

In the future I will use corpora in my classroom: is it really applicable?¹

AWATIF ALRUWAILI

Department of Languages and Translation, College of Humanities and Social Sciences, Northern Border University, Arar, Saudi Arabia

Received: 2023-03-08 / Accepted: 2023-05-24

DOI: <https://doi.org/10.30827/portalin.vi42.27574>

ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

ABSTRACT: The aim of this study was to assess in-service teachers' pedagogical skills that enable them to exploit corpora successfully in their daily teaching of English as a foreign language. The teachers attended an intensive training course and were then asked to submit end-of-course projects. The projects involved designing lesson plans based on the teaching contexts and implementing corpora into these plans according to their contexts. The project outcomes revealed that the submitted plans were preliminary and basic in terms of using corpus resources and activities. However, the teachers were genuinely keen to implement corpora into their daily teaching. The teachers proposed to apply corpora as an additional teaching methodology and in supplementary activities after explaining the learning units' aims, periods, foci, and so on. They used corpora as hands-on supplementary sources of information, which was understandable for early application. The teachers seemed willing to implement corpora in their future teaching and needed no further training courses, but they might still need more guidance and feedback on the first-time application, which could lead to successful implementation in the classroom.

Keywords: corpora; teacher-training; implementation; lessons; pedagogical skills

En el futuro utilizaré corpus en mi clase: ¿es realmente adecuado?

RESUMEN: El objetivo de este estudio era evaluar las capacidades pedagógicas de los profesores que están empezando a trabajar y que les permiten utilizar con éxito corpus en su enseñanza diaria del inglés como lengua extranjera. Los profesores asistieron a un curso intensivo de capacitación y después se les pidió que presentaran proyectos de fin de curso. Los proyectos consistían en diseñar planes de clase basados en los contextos de enseñanza e implementar corpus en dichos planes según sus contextos. Los resultados del proyecto revelaron que los planes presentados estaban en una fase inicial y en una fase básica en cuanto al uso de los recursos y las actividades de corpus. Sin embargo, los profesores se mostraron realmente interesados en aplicar corpus en su enseñanza diaria. Los profesores propusieron aplicar corpus como metodología didáctica adicional y en actividades complementarias tras explicar los objetivos, periodos, enfoques, etc. de las unidades didácticas. Utilizaban corpus como fuentes de información complementarias y prácticas, lo que resultaba comprensible para una aplicación en una fase inicial. Los profesores parecían estar dispuestos a aplicar corpus en su futura labor docente y no necesitaban más cursos de capacitación, pero podrían

¹ The author extends their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research through the project number NBU-FFR-2024-458-01.

seguir necesitando más orientación e información sobre la primera aplicación, lo que podría conducir a una aplicación satisfactoria en el salón de clases.

Palabras clave: corpus; capacitación de profesores; aplicación; lecciones; habilidades pedagógicas

1. INTRODUCTION

Research findings in task-based language teaching (TBLT) have consistently proven that Corpus linguistics (CL) is a revolutionary approach to language analysis and learning. Thus, a shift has taken place in the focus of training to help students/teachers implement CL in language learning and teaching. Nevertheless, CL is rarely used in teaching communities or by teachers (Boulton, 2017; Callies, 2019; Chambers, 2019; Ma et al., 2022). Increasingly, corpus training programmes have been integrated into teacher education (Breyer, 2009; Ebrahimi & Faghih, 2016; Farr, 2008; Heather & Helt, 2012; Leńko-Szymańska, 2017; Zareva, 2017), but few studies have addressed the training needs of in-service teachers (Alruwaili, 2023; Chen et al., 2019; Ma et al., 2022, 2023). In view of the shortage of research targeting teachers, Leńko-Szymańska (2017) argued that other teachers have difficulty copying or directly applying what pre-service teachers learn in courses, as the learners in these courses are more advanced and/or more interested in CL than other types of learners (p. 218). Thus, an urgent need exists to “[institutionalize] teacher-training courses dedicated to using CL in language instruction” (Leńko-Szymańska, 2014, p. 261).

In this study, the participants received training according to their needs and teaching contexts which was teaching English as a foreign language in public schools. The training was designed in a way that would encourage and challenge them with tasks and examine their readiness to implement CL in their classrooms. The tasks included end-of-course projects (lesson plans) that the teachers were required to submit after the end of the course. The tasks evaluated the teachers’ abilities to implement a corpus-based approach and tools in daily teaching and how they would do so. The main objective of this paper was to determine whether the participants were able to acquire the pedagogical skills necessary to exploit corpora in their future teaching, and this was achieved by investigating how teachers implemented and integrated corpora into their lesson plans and by exploring their perceptions of using corpora to develop their language pedagogy.

2. LITERATURE REVIEW

Considerable research has shown that corpora can be used both inside and outside classrooms (i.e. Leńko-Szymańska & Boulton, 2015), and some training courses instruct teacher trainees on introducing and exploiting corpora (Chen et al., 2019; Leńko-Szymańska, 2014, 2015; Lin, 2019). However, this research is inadequate, as pointed out by Leńko-Szymańska (2017), and a need exists to institutionalise such training courses for both pre- and in-service teachers to address their respective needs. In addition, these courses depend heavily on CL, and learners’ profiles vary across studies, which poses difficulties in taking advantage of or benefitting from such courses in terms of spreading and generalising the training course. From another angle, such courses differ in their aims, influencing the perceived usefulness and direct application of CL in different teaching contexts.

Some studies have focused on raising awareness of the use of CL in classrooms (Callies, 2016; Mukherjee, 2004), some have considered the development of corpus literacy skills (Ma et al., 2023; Zareva, 2017), some have aimed to improve training courses by evaluating them (Chen et al., 2019; Leńko-Szymańska, 2014, 2015), while others have investigated how to integrate CL into classroom teaching according to the teaching materials and lesson plans designed by trainees (Leńko-Szymańska, 2017; Ma et al., 2022). Studies have also described CL training and learners' attitudes towards CL at the tertiary level following training (Chen et al., 2019; Ebrahimi & Faghih, 2016; Farr, 2008; Heather & Helt, 2012; Leńko-Szymańska, 2014, 2017; Zareva 2017). There is growing research on the CL and mainly DDL approaches, which target languages other than English (Abad Castelló, 2019; Froti, 2019). Abad Castelló (2019) explored whether teachers use CL in teaching Spanish as a foreign language and found that they underuse CL in teaching Spanish. Thus, there is a need to raise awareness among teachers who teach languages other than English.

2.1. Corpora and Language Learning and Teaching

CL has repeatedly been mentioned as a revolutionary language learning and teaching method and a rich source of authentic data that supports inductive learning (Boulton & Cobb, 2017; Sinclair, 1991). CL has proven to be of value in almost all language systems, such as lexicography, grammar, language variation/change, contrastive and translation studies, discourse analysis, and language learning (McEnery & Xiao, 2011). In language learning, CL's best quality is that it does not depend on native intuition but on authentic data regarding observations of language patterns, thus encouraging inductive and discovery learning that supports learners' autonomy (Boulton, 2017; Boulton & Cobb, 2017). However, it is still not widely used in classrooms at the pre-tertiary level; instead, it is used marginally at the tertiary level (Boulton, 2017; Callies, 2019; Chambers, 2019; Ma et al., 2021). Despite the positive outcomes of studies revealing that CL improves different aspects of language, such as vocabulary (Ackerley, 2017; Ma & Mei, 2021), collocations (Fang et al., 2021; Vyatkina, 2016), grammar (Lin & Lee, 2015), and writing (Crosthwaite, 2020;), instructors still resist using it. This resistance may be due to a lack of knowledge, a lack of the ability (skills) to implement it even after the knowledge is acquired, or a mere rejection of CL in classrooms (Boulton, 2009). The lack of CL knowledge has been addressed by training courses for pre-service teachers and, in a few cases, in-service teachers. However, the lack of ability (skills) to put CL knowledge into practice and use it in language classrooms has still not been thoroughly investigated (Lin, 2019; Ma et al., 2022).

2.2. Corpus-Based Language Pedagogy and Training

Studies have investigated teacher trainees' attitudes towards and perceptions of CL in classrooms (e.g. Abdel Latif, 2021; Breyer, 2009; Ebrahimi & Faghih, 2016; Leńko-Szymańska, 2014; Ma, 2023). These studies offered participants training that mainly aimed to introduce them to CL and familiarise them with corpus tools. The results were promising in that the participants had positive attitudes and were able to master some corpus skills (Abdel Latif, 2021; Callies, 2019; Ebrahimi & Faghih, 2016; Leńko-Szymańska, 2014, 2017), but a lack

of pedagogical skills to facilitate the application of CL remains in real language classrooms (Leńko-Szymańska, 2017). Little research has considered how in-service teachers can apply CL skills to their teaching contexts (Lin, 2019; Ma et al., 2022).

Empirical studies on teachers' development of CL use in language teaching are scarce and have mainly focused on attitudes and perceptions. Some of these studies required the participants to design teaching materials (Heather & Helt, 2012; Leńko-Szymańska, 2017; Lin, 2019; Ma et al., 2022). Heather and Halt's (2012) study focused not on the teaching materials as ends in themselves but on the skills required for using corpora in language learning and instruction. Therefore, the investigation aimed to develop the participants' expertise in corpus literacy skills, with only marginal attention paid to pedagogical skills in designing teaching materials. However, Leńko-Szymańska (2017) focused on both corpus literacy and pedagogical skills and found that the participants lacked both types of skills, but their pedagogical skills were weaker than their corpus literacy skills. Leńko-Szymańska (2017) examined the participants' proficiency in three sets of skills—technical, corpus linguistic, and pedagogical—and considered whether they were essential for the successful application of corpora in language teaching. The analysis of all of the participants' corpus-based lessons revealed that they failed to reach a satisfactory level of expertise due to limited experience with corpus analysis and the data-driven learning (DDL) approach. Technical skills seemed the easiest to acquire, since the participants were able to use and manipulate corpora even at a basic knowledge level. However, regarding corpus linguistic skills, the participants were unable to select linguistic points that could be analysed using corpora, indicating only a rudimentary skill level. Regarding pedagogical skills, judging the degree of expertise was difficult, as the participants mostly selected collocations as their teaching targets; however, they did not demonstrate expertise in designing the activities, which was expected because they were pre-service teachers with no experience with in-service teaching.

Lin (2019) investigated one in-service teacher's perceptions and use of DDL in an actual language classroom. Despite some challenges due to the teacher being a first-time user, such as technical difficulties and anxiety, the investigation revealed promising results for the successful application of the DDL approach. The issue of time and effort appeared in Lin's study (2019) which was discussed by Boulton (2010) that the challenge of time and effort will continue to arise until ready-made materials become widely available for immediate use.

Ma et al. (2022) investigated how two in-service teachers from diverse backgrounds at the tertiary level developed pedagogical skills and examined the effectiveness of corpus-based teaching in actual classes. The study documented how the two teachers developed their corpus-based language pedagogy skills and achieved various levels of "success" (p. 19). The participants designed corpus-based lessons and achieved high levels of corpus literacy, pedagogical skills, and implementation in real classes.

3. METHODS

3.1. Background

The present study assessed the pedagogical skills of in-service teachers that enabled them to successfully exploit corpora in their daily teaching, investigated their willingness

to apply CL in their classrooms at the pre-tertiary education² (pre-tertiary education usually consists of three levels: primary, intermediate, and high schools), and evaluated their readiness to implement CL in their everyday teaching contexts. This research is an exploratory case study. The participants attended an online intensive training course on CL in the teaching arena. They were asked to submit lesson plans that integrated CL into the design of lesson plan activities and/or teaching materials. The format of the lesson plans was based on Leńko-Szymańska (2014; 2017). The participants who attended the intensive training course (see Alruwaili, 2023) were asked to submit full lesson plans that considered corpora tools and resources. The course consisted of eight sessions, all of which comprised reading materials and videos, along with discussion questions (through online forums). In short, the course introduced CL concepts and theories, demonstrated a wide range of CL software and tools, discussed how to develop appropriate teaching materials, and introduced and discussed corpus literacy. The course also covered learner corpora concepts and their usefulness in a teaching context. The course instructor introduced the DDL approach and demonstrated how to create materials using both direct and indirect DDL approaches. One of the course's aims was to present the DDL approach to in-service teachers and show them the different applications of DDL in foreign language classrooms. The training course established online and current study targets for pre-tertiary (primary, intermediate, and high school) contexts.

The study was guided by the following main research questions:

1. To what extent are teachers able to acquire pedagogical skills for exploiting corpora in their own teaching contexts?
2. How can teachers implement corpora in their classrooms?
3. What challenges do teachers face in designing lesson plans?

3.2. Participants and Context

The participants were in-service female teachers who teach English as a foreign language at the pre-tertiary level and who voluntarily participated in the study. The selection of female participants was accidental, as the course was announced through formal channels and those who applied to attend were all females. Eighty-five teachers attended the full course, but only 64 (75%) submitted lesson plans. This submission rate was quite good because participation was voluntary and creating the lesson plans required effort. This suggests that the in-service teachers were genuinely interested in applying CL in their classrooms. The participants' demographic information varied in terms of experience, age, and teaching level. Only the 64 teachers who submitted lesson plans were included in the final analysis. Most of the teachers (81.25%) were aged 31–40 years, 7.81% were under 30, and 10.93% were over 41. The majority of the teachers held bachelor's degrees (96.87%). Of these, 56.25%

² The pre-tertiary level refers to education below university, college, polytechnic and diploma levels. UNESCO recognises nine stages of education (Early childhood education, Primary education at primary school or elementary school, Secondary education at secondary school or high school, and Higher education or vocational education). In the current study, the target was Primary and Secondary education which consist of three stages: primary, intermediate, high school. (UNESCO, n.d.).

had postgraduate certificates in education, and 35.93% received teacher training as part of their bachelor's degrees. Table 1 shows that their teaching experience varied.

Table 1. *Participants' teaching experience*

Years of Teaching	Percentage	No. of teachers
2–5 years	7.81	5
6–10 years	45.31	29
11–15 years	21.87	14
16–20 years	18.75	12
< 21 years	6.25	4

Twenty-five per cent of the teachers taught at the primary level, 25% at the intermediate level, 26.56% at the high school level, and 23.43% at diverse levels simultaneously. Most teachers had full teaching schedules (59.37%), which equated to about 17–24 periods per week (each 45 minutes long), and 40.62% had fewer than 16 periods. Moreover, 59.37% were unfamiliar with the word 'corpus'; 20.31% had heard of it but had no idea what it meant, 7.81% had already done some work with corpora, and 6.25% had already had opportunities to use corpora in their classrooms.

3.3. Data Collection and Procedure

The researcher collected three types of data using a pre-course questionnaire, a lesson plan task, and open-ended questions. Before starting the training course, the participants received the pre-course questionnaire, which gathered their background information (i.e. age, teaching level, teaching load, teaching experience, qualifications, teacher training, and familiarity with the word 'corpus'). During the intervention phase (the intensive training course), they were asked to consider working on developing lesson plans, and after completion of the course, they were given a week to submit the requested plans. They were asked to design lesson plans that included the stages of the education system, a relevant topic, clear aims, required resources, the stages of the lessons (in detail), lesson activities and teaching materials, and homework. After submitting their plans, they were asked for their opinions about the advantages, disadvantages, and difficulties of designing lesson plans that integrated CL and whether they had any additional comments.

3.4. Data Analysis

The analysis of the data was performed quantitatively for the pre-course questionnaire and qualitatively for the lesson plans and open-ended questions. The qualitative data were analysed using the content analysis method, where a coding frame was created, as guided by the research questions and a set of parameters (mentioned below). The validity of the analyses was assessed by an inter-rater (a coder), who analysed lesson plans and whose analyses were then compared with the researcher's analyses to check for any disagreement.

The researcher used frequency scores for some parts of the qualitative analysis and qualitatively analysed the lesson plans in terms of the following:

- the lesson plan's methodology was sound in terms of clear aims, focus, and so on,
- whether it included varied language teaching materials that aligned well with the lesson aims,
- whether corpus-based activities were integrated effectively with other non-corpus activities, and
- whether the created corpus-based materials and tasks targeted all/any language systems (lexis, phraseology, grammar, discourse) or skills (writing, reading, speaking, listening).

4. FINDINGS

The researcher conducted qualitative analyses of the lesson plans to determine whether they were methodologically sound and included various teaching materials, whether corpus-based and non-corpus activities were integrated into the lesson plans, and whether the focus of the lesson plans and teaching materials were corpus- or non-corpus-based. The researcher also checked the focus of the lesson plans in terms of language systems or skills. It should be pointed out that 80% of the participants did not know what 'corpus' meant, and only 6.25% had already had opportunities to use corpora in their classrooms. Thus, testing the participants skills before the training course would have been pointless.

First, the lesson plans were reviewed in terms of whether they were complete and submitted on time. All 64 participants submitted their plans online within one week of the course's end. Four participants made only general statements about their lesson plans (e.g. the title, the students' target learning level, and topics) and were consequently excluded from the analysis.

4.1. Pedagogical Skills for Exploiting Corpora

Sixty lesson plans were analysed to determine the extent to which the participants were able to acquire pedagogical skills for exploiting corpora in their teaching context.

The analysis showed that all 60 were clear lesson plans that mostly included the following: the aims, the target level of the students, the topic, the stages of teaching (including warming-up, pre-teaching, introducing the items or targets, activities, discussion, and time required for each stage of teaching), number of periods (for completing the whole lesson), the target skill and/or language system, teaching materials, type of corpus (in any), activities, and homework.

Most of the participants (85%) identified clear aims at the beginning of their plans, but nine did not. However, a few of those nine roughly pointed out the aims within the stages of the lesson plans.

Twenty-five per cent of the participants targeted primary-level students, 48.33% targeted intermediate-level students, 11.66% targeted high school-level students, and 15% did not

select a specific level. Regarding the topics, the participants mostly chose them based on the levels at which they were teaching (see Table 2 for the most frequently chosen topics).

Table 2. *Frequency of selected topics*

Topics	Frequency
Grammar themes	5
Body	5
Food	6
Festivals	5
Sports	3
Animals	3
Friends	4

More than a half of the participants (53%) set one period for their lesson plans, while 10% set two periods for their lessons. The focus of the lessons varied in terms of the language skill and system; most of the participants combined the skill and language system into one plan, perhaps because they all taught an integrated syllabus (i.e. the use of the skill within the language system, such as phraseology in speaking). At the beginning of the lesson plans, the selected foci were as follows: 81.66% focused solely on vocabulary, 5% focused on grammar alone, 5% focused on vocabulary and grammar, 3.33% focused on grammar and reading skills, 1.6% focused on alphabet letters (a traditional lesson), and 3.33% did not specify the focus. However, the foci changed a bit during the stages of the lesson plan, with 58% focusing on vocabulary solely, 28% focusing on vocabulary and grammar, 10% focusing on grammar only, one on reading, and another on alphabet letters. Thus, there was a decrease in the focus on vocabulary (from 81.66% to 58%), an increase in the focus on grammar (from 5% to 10%), and an increase in the focus on vocabulary and grammar (from 5% to 28%).

Not all of the 60 participants who submitted the lesson plans integrated corpora into their plans, but a good number did, with 66.6% implementing corpora either directly or indirectly. The plans were mostly divided into 3–6 stages: warming-up, pre-teaching (practising target items), introducing the target items, and studying the grammatical features if necessary. Then, the corpora were used to complete activities or to study (analyse) the target items and perform exercises.

4.2. Implementation proposals of Corpora in the Classroom

The participants varied in their implementation in terms of how many stages and regarding the activities (deductive method to deliver the answers directly) or the discussions of the targeted items (inductive method to let the students deduce the rule or the answers). Out of the 60 participants, 40 implemented corpora into their plans, while 30 participants (out of 40) asked the students to use corpora in only one stage as a way to discuss the targeted items and raise awareness of the item collocations. The participants who included the use of

corpora in one stage described this in terms of asking students to search the corpora for the target items and then discuss the results. Only 10 out of 40 used corpora in two stages—to study the targeted items and their collocates and/or the related meanings to the items and/or the grammatical structures, and then to do the activities. Those who included the use of corpora in two stages mentioned that they would ask students to search the corpora for the target items, discuss the results, and then complete some activities. Therefore, they added another level of complexity to the implementation of corpora.

Most of the participants, about 55% (out of the 40), selected the British National Corpus (BNC) to use in their lesson plans, while only one chose the Corpus of Contemporary American English, and another chose Compleat Lexical Tutor. Some just mentioned ‘computer corpus’ or ‘corpus’ without specifying the corpus resources they intended to use in their lessons.³

Only two participants clearly mentioned the need to train students on the use of corpora before class. Overall, 33.3% planned traditional stages for their teaching, while the rest (66.6%) planned stages in which they included corpus use either directly, hands-on (searching the corpus), or indirectly by studying corpus extracts prepared by the teacher. Meanwhile, 87.5% used corpora directly in their lesson plans by asking the students to search the corpus, while 12.5% indirectly included corpus use, as they said they would prepare the materials for the students instead of the students accessing the corpora. Sixty-three per cent of the 60 participants wrote detailed lesson plans, including the topic, aims, target level and skill, focus, resources, timing, stages, and so on. Thirteen per cent of the 60 participants wrote detailed lesson plans to some extent. However, 23.33% of the lesson plans were not as detailed as the rest and missed information, such as the target level or detailed explanations of the stages.

Two participants did not include any activities. Therefore, the analysis of the activities included only 58 participants who designed activities. The analysis of the activities covered the number of activities, the types of activities (match, complete, or search), and whether the activities were clear in terms of the requirements (corpus-informed or traditional use of textbooks).

The number of activities per lesson varied from one to eight, with some participants including three activities per lesson (about 17.2%) and others including two activities (about 19.9%). The types of activities also varied, but most of the participants (about 46.5%) designed activities based on traditional methods, such as textbooks, while 46.5% designed activities that combined both corpus resources and workbooks (e.g. ‘Use the BNC to study the sentence, concentrating on learning, for example, a skill, network, form of communication, etc.’). In terms of designing activities, 6.8% of participants used corpus resources only. Of the 58 participants who presented activities, 81% offered clear designs and included direct requests that aligned with the aims of the lessons. See Appendix for examples of activities.

The same 60 participants who submitted end-of-course projects were also included in the homework analysis, which was based on whether homework was requested, whether the homework was corpus-based or traditional, and whether it targeted a specific skill and/or language system. Of these participants, 36.6% did not assign any homework, but 63.3% did assign homework in their lesson plans. Of the participants who assigned homework,

³ COCA and BNC were the main corpora used in the demonstrations in the course (all the details of the training materials and resources are available in Alruwaili (2023)).

48.3% included clear questions and requests, while eight participants did not formulate the questions clearly. Regarding the use of corpora in homework, most of the participants did not request its use; only three participants stated that a corpus should be used to complete the homework. See Appendix for examples of homework.

4.3. Challenges for Teachers in Integrating Corpora into Lesson Plans

The challenges of integrating corpora into lesson plans were explored through the open-ended questions that the participants answered after submitting their end-of-course projects. They were asked to evaluate their experiences after submitting the final projects. Their answers were analysed qualitatively by grouping them into themes and then calculating the frequency of each theme. The participants acknowledged the benefits of the experience and stated that lesson planning made the teaching easier and more organised. The participants also appreciated the use of corpora in the lesson plans as a first step towards using them in classrooms, and they said that corpora can support discovery learning through the search for information, helping with correcting information, building students' knowledge, and increasing students' active learning. The participants also pointed out challenges in using corpora in their classrooms, such as the need for more preparation, the need to train students, the shortage of time for application in periods, work pressure, and infrastructure (computer, internet, etc.). Others emphasised that students' proficiency levels affected the application of corpora, as corpora are not suitable for primary students due to these students' lack of language and computer skills. Some participants mentioned other factors that play a role in the application of corpora in classrooms, such as students' willingness and motivation, the education system, and the availability of consultants.

5. DISCUSSION

The teachers who took part in this study were motivated to learn about corpora and how they can be used in classrooms; hence, they attended an online training course and submitted lesson plans as end-of-course projects. Based on Callies's (2016) suggestion that online training courses "seem promising options" (p. 411), the training course was offered online to target a larger sample of teachers and to accommodate their schedules and commitments. One of the main contributions of this study is that one training course can be sufficient as a starting point for including corpora in the classroom, but there is still a need for hands-on training to master corpus linguistics and pedagogical skills, as technical skills do not seem an issue, particularly after the pandemic. A significant outcome has been the need to change the types of training that target in-service teachers, especially those with teaching experience, so the training offered to in-service teachers can be more guided in terms of direct application in the language classroom and the availability of readymade materials and activities. Teachers across the world are overwhelmed with workloads, so when they experience a new teaching method, they require consultation and support to hasten the process of implementing it; thus, guided and supported training is good way to encourage the application of corpora in everyday teaching.

The training course was useful and successful, according to the teachers (Alruwaili, 2023). Nevertheless, their evaluations were not sufficient to indicate the successful implementation of CL in the classroom. Thus, there is a need to challenge them to use corpora in their teaching contexts and not depend on their good intentions to use corpora in their future teaching, which supports the findings of previous studies (i.e. Abdel Latif, 2021; Ebrahimi & Faghih, 2016; Farr, 2008). Teachers need to be trained in corpus use and corpus-driven teaching and then either be challenged with relevant tasks, such as those in the current study, or observed during their teaching and first-hand implementation.

Leńko-Szymańska (2017) claimed that training should start with a “particular teaching context” (p. 234). This proved to be inadequate, since the current study revealed that, even when the teachers started by aligning their training with their current teaching contexts which was teaching English as a foreign language at the pre-tertiary education, they still required more training and time to comprehend the concepts of corpus-informed and corpus-driven teaching or instruction. The course was designed to help the teachers acquire the three skills that have frequently been emphasised in the literature: technical, corpus linguistic, and pedagogical skills (Callies, 2016; Leńko-Szymańska, 2017; Mukherjee, 2006). These skills were expected to be partially or fully reflected in the end-of-course projects. The teachers were able to design and submit full lesson plans for their teaching contexts. However, as Lenko-Szymanska (2017) pointed out in her study, one training course is not sufficient to enable pre-service teachers to acquire such skills or reach a satisfactory level of corpus use. The same result was obtained in the current study, which found that, even for the participants who were in-service teachers with considerable teaching experience, one training course did not develop the three skills to the level of expertise. Thus, to improve the application and implementation of CL in classrooms, long-term training should be established.

The analysis of the end-of-course projects demonstrated that the participants were able to design lesson plans that aligned with their teaching contexts and fit their teaching objectives. They clearly stated the aims of the lessons and identified the target level, revealing the teachers’ grasp of how to apply corpora in their classrooms. The participants selected the topics according to their students’ proficiency levels, which indicated that they were genuinely keen to implement CL in their teaching. The selection of topics followed common themes, such as ‘food’. This might have posed challenges in searching corpora to find relevant corpus-informed materials, leading to the participants not reaching a satisfactory level of corpus linguistic skills (i.e. the participants possibly did not consider the high number of hits they would obtain for such common themes and/or know how to control such big data or how to apply appropriate filtering options). Regarding technical skills, it was extremely difficult to judge whether the participants acquired proficiency in technical skills; most of the presentations emphasised that technical skills were not an issue because the pandemic had forced the use of online learning.

In designing their lesson plans, most of the participants chose vocabulary as a target for the use of corpora, which, as explained by Lenko-Szymanska (2017, p. 233), meant that they might not have gained enough confidence in analysing corpus data or, as Heather and Helt (2012) found, that the participants focused on vocabulary in their projects because they did not have sufficient skills in using concordance lines to identify lexico-grammatical patterns. However, it could be that vocabulary was the easiest target to search for in corpora because

the participants could look for different uses of the target vocabulary according to their lesson aims. This finding agrees with Zareva's (2017) study, in which the participants preferred to use corpus research knowledge to teach vocabulary and grammar (mostly collocations). Although most of the participants in the current study focused on vocabulary, some targeted lexico-grammatical features (about 27%) and grammar (about 10%). It should be noted that the participants possibly depended on the training they received during the sessions; either they chose target items that retrieved substantial amounts of data but did not use filters, or they simply did not acquire sufficiently sophisticated skills to enable them to manipulate the data. Thus, as Lenko-Szymanska (2017) and Ma et al. (2023) have claimed, one training course that covers various aspects of CL is not enough to develop corpus linguistic and pedagogical skills; instead, training should comprise a series of courses/sessions that cover subskills, such as analysis and search skills.

Regarding pedagogical skills, the participants reached a satisfactory level to the extent that they were able to design and create suitable corpus-informed materials based on information retrieved from corpora (extracts from the corpus, which were used by only a few participants) and corpus-based materials that incorporated the use of corpus data. Lenko-Szymanska (2017) argued that the acquisition of pedagogical skills depends on varied teaching materials that are methodologically sound and integrated with other teaching techniques and methods in specific teaching contexts. The participants in the current study achieved this to some extent and, contrary to Ma et al.'s (2023) findings, had no difficulty integrating corpora resources/activities with other resources, such as textbooks. Most of them thought that integrating traditional activities with corpus-based activities was the safest option to start with and had no difficulty doing so; however, others decided that using corpora as part of the teaching stages (introducing the targeted items) was enough and employed workbooks as the basis for activities. In addition, most of the activities focused on vocabulary, and few focused on grammar. The participants followed the same approach in assigning homework as they did for the activities in that they depended on traditional methods, such as textbooks, and only a few requested that corpora be used for doing the homework. Ma et al. (2021) suggested that, in designing a corpus-based lesson, there should be "a balance between the use of corpus resources and other concurrent resources/activities" (p. 24), and the participants in the current study instinctively followed this recommendation, possibly because of their teaching experience and pedagogical knowledge. The participants attempted to integrate corpus activities into lessons without changing the usual flow of their teaching process.

The teachers who implemented corpus-based teaching specified that they needed more time to finish such a lesson than the typical single period required for traditional teaching. Ebrahimi and Faghih (2016) and Farr (2008) found that time constraints were a challenge in exploiting corpora in classrooms, which is consistent with this study's finding that the participants integrated CL into specified periods of their lesson plans and pointed out time constraints as a disadvantage in their comments.

To summarise, the participants were good at grasping the concept of CL in classrooms and were willing to use it, as their submitted lesson plans showed. Nevertheless, gaps remained in their acquired skills, which may indicate a need for more hands-on training, since the teachers in this study received online training to save time and effort. Some sessions could be conducted on site with real students to provide teachers with opportunities

to use CL in authentic classes, thus helping them gain confidence and realistic views of the process before transferring their learning to other schools/classes. Chen et al. (2019, p. 9) emphasised the need for long-term training and follow-up support and stated that providing free teaching materials is essential for enabling teachers to successfully apply DDL in their teaching. Ebrahimi and Faghih's (2016) study revealed that their participants required considerable instructor support for online courses; however, in the current study, the participants asked for support on site through in-person availability of consultants in schools, which is contradict Boulton's (2011) finding that the participants were able to gauge the tools and applications by themselves. Ma et al. (2021) argued that "providing teachers with a variety of learning opportunities, including online collaborative learning, can energise and empower corpus-based teacher training" (p. 23). This was reflected in the current study through most of the participants collaborating in designing the lessons, possibly because they were familiar with each other (i.e. from the same schools). The results of the current study contradicted other research (Farr, 2008; Heather & Helt, 2012) reporting that the participants disliked the technical side of corpus work, possibly because the shift to online learning caused by the pandemic has enhanced the participants' technical skills. However, the participants reported other (more practical and profession-related) challenges, such as workloads or students' proficiency levels. The heavy workload was reported as a particular hindrance in implementing corpora in classrooms, which is consistent with other studies (Leńko-Szymańska, 2017; Lin & Lee, 2015; Ma et al., 2023).

6. CONCLUSION

The study's results revealed that the participants were, to some extent, able to acquire the skills needed to exploit corpora in classrooms, as posited by Callies (2016) and Mukherjee (2004): technical, corpus linguistic, and pedagogical skills. An evaluation of the participant-designed lesson plans revealed that most of them could design an appropriate corpus-based lesson, demonstrating good pedagogical awareness of how to integrate corpus activities and resources into their current teaching contexts. The results of this study are encouraging, suggesting that the continued integration of corpus-based instruction into language teacher education programmes should be promoted despite some identified challenges. The participants were able to implement corpora in their teaching contexts, but further training and guidance was needed. They were also able to integrate corpus-based activities into their teaching plans. Although some of them designed lessons using a traditional methodology, they used corpora to create activities and assign homework. The results were promising and provided significant information on how teachers can implement corpora in their daily teaching. However, the results also showed that, even when teachers are well versed in language teaching methodologies, they may still lack the ability to design suitable corpus-informed and corpus-based teaching materials, which could be remedied by long-term hands-on training. There is a need to train teachers using hands-on methods, possibly across a semester, and to give teachers the chance not only to attend the training but also to apply their learning in their classrooms.

7. REFERENCES

- Abad Castelló, M. (2019). Uso de corpus lingüísticos por y para profesores de español como lengua extranjera [Uses of corpora in Spanish language teaching]. *Revista Electrónica de Didáctica del Español Lengua Extranjera*, 31(7), 1–20. <https://www.culturaydeporte.gob.es/dam/jcr:39778ef0-dddf-422f-b3c2-a8f08730225b/07-uso-de-corpus-linguisticos.pdf>
- Ackerley, K. (2017). Effects of corpus-based instruction on phraseology in learner English. *Language Learning & Technology*, 21(3), 195–216.
- Alruwaili, A. (2023). An Online training course on the use of corpora for teachers in public schools. *The JALT CALL Journal*.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- Boulton, A. (2009). Data-driven learning: reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 35 (1), pp.81-106.
- Boulton, A. (2017). Corpora in language teaching and learning. *Language Teaching*, 50(4), 483–506. <https://doi.org/10.1017/S0261444817000167>
- Boulton, A. (2010). Learning outcomes from corpus consultation. In M. Moreno Jaén, F. Serrano Valverde, & M. Calzada Pérez (Eds.), *Exploring new paths in language pedagogy: Lexis and corpus-based language teaching* (pp. 129–144). Equinox.
- Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Goźdz-Roszkowski (Ed.), *Explorations across languages and corpora* (pp. 563–580). Peter Lang.
- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–172. <https://doi.org/10.1080/09588220902778328>
- Callies, M. (2016). Towards corpus literacy in foreign language teacher education: Using corpora to examine the variability of reporting verbs in English. In R. Kreyer, S. Schaub, & B. A. Güldenring (Eds.), *Angewandte Linguistik in Schule und Hochschule [Applied linguistics in secondary school and at university]* (pp. 391–415). Peter Lang.
- Callies, M. (2019). Integrating corpus literacy into language teacher education: The case of learner corpora. In S. Götz, & J. Mukherjee (Eds.), *Learner corpora and language teaching* (pp. 245–263). John Benjamins. <https://doi.org/10.1075/scl.92.12cal>
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 52(4), 460–475. <https://doi.org/10.1017/S0261444819000089>
- Chen, M., Flowerdew, J., & Anthony, L. (2019). Introducing in-service English language teachers to data-driven learning for academic writing. *System*, 87. <https://doi.org/10.1016/j.system.2019.102148>
- Crosthwaite, P. (2020). Taking DDL online: Designing, implementing and evaluating a SPOC on data-driven learning for tertiary L2 writing. *Australian Review of Applied Linguistics*, 43(2), 169–195. <https://doi.org/10.1075/ara1.00031.cro>
- Ebrahimi, A., & Faghih, E. (2016). Integrating corpus linguistics into online language teacher education programs. *ReCALL*, 29(1), 120–135. <https://doi.org/10.1017/S0958344016000070>
- Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation use in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning*, 1(1), 80–109. <https://doi.org/10.1515/jccall-2021-2004>
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25–43. <https://doi.org/10.2167/la414.0>

- Forti, L. (2019). Evaluating Data-Driven Learning Effects in the Italian L2 Classroom: Etic and Emic Perspectives Combined. *Educazione Linguistica Language Education*, 8(2), 363–378. DOI 10.30687/ELLE/2280-6792/2019/02/006
- Heather, J., & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, 20(4), 415–440.
- Latif, M. M. M. A. (2021). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices. *ReCALL*, 33(1), 34–48. <https://doi.org/10.1017/S0958344020000129>
- Leńko-Szymańska, A., & Boulton, A. (Eds.), (2015). *Multiple affordances of language corpora for data-driven learning* (pp. 1–14). John Benjamins. <https://doi.org/10.1075/scl.69>
- Leńko-Szymańska, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL*, 26(2), 260–278. <https://doi.org/10.1017/S095834401400010X>
- Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. In A. Turula, & B. Mikołajewska (Eds.), *Insights into technology-enhanced language pedagogy* (pp. 129–144). Peter Lang.
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217–241
- Lin, M. H., & Lee, J.-Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, 69(3), 264–274. <https://doi.org/10.1093/elt/ccv010>
- Lin, M. H. (2019). Becoming a DDL teacher in English grammar classes: A pilot study. *The Journal of Language Learning and Teaching*, 9(1), 70–82.
- Ma, Q., Chiu, M., Lin, S., & Mendoza, N. (2023). Teachers' perceived corpus literacy and their intention to integrate corpora into classroom teaching: A survey study. *ReCALL*, 35(1), 19-39. doi:10.1017/S0958344022000180
- Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, 1(1), 177–190
- Ma, Q., Tang, J., & Lin, S. (2021). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, 35(9), 2731–2760. <https://doi.org/10.1080/09588221.2021.1895225>
- Ma, Q., Yuan, R., Cheung, L. M. E., & Yang, J. (2022). Teacher paths for developing corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*, 1–32. <https://doi.org/10.1080/09588221.2022.2040537>
- McEnery, T., & Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2) (pp. 364–380). Routledge.
- Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor, & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 239–250). Rodopi. doi: 10.1163/9789004333772_014.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5–24). Peter Lang,
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- UNESCO. (2023, April). Education & Literacy. UNESCO Institute for Statistics. <http://uis.unesco.org/en/topic/international-standard-classification-education-isced>

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, 20(3), 159–179.

Zareva, A. (2017). Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1), 69–79. [https://doi.org/ 10.1093/elt/ccw045](https://doi.org/10.1093/elt/ccw045)

8. APPENDIX

Examples of activities include the following:

Vocabulary activities

1. Search the BNC corpus to study sentences concerned with *food, drink, eating (rice, juice)*. Then answer the following questions:

1.a. Rearrange these letters for each of the following words:

niasrg- emloehwla- cinagro- urofl. (grains- wholemeal- organic- flour)

2.b. Fill in the gaps with suitable words from the following:

Brown rice- Husk- Arborio (risotto rice) – Basmati

- sometimes referred to as Italian rice, and is really a short grain rice, although slightly longer than the traditional round grain pudding rice.
- is the everyday luxury rice. It is grown on the slopes of the Himalayas in Pakistan.
- unpolished and unprocessed grains of rice that still contains the bran.
- are the hard protecting coverings of grains of rice.

2. Search the corpus to study collocates of the word *eat*. Can you notice any similarities.

- ▶ Show the students that the words occurring after the verb is nouns (names of food).
- ▶ Show and explain to the students the two types of meanings (literal and figurative).

2.a. explain the difference between the two sentences:

- I have a mouthful of cider from my water-bottle and eat the remains of a chicken.
- It's a killing job. Little kids eat you alive, they want so much physically and emotionally.

3. Search the corpus to study the sentences with some adjectives and to look for any interesting features.

- Show the students that the suffixes occurring after the nouns, or the verbs can change them into adjectives. (In this activity the teachers provide the search words which are slow-slowly, quick-quickly, beauty-beautiful, create-creative, talk-talkative).

Grammar activities

1. Use computer corpus to learn about auxiliary verbs and how to use them in cases of acceptance and rejection. Search for the auxiliary verbs such as can and could and study the sentences.
 - 1.a. Classify the sentences to acceptance or rejection:
 - You can not borrow the book.
 - She could speak English.
 - You can not be serious.
2. Use computer corpus to search the verbs occurring after is and are, and the subjects before is and are. Then discuss the agreement of the verb of be.
 - 2.a. Fill in the gaps with is-are, then explain your choice:
 - Over the next decade a global approach going to be essential.
 - The team caring for growing numbers ill on the South Coast.
 - We seeking to find out what local people want.
 - Many churches responding.
3. Study the grammatical features ‘a, an’ with singular and numeric names in the following extracts (the extracts are from BNC), then summarize the rule:

You are not expected to be	a	nurse
Margaret Gillies, currently has	a	team of 20 volunteers
In this way, with	a	little money but much commitment, a great deal can be done
Catherine, who is on	a	three-month contract with ACET, is looking forward to caring for people
If I can't take him, he will have	a	two-hour wait for an ambulance.
Tony has been given	an	appointment at the local hospital.
There was	an	error in the appeal on behalf of Taiwanese prisoner Huang Hua.
Ui-gyun continues to serve	an	eight-year sentence for espionage.
Yet Prague was always	an	international city, whose culture was dependent on native and external traditions.
The content of	an	article will tend to be determined by its length

- ▶ The rule for using a is:
- ▶ The rule for using an is:

Examples of homework include the following:

Homework 1

1. Use Learner Corpora, help the students input the texts, and then ask them to correct the existing errors.

Homework 2

1. What is the correct form of the verb that comes after the pronouns? (The agreement between the verb to be and the subject)