extended over the course of two semesters with two different groups of students. The first group comprised twenty students (fifteen females and five males) in the age range 21–51 and the second group consisted of twelve students (eight females and four males) in the age range 21–40. They were selected with a non-random purposive sampling method, since they had already been allocated into classes. Examining the demographics of the participants revealed that attending an online course was a first experience for all; except for two participants in each group, none had any experience with CL. So, both the course and the delivery mode were new for almost all the STs. With regard to computer literacy and use, most STs had long years of experience in working with computers. Some were teaching or had teaching experience (80% vs. 100% in the first and second group, respectively).

2.2 The online CL course

Following Bird's (2007) "3 'C'" model of course design, a seven-week online introductory course in CL was conceived. The model encompasses three major concepts: (1) content or declarative knowledge; (2) construction, i.e. the social construction of knowledge; and (3) Consolidation, which includes reflection. The aim was to create a vibrant learning community by encouraging weekly reflections and discussions on the part of the STs. The instructor (the first author) would start the discussions by posting two to three questions after each lesson; to encourage more student involvement, her intervention in discussions

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

Table 1 *Resources used*

BNC Online	http://www.natcorp.ox.ac.uk/
COCA	http://corpus.byu.edu/coca/
MICASE	http://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=simple
Just the word	http://www.just-the-word.com/
Compleat Lexical Tutor	http://www.lextutor.ca/
AntConc	http://www.laurenceanthony.net/software/antconc/
Tagcrowd	http://www.tagcrowd.com/

was kept to a minimum. The CL course was an elective addendum to a face-to-face CALL course and an English for specific purposes course in the first and second group, respectively. Additional bonus points were assigned for STs' active participation in the CL course.

The course introduced some free, easily available corpora resources suitable for pedagogical or research purposes, which are presented in Table 1.

The main instructional materials were instructor-made reading texts of about ten to twelve pages for each lesson alongside four or five hands-on activities, supplementary reading texts, and seventeen screen captured videos of about five minutes – made either by the software developers (10) or by the instructor (7). The software used for making the videos was Jing (available at: www.jingproject.com).

The sources and topics were chosen based on the first author's previous experience of conducting four workshops on the use of corpora in ELT, the feedback gained from the 50 participants of the workshops, as well as the ideas suggested in the literature by experts in the field. The main aims of the course were: to give STs some insight into the basic concepts in CL and a range of free corpora, corpus-based resources, and tools; to familiarize STs with the principles of compiling corpora and interpreting some statistical measures; to prepare STs to understand the various applications of corpora in language pedagogy and research; and to empower STs with the knowledge and skills to develop corpus-based teaching materials and class activities. At the end of the course, STs were required to design three corpus-based activities as final course projects. The course was delivered on a weekly basis through a virtual learning environment called Nicenet (available at: www.nicenet.org). Table 2 displays the course content.

2.3 Data collection and analysis

The main sources of data included: a pre-course questionnaire on STs' demographics, STs' weekly written reflections on each lesson, discussion forum entries, and post-course semi-structured email interviews. Since one of the main aims of the study was to engage STs in critical reflection of the course content, participants were asked on a weekly basis to reflect on the strengths and weaknesses of each lesson, assess the contents related to their profession as an MA student, a teacher/would-be teacher, and a researcher, and make suggestions on how it could be improved. The web-based interviews asked four open-ended, essay-style questions about STs' attitudes toward attending a fully online CL course, the affordances and constraints of CL, and the suitable level for introducing CL into ELT programs. To analyze these qualitative data, we used MaxQDA (version 10) – a piece of computer-assisted qualitative data analysis software (CAQDAS). The data were coded

Table 2 *Course content*

Lessons	Topics covered
1. Introduction to Corpora	What is a corpus? What is a concordancer? What is corpus linguistics? How big should a corpus be? What are corpora used for? Different types of corpora; Examples of free online corpora (BNC, COCA, MICASE); Key terms (hapax, keywords, lemma, lexical density, span, token, type, type/token ratio, word lists).
2. A Web-based Concordancer: The Compleat Lexical Tutor	Concordance; Multi-concordance; Text-Based Concordances; Range; Text-Based Range; Text-Lex Compare; N-Gram Phrase Extractor; VocabProfile; Keywords Extractor.
3. The Corpus of Contemporary American English (COCA)	Conducting searches using List, Chart, KWIC and Compare functions of COCA; Finding the collocates; Working with POS list; Searching for lemmas; Searching for synonyms; Searching in sub-registers.
4. Data Driven Learning (DDL)	What is DDL? The role of teachers in DDL; Using corpora in language classrooms; Preparing students to work with concordance lines; Sample DDL activities.
5. Teaching & Researching Vocabulary and Grammar (Lexico-grammar)	Frequency lists; Raw frequency and normalized frequency; Collocation patterns; Weak and strong collocations; Statistics for collocations; Colligations; Fixed phrases: lexical chunks; Frequency-based wordlists; Teaching and researching grammar; Spoken grammar vs. written grammar; Just the Word.
6. Using Corpora to Teach Reading and Writing	Reading challenges; Writing challenges; Benefits of corpora; Creating your own corpus; Selecting and grading of texts for reading: Lexical text analysis; Text Clouds.
7. AntConc	Uploading files; Concordance tool; Concordance plot tool; File view tool; Clusters; Collocates; Word List, Keyword List.

http://www.laurenceanthony.net/software/antconc/

inductively by the first author following the 3 'C' approach proposed by Lichtman (2006), where the three Cs stand for codes, categories, and concepts. The data were coded and grouped with special reference to the research goals and questions.

To establish study reliability, as a first step, the first author coded 20% of the data and then asked another coder to code the same data. The inter-coder reliability was calculated as 0.85 for interview data and 0.83 for reflection data. After discussing the points of disagreement, the first coder continued the coding procedure. To ensure consistency of analysis, a code-recode approach was employed (Oriogun & Cave, 2008). After initial coding, the same researcher recoded the same data after an interval of three months which yielded 90% consistency between the first and second content analyses for interview data, and 85% for STs' reflection data. In reporting the results, to observe the ethics of qualitative research, the

participants were given numbers instead of their true names, i.e. STs in the first group were given the numbers from 1 to 20 and STs in the second group, from 21 to 32.

3 Results

Using the 3 'C' model of data coding, the main concepts repeatedly mentioned in the data were classified into two main categories: (1) CL integration in teaching/learning; and (2) training demands.

3.1 CL integration in teaching/learning

This category encompasses the two following subcategories: (1) opinions on CL tools and techniques; and (2) affordances and constraints of the application of CL in ELT.

3.1.1 Opinions on CL tools and techniques. Several sources were presented in the course which provoked different reactions from the STs. Their evaluations of CL tools can be categorized in two subcategories: (1) CL tools mainly useful for teachers and researchers; and (2) CL tools useful for learners.

CL tools mainly useful for teachers and researchers: Interestingly, while reflecting on different CL tools, STs found some tools more applicable for language teachers and researchers than learners. For example, as regards Lextutor, except for the 'Corpus-based Concordances' and 'Multi-Concordance' sections, most STs considered the other tools, e.g. Text-based Concordances, Text-based Range, Text-Lex Compare, N-Gram Phrase Extractor, Vocabprofile, and Keywords Extractor, as mainly proper for teachers/researchers. They thought that working with these tools required special training and understanding of some jargon related to CL such as type, tokens, type/token ratio, collocation, and range. In addition, five STs found the appearance of the website not very attractive because of its dark background. For example, ST 15 thought that Lextutor was not user-friendly and attractive, yet she liked the availability of different types of corpora there:

> *In my opinion working with Lextutor does not seem user-friendly and I do not like to work with it. The interface is somehow messy and not attractive because of the black background.* (ST 15)

Of the Lextutor tools that attracted the STs' attention, mention can be made of N-Gram Phrase Extractor and Vocabprofile (VP). N-Gram Phrase Extractor enables its users to search for chunks of words in any text they submit. VP also analyzes a text's vocabulary and organizes it into four categories – K1, K2, academic words (AWL), and off-list words. Seven STs were of the opinion that lexical chunks were very important in vocabulary learning and hence valued N-Gram Phrase Extractor. Moreover, some STs (twelve participants) appreciated the VP section and considered it a useful tool for teachers and researchers for analyzing the learners' writings or examining the level of complexity of different texts. ST 16's reflection is provided here:

> *As a teacher, I found the vocabprofile part the most interesting…Checking the variety of words used by students in writing is not something that a teacher can do without the help of technology*. (ST 16)

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

The STs had similar feelings toward COCA. Most of them appreciated the website for its features; however, they found it too complicated to be used in language classes because of its strict search rules. Some STs even regarded the tool as being difficult for teachers, and thought they would need to put in more time and effort to fully explore COCA's features. Another issue complicating the process was the analysis of the output, which needed a lot of practice and experience. The quotations below from the students speak for themselves:

The capabilities of this tool seemed extensive but complicated and I didn't quickly see how they could be easily used by the actual teacher or student. (ST 14)

I understand how I should search but I think I need more time to work with it and explore it. It is somehow difficult for me to analyze statistics. (ST 6)

STs reacted to COCA differently; some of them were very optimistic while others had moderate or even negative views. Some STs with more teaching experience were looking at the possible pedagogical applications of COCA with more caution; they took into account the constraints of the classroom. For example, ST 16 with five years' teaching experience was able to see the realities of the classroom more accurately. Therefore, unlike other STs with little or no teaching experience, she warned that COCA could not be used with all types of learners because of its strict search rules:

I think that COCA is a little bit complicated and time consuming... Now I think that we cannot use corpora for the whole classes that we have and not even for the all the students in a class or we may need to provide facilities and more instructions. (ST 16)

Another teacher, ST 14, who had been teaching for eight years, was of the same opinion. She referred to two prerequisites vital for learners to be able to use COCA: technical proficiency and grammar knowledge:

Unfortunately, I doubt that I would use this particular tool as it seems aimed at users of a higher degree of technical proficiency and grammar knowledge. (ST 14)

As a native speaker of English, ST 14 herself lacked formal grammar knowledge and thus was unable to conduct grammatical searches. This implies that in order to do successful searches with concordancers, one should have some formal grammatical knowledge. This is especially true in working with parts-of-speech-tagged corpora like COCA where one has to specify the parts of speech (POS) of the words. Lack of sufficient grammatical knowledge may hinder natives or learners with little grammatical knowledge in some aspects of corpus work.

Interestingly, STs with less or no teaching experience were much more optimistic about using the tool in language classes. ST 5, with no teaching experience, noted:

As a would-be teacher, I definitely will use COCA in my classes. I will introduce it to my students in the classroom and teach them how to use them step by step. (ST 5)

Unlike teachers at the two extremes, some STs were more moderate. There were six STs, with three to five years' teaching experience, in favor of *indirect use* of COCA in language classes, who suggested using printouts of search results in classrooms. On the other hand,

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

most STs were pessimistic about introducing COCA to learners at lower levels of English proficiency. They suggested that the tool should be introduced to advanced learners only:

> *I think lower levels of students cannot work with COCA. It would be highly confusing for them. However, it can be a useful reference tool for students at advanced levels to improve their vocabulary and grammatical knowledge.* (ST 31)

Another tool considered convenient mainly for language teachers and researchers was AntConc – a free, downloadable piece of concordancing software which includes a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a tool for showing word distribution plots. The STs were in agreement that the tool could assist teachers to correct their learners' writing. However, its use in language classes was treated with skepticism. The view that AntConc was suitable only for adult learners with high level of proficiency in English was a shared concept to which many participants agreed. According to two STs, AntConc was not a user-friendly tool for learners, especially learners at beginner or intermediate levels, as its use requires a lot of training. On the other hand, ST 5 suggested training adult learners in the use of some sections of AntConc, such as the concordance tool, file view, and concordance plot. ST 29 made a good observation regarding the use of the tool in language classrooms: while accepting the applicability of all the functions of AntConc, she rightly acknowledged that in applying the tool, different factors should be considered such as class time, students' level and expectations, and the teacher's creativity:

> *All parts are applicable to the language classroom. It depends on teacher's creativity, class time, students' level and expectations.* (ST 29)

Two other STs found AntConc especially effective in ESP classes where both the learners and teachers can compile a corpus related to a specific field of study and create keywords lists by the use of a reference corpus.

CL tools useful for learners: The participants considered some tools such as the 'corpus-based concordances' and 'multi-concordance' sections of Lextutor, text cloud generators like 'Tagcrowd', and collocation finding tools like 'Just the word', most applicable for learners. These tools were regarded as both user-friendly and interesting for language learners at different levels.

As regards the 'Corpus-based concordances' section of Lextutor, the STs thought that this section could serve as a good starting point for encouraging learners to work with the different corpora available there. The two corpora of great interest for the participants of this study were 1K and 2K graded corpora. These corpora comprise hundreds of graded readers developed over ten years, so they contain a limited vocabulary comprehensible for students even at beginner levels. One problem with general corpora is that they expose learners to many unknown vocabularies which may lead to disappointment. By using 1K and 2K graded corpora, teachers can alleviate this problem. In addition, the 'Multi concordance' program had great attraction for ten STs, who considered its use in teaching near synonyms very valuable. As an illustration, ST 22 explained in the discussion forum how this tool can be used to compare the use of the three words *talk*, *speak*, and *tell*:

> *I like the multi concordance a lot. It can be used by teachers in language classes to teach near synonyms. We can write, for example, the three words 'talk', 'speak', and 'tell' and ask students to compare the results.* (ST 22)

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

Another tool welcomed by STs was 'Tagcrowd' – a tool for generating word clouds or text clouds. A text cloud is a visual display of a list of words most frequently used in a text where the font size of the words reflects their frequency of use in the text. Text clouds can be used for pre-reading, reviewing the previous lessons, writing etc. The ease of generating text clouds was appealing to most STs who viewed the tool as less intimidating for learners:

> *The most exciting tool is text-cloud. You know because I myself am highly visualized and I think most of the people are so.* (ST 12)

Finally, the last tool highly appreciated by most STs was a collocation finder site called 'Just the word'. Its easy-to-use interface was attractive for almost all participants. STs suggested that learners could use the tool at home.

3.1.2 Affordances and constraints of the application of CL in ELT. To examine the pros and cons of applying CL in ELT, the STs were asked to express their views in this regard. To sum up their viewpoints, the STs' comments are arranged in the following list with the numbers in parentheses indicating the number of times each was mentioned.

a. Affordances of CL

- It gives the opportunity to have access to a rich, varied, and authentic language database. (12)
- It is an effective way to learn and teach vocabulary. (10)
- It is motivating, attractive and useful. (7)
- It is a valuable tool for conducting linguistic research. (5)
- Learners can have active control of their learning. (3)

b. Constraints of CL

- It needs access to computers and the internet. (12)
- Learning to work with it is very time-consuming. (11)
- Designing corpus-based activities are very time-consuming and difficult. (9)
- Interpreting concordance lines is difficult. (7)
- Learners should follow strict search rules, which are frustrating and difficult. (6)
- Data are short and incomplete. (5)
- It needs inductive learning, which is not appealing to some students. (3)

As can be seen, accessing rich, authentic language and exploiting corpora in teaching and learning vocabulary were viewed by many as advantageous; however, the need for access to computers and the internet, the time-consuming nature of the approach, and the difficulty in interpreting concordance lines were found to be most disadvantageous.

3.2 Training demands

The themes that emerged in this category are classified into five subcategories: (1) emphasis on practical applications of CL; (2) emphasis on theoretical and research applications of CL; (3) instructor support; (4) instructional materials; and (5) early exposure to CL.

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

3.2.1 Emphasis on practical applications of CL. Eighteen participants showed a high interest for pedagogical applications of corpora. Mainly, they were STs who were teaching at that time or had had some experience in teaching. They were eager to learn how to work with CL tools and apply them efficiently in language classrooms. For example, after being introduced to second and third lessons, ST 14, ST 3 and ST 13 complained that there were few examples on how to use the tools in real classrooms. ST 14 wrote:

> *Unfortunately, there weren't many examples of how you could actually use these applications in a real way, say in the classroom or for homework assignments.* (ST 14)

However, when lesson four was offered – a lesson on the use of DDL in language classrooms – and STs became familiar with the design of corpus-based activities, some STs confessed that this was the very lesson they were longing for from the start of the course. For some students, the practical aspects were so prominent that three of them suggested that in introducing the CL tools, examples from textbooks taught in language classrooms should be provided. This highlights the fact that STs should be constantly given examples of real-life applications of the tools; merely working with different interesting CL tools is of little value if there is no effort to link it to practice.

3.2.2 Emphasis on theoretical and research aspects of CL. Seven STs were highly research-oriented and repeatedly insisted on receiving more information on research-related applications of CL. Maybe, conducting research for their MA thesis was the main reason for their insistence. These STs appreciated different research sections of Lextutor as well as the statistical information introduced in Lesson 5. Interestingly, three of these STs pursued their interest after the course was over and did some research based on the instruction obtained in the course. Unlike this group, others regarded the sections related to research and statistics to be irrelevant. Comparing the number of STs interested in the practical applications of CL (Eighteen participants) and those interested in research applications (Seven participants), it can be concluded that training programs in CL should focus more on those practical aspects of CL immediately applicable in language teaching.

3.2.3 Instructor support. In this study, the role of the instructor was significant: the STs heavily relied on her to gain answers to their problems, both technical and subject-related. She would constantly receive emails from the STs requesting her help to solve their problems. Everything had to be explained in detail by the instructor.

As an illustration, after introducing different sections of Lextutor in Lesson 2, the instructor did not go into detail for two sections, i.e. Keywords Extractor and Text-Lex Compare; she just provided the related links along with necessary explanations to help STs further their explorations of the website on their own. It appeared that for some STs (five) this was not sufficient. In the words of ST 2:

> *You didn't explain some sections in detail. They were not screenshots for keywords extractor and Text-Lex compare. It would be better if like other section you could provide more explanations and related screenshots for these sections as well.* (ST 2)

On another occasion, when the procedure for registering in COCA was not explained, some STs got into trouble. To register in COCA, one has to choose from among

three categories: (1) researcher; (2) semi-researcher; or (3) not a researcher. For the first two categories, one has to provide a link to a page on their university website where their names and information are mentioned. Most of the students got into trouble by choosing one of the first two categories. ST 13 wrote:

> *The process of registering is confusing and not explained adequately. So it's better that you talk about it more when it comes to specifying the category options.* ST13

It is worth noting that there is a comprehensive 'help' section in most websites which provides assistance to the users; however, it appeared that most STs did not refer to this section to get help; rather, they preferred to consult with their instructor on problematic matters.

3.2.4 Instructional materials. Examining the STs' reflections on instructional materials revealed that the most frequent statements were related to providing hands-on practice and screencasts. The STs highly valued the hands-on activities offered in the course and asked for more activities:

> *The hands on exercises are very helpful. They help teachers to realize the real use of corpora and will both enhance their understanding and manifest the way to use it in a right place.* (ST 3)

This denotes that reading a text on the use of CL tools is of little worth if it is not supported by hands-on experience; exploring the tools personally can foster STs' learning. An efficient way to encourage this practice is to provide links to online corpus resources in the reading materials. Another possible way is to design appropriate assignments: most STs constantly asked for more exercises to explore the tools thoroughly. Another option is to ask STs to conduct final course projects. One of the STs wrote:

> *The experience which I will never forget is our mini project. I just felt excited that I am designing activities for a lesson which I did not have any experience about it before the course. I had a very good feeling toward it.* (ST 5)

Besides hands-on activities, the STs' reflections on course materials highlighted the importance of using screencasts along with screenshots (computer screen saved as a picture). A screencast – also known as a video screen capture – is a digital recording of computer screen output often along with a narration. The use of screencasts would serve to give students – especially those with audio-visual learning preferences – more options for gaining knowledge.

Though most STs were satisfied with the screenshots, some preferred screencasts. For example, after introducing the second lesson on the use of Lextutor, three suggestions were made to add relevant videos. ST 1 complained that in the absence of videos, an online course would resemble the same boring traditional way of teaching which could demotivate some students. ST 21 and ST 15 were also of the opinion that supplementing the reading texts with screen capture videos would make the topics more understandable and interesting:

> *It was described just through text and screenshots. It was better to include an audio or video file in this lesson.* (ST 15)

The suggestions motivated the instructor to add some videos in subsequent lessons, featuring records of everything done on the computer screen along with the instructor's

https://doi.org/10.1017/S0958344016000070 Published online by Cambridge University Press

narration explaining the procedures in a step-by-step manner. Most STs appreciated the addition, as reflected in the statement made by ST 14:

> *Before watching the videos I was not impressed with the tool, but after the videos I was able to really see how we could use COCA to help our students.* (ST14)

However, due to individual differences, some STs expressed their preferences for more detailed written texts than videos. As an example, after presenting the third lesson, ST 23 asked for a more comprehensive written text on searching in COCA.

3.2.5 Early exposure to CL. All participants in the present study were of the opinion that CL should definitely be added to the syllabus for TEFL. However, there was a diversity of opinions on when and how it should be done. Six STs considered the MA level as suitable, while others regretted that they had not been introduced to CL at the bachelor of arts (BA) level and strongly suggested that the syllabus of ELT programs should contain this element earlier at the BA level – either embedded in a writing course or as a linguistics course. This would give the STs the opportunity to work with CL tools personally and would help them gain the necessary competence and confidence to apply the tools and techniques in their teaching. Teachers should be competent corpus users capable of handling the tools adeptly so as to guide their students to use the approach efficiently. This initial exposure to corpora at the BA level could help students gain the necessary technical knowledge and skills; later, STs could explore CL more deeply at the MA level. At this level, besides providing STs with more advanced technical skills on the use of corpora for conducting research, more attention could be paid to equipping STs with *pedagogical* knowledge and skills for applying CL in language classrooms. ST14 suggested that at the MA level the course could be offered not as a core course, but as an optional course for those interested in exploring CL further.

4 Discussion and conclusion

As stated earlier, studies describing CL courses offered to MA students, especially in virtual environments, are scarce. Therefore, while discussing the results, in some cases, comparisons are made between the findings of this study and those of other online courses offered to prepare pre-service teachers to integrate technology in education.

The findings about STs' views on the affordances of CL are consistent with the results reported in Farr (2008). In her study, STs found novelty, fun, and interest to be important aspects alongside accessing data of real language use and ascertaining information about frequency, cluster, and collocations as the obvious benefits of corpus use. The STs in the current study also appreciated CL mainly for the opportunity it provides for its users to have access to rich, varied and authentic language as well as for its contribution to learning vocabulary.

On the downside, time constraints were regarded as problematic in both studies. The STs felt that introducing CL to learners as well as exploiting the approach in language classrooms would require a lot of time. It is a fact that time constraints may hinder or discourage the use and integration of any innovation by teachers. Therefore, in selecting CL tools for use in classrooms, priority should be given to introducing easy-to-use tools. Also, considering the fact that designing corpus-based activities is time-consuming, more corpus-based materials should be developed by materials developers to help teachers in this regard.

Unlike Farr's (2008) observation, however, for participants of this study the most disadvantageous aspect of using corpora was the necessity to have access to computers and the internet. This is due to the fact that most educational settings in Iran lack the necessary technical infrastructure to apply CL directly in language classrooms (Hedayati & Marandi, 2014); this indicates that for introducing CALL and hence CL in technology-poor educational settings like Iran, providing computer facilities in educational centers should be regarded as a high priority. For now, the training courses for teachers should attend more to issues related to *indirect* use of corpora for pedagogical purposes. Indirect applications of corpora could be done in various ways. For example, teachers can use wordlists to establish vocabulary targets for learners at different levels, or they can employ corpora tools such as VocabProfile to analyze students' writing, as well as analyze the difficulty level of different instructional texts. Corpora can also be used indirectly by language teachers for designing classroom activities or developing tests.

As regards training demands, the results of this study showed that eighteen participants were in favor of gaining more information on the practical applications of CL in language pedagogy rather than research applications. Other researchers have reported similar results. For example, Goktas, Yildirim, and Yildirim (2008) found that pre-service teachers in a technology course were more interested to learn the practical applications of the tools rather than the theoretical aspects. The researchers suggested that all examples and applications used in the course should be related to the participants' future profession as teachers. Jang (2008) also suggests that conceptual or theoretical information should be linked to practice so that pre-service teachers can understand the reasons behind using technology. With regard to CL, as the results revealed, attention should be paid to introducing more user-friendly tools immediately applicable in language classrooms. According to Rogers (1995), innovations that are perceived by key users as useful and simple to use are more easily adopted.

As the analysis revealed, the STs in this study sought the support of their instructor to a considerable extent. This finding is consistent with Kop's (2008) study, where nearly all students participating in an online course preferred the help and support of the local or online tutor. This highlights the pivotal roles of online instructors especially in teaching fully online technical courses like CL. They should guide the STs through resources and activities, critically engage them in the course content, and provide constant encouragement and affective support.

Interestingly, Boulton's (2011a) observation is in contradiction with this finding. He offered a course in CL entirely at a distance, as part of an MA program. He reports that even with minimal explanation the students were able to explore the tools themselves to discover their features and applications. The current study does not support this claim. Unlike Boulton's participants who were relatively autonomous, STs of this study were very dependent on their instructor. It seems that most participants were not ready for personalized and autonomous learning. They would repeatedly ask for more examples and explanations on various points. This was done despite the fact that all lessons were designed in a way which offered sufficient explanations with accompanying screencasts on each issue. Moreover, STs had the opportunity to further address the issues in the discussion forums. Yet these were not satisfactory for some of the STs.

Various factors can account for this controversy. First, nearly all of the STs in this study were unfamiliar with CL and the use of concordancers, and so were unaware of the plethora

of resources available on the topic. They thought of themselves as less knowledgeable and therefore felt safe relying more on their instructor for gaining the necessary information. Sometimes, the questions posed in the discussions were highly technical and needed the direct involvement of the instructor. Second, they were all new to online education and had less experience in attending an academic discussion forum. In fact, six STs expressed their preferences for attending a hybrid course where some of the sessions are offered face-to-face. Third, most STs, especially in the second group, were busy working and studying. Simply, the time pressure and constraints may have prevented them from exploring the materials individually. Fourth, as acknowledged by most STs, they were all accustomed to teacher-centered classes; Noora (2008) also maintains that the nature of teaching in Iran is mainly teacher-centered. Hence, some STs may hesitate to change their established traditional roles and practices and act as autonomous learners in online environments. Attending a hybrid course in CL may satisfy the learning needs of these students.

The results related to the instructional materials highlight the importance of hands-on practice and screen capture videos. The present study suggests that CALL-related subject matters like CL should encompass a considerable number of hands-on activities. This finding supports the guidelines of Goktas, Yildirim, and Yildirim (2008: 177) indicating that technology-based training for teachers should "provide hands-on practice so that teachers become comfortable with it". Wu and Sankar (2013) also report that participants in their study expressed improved satisfaction with hands-on activities.

Moreover, most participants in the current study expressed their satisfaction with the screen capture videos. Since the course content required working with different tools, for some STs, especially the audio-visual-oriented participants, the mere text-based instruction was tedious and sometimes not satisfactory. The STs acknowledged that combining reading texts with related videos helped them grasp the content more profoundly. This finding corroborates the results reported by other researchers. For example, Rice, Hiltz, and Spencer (2005) also noticed that combining one or more media with text-based documents might result in a better learning outcome. They argue that "text-based communication may not create an optimal learning environment for some learners, given that learners have different learning styles and preferences in terms of type of medium of information" (p. 227). This finding, however, refutes the results reported by Clair (2009). In his experience in teaching two online courses, about 80% of the students considered videos "to be more of a nuisance and a distraction… Many students indicated that they often ignored videos that had been assigned in their previous classes" (p. 172).

Of course, it should not go unnoticed that students differ from each other in their learning styles and preferences as well as access to high speed internet; thus, the present study suggests that videos should be exploited as supplemental to course text-based materials; the course text transcripts should contain all the necessary information for those who do not like to watch videos or do not have access to high speed internet. However, it should not be ignored that the utilization of video technology strongly depends on technological infrastructure and bandwidth (Lazarevic, 2011). Thus, in choosing or developing the videos, consideration should be given to make short videos which could be easily downloaded by all STs.

Another finding of the present study was most STs' preferences for being introduced to CL earlier at the BA level. This suggests that STs value the CL instruction to improve their own language awareness. Of course, as Leńko-Szymańska (2014) rightly acknowledges, CL instruction should not be restricted to a one-semester-long course.

In sum, insights obtained from the data pinpoint some procedures that can be employed for successful integration of CL in online teacher education programs, such as providing the required technological infrastructure in educational settings; integrating CL instruction in initial stages of language teacher education degree programs and extending it throughout the whole ELT curriculum; focusing more on the practical aspects of CL with more emphasis placed on the necessary pedagogical knowledge and skills for successful exploitation of CL by language teachers; introducing user-friendly tools and encouraging indirect use of corpora in the absence of necessary technological facilities, providing adequate and effective instructional materials (text-based reading materials, screen capture videos, hands-on activities, etc.) alongside sufficient instructor support, and encouraging the STs to reflect on the approach critically.

As suggestions for future studies, other researchers could offer more comprehensive CL courses including topics not covered in this course, such as examining learner corpora, using the web as a corpus, and tagging and annotating corpora. Studies may also be replicated with a larger number of participants within and across different universities so that the results could be generalized more broadly.

References

Bird, L. (2007) The 3 'C' design model for networked collaborative e-learning: A tool for novice designers. *Innovations in Education and Teaching International*, **44**(2): 153–167.

Boulton, A. (2010) Data-driven learning: Taking the computer out of the equation. *Language Learning*, **60**(3): 534–572.

Boulton, A. (2011a) Bringing corpora to the masses: Free and easy tools for language learning. In: Kübler, N. (ed.), *Corpora, language, teaching, and resources: From theory to practice*. Bern: Peter Lang, 69–96.

Boulton, A. (2011b) Data-driven learning: The perpetual enigma. In Goźdź-Roszkowski, S. (ed.), *Explorations across languages and corpora*. Bern: Peter Lang, 563–580.

Breyer, Y. (2011) *Corpora in language teaching and learning: Potential, evaluation, challenges*. Frankfurt: Peter Lang.

Chapelle, C. A. (2006) Foreword. In: Hubbard, P. and Levy, M. (eds.), *Teacher education in CALL*. Amsterdam: John Benjamins, vi–viii.

Clair, D. J. S. (2009) My experience with teaching online: Confessions and observations of a survivor. *MERLOT Journal of Online Learning and Teaching*, **5**(1): 166–175.

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.

Flowerdew, L. (2012) *Corpora and language education*. Basingstoke: Palgrave Macmillan.

Goktas, Y., Yildirim, Z. and Yildirim, S. (2008) A review of ICT related courses in pre-service teacher education programs. *Asia Pacific Education Review*, **9**(2): 168–179.

Granger, S. (2013) Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, **20**(3): 465–480.

Heather, J. and Helt, M. (2012) Evaluating corpus literacy training for pre-service language teachers: Six case studies. *International Journal of Technology and Teacher Education*, **20**(4): 415–440.

Hedayati, H. F. and Marandi, S. S. (2014) Iranian EFL teachers' perceptions of the difficulties of implementing CALL. *ReCALL*, **26**(3): 298–314.

Hewings, M. (2012) Using corpora in research, teaching, and materials design for ESP: An evaluation. *Taiwan International ESP Journal*, **4**(1): 1–24.

Hable, A. (2010) *Implementing computer corpora as a source in English language teaching (ELT)* (Unpublished doctoral dissertation). University of Vienna: Vienna, Austria.

Jang, S. J. (2008) The effects of integrating technology, observation and writing into a teacher education method course. *Computers & Education*, **50**(3): 853–865.

Jones, C. M. and Youngs, B. L. (2006) Teacher preparation for online language instruction. In: Hubbard, P. and Levy, M. (eds.), *Teacher education in CALL*. Amsterdam: John Benjamins, 267–280.

Kop, R. (2008) Web 2.0 technologies: Disruptive or liberating for adult education? Gateway to the Future of Learning. *Proceedings 49th Adult Education Research Conference*. St. Louis: University of Missouri, 222–227.

Lazarevic, B. (2011) *Examining the role of the introductory video in the development of teaching presence in online instruction* (Unpublished doctoral dissertation). University of Nebraska-Lincoln: United States.

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.

Lichtman, M. (2006) *Qualitative research in education: A user's guide*. Thousand Oaks, CA: Sage.

Oakey, D. (2011) Using corpus and web language data to create EAP teaching materials. In: Ho, C., Anderson, M. L. K. T. and Leong, A. P. (eds.), *Transforming literacies and language*. London: Continuum, 167–186.

Noora, A. (2008) Iranian undergraduate non-English majors' language learning preferences. *GEMA Online Journal of Language Studies*, **8**(2): 33–44.

Peters, M. (2006) Developing computer competencies for pre-service language teachers: Is one course enough? In: Hubbard, P. and Levy, M. (eds.), *Teacher education in CALL*. Amsterdam: John Benjamins, 153–165.

Oriogun, P. K. and Cave, D. (2008) Using code-recode to detect critical thinking aspects of asynchronous small group CMC collaborative learning. *Journal of Interactive Online Learning*, **7**(3): 208–226.

Rice, R., Hiltz, S. R. and Spencer, D. (2005) Media mixed and learning networks. In: Hiltz, S. R. and Goldman, R. (eds.), *Learning together online: Research on asynchronous learning network*. Mahwah, NJ: Lawrence Erlbaum Associates, 215–237.

Rogers, E. M. (1995) *Diffusion of innovation*. New York: Free Press.

Römer, U. (2010) Using general and specialized corpora in English language teaching: Past, present and future. In: Campoy-Cubillo, M. C., Bellés-Fortuño, B. and Gea-Valor, L. (eds.), *Corpus-based approaches to English language teaching*. London: Continuum, 18–35.

Wu, Y. and Sankar, C. S. (2013) Impact of hands-on research experience on students' learning in an introductory management information system course. *Decision Sciences Journal of Innovative Education*, **11**(4): 335–358.