



Corpus linguistics and continuous professional development: participants' prior knowledge, motivations and appraisals

Linguística de corpus e formação profissional contínua: conhecimento prévio, motivações e avaliações dos participantes

Vander Viana

University of East Anglia, Norwich / United Kingdom

vander.viana@uea.ac.uk

<http://orcid.org/0000-0003-3079-4393>

Lu Lu

Hong Kong Polytechnic University, Hong Kong / Hong Kong

lu-cbs.lu@polyu.edu.hk

<http://orcid.org/0000-0002-6049-6154>

Abstract: Previous studies on the application of corpus linguistics (CL) to education have primarily examined language-related contexts where students are pursuing a formal degree (e.g. undergraduate and Master's programs). Little do we know about the informal learning of CL especially by (but not limited to) academics/professionals who are not educated and/or do not work in language-oriented fields. The present study addresses these research gaps by examining the perspective of participants in a non-credit-bearing continuous professional development (CPD) project aimed at academics/professionals in a range of disciplines, who did not need to have any prior knowledge of CL. More specifically, we administered a questionnaire to 28 participants of a UK-based CPD project on CL with a view to researching four main aspects: (i) these participants' CL background; (ii) their motivations to participate in this type of project; (iii) the advantages and barriers of employing CL in their teaching practice; and (iv) their appraisal of corpus analysis integration in their research practice. The results point out to the role of CPD projects in democratizing access to CL education both to language-oriented and non-language oriented academics/professionals and in potentially raising their interest in CL learning. Lack of knowledge is perceived to

be the main barrier in embedding corpus approaches to teaching and research, thus reinforcing the relevance of developing formal and informal CL learning opportunities for academics/professionals in different fields.

Keywords: corpus linguistics; continuous professional development; educational corpus integration; evaluation of corpus use in professional practices; corpus application to teaching and research; language teacher education; translator education; interdisciplinarity.

Resumo: Estudos sobre a aplicação da linguística de *corpus* (LC) à educação examinaram uma série de contextos diferentes – principalmente aqueles em que os alunos recebem um diploma de colação de grau (por exemplo, cursos de graduação e mestrado). No entanto, pouco se sabe a respeito da aprendizagem informal da LC, especialmente por (mas não se limitando a) acadêmicos/profissionais que não tem uma formação educacional e/ou não trabalham em áreas relacionadas aos estudos da linguagem. A presente pesquisa preenche essas lacunas, examinando a perspectiva dos participantes de um projeto de formação profissional contínua destinado a acadêmicos/profissionais de várias disciplinas, que não precisavam ter conhecimento prévio de LC. Mais especificamente, administramos um questionário a 28 participantes de um projeto de formação profissional contínua na área de LC realizado no Reino Unido com o objetivo de pesquisar quatro aspectos principais: (i) a formação educacional em LC dos participantes; (ii) suas motivações para participar desse tipo de projeto; (iii) as vantagens e barreiras de empregar a LC em suas práticas pedagógicas; e (iv) suas avaliações sobre a integração da análise de *corpus* em suas práticas de pesquisa. Os resultados apontam para o papel dos projetos de formação profissional contínua na democratização do acesso à educação em LC para profissionais tanto da área de estudos da linguagem quanto de outras áreas e no potencial aumento do interesse desses profissionais na aprendizagem de LC. A falta de conhecimento é percebida como a principal barreira para a incorporação de abordagens de corpus para o ensino e a pesquisa, reforçando assim a relevância do desenvolvimento de oportunidades de aprendizagem formal e informal para acadêmicos/profissionais em diferentes áreas.

Palavras-chave: linguística de *corpus*; formação profissional contínua; integração educacional de *corpora*; avaliação do uso de *corpora* em práticas profissionais; aplicação de *corpora* no ensino e na pesquisa; formação de professores de línguas; formação de tradutores; interdisciplinaridade.

Submitted on September 10th, 2020

Accepted on October 21th, 2020

1 Introduction

This special issue of *Revista de Estudos da Linguagem* aims to take stock of the achievements and challenges of corpus linguistics (henceforth CL) over the years. While it would be challenging to precise exactly when CL started (see VIANA; ZYNGIER; BARNBROOK, 2011), Johansson (2008) clarifies that Jan Aarts first proposed the term *corpus linguistics* in the 1980s. In this decade, we also start to observe the academic uptake of corpus studies mainly due to the popularization of personal computers. In all these past years, CL has considerably evolved and has afforded new perspectives to our understandings of language use.

Corpus approaches have been used to examine different languages and their specific uses; however, the educational impact of CL has not been explored to the same extent. Naturally, it would be factually inaccurate to claim that there is little research on this topic: previous studies have investigated the integration of corpus analysis in numerous classroom settings. These settings include different languages being taught/learned (e.g. O'SULLIVAN; CHAMBERS, 2006 on French), educational levels (e.g. FRANKENBERG-GARCIA, 2015 on Master's students), countries (e.g. TODD, 2001 on Thailand), and disciplines (e.g. HAFNER; CANDLIN, 2007 on law students).

A review of the literature, however, reveals that much of the work conducted to date focuses on language-oriented educational contexts (e.g. FARR 2008; GAN; LOW; YAAKUB, 1996; HEATHER; HELT 2012; ZAREVA 2016) and degree-awarding settings where CL is taught in a compulsory or an optional module (e.g. BUENDÍA-CASTRO; LÓPEZ-RODRÍGUEZ, 2013; FRANKENBERG-GARCIA, 2015; GALLEGO-HERNÁNDEZ, 2015b). In other words, disciplines other than language-related ones and educational programs which are not credit-bearing remain underexplored in the research literature on educational applications of CL. To address these two research gaps, the present study innovates by investigating the perspective of participants from a range of disciplines in a non-credit-bearing continuous professional development (CPD) project. More specifically, it focuses on four main aspects here: (i) the CL background of participants who are drawn to CPD opportunities like this one; (ii) their motivations to participate in it; (iii) the advantages and challenges of employing CL in their teaching practice; and (iv) their evaluation of the integration of corpus analysis in their research practice.

To this end, a questionnaire was administered to the participants of a CL CPD project in the UK. The empirical data were analyzed in a bottom-up way which combined quantitative and qualitative analytical methods.

The present paper is divided into seven sections. Following this introduction, Section 2 reviews the literature on the integration of CL in two professional fields – language teaching and translation. In Section 3, we describe the CPD project that was offered to the research participants. Section 4 presents the methodological procedures adopted in this study. In Section 5, we describe our research participants, clearly indicating how they differ from the population samples in most of the studies conducted to date. The results of our analysis are presented and discussed in Section 6 before some final remarks are made in Section 7.

2 Literature review

CL has revealed its potential to contribute to several occupations – from language-oriented ones such as lexicographers and materials developers (FLOWERDEW, 2012; O'KEEFFE; MCCARTHY, 2010) to those which do not necessarily have a direct language component such as healthcare practitioners (CRAWFORD; BROWN, 2010) and lawyers (HAFNER; CANDLIN, 2007). In the present paper, we focus our attention on the embedding CL into teacher education (especially language teacher education) and translators, the two occupations that have received most attention in the research literature. The following subsections review the available research literature on corpus embedding in the education of these two professional groups.

2.1 CL in language teacher education

Many publications have highlighted the contribution that CL can bring to the field of language teaching (for useful summaries and overviews, see BIBER; REPPEN, 2015; O'KEEFFE; MCCARTHY, 2010). However, language teachers' use of corpus approaches in their language classrooms is far from being the mainstream practice (BOULTON 2010; RÖMER, 2010). The proponent of data-driven learning, Johns (1991) writes about two challenges in the integration of CL into language teaching. One challenge relates to teachers' roles, which need to change to 'a director and coordinator of student-initiated research' (p. 3). The other challenge relates to the use of traditional

teaching materials, which may have their accuracy questioned when concordancing tools take central stage in the classroom, and actual language use is analyzed. These two challenges (see also ASTON, 2011; CONRAD, 2011; VIANA, 2011) may help to explain why teacher education programs have not extensively incorporated modules on corpus analysis (CALLIES, 2019; FARR, 2010; GRANATH, 2009; MCCARTHY, 2008). The next two subsections will be dedicated to the integration of CL in the professional development of, respectively, pre-service and in-service teachers.

2.1.1 Pre-service language teacher education

Empirical studies on the effectiveness of using corpora in professional development have underlined the potential integration of CL in teacher education programs. As argued in Breyer (2009), the dual role – as a student and a future teacher – of pre-service teachers enables them to build a strong knowledge base in corpus queries and analyses as a learner before they expand what they have learnt to their workplace.

Several studies have focused on the use of CL in the teaching of vocabulary and grammar in pre-service English language teacher education (e.g. HEATHER; HELT, 2012; FARR, 2008; ZAREVA, 2016). Gan, Low, and Yaakub (1996) contrasted corpus approaches with traditional teacher-centered pedagogy at a Malaysian university regarding the teaching of vocabulary skills. Students in a pre-service teacher education program in Teaching English as a Second Language (TESL) were divided into an experimental group, which had five two-hour sessions on computer-based concordancing exercises, and the control group, which followed the teacher-centered approach with consultation from dictionaries. Pre- and post-tests revealed the experimental group excelled in the use of words in context. The benefits of CL in grammar teaching are well observed in Farr's (2008) sample of postgraduates in English Language Teaching (ELT), Zareva's (2016) survey on trainee teachers in the field of Teaching English to Speakers of Other Languages (TESOL), and Heather and Helt's (2012) grammar course for teachers of English as a Second Language (ESL). Difficulties and problems that may constrain the implementation of corpus approaches were also identified. One of the constraining factors is student teachers' language proficiency levels. In Heather and Helt's (2012) semester-long English grammar

course for pre-service teachers, the researchers collected students' questionnaire responses, their critique of prescriptive grammar rules and their design of supplementary teaching materials with corpus approaches. The findings revealed that pre-service teachers' grammatical knowledge affects the interpretation of corpus results to students. For instance, one student teacher with weak class performance wrongly categorized the auxiliary use of *be* as the main verb.

Apart from the above studies focusing on general English, the educational application of CL has been explored in ESP settings as well (e.g. VIANA; BOCORNY; SARMENTO, 2018). Hüttner, Smit and Mehlmauer-Larcher (2009) implemented corpus tools to ESP teaching: a small and specialized English corpus was built and contrasted with general reference corpora such as the BNC. This comparison offered a handy and reliable approach for students to identify obligatory and optional moves as well as formulaic expressions. As regards the tools for ESP teaching, two of the 32 pre-service teachers in Ebranhimi and Faghih's (2017) research believed that the free corpus software AntConc was useful in ESP education. This is because AntConc enables both learners and teachers to build a specialized corpus and generate a keyword list. However, student teachers also face challenges in CL-informed ESP classes. In Leńko-Szymańska's (2017) semester-long CL course, she collected students' end-of-semester assignments (e.g. self-compiled ESP corpora and corpus-based lesson plans) and argued that pre-service teachers only mastered basic CL technical skills at the lexical and phraseological levels, leaving other language features (e.g. register differences) barely untouched. Her study demonstrates that a semester-long course is not sufficient for pre-service teachers, who may lack the confidence and expertise to design CL-based activities for their own students in the future.

2.1.2 In-service language teacher education

Previous research on the interface between CL and in-service language teacher education has examined current professionals' use of corpora through questionnaires. Mukherjee (2004) investigated the actual use of CL in German secondary-school language teaching practice. His research results revealed that English language teaching had been hardly influenced by language features attested in corpora. Also drawing on the German context, Römer's (2009) survey of 78 secondary-school

English language teachers uncovered some of their desires and problems that could be well addressed by CL such as better teaching materials and reference resources other than non-corpus-based descriptions. More recently, Chen *et al.* (2019) analyzed the questionnaire responses provided by 54 in-service teacher participants of a data-driven learning workshop in Hong Kong. Among other findings, the results revealed correlations between teachers' prior knowledge of CL and their evaluation of the difficulty of corpus tools and between teachers' motivation for professional development and their adoption of data-driven learning. The types of investigation reviewed so far in this section are similar to the one that we have conducted in that we also adopted a questionnaire as our research instrument (cf. Section 4); however, our population sample is distinct as will be discussed in Section 5.

Another area that has been explored in the interface between CL and in-service teacher education is current professionals' use of corpora and corpus resources. One example of these resources is the Teachers of English Education Nexus (TeleNex), a website that is aimed at supporting the work of primary and secondary English language teachers in Hong Kong. Based on the analysis of 1,294 teacher-generated questions over eight years, Tsui (2005) advocates that schoolteachers' use of corpora is an effective way to help them understand the meaning and usage of a linguistic item. Corpus data were preferred over dictionary definitions, especially regarding queries on synonyms (such as *finally* vs. *lastly*) and stylistic patterns that seem to go against traditional prescriptive rules (e.g. whether sentences can start with conjunctions like *because*, *but* and *and*).

CL research has similarly investigated teachers' language use (e.g. CHAMBERS; O'RIORDAN, 2007; FARR, 2006). For instance, Vaughan (2007) examined how teachers use jargons and humor to maintain their membership identity in teacher-teacher talks at staff meetings. Farr (2006) researched trainer-trainee interactions in language teaching to uncover some of the linguistic and communicative features of this type of professional discourse such as self-disclosure strategies (e.g. 'all of us would...') and the high-frequency use of compliments (e.g. 'good') (see also FARR, 2005, 2010). Drawing on a US context, Reppen and Vásquez's (2007) and Vásquez and Reppen's (2007) action research over two semesters explored the spoken interaction between four pairs of teachers and their respective supervisors in post-observation meetings. Results from the first semester highlighted supervisors' higher talking time

in the collected data and prompted a change in supervisory practice. More specifically, supervisors created questions to give teachers the opportunities to ponder and generate more talk. The comparison of data in Semesters 1 and 2 revealed a change in the post-observation meetings: there was an increase in the total number of words produced by teachers whereas supervisors' amount of talk decreased or remained approximately the same. Reppen and Vásquez's (2007) and Vásquez and Reppen's (2007) studies illustrate the key role that corpus research can play in the reconsideration of professional practices. In their case, corpus findings triggered a change in supervisors' and teachers' roles in post-observation meetings, encouraging the teachers to take center stage in these meetings.

2.2 CL in translator education

Translation is another area that CL has influenced over the years potentially due to the facilitating role that corpora can have in translation tasks. For example, parallel corpora can facilitate the search for translation correspondences; and corpora of trainee translators and monolingual target-language corpora can be examined to identify and evaluate trainee professionals' translation solutions. The following sections will review empirical studies on the integration of CL in, respectively, pre- and in-service translator education.

2.2.1 Pre-service translator education

Previous studies on pre-service translator education have examined a few different contexts. From a geographical perspective, these studies have taken place, for example, in Denmark (e.g. LAURSEN; PELLÓN, 2012), Germany (e.g. KRÜGER, 2012), Spain (e.g. GALLEGO-HERNÁNDEZ, 2015b; MONZÓ-NEBOT, 2008; RODRÍGUEZ-INÉS, 2009), and the UK (e.g. FRANKENBERG-GARCIA, 2015). From an educational perspective, these studies have primarily examined degree-awarding courses – either at the undergraduate (e.g. BUENDÍA-CASTRO; LÓPEZ-RODRÍGUEZ, 2013; GALLEGO-HERNÁNDEZ, 2015b; ZANETTIN, 2001) or postgraduate level (e.g. FRANKENBERG-GARCIA, 2015).

The benefits of corpus introduction in pre-service translator education have been well argued and advocated. Working with a specialized monolingual first-language corpus, Bowker (1998), for

example, conducted a pilot study in a French-to-English translation classroom and showed that corpus analysis helped trainee translators understand the translation subject, terminology and idiomatic expressions. Zanettin (2001) drew on undergraduates' Italian-to-English translation of a newspaper text on the Olympic Games to illustrate how corpus exploitation can enable student translators to contrast source and target languages and to facilitate the selection of translation correspondences. Similar benefits have been reported in Rodríguez-Inés's (2009) Spanish-English translation class for 26 final-year Spanish students and Monzó-Nebot's (2008) legal translation courses for third- and fourth-year undergraduates in Spain.

The use of the web as a corpus (e.g. GATTO, 2014) in pre-service translator education has also been examined, and its advantages have been identified. The MA Specialized Translation students in Krüger's (2012) research believe that web concordances such as WebCorp can provide immediate solutions to language-related doubts that do-it-yourself (DIY) corpora (i.e. ad-hoc self-compiled corpora) may fail to solve. In another web-as-corpus study, Buendía-Castro and López-Rodríguez (2013) conducted an experiment with third-year undergraduate students in Translation and Interpreting at a university in Spain. These students were tasked with the translation to English of a research article excerpt on swine flu originally written in Spanish. It showed that the use of automatically built specialized corpora compensate for pre-service translators' lack of discipline-specific knowledge.

In addition to advantages, the literature on corpus integration in pre-service translator education has identified challenges. For example, Rodríguez-Inés (2010) highlighted the amount of time required in learning about corpora and their use. She pointed out that this may result in a time loss for pre-service translators to develop their translation skills and competence in a strict sense. In Gallego-Hernández's (2015b) study, the pre-service translator participants were split in their evaluation of the difficulty (N=14) or easiness (N=11) of corpus methods, and they indicated that time was a factor in their engagement with CL. Mixed feelings towards corpus work are also observed in Frankenberg-Garcia's (2015) study on 13 MA students in Translation at a UK university. While students appreciated exploiting corpora to assist them with the handling of unfamiliar terminologies and idiolects, they had some trouble in choosing appropriate corpora or corpus tools.

2.2.2 In-service translator education

The central conundrum in the integration of CL in in-service translator education is akin to the one observed in language teacher education (cf. Section 2.1). Although translators have begun to become aware that corpus approaches may support their day-to-day professional practice (e.g. VARELA-VILA, 2009), the uptake of these approaches is still relatively reduced (see, for instance, BOWKER, 2004; FRÉROT 2016; GALLEGO-HERNÁNDEZ, 2015a; JÄÄSKELÄINEN; MAURANEN, 2006; MELLANGE, 2006). Similar to pre-service translator education (cf. Section 2.2.1), time appears as one of the major barriers in translators' corpus uptake (ASTON, 2009; WILKINSON, 2006). This barrier is often noticed if translators have to compile their corpora (GALLEGO-HERNÁNDEZ, 2015a). It seems that translators would be more open to corpus approaches if the translation task involves a very large or interdisciplinary text, or if the translators themselves work full time.

A vicious circle can be identified in the integration of CL and professional translation. Perhaps because corpus uptake is not widespread, corpus skills are not mentioned in person specifications for translation posts as observed in Bowker's (2004) investigation of Canadian job adverts. At the same time, the lack of recognition of corpus skills as a sought-after ability in job ads does not encourage professional translators to acquire these skills (see also FRÉROT 2016). For example, Jääskeläinen and Mauranen's (2006) survey into Finnish timber industry found that corpora and concordance tools were not widely used, especially among freelance translators.

There are, however, some positive prospects in the integration of CL in in-service translator education. The survey of 1,015 translators and interpreters around the world in the Multilingual eLearning in LANGUAGE Engineering Project (MELLANGE, 2006) revealed some promising aspects: 20.2% of the respondents had used concordancers, and 82.0% would be interested to learn more about corpus-based translation skills. Gallego-Hernández's (2015a) survey equally indicates promising results: nearly half of the 526 participants indicated that they engaged in corpus exploration in their translation practice at a frequency that varied from 'sometimes' to 'very often'.

The present literature review indicates that the potential CL has to offer to professional education has not been fulfilled yet. In other words,

corpus approaches do not seem to have entered mainstream education or professional education courses. In the present study, we examine the perspectives of academics/professionals from different backgrounds on the integration of CL in their practices. The following section describes the CPD project that was offered to these participants.

3 CPD project on CL

This research investigated the perspectives of participants in a blended CPD project on CL funded by the British Academy. Merging research, teaching and learning perspectives, the project aimed at showing participants how to develop their CL skills and their students'/supervisees'. The face-to-face element consisted of three day-long events spread over one year (i.e. June, September and December) with sessions delivered by experts in the field (e.g. Marina Bondi, Paul Thompson, Ute Römer). While Chen *et al.*'s (2019) research is also on a non-credit-bearing CL workshop, their target participants were limited in professional terms (i.e. it was aimed at English language teachers) and the length of their sessions was shorter (i.e. two three-hour long workshops, totaling six contact hours). The online space in our CPD project provided a further means for interaction among participants and for their learning to be consolidated over time with asynchronous input from the same team of experts.

This CPD project did not assume any prior knowledge of CL. The face-to-face and online activities were planned in such a way that participants would be introduced to the main concepts in CL before putting these concepts into practice in hands-on sessions and exploring the application of CL to their teaching and research.

The three face-to-face events had different but complementary foci. Participants were first introduced to the basics of language education and CL before they had two full days examining how this could be applied to language in general and to language for academic purposes. A decision was made to focus on English since this was the only shared language among all attendees, but the transferable nature of corpus skills was stressed.

4 Methods

We decided to use a questionnaire to collect data for the present study. Despite its inherent limitations (e.g. the potentially thin data to be collected), a questionnaire was the most appropriate option for our data collection plans. It required a reduced time commitment from participants, thus potentially increasing the final volunteer sample. This can be seen in our response rate, which will be discussed in Section 5.

In addition to the cover sheet, the questionnaire consisted of 25 questions divided into three parts. The first one contained questions about personal matters (e.g. sex, age, home country), participants' work experience, their educational background, and language knowledge and proficiency. Part 2 contained questions on participants' prior knowledge of CL as well as of related matters such as discourse analysis and statistics. Part 3 was dedicated to participants' reasons for registering for the CPD project, their expectations of it, and their appraisals of CL application to teaching and research.

At the beginning of the first event, participants were invited to complete the questionnaire anonymously. From an ethical perspective, our decision to ask participants to answer the questionnaire in the first face-to-face event could be challenged. While this is not unusual (e.g. FRANKENBERG-GARCIA, 2015; GAN; LOW; YAAKUB, 1996), we thoroughly considered whether the questionnaire should be answered online before the event or in person at the first event. We opted for the latter option because of two main reasons. Firstly, our target participants were primarily academics and/or professionals, who would probably struggle to find the time to answer the questionnaire before the event. Secondly, we felt it was essential for us to get to know the participants and to introduce the project to them in person before making any requests.

We were, however, aware that our request to answer the questionnaire in the first face-to-face session could be seen as a potential imposition by our participants, which would limit their perceived scope for declining to do so. This potential imposition is lower than in previous studies involving students where the researcher is also the teacher in charge of assessing the student participants (e.g. FRANKENBERG-GARCIA, 2015). Our relationship with the participants did not take place in any formal educational context where they would be evaluated for a credit-bearing module, for example. This was an optional CPD project

for which the participants had decided to register and to which they had already been accepted.

We followed four main steps in order to reassure participants of their freedom to decide whether or not to answer the questionnaire.

1. We explained to them the voluntary nature of their participation, and they had an opportunity to ask any questions before the questionnaire was distributed. They could naturally ask further questions at any point in time as well.
2. Participants completed the questionnaires anonymously. While we asked for some background information, it does not allow us to identify them – nor is it important to our research either.
3. Both researchers kept a physical distance from the participants during questionnaire completion so that they did not feel coerced to complete it. We would only approach specific participants if they called us to clarify any questions that they had.
4. The questionnaires were returned anonymously: we asked the participants to put their questionnaires in a manila envelope, which was placed at the back of the room. The envelope was only opened at the end of the first day after the participants had already left the venue. As no other writing sample was collected from participants throughout the CPD project, they were reassured that their identities were never disclosed to us. This procedure means that, once the participants returned their questionnaire, they could not withdraw from the research anymore. However, we felt that this was a fair compromise to ensure their anonymity, which we believed to be of higher importance in this research.

Following the completion of the paper questionnaires, participants' responses were digitized verbatim so that we could investigate the data electronically independently. The data collected through closed questions were analyzed quantitatively while the participants' answers to open-ended questions were analyzed qualitatively and quantitatively. Our approach to open-ended answers meant that they were initially studied in a bottom-up manner to identify themes, which were then quantified based on the number of occurrences.

Both of us were involved in the analysis. The first author analyzed all the data independently initially. He then shared the results with the

second author, who compared the results with her original analyses. She checked the quantitative results for accuracy and the qualitative results for thoroughness. There were only minor discrepancies in the qualitative analyses, which were resolved by discussing each of the relevant cases. Before the results are presented in Section 6, the following section will detail the participant sample in the present study.

5 Participants

A total of 36 registered participants were expected to attend the face-to-face events. Out of this total, three had expressed their impossibility in attending the first event, three were speakers who had to either arrive late or leave early, and two were the CPD project organizers, who are also the authors of this paper. This resulted in a pool of 28 potential participants, all of whom agreed to contribute to the study and answer the questionnaire. While the sample may be considered small, we worked within a non-interventionist research paradigm with the participants of a specific, real-life educational CPD project. As reviewed in Section 2, other pedagogical studies have researched a similar or even smaller number of participants. For example, Farr (2008) examined a sample of 25 MA student teachers in her questionnaire-based evaluation on participants' perception of corpus-assisted courses; Frankenberg-Garcia's (2015) study drew on the data provided by 13 Master's students in Translation at a UK university; Zareva (2016) analyzed 21 trainee teachers' responses to a questionnaire aimed at evaluating a corpus-based course design.

Our study had a 100% return rate, which is high for non-course/degree-based questionnaire studies. In Römer's (2009) research with in-service teachers, for instance, 78 out of 120 questionnaires were completed and returned. However, the difference in the overall population sample must be acknowledged. Our decision to request participants to complete the questionnaire in the first face-to-face event after we had initially established rapport with the participants (cf. Section 4) may have contributed to this high return rate.

5.1 Personal characteristics

Our participant sample is varied. We had a total of 16 female participants and 12 male ones, which is a somewhat even split. This differs

from many previous studies in which females considerably outnumber males (cf. 16 female vs. 5 male student teachers in ZAREVA, 2016) or the distribution of sex is not disclosed (cf. CHEN *et al.*, 2019; HEATHER; HELT, 2012; LEŃKO-SZYMAŃSKA, 2014).

Participants' ages vary from 23 to 55 with the mean being 41 years old. The age in our sample is older than in previous studies: participants' mean age in Zareva (2016) is 35.4 years old, and the participants' ages in Vásquez and Reppen (2007) range from the mid-20s to mid-30s. This difference is not surprising given the CPD project (cf. Section 3) and the primary occupation of the target participants (see Section 5.3).

5.2 Countries of residence/origin and language knowledge

Participants were asked about their countries of residence and origin. The answer to the former question indicates that most of them (N=26) lived in the UK (i.e. two participants decided not to answer this question) at the time of data collection. This is understandable and unsurprising given that the project entailed three face-to-face events in this country in one calendar year (i.e. June, September and December). This set-up would make it difficult for overseas participants to join the on-site events. In relation to participants' home countries, while we observe that most respondents are from the UK (N=20), there is more diversity with two participants from China and one participant from each of the following countries: Canada, France, Germany, India, Malaysia and Russia.

The range of nationalities described above helps to explain participants' language knowledge. English is the first language of most participants (N=15). Participants also reported having German (N=2), Latin (N=1), and Mandarin (N=2) as their first language. Altogether, nine participants decided not to answer this question.¹ English language command was not an issue: most participants (N=22) self-assessed themselves as proficient in the Common European Framework of Reference for Languages (CEFR), that is, either at C1 or C2 level. Only four participants declared to be at independent user levels (B1 and B2), and two decided not to answer this question. In terms of other languages, participants indicated that they knew 16 other languages to varying

¹ The total of 29 first languages is due to one participant's reporting to speak both German and Latin as first languages.

degrees of proficiency. The most recurrent additional languages were French (N=11), German (N=7), Japanese (N=3) and Russian (N=3). There were also mentions to Danish, Dutch, Hindi, Italian, Malay, Mandarin, Spanish, Telugu, Thai and Turkish to cite some examples.

5.3 Educational and professional background

Our project targeted a specific group of academics/professionals. All of the participants held at least a first degree and either had or were working towards higher degrees. One participant had a Diploma, 17 were educated to Master's level or were reading for one, and 10 were doctoral degree holders or were taking such a course. Most of these participants were affiliated with a higher education institution (N=27), encompassing both universities (in most cases) and colleges. Considerably smaller numbers worked at other educational institutions – e.g. schools and local authorities (N=4), and publishers (N=2).²

With regard to their occupation, most participants were based in an educational environment: language teachers (N=13), university lecturers (N=10), students (N=9). There were two other professions represented in the sample (i.e. a publisher and a corpus developer), and one participant decided not to answer this question.

The educational and professional profile outlined above coheres with our participants' ages. As their average age is 41 years old and as most participants are in their 40s and 50s (N=17), they have higher educational degrees (i.e. at postgraduate level) and have considerable work experience in their respective fields. The profile is also aligned with the target participant group for the CPD project, namely, academics and/or professionals.

5.4 Distinctive features of our population sample

Our participant sample stands out from the samples in previous studies due to our focus on a CPD project. Our participants held or were pursuing higher degrees – generally Master's or a doctorate. While there are studies with participants at postgraduate levels, they are generally with

² The total in this case is higher than the overall number of participants (N=28) because some of them declared more than one affiliation. The same is the case for the participants' reported occupations.

student cohorts (e.g. FARR, 2008; FRANKENBERG-GARCIA, 2015; KRÜGER, 2012), who learn about CL as part of their degrees – either on a compulsory or on a voluntary basis. Our CPD project differed from these courses in that it did not lead nor contribute to any educational degree. Participants' decision to register for this project was probably not because they may have felt it was compulsory to do so nor because they would be awarded a certificate at the end of it (see also CHEN *et al.*, 2019). Instead, as this was a voluntary CPD project, they were potentially intrinsically motivated to do so. After all, they had to make several commitments in relation to, for example, time (e.g. travelling to the venue and attending the three full-day events) and money (e.g. paying for their travel expenses).

In relation to their occupation, most participants worked in educational environments – be they language teachers or university lecturers. While nearly one-third of our participants were students (N=9), only five of them were exclusively students. The other four were either students who worked as teachers (N=3) or a student who was a lecturer (N=1). There have been studies conducted with language teachers, especially schoolteachers (cf. MUKHERJEE, 2004; RÖMER, 2009; TSUI, 2005). However, there seems to be a research gap concerning university lecturers – a notable exception is Chen *et al.*'s (2019) study.

Another stark difference between our study and the available literature has to do with our participants' most recent teaching experience. Understandably, most of them either teach English language (N=15) or work with English language teacher education (N=5). While participants similar to both groups have already been investigated in other educational and national contexts (e.g. CHEN *et al.*, 2019; FARR, 2008; LENKO-SZYMAŃSKA, 2014; RÖMER, 2009; TSUI, 2005), we have a more comprehensive range of disciplines being represented in our sample, which includes Applied Linguistics, Dementia, Education, History, Japanese, Russian and Sociology. One participant represented each of these disciplines except for Sociology, which was taught by two participants. The thinly spread disciplinary representation in this study does not allow us to make any specific points about them individually. However, the participant sample as a whole helps us to advance our current knowledge and understanding of the appeal of CL beyond the exclusively language-related disciplines, and it opens up an exciting new area of exploration in order to help us deepen the impact of CL across disciplines.

6 Results

The results of our study are presented and discussed in the following subsections. These subsections focus on four topics: (i) participants' background knowledge of CL before the start of the CPD project, (ii) their motivation to join this CPD project, (iii) their pre-project appraisal of the actual or potential application of CL to their teaching, and (iv) the same appraisal in relation to their research practice.

6.1 Previous education on CL

The publicity materials for the CPD project clearly stated that no prior knowledge of CL would be assumed or required from the participants. Instead, everyone with a keen interest in learning about corpus applications to education was welcomed and encouraged to apply. We were interested in finding out whether the call had a circular effect by appealing just to those who already had some knowledge of CL or whether it had been successful in drawing the attention of colleagues who had no or little knowledge of this field.

The findings reveal almost a split in participants' educational background on CL: 15 had studied it while 13 had never done so. Nearly half of the respondents who had previously studied CL (N=7) indicated that they had undertaken a CL module as part of their Master's in TESOL (N=5), Applied Linguistics (N=1) and Linguistics (N=1). Two participants developed their CL knowledge during their PhD since they employed corpus methods in their thesis research. Apart from one participant who learned about CL in his/her Diploma course, all the remaining eight indicated that they learned about CL through routes which did not lead to the award of a formal degree. These routes include a massive open online course, a workshop, and the informal and voluntary auditing of CL modules.

The 13 participants who indicated having had no prior study of CL before the CPD project were asked to explain why this was the case. Five participants referred to their lack of opportunities to do so.

1. “I have not had the opportunity up until now”³ [F; 23; S; PhD (Social Policy)]⁴
2. “I’ve got limited chance to learn.” [F; 27; S; Master’s (TESOL)]

As these two examples show, this lack of CL learning opportunities is not confined to non-language-related areas such as Social Policy (cf. Example 1) and History, but it is also observable in the previous experience of participants who specialize in areas like TESOL (cf. Example 2) and Translation (see ASTON, 2009 on the major barriers of CL among translators). Another explanation for the lack of formal CL education was participants’ lack of interest in it. This explanation was given only by participants with a language-oriented educational background (i.e. Linguistics and TESOL): they acknowledged that they could have learned about CL during their formal studies, but they decided not to pursue this option. The other reasons mentioned by individual respondents referred to CL being claimed to be underdeveloped in a participant’s field of research (i.e. English for Academic Purposes), a Social Work Lecturer’s lack of CL awareness and long-standing focus on qualitative methods, and a participant’s existing working knowledge on the use of corpora.

All of these responses provide useful pointers to help us pave the way for the future educational CL expansion. It seems vital for us to provide students with the opportunity to learn about CL as part of their degrees and/or in CPD projects like the one reported here. The introduction of CL in formal and informal educational programs is not a new recommendation: Renouf, back in 1997, argued in favor of introducing CL to postgraduate students of applied linguistics in the UK; Römer (2009) further recommended universities to reach out to teachers on problems and needs directly related to teaching through lectures and workshops. However, these recommendations do seem to have been

³ Participants’ responses have not been edited, and they are here reproduced verbatim.

⁴ The code adopted in this study consists of four parts: the first letter indicates sex (F=female, M=male); the following numbers reveal participants’ ages; the subsequent letter(s) stand for participants’ occupations (D=Developer, L=Lecturer, P=Publisher, S=Student, T=Teacher, ?=no answer); and the final part indicates participants’ highest degree (either completed or in progress) and its corresponding field. Therefore, “F; 23; S; PhD (Social Policy)” refers to a 23-year-old female student participant who either holds or is studying towards a PhD in Social Policy.

fully implemented. If, for the sake of illustration, we focus on TESOL Master's in the UK (cf. PAPAGEORGIU *et al.*, 2017), only 35 out of 141 programs offer CL as a standalone module. These 35 programs are found in 17 universities (15 in England and 2 in Scotland), and the CL modules are nearly all optional with only two exceptions (VIANA, 2017). The provision of CL modules in the UK has indeed increased over the years, and this can be seen in the findings from the present study where most participants who already knew about CL had taken a specific module in their Master's. However, there is still scope for further improvement.

Not only should students be provided with opportunities to learn CL, their awareness and interest in it should also be raised. If we return our attention to TESOL Master's in the UK (cf. COPLAND *et al.*, 2017), we will see that the finding reported here is not a one-off occurrence. TESOL Master's students were asked to appraise 15 modules on a 6-point Likert scale, ranging from 'not at all important' to 'extremely important'. The results show that CL had a mean of 4.58 with a standard deviation of 1.11 (VIANA, 2017). While this seems a somewhat encouraging result since it is above the 3.5 threshold, the mean for CL is the second last when all the 15 modules are considered (VIANA, 2017).⁵

The provision of more CL opportunities and the increase in their uptake are two related action points. The former is perhaps easier to achieve since it depends primarily on teaching staff to change the curriculum. The latter, however, will take more time because it will possibly require an attitudinal change among future generations.

6.2 Motivational drivers for CPD project participation

In order to help foster engagement with future CPD projects on CL, we must understand the participants' motivational drivers. Table 1 indicates participants' reasons for registering for the focal CPD project.

⁵ The module perceived as least important by the participants in Copland *et al.*'s (2017) study is Translation with a mean of 3.83 and a standard deviation of 1.57.

TABLE 1 – Participants’ reasons for registering for the CPD project on CL

Category	Frequency	Example
Application of CL	16	3. “With a background in corpus use for a very special purpose (machine translation & lexicography), I wanted to learn how corpora can be used for language teaching.” [F; 26; D; Master’s (Artificial intelligence)]
Research development	11	4. “Learn new research methods in stylometrics using burrows delta and nearest neighbour drivers” [M; 54; L; PhD (History)]
Knowledge of CL	9	5. “To get a better understanding of the software use in analysing corpora To get a handle on statistics” [F; 49; S&T; EdD (Education)]
Personal interest	6	6. “I already had an interest but hadn’t studied/read much so this project seemed to be something that could develop my interest and help me use corpora practically.” [F; 55; T&L; Master’s (Chinese)]
Financial reasons	1	7. “To Save money!” [F; 54; L; PhD (Educational Linguistics)]

Participants were primarily drawn to this CPD project as a way of developing their knowledge and understanding of CL applications. Most of the answers refer to educational applications of CL, which is understandable given that this was the main project focus. Although this motivation was observed in previous research (e.g. HEATHER; HELT, 2012; ZAREVA, 2016), participants’ educational application needs were varied in this study. In Example 3, the participant’s focus lies on the application of her previous knowledge as a corpus developer to the creation of outputs of relevance to language teachers. The applications were not exclusively related to education: there was also a reference to the application of the knowledge acquired in the UK to one participant’s home country, for example.

Participants’ wish to develop as researchers was the second most frequent reason. This finding should be interpreted alongside the unique population sample for this study (cf. Section 5): several of the participants were experienced professionals who had been working for a considerable number of years and/or who held academic positions. While research-related motivations can be found in the literature (e.g. ZAREVA, 2016), the studies are generally confined to language-related disciplines. Here, we notice that, in addition to the language-related

connections, our interdisciplinary participant sample establishes links to other fields. These areas include the investigation of authorship of historical materials (cf. Example 4) and reports in Accounting. There is an ample area for future exploration of CL: not only do we need to sediment the relationship between CL and language-related areas such as language teaching and translation, but we also need to capitalize on the impact that corpus studies may have outside our field of research. While we cannot claim how widespread the interest in CL across disciplines is or predict if this interest will eventually translate into real applications, our findings suggest that some colleagues from other areas are already (at least initially) willing to learn about CL. This means that they would not need to be convinced to do so and that they are open to this learning, which is an inspiring starting point.

Learning about CL was identified as a motivational driver by nearly one-third of the participants. In Example 5, the participant specifies CL-related matters such as statistical knowledge. This is not always the case, however. Sometimes participants provide a rather general indication of what they would like to learn by just referring to CL as a whole.

As expected, participants' intrinsic motivation played an essential role in their decision to register for the event (see also HEATHER; HELT, 2012; LENKO-SZYMAŃSKA, 2014). Example 6 indicates the mutual, two-way relationship between the participant's interest and her project participation. At the same time that the project was a way for this participant to undertake an activity that she was already willing to engage in, her participation also helped to foster her interest in the field.

A final reason mentioned by a single participant had to do with finances. There is not much contextual information to help the interpretation of this reason, but it could be assumed that this was possibly an allusion to the free-of-charge nature of the CPD project to the participants since the British Academy had fully sponsored it.

To further our understanding of the motivational drivers, participants were asked to indicate their three main expectations for the CPD project. This allowed us to check the extent to which their expectations matched their reasons for joining the project and whether any other relevant aspect had not been captured in the previously reported open-ended question. Their expectations were thematically analyzed, and the final categories are presented in Table 2.

TABLE 2 – Participants’ expectations for the CPD project on CL

Category	Frequency	Example
Knowledge enhancement	31	8. “understanding technical aspects of building corpora” [F; 55; L; PhD (Linguistics)]
Application of CL	21	9. “to develop a better understanding of how learners can use corpora to improve their own linguistic competencies” [M; 35; T; Master’s (International Relations)]
Research development	16	10. “develop ideas in how can incorporate corpus linguistics into an application for a research project” [F; 43; L; PhD (Social Work)]
Networking	6	11. “opportunity to meet new contacts” [F; 46; P; Master’s (Languages)]
Practical skills improvement	6	12. “Learn hands-on skills” [F; 54; L; PhD (Educational Linguistics)]
Motivation increment	2	13. “enthuse me into world of corpora which I find quite dry just now” [F; 42; T; Master’s (Applied Linguistics)]
CV	1	14. “CV building” [M; 25; ?; Master’s (History)]

Participants’ top three expectations were coherent with their reasons for having registered for the project – the only difference is the order in which they appear. ‘Knowledge enhancement’ features as the most frequent expectation while ‘knowledge of CL’ appeared as the third most frequent reason in Table 1. The former is wider encompassing than the latter; however, most of the expectations included in this category (N=24) referred to CL in general or specific CL matters like corpus compilation (cf. Example 8). The few remaining expectations included in this category (N=7) dealt with learning about quantitative methods and/or language use.

‘Application of CL’ and ‘research development’ were both main reasons to participate in the project and top expectations for it (cf. TABLES 1 and 2). Most participants’ answers grouped in ‘application of CL’ refer to Education (cf. Example 9), which was the focus of the CPD project. Although research development has not been identified as a motivational driver in previous studies (see Section 2), it is one of the major driving forces to attend this project. Another interesting point is that, in our study, research development is not restricted to the usual language-oriented knowledge areas: it also concerns the other knowledge areas represented in the study as indicated in Example 10 from a Lecturer in Social Work.

The category of ‘motivation increment’ in Table 2 could be linked to ‘personal interest’ in Table 1. Despite the difference in their foci, ‘motivation increment’ encompassed examples where the participants indicated that the project could help to increase their interest in CL. Participants’ reduced motivation level is more explicitly conveyed in the ‘motivation increment’ category as is evident in Example 13: the teacher participant expresses her lack of excitement with corpus work.

The list in Table 2 contains three new expectations that had not been mentioned in the previous analysis. The need to be in contact with like-minded CL researchers and practitioners was a top expectation to six participants. It is important to note that this expectation came from participants primarily in language-related fields (i.e. Educational Linguistics, Languages, Linguistics, TESOL) where there are notably more corpus experts and where it is easier to establish such networks. A few initiatives on this front can be seen in different parts in the UK such as ‘Corpus Linguistics in the South’ and ‘Corpus Linguistics in Scotland’.

The category of ‘practical skills improvement’ is linked to the hands-on nature of CL. This can be interpreted in relation to Fligelstone’s (1993, p. 98) well-known taxonomy of corpus-related activities: “teaching about”, “teaching to exploit” and “exploiting to teach”. Participants have shown their willingness to develop their knowledge and understanding of these three categories. They want primarily to be taught about CL (cf. ‘knowledge enhancement’) and to acquire or sharpen their skills in relation to exploiting to teach (cf. ‘application of CL’). Some of them expect the workshop to teach them to exploit corpora (cf. ‘practical skills improvement’).

Finding the right balance among these three categories must be considered in the planning of pedagogical projects on CL. The data do not show any difference between those who had previously studied CL and those who had not had that experience. This means that the two groups want to learn about CL in the first place and they are also similarly interested in its applications. Learning how to do corpus analysis seems less of a priority for both groups. This may be because they need to understand whether acquiring these skills is a worthwhile investment of their time in the first place. An alternative explanation might be the participants’ belief that they can develop their practical skills at a later stage perhaps in a more independent way. The results indicate that, for the type of target participants that we had envisaged (see Section 5), general aspects of CL should be prioritized over corpus practicalities.

The last new category observed in the project expectations relates to a concern with CV building, which was mentioned by a single participant. While pragmatic reasons such as this one are found in research projects examining student perspectives on their education (e.g. COPLAND *et al.*, 2017), this was not relevant in this study. We believe that this is probably because of the population sample consisting primarily of academics/professionals working in the UK and of the non-credit-bearing aspect of the CPD project.

The findings indicate a roadmap for future pedagogical CPD projects to upskill academics'/professionals' knowledge and understanding of CL. A focus on both the core theoretical and research content as well as on CL applications seems to be necessary to meet participants' interest and to fulfill their expectations. This way, the CPD projects will act in a two-way relationship: appealing to participants' interest and raising their motivation. Although they do not seem to be essential, providing ways to develop participants' practical skills and networks should be given some consideration as well.

6.3 Appraisal of corpus applications to teaching

Participants were asked to identify the advantages and barriers of applying CL to their teaching practice. In both cases, approximately one-third of the participants (N=8 for the question on advantages and N=9 for the one on barriers) decided not to answer these questions, and they have not been included in the results reported in this section. The blank responses could be interpreted in several ways: the questions were open-ended for which response rates are usually lower than closed questions; participants were provided with the most generous space for their answers in this part of the questionnaire, thus suggesting that these answers would potentially be the longest ones; and/or the participants might be less motivated to complete this question since it was the penultimate one in the research instrument. In addition to these blank responses, there were a couple where the participants declared not to be applicable to them. For instance, one of the participants was a corpus developer and did not have any teaching experience. These answers were not included in the final analysis either.

Table 3 summarizes the categories created after a bottom-up analysis of the empirical data.

TABLE 3 – Advantages of applying CL to teaching practice

Category	Frequency	Example
Language use	7	15. “get students to be exposed to the real language” [M; 29; S; Master’s (TESOL)]
Pedagogical improvement	7	16. “I have done some level of autonomy over ‘how’ I structure my teaching, and CL is going to make me a better practitioner” [M; 45; T; Master’s (TESOL)]
Student autonomy	4	17. “students can use their own language as the basis for corpus searches” [M; 35; T; Master’s (International Relations)]
Big data	1	18. “A way of analysing large amount of data” [F; 54; L; PhD (Educational Linguistics)]

The advantages can be related to either corpus or educational matters. Answers included in ‘language use’ or ‘big data’ reiterate points that are usually associated with corpus work. This is especially the case in relation to the first category, which is one of the main advantages of corpus investigations (e.g. SINCLAIR, 1991; TOGNINI-BONELLI, 2001). The association of CL with the investigation of large textual datasets is not uncommon (e.g. BOWKER; PEARSON, 2002; CONRAD, 2002), but corpus work is not restricted to them. The exploration of small, specialized corpora is also relevant in CL (e.g. FLOWERDEW, 2004; GHADESSY; HENRY; ROSEBERRY, 2001).

The two remaining categories establish a link between CL and Education. Participants believe that the exploration of corpora in their pedagogical practice will contribute to their professional development (cf. Example 16), their materials design skills, their skills in increasing student motivation, and their enhanced language explanations, to cite just some examples. Those advantages are observed in CL-informed classrooms such as the ones investigated by Heather and Helt (2012), Leńko-Szymańska (2014) and Zareva (2016). Participants also foresee a link between corpus work and student autonomy, a point that is recurrent in the literature (e.g. ASTON, 2011; CHARLES, 2014; GAVIOLI, 2009). As Example 17 indicates, the reference to student autonomy development is not restricted to those from a language-oriented educational background; it also made by participants with degrees in other fields.

Our analysis of the barriers of applying CL to teaching practice reveals that the most frequent issue faced by the participants is their lack

of relevant knowledge. This may relate to research (cf. Example 19), CL, IT skills and/or hands-on practice (see HEATHER; HELT, 2012; ZAREVA, 2016 for similar difficulties). This barrier is coherent with participants’ reasons for enrolling in the project and their expectations of it: knowledge, research and practical skills development featured as important factors (cf. Section 6.2).

TABLE 4 – Barriers of applying CL to teaching practice

Category	Frequency	Example
Lack of knowledge	6	19. “Knowledge of the research process. I’m also a luddite so not very [unclear word] with computers.” [M; 39; S&T; Master’s (TESOL)]
Resources	4	20. “too much resources to select and summarise” [F; 27; S; Master’s (TESOL)]
Time	4	21. “needing to spend a lot of time explaining what corpus linguistics is to my students before being able to use it in my teaching” [F; 55; L; PhD (Linguistics)]
Student-related issues	3	22. “Can students understand this or will they be interested?” [F; 42; T; Master’s (Applied Linguistics)]
Lack of support from colleagues	2	23. “I am the only person in my workplace who engages with CL, and the general attitude is one of skepticism toward it” [M; 45; T; Master’s (TESOL)]
Teaching-related challenges	2	24. “the staging of the lesson plan should be carefully prepared” [M; 29; S; Master’s (TESOL)]

Two other barriers – ‘student-related issues’ and ‘teaching-related challenges’ – could be linked to participants’ motivational drivers for participating in the project. As discussed in Section 6.2, learning about the corpus application (especially to education) was among the top factors. Interestingly, one participant questions students’ ability to understand CL or to be interested in it (cf. Example 22). While it might be more challenging to address the question about students’ interest, there is plenty of evidence in the literature that students are able to profit from data-driven learning (e.g. BOULTON, 2012; CHARLES, 2014; TODD, 2001).

Participants’ reported lack of support from colleagues could perhaps explain their expectation to capitalize on their project participation

for networking purposes (see Section 6.2). In this sense, professionals differ from students, who reportedly receive sufficient support from their instructors (see ZAREVA, 2016). Professionals need to make the best use of the opportunity presented at such CL CPD projects to reach out to speakers and other participants for mutual support and development if they cannot find the support they need at their respective workplaces.

Two of the barriers are of a more practical nature: ‘resources’ and ‘time’. The former category includes comments about computer access and burdensome programs. It also encompasses a comment on the large availability of resources (cf. Example 20), which is seen negatively because it requires one to select the most appropriate resource. A similar problem is evident in Frankenberg-Garcia (2015): student translators express difficulties in choosing the right corpora or the most effective query tools. Our research participants also mention the lack of time as a deterrent of corpus use in their teaching practice. This issue is approached from a range of angles: remarks are made in relation to the time required to prepare a corpus-based lesson and to explain some of the CL basics to students. In Example 21, the participant refers to the learning time required before reaping any positive outcomes (see also ASTON, 2009; WILKINSON, 2006).

When Tables 3 and 4 are compared, we notice that there is a similar overall number of advantages and barriers: 19 vs. 21, respectively. These come from the same number of participants. All the participants who identified at least one advantage also listed one barrier. The only exceptions lie with Participant F; 55; T&L; Master’s (Chinese), who only focused on the positive side; and with Participant M; 39; S&T; Master’s (TESOL), who did the opposite and commented on the challenges of integrating CL to his teaching practice.

6.4 Appraisal of corpus applications to research

Given the project’s two foci on corpus applications to teaching and research, participants were additionally asked to identify the advantages and barriers of applying CL to their research practice. Zareva’s (2016) study reveals an optimistic level of research enthusiasm ($M=3.7$ on a five-point Likert scale) among graduate students of TESOL in learning how to do corpus research. However, there seems to be a dearth of studies investigating how academics and/or professionals evaluate the use of CL in their research. The results reported in this section address this gap.

Similar to the procedure described in Section 6.3, the instances where the participants decided not to answer these questions and/or they felt they were unable to answer them were discarded. There were slightly fewer instances of non-completion in this case (N=7 for advantages and N=7 for barriers) when compared to the question on participants' teaching practice (N=8 for advantages and N=9 for barriers). Some participants indicated that this would not apply to their circumstances, and these answers were not included in the results either.

Table 5 summarizes the advantages that participants see in their adoption of a corpus approach to their research practice. Two of the categories are the same from Table 3: 'language use' and 'big data'. These reasons have been mentioned by different participants with a single exception. Participant M; 29; S; Master's (TESOL) referred to 'language use' twice in his answers. However, his answers are clearly distinct. Pedagogically, he commented on the introduction to students to real-life language use; research-wise, he singled out the role of corpora in discourse-related research and investigations of language use in different contexts.

TABLE 5 – Advantages of applying CL to research practice

Category	Frequency	Example
Methodological approach	12	25. "Learning to apply new research methods" [M; 54; L; PhD (History)]
Language use	3	26. "Corpus analysis can help with research of terminologies, collocations." [F; 45; S; Master's (Translation)]
Interdisciplinarity	2	27. "I think it could be a useful method to offer new/future insights into accounting research" [M; 25; S; Master's (Research)]
Big data	1	28. "Analysis of a lot of data with relative ease" [M; 27; S&T; PhD (Linguistics)]
Collaborative work	1	29. "I am working with a corpus expert on a research." [F; 42; L; PhD (Linguistics)]
None	1	30. "None as yet" [M; 35; T; Master's (International Relations)]

The most frequently mentioned advantage relates to the methodological affordances provided by corpus work. There is a long-standing debate in the field whether CL is a science or a method

(see VIANA; ZYNGIER; BARNBROOK, 2011). It seems that the methodological advantages of CL correspond to its biggest advantage for the participants. For example, they commented on corpus techniques, the complementary/supplementary role that CL may play to qualitative research, and the new perspectives that corpus analysis may open up (see Example 25). We could potentially interpret this result in light of our participants' educational background. Because only five of them have their highest degree in Linguistics (here conceived in a strict way to include 'Linguistics', 'Languages' and 'Chinese'), most of the participants may be primarily more interested in the practical application of CL, thus seeing it as a way of assisting them in their research practice.

Other advantages included 'interdisciplinarity' and 'collaborative work'. The former highlights once more the nature of our participant sample: the project was successful in gathering colleagues from other non-language-related fields, who were interested in learning and applying CL to their research practice (see, for instance, Example 27 about Accounting). The category of 'collaborative work' reinforces points that had been made earlier in this paper: participants value opportunities to network (cf. Section 6.2) and they resent lack of support from teaching colleagues (cf. Section 6.3).

Finally, one participant claimed that CL had no advantages to his research practice. His concise reply does not allow for further discussion of his answer. However, because he gave the same answer as to the barriers (see TABLE 6) and because he provided full answers to the section on teaching practice, he may have meant that he was not able to answer this question due to, for instance, lack of CL research knowledge and/or lack of research experience. Alternatively, he may have meant that this question did not apply to his circumstances since he was working as a teacher at the time.

In relation to the participants' identified barriers as to the use of CL in their research, Table 6 reveals that there is some similarity to the pedagogical barriers (cf. TABLE 4). Three of the categories are the same: 'lack of knowledge', 'time' and 'resources' are equally impeditive to their undertaking of corpus research. Because of the physical proximity of these questions in the questionnaire, we checked whether participants had potentially given the same answers. This occurred in a limited number of cases: only five answers from four participants were equal, suggesting that they did not see any difference in the barriers they faced with the integration of CL in teaching or research.

The category entitled ‘none’ had also appeared in relation to the advantages (see TABLE 5). One case relates to Participant M; 35; T; Master’s (International Relations) and has been discussed earlier in this section. The other instance is from Participant M; 48; T; Master’s (TESOL), who has identified advantages and barriers in all the other three cases. This seems to suggest that he holds a positive perspective as to the use of corpus in his research practice.

TABLE 6 – Barriers of applying CL to research practice

Category	Frequency	Example
Lack of knowledge	10	31. “Unfamiliarity with verification procedures of research results” [M; 54; L; PhD (History)]
Time	4	32. “time consuming” [F; 42; T; Master’s (Applied Linguistics)]
Copyright	2	33. “I think copyright can be an issue, with some document that I can use for corpus analysis” [F; 45; S; Master’s (Translation)]
Lack of patience	2	34. “I’m also quite impatient- a bit problem perhaps for CL” [M; 39; S&T; Master’s (TESOL)]
None	2	35. “None as yet” [M; 35; T; Master’s (International Relations)]
Resources	2	36. “Computers!” [F; 45; S&L; PhD (Education)]
Avoidance of bias	1	37. “avoid researcher bias” [M; 29; S; Master’s (TESOL)]

Table 6 contains three new categories. Two participants mention ‘Copyright’ as a hurdle that they may have to overcome in order to gain access to the texts for corpus analysis. This issue has been discussed in the literature, especially in relation to corpus compilation (e.g. BOWKER; PEARSON, 2002; MCENERY; XIAO; TONO, 2006; WYNNE, 2004). Example 33 comes from a student participant in Translation where copyright makes it extremely difficult or virtually impossible to investigate different translated versions of the same recent literary work of art, for example.

Interestingly enough, two participants identified their lack of patience to conduct corpus research as a barrier. Their answers to this specific question (see Example 34) do not provide much information for us to understand why their impatience would be an issue. A close investigation of their answers to the questionnaire indicate that they have a similar profile: both are from the UK, hold a Master's in TESOL and worked as language teachers at the time. They were not CL beginners: these two participants had studied CL in their Master's before their CPD project participation. In their responses to the barriers about teaching, both alluded to computer skill issues (see Example 19), a result that is supported by the self-evaluation of their familiarity with computer use (i.e. 'slightly familiar' for one participant and 'moderately familiar' for the other). It could be hypothesized that their unfamiliarity with technology might be a key factor leading to their impatience when using computer tools to undertake corpus research.

The other new category in Table 6 is 'avoidance of bias', which was mentioned by a single participant (cf. Example 37). The lack of any additional information in the questionnaire makes it difficult to make sense out of the participant's answer. It might have been the case that he wanted to include it in the box on the advantages and misplaced it in the barriers box. This misunderstanding would be coherent with the literature in that a corpus approach may reduce researcher bias (e.g. BAKER, 2006).

A comparative analysis of Tables 5 and 6 reveals that, while slightly more barriers than advantages have been identified, the difference is small to support any conclusions. Similar to what was observed in relation to teaching, the same number of participants identified advantages and barriers, meaning that the positive and negative aspects were equally dispersed.

7 Conclusion

This original study examined the perspectives of participants from different disciplinary backgrounds on a CL CPD project. It therefore addressed two research gaps in the educational application of CL: it researched the experience of participants from several disciplines (rather than only those from language studies) and investigated an underexplored educational context – a non-degree-awarding CPD one.

Methodologically, the study drew on a project funded by the British Academy that aimed to introduce academics and/or professionals to corpus research and applications. We administered a questionnaire to the 28 CPD project participants at the first face-to-face event, all of whom voluntarily decided to participate in this study. The rigorous analysis was conducted by integrating qualitative and quantitative approaches with both researchers carefully and thoroughly checking the analysis that had been undertaken by each other.

The findings revealed that the successful nature of the focal CPD project in drawing the attention of two types of participants: those who had never studied CL before, the main target participant group, and those who had already studied it previously. In this sense, the CPD project was successful in that it did not only appeal to those somehow versed in CL. Instead, complete novices in CL were reached, including those from non-language-related educational backgrounds.

Among the participants who had previously studied CL, half of them learned about it formally (primarily through Master's modules), and the other half developed their learning through routes that did not lead to the award of a degree. The former reinforces the relevance of including CL in the curriculum while the latter highlights the importance of informal learning initiatives such as CPD projects in making CL knowledge accessible to a larger number of (current/future) academics/professionals. As the participants indicated that lack of opportunities was one of the top barriers for their prior study of CL, more formal and informal CL learning opportunities should be provided in the years to come.

Increasing CL learning provision does not seem to suffice, though. As some participants pointed out, they had had the chance to learn about CL before this CPD project, but they lacked the interest to engage in it. We should therefore foster academics'/professionals' interest in these learning opportunities. The need to work on this personal aspect is reinforced by participants' indication that one of the drivers for their registration in this CPD project was to increase their motivations to undertake corpus analysis.

Participants' expectations of this CL CPD project foregrounded three main aspects. These academics and/or professionals wanted to (i) enhance their knowledge, primarily in relation to CL matters; (ii) learn about corpus applications, especially to the field of Education, which was the focus of the project; and (iii) further their research development

– mainly with regard to corpus analysis. To a certain extent, these expectations match their perceived barriers on the use of CL in their teaching and research practices: lack of knowledge was the main factor, thus potentially explaining why they had decided to register for this CPD project.

Our findings reveal participants' appraisals of the embedding of CL in their professional practices. When it comes to teaching, they indicate that corpus approaches allow them and their students to analyze language in use and that these approaches improve their pedagogical practice. Research-wise, the participants believe that the main advantage is the methodological approach afforded by CL such as in the case of interdisciplinary research. Naturally, the application of CL to teaching and research is not only seen from a positive angle. In both cases, the participants identify their lack of knowledge as the main barrier to the embedding of corpus approaches in their professional practices.

While the present study is based on a small participant sample, it is similar in size or even more extensive than some of the previous studies on the application of CL in educational contexts (e.g. FARR, 2008; FRANKENBERG-GARCIA, 2015; ZAREVA, 2016). Most importantly, we worked with all the participants who joined a specific CPD project on CL. It would not be sensible to increase the participant sample for research purposes since it would not reflect the real-life educational initiative being investigated here.

The significance of this study is two-fold. Research-wise, it advances our knowledge of academics'/professionals' perceived advantages and barriers of embedding CL in their respective workplaces, a context that is underexplored and differs from the one reported in previous studies (e.g. ZAREVA, 2016; RODRÍGUEZ-INÉS, 2013). The research findings lead to the study's practice-related significance: the need to increase the number of CL modules on offer and to develop more CPD projects like the one funded by the British Academy.

As educational initiatives would generally aim to have a long-lasting impact, a follow-up longitudinal study should be conducted in order to capture any potential changes in participants' perspectives on the use of CL in their teaching and/or research activities. It would also be worthwhile going beyond an investigation of their perspectives and examining their actual practices. However, we acknowledge that, although not impossible, this would be more challenging to accomplish

on many fronts such as having access to the participants' workplaces and ensuring the comparability of observed practices given participants' diverse professional backgrounds and activities.

The present study on the educational pedagogical application of CL has explored the significantly under-researched topic of CPD projects. The investigation of CPD participants' perspectives is much needed for furthering the impact of CL: it helps us understand what can be done to engage participants in CL work. Researching CPD participants' prior knowledge, motivations and appraisals are vital in tailoring future CPD projects accordingly, thus fulfilling their aim of democratizing access to CL education in informal settings.

Acknowledgements

The authors are grateful to the British Academy, which supported this project through a Skills Innovator Award (grant number SK150041), and to the participants, who kindly agreed to contribute to this research.

Authors' contributions

Viana had the idea for the study and was the lead researcher for most tasks. Both authors worked collaboratively in the design of the research instrument, which was administered by Lu. The data were digitized by Lu and analyzed by Viana and Lu. Viana was in charge of the overall structure of the paper, wrote most of the sections and thoroughly revised the entire paper. Lu was responsible for the literature review, drafted an initial version of the section on methods and the conclusion, contributed to the discussion of the findings, and read the entire paper critically.

References

ASTON, G. Foreword. *In*: BEEBY, A.; RODRÍGUEZ-INÉS, P.; SÁNCHEZ-GIJÓN, P. (ed.). *Corpus use and Translating*. Amsterdam/Philadelphia: John Benjamins, 2009. p. ix-x.

ASTON, G. Applied Corpus Linguistics and the Learning Experience. *In*: VIANA, V.; ZYNGIER, S.; BARNBROOK, G. (ed.). *Perspectives on Corpus Linguistics*. Amsterdam: John Benjamins, 2011. p. 1-16. DOI: <https://doi.org/10.1075/scl.48.01ast>

BAKER, P. *Using Corpora in Discourse Analysis*. London: Continuum, 2006.

BIBER, D.; REPPEN, R. *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 2015. DOI: <https://doi.org/10.1017/CBO9781139764377>

BOULTON, A. Data-Driven Learning: On Paper, in Practice. In: HARRIS, T.; MORENO JAÉN, M. (ed.). *Corpus Linguistics in Language Teaching*. Bern: Peter Lang, 2010. p. 17-52.

BOULTON, A. Corpus Consultation for ESP: A Review of Empirical Research. In: BOULTON, A.; CARTER-THOMAS, S.; ROWLEY-JOLIVET, E. (ed.). *Corpus-Informed Research and Learning in ESP: Issues and Applications*. Amsterdam: John Benjamins, 2012. p. 261-292. DOI: <https://doi.org/10.1075/scl.52.11bou>

BOWKER, L. Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study. *Meta*, [S.l.], v. 43, n. 4, p. 631-651, 1998. DOI: <https://doi.org/10.7202/002134ar>

BOWKER, L. Corpus Resources for Translators: Academic Luxury or Professional Necessity? *TradTerm*, São Paulo, n. 10, p. 213-247, 2004. DOI: <https://doi.org/10.11606/issn.2317-9511.tradterm.2004.47178>

BOWKER, L.; PEARSON, J. *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge, 2002. DOI: <https://doi.org/10.4324/9780203469255>

BREYER, Y. Learning and Teaching with Corpora: Reflections by Student Teachers. *Computer Assisted Language Learning*, [S.l.], v. 22, n. 2, p. 153-172, 2009. DOI: <https://doi.org/10.1080/09588220902778328>

BUENDÍA-CASTRO, M.; LÓPEZ-RODRÍGUEZ, C. I. The Web for Corpus and the Web as Corpus in Translator Training. *New Voices in Translation Studies*, [S.l.], n.10, p. 54–71, 2013.

CALLIES, M. Integrating Corpus Literacy into Language Teacher Education. In: GÖTZ, S.; MUKHERJEE, J. (ed.). *Learner Corpora and Language Teaching*. Amsterdam: John Benjamins, 2019. p. 245-263. DOI: <https://doi.org/10.1075/slcs.201.12cal>

CHAMBERS, A.; O'RIORDAN, S. Learning to Teach through French from a Corpus of Classroom Discourse: Giving Corrective Feedback. In: CONACHER, J. E.; KELLY-HOLMES, H. (ed.). *New Learning Environments for Language Learning*. Frankfurt: Peter Lang, 2007. p. 87-101.

CHARLES, M. Getting the Corpus Habit: EAP Students' Long-Term Use of Personal Corpora. *English for Specific Purposes*, [S.l.], n. 35, p. 30-40, 2014. DOI: <https://doi.org/10.1016/j.esp.2013.11.004>

CHEN, M.; FLOWERDEW, J.; ANTHONY, L. Introducing In-Service English Language Teachers to Data-Driven Learning for Academic Writing. *System*, [S.l.], n. 87, p. 102-148, 2019. DOI: <https://doi.org/10.1016/j.system.2019.102148>

CONRAD, S. Corpus Linguistic Approaches for Discourse Analysis. *Annual Review of Applied Linguistics*, Cambridge, n. 22, p. 75-95, 2002. DOI: <https://doi.org/10.1017/S0267190502000041>

CONRAD, S. Variation in Corpora and Its Pedagogical Implications. In: VIANA, V.; ZYNGIER, S.; BARNBROOK, G. (ed.). *Perspectives on Corpus Linguistics*. Amsterdam: John Benjamins, 2011. p. 47-62. DOI: <https://doi.org/10.1075/scl.48.04con>

COPLAND, F.; VIANA, V.; BOWKER, D.; MORAN, E.; PAPAGEORGIOU, I.; SHAPIRA, M. *ELT Master's Courses in the UK: Students' Expectations and Experiences*. London: British Council, 2017.

CRAWFORD, P.; BROWN, B. Health Communication: Corpus Linguistics, Data Driven Learning and Education for Health Professionals. *Taiwan International ESP Journal*, Taiwan, v. 2, n. 1, p. 3-28, 2010.

EBRAHIMI, A.; FAGHIH, E. Integrating Corpus Linguistics into Online Language Teacher Education Programs. *ReCALL*, Cambridge, v. 29, n.1, p. 120-135, 2017. DOI: <https://doi.org/10.1017/S0958344016000070>

FARR, F. Relational Strategies in the Discourse of Professional Performance Review in an Irish Academic Environment: The Case of Language Teacher Education. In: SCHNEIDER, K. P.; BARRON, A. (ed.). *Variational Pragmatics: The Case of English in Ireland*. Berlin: Mouton de Gruyter, 2005. p. 203-234.

FARR, F. Reflecting on Reflections: The Spoken Word as a Professional Development Tool in Language Teacher Education. In: HUGHES, R. (ed.). *Spoken English, Applied Linguistics and TESOL: Challenges for Theory and Practice*. Hampshire: Palgrave Macmillan, 2006. p. 182-215. DOI: https://doi.org/10.1057/9780230584587_9

FARR, F. Evaluating the Use of Corpus-Based Instruction in a Language Teacher Education Context: Perspectives from the Users. *Language Awareness*, [S.l.], v. 17, n. 1, p. 25-43, 2008. DOI: <https://doi.org/10.2167/la414.0>

FARR, F. How Can Corpora Be Used in Teacher Education? In: O'KEEFEE, A.; MCCARTHY, M. (ed.). *The Routledge Handbook of Corpus Linguistics*. London/New York: Routledge, 2010. p. 620-632.

FLIGELSTONE, S. Some Reflections on the Question of Teaching, from a Corpus Linguistics Perspective. *ICAME Journal*, [S.l.], n. 17, p. 97-109, 1993.

FLOWERDEW, L. The Argument for Using English Specialized Corpora to Understand Academic and Professional Settings. In: CONNOR, U.; UPTON, T. (ed.). *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: John Benjamins, 2004. p. 11-33. DOI: <https://doi.org/10.1075/scl.16.02flo>

FLOWERDEW, L. *Corpora and Language Education*. Basingstoke: Palgrave Macmillan, 2012. DOI: <https://doi.org/10.1057/9780230355569>

FRANKENBERG-GARCIA, A. Training Translators to Use Corpora Hands-On: Challenges and Reactions by a Group of Thirteen Students at a UK University. *Corpora*, Edinburg, v. 10, n. 3, p. 351-380, 2015. DOI: <https://doi.org/10.3366/cor.2015.0081>

FRÉROT, C. Corpora and Corpus Technology for Translation Purposes in Professional and Academic Environments: Major Achievements and New Perspectives. *Cadernos de Tradução*, Florianópolis, n. 36, p. 36-61, 2016. DOI: <https://doi.org/10.5007/2175-7968.2016v36nesp1p36>

GALLEGO-HERNÁNDEZ, D. The Use of Corpora as Translation Resources: A Study Based on a Survey of Spanish Professional Translators. *Perspectives*, [S.l.], v. 23, n. 3, p. 375-391, 2015a. DOI: <https://doi.org/10.1080/0907676X.2014.964269>

GALLEGO-HERNÁNDEZ, D. Business Translation Training and ad Hoc Corpora. In: SÁNCHEZ-GIJÓN, P.; TORRES-HOSTENCH, O.; MESA-LAO, B. (ed.). *Conducting Research in Translation Technologies*. Oxford/New York: Peter Lang, 2015b. p. 119-140.

GAN, S. L.; LOW, F.; YAAKUB, N. F. Modeling Teaching with a Computer-Based Concordance in a TESL Preservice Teacher Education Program. *Journal of Computing in Teacher Education*, [S.l.], v. 12, n. 4, p. 27-32, 1996.

GATTO, M. *The Web as a Corpus: Theory and Practice*. London: Bloomsbury, 2014.

GAVIOLI, L. Corpus Analysis and the Achievement of Learner Autonomy in Interaction. In: LOMBARDO, L. (ed.). *Using Corpora to Learn about Language and Discourse*. Bern: Peter Lang, 2009. p. 39-72.

GHADESSY, M.; HENRY, A.; ROSEBERRY, R. L. (ed.). *Small Corpus Studies and ELT: Theory and Practice*. Amsterdam: John Benjamins, 2001. DOI: <https://doi.org/10.1075/scl.5>

GRANATH, S. Who Benefits from Learning How to Use Corpora? In: AIJMER, K. (ed.). *Corpora and Language Teaching*. Amsterdam/Philadelphia: John Benjamins, 2009. p. 47-65. DOI: <https://doi.org/10.1075/scl.33.07gra>

HAFNER, C. A.; CANDLIN, C. N. Corpus Tools as an Affordance to Learning in Professional Legal Education. *Journal of English for Academic Purposes*, [S.l.], v. 6, n. 4, p. 303-318, 2007. DOI: <https://doi.org/10.1016/j.jeap.2007.09.005>

HEATHER, J.; HELT, M. Evaluating Corpus Literacy Training for Pre-Service Language Teachers: Six Case Studies. *Journal of Technology and Teacher Education*, [S.l.], v. 20, n. 4, p. 415-440, 2012.

HÜTTNER, J.; SMIT, U.; MEHLMAUER-LARCHER, B. ESP Teacher Education at the Interface of Theory and Practice: Introducing a Model of Mediated Corpus-Based Genre Analysis. *System*, [S.l.], n. 37, p. 99-109, 2009. DOI: <https://doi.org/10.1016/j.system.2008.06.003>

JÄÄSKELÄINEN, R.; MAURANEN, A. Translators at Work: A Case Study of Electronic Tools Used by Translators in Industry. In: BARNBROOK, G.; PERNILLA, D.; MAHLBERG, M. (ed.). *Meaningful*

Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora. London: Continuum, 2006. p. 48-53.

JOHANSSON, S. Some Aspects of the Development of Corpus Linguistics in the 1970s and 1980s. In: LÜDELING, A.; KYTÖ, M. (ed.). *Corpus Linguistics: An International Handbook*. Berlin: Mouton de Gruyter, 2008. p. 33-54.

JOHNS, T. Should You Be Persuaded: Two Examples of Data-Driven Learning. In: JOHNS, T.; KING, P. (ed.). *ELR Journal 4: Classroom Concordancing*. Birmingham: CELS, The University of Birmingham, 1991. p. 1-16.

KRÜGER, R. Working with Corpora in the Translation Classroom. *Studies in Second Language Learning and Teaching*, Kalisz, Poland, n. 4, p. 505-525, 2012. DOI: <https://doi.org/10.14746/ssl.t.2012.2.4.4>

LAURSEN, L. A.; PELLÓN, A. I. Text Corpora in Translation Training: A Case Study of the Use of Comparable Corpora in Classroom Teaching. *The Interpreter and Translator Trainer*, [S.l.], n. 61, p. 45-70, 2012. DOI: <https://doi.org/10.1080/13556509.2012.10798829>

LEŃKO-SZYMAŃSKA, A. Is This Enough? A Qualitative Evaluation of the Effectiveness of a Teacher-Training Course on the Use of Corpora in Language Education. *ReCALL*, Cambridge, v. 26, n. 2, p. 260-278, 2014. DOI: <https://doi.org/10.1017/S095834401400010X>

LEŃKO-SZYMAŃSKA, A. Training Teachers in Data Driven Learning: Tackling the Challenge. *Language Learning & Technology*, Austin, TX, v. 21, n. 3, p. 217-241, 2017.

MCCARTHY, M. J. Accessing and Interpreting Corpus Information in the Teacher Education Context. *Language Teaching*, Cambridge, v. 41, n. 4, p. 563-574, 2008. DOI: <https://doi.org/10.1017/S0261444808005247>

MCENERY, T.; XIAO, R.; TONO, Y. *Corpus-Based Language Studies: An Advanced Resource Book*. London: Routledge, 2006.

MELLANGE. *Corpora and E-Learning Questionnaire*. Results Summary. 2006. Available on: <http://mellange.eila.univ-paris-diderot.fr/Mellange-Results-1.pdf>. Retrieved at: July 18, 2018.

MONZÓ-NEBOT, E. Corpus-Based Activities in Legal Translator Training. *The Interpreter and Translator Trainer*, [S.l.], n. 22, p. 221-252, 2008. DOI: <https://doi.org/10.1080/1750399X.2008.10798775>

MUKHERJEE, J. Bridging the Gap Between Applied Corpus Linguistics and the Reality of English Language Teaching in Germany. In: CONNOR, U.; UPTON, T. A. (ed.). *Applied Corpus Linguistics: A Multi-Dimensional Perspective*. Amsterdam: Rodopi, 2004. p. 239-250. DOI: https://doi.org/10.1163/9789004333772_014

O'KEEFFE, A.; MCCARTHY, M. (ed.). *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 2010. DOI: <https://doi.org/10.4324/9780203856949>

O'SULLIVAN, Í.; CHAMBERS, A. Learners' Writing Skills in French: Corpus Consultation and Learner Evaluation. *Journal of Second Language Writing*, [S.l.], n. 15, p. 49-68, 2006. DOI: <https://doi.org/10.1016/j.jslw.2006.01.002>

PAPAGEORGIU, I.; VIANA, V.; COPLAND F.; BOWKER, D.; MORAN, E. *Master's ELT Audit Document*. 2017. Available on: <https://www.teachingenglish.org.uk/sites/teacheng/files/Audit%20Final%2010.pdf>. Retrieved on: July 1, 2020.

RENOUF, A. Teaching Corpus Linguistics to Teachers of English. In: WICHMANN, A.; FLIGELSTONE, A.; MCENERY, T.; KNOWLES, G. (ed.). *Teaching and Language Corpora*. London: Longman, 1997. p. 255-266. DOI: <https://doi.org/10.4324/9781315842677-22>

REPPEN, R.; VÁSQUEZ, C. Using Corpus Linguistics to Investigate the Language of Teacher Training. In: WALÍŃKI, J.; KREDENS, K.; GOŹDŹ-ROSKOWSKI, S. (ed.). *Corpora and ICT in Language Studies, PALC 2005*. Frankfurt am Main: Peter Lang, 2007. p. 13-29.

RODRÍGUEZ-INÉS, P. Evaluating the Process and Not Just the Product When Using Corpora in Translator Education. In: BEEBY, A.; RODRÍGUEZ-INÉS, P.; SÁNCHEZ-GIJÓN, P. (ed.). *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 2009. p. 129-150. DOI: <https://doi.org/10.1075/btl.82.09rod>

RODRÍGUEZ-INÉS, P. Electronic Corpora and Other Information and Communication Technologies Tools: An Integrated Approach to Translation Teaching. *The Interpreter and Translator Trainer*, [S.l.], v.

4, n. 2, p. 251-282, 2010. DOI: <https://doi.org/10.1080/13556509.2010.10798806>

RODRÍGUEZ-INÉS, P. Electronic Target-Language Specialized Corpora in Translator Education. *Babel*, [S.l.], v. 59, n. 1, p. 57-75, 2013. DOI: <https://doi.org/10.1075/babel.59.1.04rod>

RÖMER, U. Corpus Research and Practice: What Help Do Teachers Need and What Can We Offer? In: AIJMER, K. (ed.). *Corpora and Language Teaching*. Amsterdam/Philadelphia: John Benjamins, 2009. p. 83-98. DOI: <https://doi.org/10.1075/scl.33.09rom>

RÖMER, U. Using General and Specialized Corpora in English Language Teaching: Past, Present and Future. In: CAMPOY-CUBILLO, M. C.; BELLÉS-FORTUÑO, B.; GEA-VALOR, L. (ed.). *Corpus-based Approaches to English Language Teaching*. London: Continuum, 2010. p. 18-35.

SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

TODD, W. R. Induction from Self-Selected Concordances and Self-Correction. *System*, [S.l.], n. 29, p. 91-102, 2001. DOI: [https://doi.org/10.1016/S0346-251X\(00\)00047-6](https://doi.org/10.1016/S0346-251X(00)00047-6)

TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. Amsterdam: John Benjamins, 2001. DOI: <https://doi.org/10.1075/scl.6>

TSUI, A. B. M. ESL Teachers' Questions and Corpus Evidence. *International Journal of Corpus Linguistics*, [S.l.], v. 10, n. 3, p. 335-356, 2005. DOI: <https://doi.org/10.1075/ijcl.10.3.03tsu>

VARELA-VILA, T. Córpora ad hoc en la práctica traductora especializada: aplicación al ámbito de las enfermedades neuromusculares. In: CANTOS-GÓMEZ, P.; SÁNCHEZ-PÉREZ, A. (ed.). *A Survey on Corpus-Based Research*. Panorama de investigaciones basadas en corpus. Murcia: Asociación Española de Lingüística del Corpus, 2009. p. 814-831.

VÁSQUEZ, C.; REPPEN, R. Transforming Practice: Changing Patterns of Interaction in Post-Observation Meetings. *Language Awareness*, [S.l.], v. 16, n. 3, p. 153-172, 2007. DOI: <https://doi.org/10.2167/la454.0>

VAUGHAN, E. "I Think We Should Just Accept Our Horrible Lowly Status": Analysing Teacher-Teacher Talk in the Context of Community of

Practice. *Language Awareness*, [S.l.], v. 16, n. 3, p. 173-189, 2007. DOI: <https://doi.org/10.2167/la456.0>

VIANA, V. The Politics of Corpus Linguistics. In: VIANA, V.; ZYNGIER, S.; BARNBROOK, G. (ed.). *Perspectives on Corpus Linguistics*. Amsterdam: John Benjamins, 2011. p. 229-245. DOI: <https://doi.org/10.1075/scl.48>

VIANA, V. Data-Driven Language Learning and Teaching: A Corpus Perspective. In: LINGUISTICS AND KNOWLEDGE ABOUT LANGUAGE IN EDUCATION (LKALE) BRITISH ASSOCIATION FOR APPLIED LINGUISTICS (BAAL) SPECIAL INTEREST GROUP (SIG) MEETING, 2017, Sheffield. Sheffield, 2017. [Oral presentation].

VIANA V.; BOCORNY, A.; SARMENTO, S. *Teaching English for Specific Purposes*. Alexandria: TESOL Press, 2018.

VIANA, V.; ZYNGIER, S.; BARNBROOK, G. (ed.). *Perspectives on Corpus Linguistics*. Amsterdam: John Benjamins, 2011. DOI: <https://doi.org/10.1075/scl.48>

WILKINSON, M. Compiling Corpora for Use as Translation Resources. *Translation Journal*, [S.l.], n. 101, [s.p.], 2006. Available on: <http://www.translationdirectory.com/article910.htm>. Retrieved at: Aug. 21, 2020.

WYNNE, M. Archiving, Distribution and Preservation. In: WYNNE, M. (ed.). *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow Books, 2004. p. 71-78.

ZANETTIN, F. Swimming in Words: Corpora, Translation, and Language Learning. In: ASTON, G. (ed.). *Learning with Corpora*. Bologna: CLUEB, 2001. p. 177-197.

ZAREVA, A. Incorporating Corpus Literacy Skills into TESOL Teacher Training. *ELT Journal*, Oxford, v. 71, n. 1, p. 69-79, 2016. DOI: <https://doi.org/10.1093/elt/ccw045>