



Corpus literacy development: Three teachers' stories

*Cathryn Bennett**, Trinity College Dublin

Abstract

Corpus and applied linguists have reported on the benefits that teachers gain from teaching with language corpora as early as the 1990s; however, recent studies confirm that few teachers use corpora in their classrooms. In attempts to change this reality, some researchers have called for corpus literacy training programmes to guide teachers in using corpora/corpus tools to design their typical classroom tasks. A training programme was built around this idea. This paper outlines three teachers' corpus literacy development during the training programme: a teacher with previous experience teaching with corpora, a teacher with knowledge of but no experience teaching with corpora, and a teacher who reported no knowledge or prior experience teaching with corpora. To provide an in-depth perspective, a qualitative thematic analysis was completed with themes emerging from the dataset. Findings show that teachers viewed the training positively, that they incorporated corpus-based materials in their classes, and that they were using corpora in their teaching three months following the programme. Further research is called for which highlights teachers' voices in their corpus literacy development.

Keywords: *In-service Teacher Training; Corpus Literacy Development; Materials Design; ESL/EFL Teachers*

Language(s) Learned in This Study: *English*

APA Citation: Bennett, C. (2025). Corpus literacy development: Three teachers' stories. *Language Learning & Technology*, 29(1), 1–14. <https://doi.org/10.64152/10125/73632>

Introduction

Researchers have emphasised the need to hear teachers' own voices in corpus literacy trainings (Kavanagh, 2021b). In this paper, I present the experiences of three in-service teachers from the private ELT sector who participated in a corpus training programme which acknowledges the challenges faced by this cohort as those teachers who often experience high teaching demands and little time for professional development as suggested by Mercer (2020) and Stanley (2016). Corpus literacy has been defined as understanding how to use corpus tools to analyse sentences within a corpus to make conclusions about language use (Farr & Leńko-Szymańska, 2024).

A pivotal study which prompted the current investigation was carried out by Frankenberg-Garcia (2012) in which she described the constraints of everyday teachers, including set coursebooks, syllabi and lesson plans, as being barriers towards the development of corpus literacy. She advocates for short, corpus-based activities that teachers can easily integrate into their pre-set lesson plans as a means of training teachers to learn to use corpora. She provides several corpus-based/informed exercises that teachers can use to teach language skills in the classroom. She argues that learners do not necessarily need training, or even a full explanation of what corpora are, to benefit from them as a learning tool.

To carry forward this suggestion, a training programme was designed to train in-service teachers to learn to use corpora/corpus tools and design such short, corpus-based activities with the guidance of the trainer. The guiding research question for the study:

* **Corresponding Author:** Cathryn Bennett, cabennet@tcd.ie

How do teachers demonstrate corpus literacy development in the training programme regardless of their extent of prior corpus-awareness?

Literature Review

The benefits of students becoming corpus literate have been outlined in a wealth of research to date including students becoming more autonomous learners (Braun, 2004), students being exposed to authentic language, and learning collocations and colligations (O’Keeffe et al., 2007). Several textbooks (e.g., *From corpus to classroom* by O’Keeffe et al., as well as *Using corpora in the language classroom* by Reppen) were published many years ago, yet corpora remain on the periphery of mainstream teaching practice (Leńko-Szymańska, 2022). Research has elucidated the benefits that teachers reap from learning how to use corpora/corpus tools and designing materials with corpora, such as fostering student engagement and autonomy, bringing authentic language into the classroom, and more efficient lesson planning (Bennett, 2024; Basal et al., 2024). Several corpus literacy training models have been published such as Farr (2008), Leńko-Szymańska (2015), and Bennett and Uí Dhonnchadha (2023), with similar training components including hands-on demonstrations of corpus tools, reflective tasks and materials design.

Corpus-based language pedagogy (CBLP) focuses on how teachers can learn to use corpus tools in “real classroom settings” (Ma et al., 2021, p. 2735); however, research currently lacks the guidance in not only how to use corpus tools but also how to incorporate these tools in language learning classroom activities. In their recent article, Farr and Leńko-Szymańska (2024) describe three categories of research for common integrating corpora into the language classroom, such as instruction that focuses on creating activities with corpora, of which the current study falls into. Corpus literacy studies by Abdel-Latif (2020) and Bennett (2024) demonstrate the value of planning short corpus activities when learning to use corpus tools, while Basel et al. (2024) and Bennett (2024) highlight teachers’ reported sense of self-efficacy when using corpora for language teaching. Table 1 presents previous studies with in-service, English as a Foreign Language (EFL) corpus literacy training findings that are pertinent to the present study. Three are outlined due to length constraints; however, there is a growing body of research in this topic, including Zareva (2016), Ebrahimi and Faghih (2017), Poole (2020), and Xodabande and Nazari (2022).

Participants in Kavanagh’s (2021a) study were Norwegian primary and secondary school teachers, 39% ($n = 193$) of whom reported “little or no idea” of corpora (Kavanagh, 2021a, p. 11). Following the training, teachers reported benefiting from learning to use corpora by reflecting on how well they knew the vocabulary words they were teaching, and teaching specific aspects of vocabulary (i.e., collocates and frequency). However, teachers felt corpora would be too challenging for students to use in class and that they lacked confidence in using them, which has also been noted in research (e.g., Leńko-Szymańska, 2014). The relatively short training duration may explain teachers’ lack of confidence in designing corpus activities, with teachers confirming they were able to learn the basics of using corpora. Finally, there was some disagreement amongst teachers regarding whether authentic language was beneficial or “chaotic” for students (Kavanagh, 2021a, p. 17) and that COCA was not very user-friendly.

While Ma et al.’s (2022) study focused on two in-service teachers, their case study approach enabled them to identify aspects of learning that resulted in teachers successfully building and teaching corpus-based lessons. By analysing and observing them teach with corpora in class, they found that teachers who were highly competent in pedagogy and/or proficient with corpora were best equipped for integrating corpora into their classrooms. Additionally, both teachers had been teaching for more than 10 years, leading them to believe that highly experienced teachers are more likely to try corpus tools. The authors recommend that corpus training programmes should include publicly available corpus websites, consider learners’ needs when designing materials to encourage student engagement with corpora, and use hands-on corpus activities followed by peer or individual consultation to enable teachers to reflect and learn from learning to use corpora.

Table 1*Notable Previous In-service Corpus Literacy Trainings*

Researchers, Location of study	Duration	Corpora	Relevant findings
Kavanagh (2021a), Norway	Approx. 4.5 hours (over 17 weeks)	SKeLL, COCA, Netspeak	<p>Teachers' preference towards free corpora/corpus websites such as SKeLL.</p> <p>Perceived benefits included teaching aspects of vocabulary (collocations, idioms, frequency).</p> <p>Teachers' lack of confidence using corpora, calls for trainers to show teachers how to use corpus tools to design common classroom activities.</p>
Ma et al. (2022), Hong Kong	6 hours (over 4 weeks)	LexTutor, COCA	<p>Use of publicly available corpus websites and prepare students in advance for potential technical issues.</p> <p>Identify students' needs to encourage engagement with corpora.</p> <p>Hands-on training, along with follow-up consultation, aid teachers' reflection/learning to use corpora.</p>
Basal et al. (2024), Türkiye	24 hours (over 3 days)	COCA, Sketch Engine, UAM tool, LexTutor	<p>Corpora help teachers to design materials for different levels/language skills.</p> <p>Teachers felt more effective in lessons.</p> <p>Teachers may want to explore corpora on their own, before using it with their students.</p>

Note. The Corpus of Contemporary American English (COCA)

Finally, Basal et al.'s (2024) study included 29 in-service teachers from state schools in Türkiye. Similar to Kavanagh's (2021a) study, most teachers were familiar with corpora in name but had no prior teaching experience with corpora/corpus tools. Following the training programme, teachers viewed corpora positively feeling that they were a valuable teaching tool for designing tailored materials according to students' levels and needs. They also reported that corpora helped them to design more effective lessons, promote autonomy, and bring authentic language into their classroom. Their concerns regarding corpus-use in class centred on, similar to Kavanagh (2021a), whether corpus-based activities were suitable for lower-level learners and how trainers could build teachers' confidence using corpus tools (i.e., giving them time to initially explore the tools without their students).

The findings of these studies bear importance to the present study. In particular, the purpose of this study is to demonstrate how EFL teachers - regardless of prior knowledge of/experience with corpora - learned to use corpus tools and design materials to meet their learners' needs. This is elaborated on in the Findings section wherein the experiences of three teachers will be illustrated for the reader. Next, I will outline the methods and analysis of the dataset collected.

Methodology

Participants

Ethics approval was obtained from the author's university where the research was conducted. Fifteen respondents from the scoping questionnaire who expressed an interest in the training programme and whose availability matched the trainer's availability were emailed information explaining the study and a consent form to participate. Ultimately, seven teachers participated, their demographic information is provided in [Table 2](#). Most participants in the training programme were teaching online at the time due to the Covid-19 pandemic.

Most participants had completed postgraduate education and a pre-service teacher certification. The CELTA is a pre-service teacher certification programme by Cambridge English, which offers a four-week intensive programme which costs less and is of shorter duration than formal teacher education programmes. The DELTA is an in-service teacher certification programme. A CELTA certificate is a typical requirement for teaching English overseas; therefore, it is not surprising that most participants have this certification.

Most teachers (6 out of 7) had been teaching for more than 10 years; teachers had a varying number of teaching hours per week, and most (4 out of 7) were teaching online at a university. All reported being comfortable with using technology in the classroom, (e.g., online dictionaries, Google search engine, websites with practise exercises, and using YouTube).

[Table 3](#) shows teachers' reported awareness of corpora in a scoping questionnaire which was distributed online via a Qualtrics link over social media applications in November 2021. It collected information about teachers' previous familiarity and use of corpora in their classroom including which corpora they had used and which language skills and learner levels they had used corpora with. Participants have been given pseudonyms to personalise their experiences for the reader. For this particular paper, the experiences of three cases are showcased for readers. Three teachers with differing levels of corpus experience are highlighted in [Table 3](#), so that readers at different stages of corpus literacy can imagine their own journey and identify with one of these teachers as explained in the Findings section.

As shown in [Table 3](#), although participants were familiar with corpora, only one had used a corpus in their classroom. Two teachers reported designing classroom materials with corpora.

The Training Programme

The training programme was comprised of three Teaching English to Speakers of Other Languages (TESOL) paradigms: (a) needs analysis (Long, 2005), (b) exploratory practice (Allwright, 2003; Hanks, 2017), and (c) reflective practice (Farrell, 2012, 2022). Needs analysis was selected to help teachers to design learner-centred materials based on their students' needs. In matching learners' needs to corpus tools, teachers can also begin the corpus literacy skill of understanding how to build activities using corpus data. Exploratory practice was chosen because it emphasises teachers' need to explore their curiosity with corpus tools. Last, reflective practice, a common element in teacher training programmes, was selected as it enables teachers to learn from classroom experiences (Breyer, 2009; Ebrahimi & Faghih, 2017).

Table 2*Participants' Background Information*

Background information	Teachers' target student levels	Learner level	Participants' teaching environment	Number of hours teaching per week	Number of years teaching	Highest academic qualification	Highest teaching qualification
Participant 1	Tertiary	CEFR B1/B2	No response	11-20	10+	PG	PG
Participant 2	Secondary	CEFR B1	In school	21-30	10+	UG	DELTA
Participant 3	Tertiary	CEFR B2	Online	11-20	10+	PG	DELTA
Participant 4	Secondary	CEFR B2	Online	6-10	10+	UG	DELTA
Participant 5	Tertiary	CEFR B1	Online	21-30	10+	PG: PhD	CELTA
Participant 6	Primary	CEFR A2/B1	In school	11-20	10+	UG	CELTA
Participant 7	Tertiary	CEFR B1/B2	Online	6-10	4-7	PG: PhD	DELTA

Note. Undergraduate (UG), postgraduate (PG), Common European Framework of Reference for Languages (CEFR)

Table 3*Participants' Knowledge of Corpora Prior to the Training Programme*

Questionnaire item	Participant 2: Luka	Participant 3: Aubrey	Participant 6: Makenzie
Are you familiar with the term 'corpus'?	Yes	Yes	Yes
Have you used a corpus website for language teaching purposes?	Yes	Yes	No
Which corpus websites have you used?	Sketch Engine, SKeLL	LexTutor	-
Do you use a corpus to design materials for class?	Yes	Yes	-

Teachers in the training programme learned to use two corpus websites: the Corpus of Contemporary American English (COCA) and corpora available in Sketch Engine's free trial. These corpus websites were chosen because they are publicly available.

The materials in the programme included a trainee handbook and weekly PowerPoint slides, designed by the trainer (and author), which included a stepwise account of how to use corpus tools and design optional corpus activities based on their learners' needs (see [Appendix A](#)). Participants also received a reflective journal word document to record their reflections before and after teaching with a corpus in their classroom.

Implementation

The training programme took place over a 5-week period in Spring 2022 on the videoconferencing platform Zoom. The first week was introductory, outlining how to collect their students' needs for materials design with the subsequent four weeks presenting a corpus tool to design activities. See [Appendix B](#) for a sample corpus activity. Training sessions were approximately 50 minutes, with an optional additional 30 minutes for building activities with the trainer. Training sessions were recorded and sent to participants each week.

Interviews occurred one week after the training programme concluded and were conducted by an independent scholar who also worked at the university where the research was conducted to avoid researcher bias, as suggested by the ethics committee. A link to the follow-up survey on Qualtrics was emailed to participants three months after the training programme concluded (see [Appendix C](#) for questions). The design and questions of the follow-up survey closely followed the scoping questionnaire to determine whether the training programme was effective in developing teachers' corpus literacy.

Data Analysis

Qualitative data in the form of weekly reflective journals and post-training interviews were used to investigate teachers' experiences in the training programme. A thematic analysis as outlined in Braun and Clarke (2022) was carried out. Reflective journals were coded by data-driven methods wherein they were read several times and initial child codes were created. This method of coding was chosen as the structure of the reflective journals did not explicitly mention theories within the training programme. Secondly, data from post-training interviews were coded using theory-driven coding (i.e., interview questions were grouped around needs analysis, exploratory practice, and reflective practice). Thirdly, all child codes were

examined and combined under larger parent codes to establish potential themes. Codes were verified with another colleague in the university where the research was conducted. Ultimately, 58 child codes were identified, and seven parent codes were established within the dataset. As shown in Table 4, three overarching themes encapsulated the seven parent codes: *training matters*, *learning to use corpora*, and *the (slow) learning curve*.

Table 4

Parent Codes and Themes Identified

Data instruments	Themes	Parent Codes	No. of Child Codes
Reflective journals and post-training interviews	Training matters	Training evaluation	18
		Training objectives	6
	Learning to use corpora	Designing activities and teaching with corpora	17
		First impressions of corpora	5
		The (slow) learning curve	4
	The (slow) learning curve	Corpus literacy behaviours	4
		Teacher beliefs regarding teachers and teaching	4
		Ideas for future training sessions	4

Broadly speaking, the largest theme, *training matters*, includes participants' comments of being supported in the training sessions when learning to use a new tool and whether teaching corpora activities met their lesson objectives. The second theme, *learning to use corpora*, contains teachers' initial impression of corpora and how they designed and taught corpus activities in their classroom. The final theme, *the (slow) learning curve*, reflects teachers' acknowledgement of the slow process of learning to use new corpus tools.

Additionally, data from the scoping questionnaire and follow-up survey were analysed using descriptive statistics, such as percentages derived from numerical data of participants' levels of comfort using technology and confidence in using corpora before and after the training programme. In the Findings section that follows, I provide examples of these themes and numerical data which illuminate teachers' experiences through the three focal case studies.

Findings

This section contains the data of three EFL teachers to answer the research question: *How do teachers demonstrate corpus literacy development in the training programme regardless of their extent of prior corpus-awareness?* It begins with the teacher who reported having the most experience with using corpora and concludes with the teacher who reported the least.

Luka: Previous Experience with a Corpus

Luka reported in the scoping questionnaire his highest academic qualification was an undergraduate degree, but his highest teaching qualification was a DELTA certificate. He reported teaching close to 30

hours per week in a language school. He mentioned being shown Sketch Engine as a part of his DELTA qualification but mainly used SKeLL¹. He reported his level of confidence in designing materials with a corpus at 66% (more confident). His students' needs were based on grammatical points such as continuous and perfect tenses and linking devices.

In the second week of his reflective journal, Luka commented that using the Chart tool in COCA provided new opportunities for his students to learn new vocabulary, similar to Kavanagh (2021a). "They were able to pick up new vocabulary relevant to their area of interests ... or review known vocabulary items that they do not frequently use." This demonstrates his corpus literacy development as he had no prior exposure to COCA prior to the training programme.

In Luka's post-training interview, he was asked about his *First impressions of corpora* and replied:

"It is quite technical...I think it takes a while, until you can say 'I can use corpora to answer your question.'"

When asked if learning to use a corpus has affected his teaching, a sub-theme of the data, he commented, "Not yet, but I think it can grow on me. Yeah, I think it has potential to change the way I teach." Here is the theme *learning to use corpora* as Luka becomes more familiar with new corpus tools and how they can support his teaching.

Luka reported continuing to use corpora in the follow-up survey three months later. His confidence in designing corpus materials decreased slightly from 66% to 60% following the training sessions. This could be due to being exposed to more corpus tools with not enough time to practice, given his high teaching demands, as suggested in *the (slow) learning curve* theme. When asked if he would be interested in learning to use corpora more in class, Luka commented that he wanted to learn additional ways to incorporate corpora in authentic materials design.

Aubrey: Some Experience with a Corpus-Based Website

Aubrey reported in the scoping questionnaire that her highest academic qualification was a postgraduate degree and her highest teaching qualification was a DELTA certificate. She worked 11-20 hours per week teaching a translation class online at an Italian university. She taught herself to use corpora through various online tutorials and noted experience with only LexTutor. She rated her confidence level in using corpora at 37% (slightly confident) in the scoping questionnaire. Her learner needs consisted of mainly grammatical points such as conditional use, language for conceding and comparisons.

In her reflective journal, Aubrey felt corpora exposed her students to authentic language. By the end of the training programme, she recognised she was using corpora with her students more autonomously, asking them to check new vocabulary words in COCA. This reflects the theme of *Training matters*.

In her post-training interview, Aubrey was asked whether collecting her learners' needs helped her learn to use corpora, corresponding to the parent code of *Training evaluation*, which she firmly agreed:

"I think so because it gave us a starting point. Otherwise, without a need, I wouldn't have known what to look for in the corpus."

Aubrey mentioned using corpora to achieve multiple learning aims, such as encouraging autonomy and lexical development, as well as checking her own intuition when marking her students' assignments, demonstrating use of corpora inside and outside of her classroom. This supports the theme of *Training matters* as she integrated corpora more into her teaching practice after training.

In the follow-up survey, Aubrey continued to use a corpus for language teaching after training sessions, and her confidence level in designing corpus materials rose from 40% to 60%. The theme that *Training matters* is evident in her response to the last question, as she reports she is interested in more training with corpora to build her confidence.

Makenzie: Corpus Awareness, No Prior Corpus Experience

Makenzie reported in the scoping questionnaire that her highest academic qualification was an undergraduate degree, and her highest teaching qualification was a CELTA certificate. She reported teaching approximately 11-20 hours per week in a primary school in France and that she was familiar with the term ‘corpus’ but had not used a corpus for language teaching. Her learners’ needs consisted of teaching comparisons, superlatives, and phrases to express similarities and differences.

As Makenzie reported no previous experience using corpora or corpus websites, her reflections on the training programme were critical. In the first week, it was not surprising that Makenzie reported struggling with performing corpus searches. Makenzie’s corpus literacy grew steadily over the following weeks as she used the tools presented in the training sessions which boosted her confidence with corpora suggesting the larger theme of *Learning to use corpora* which is both exciting and challenging at the outset.

In her post-training interview, Makenzie commented on her *First impression of corpora*, stating that she was “blown away that I hadn’t come across it before”. She commented on corpora’s benefit of bringing authentic language into her lessons:

“Seeing a curse word in the sentences...you can’t always control that,
...but it’s good because it’s still the language that we use, isn’t it?”

She mentioned she could censor this by choosing the sentences herself beforehand to “better prepare” herself, demonstrating the *Learning to use corpora* theme. Makenzie added that corpora was a new tool to use to keep her young students’ attention.

Although Makenzie had no prior experience using corpora and corpus websites, she adapted quickly to using corpora in her classroom. While originally reporting “no” to having used a corpus website for teaching and using a corpus to design class materials, Makenzie used COCA and Sketch Engine following the training sessions and placed her level of confidence in designing corpus materials at 80% in the follow-up survey. This supports the theme that *Training matters*, as Makenzie had no prior experience to draw from. She responded “yes” to learning more about corpora, suggesting the theme *The (slow) learning curve*.

Discussion

In this study, all three teachers demonstrated corpus literacy skills development, though logically, this depended on their respective starting points of corpus-awareness. For example, Luka and Aubrey began the training programme with basic corpus literacy skills such as understanding what a corpus is and corpora’s limitations and strengths.

Luka demonstrated his corpus literacy development in learning to perform searches with a new corpus: COCA. In using the Chart tool, he found it provided more authenticity and autonomy to the lesson by enabling students to look deeper at collocations and identify words that interested them. Indeed, in having some prior exposure to corpora, participating in the training programme likely helped to solidify corpora’s limitations for language teaching. He explained this by saying that he felt corpora was a tool that was effective in teaching some things, such as vocabulary, but that other teaching tools, such as timelines, would be more effective when teaching tenses. In this way, he acknowledges that corpora can be a helpful tool but not the only tool in the classroom. In-service teachers in Kavanagh’s (2021a) study also reported preferring to use corpora to teach vocabulary, suggesting that this finding aligns with previous research.

For Aubrey, one means of demonstrating her corpus literacy development was the decreasing time she spent planning corpus activities. In week one of the training programme, she reported that planning a corpus activity took her around 20 minutes; however, this decreased to 10 minutes by the end of the

training programme. This suggests she became faster at the more complex corpus literacy skills of reading, analysing, selecting, and critically performing corpus searches in her free time. She used both COCA and Sketch Engine, suggesting that she learned to use two corpus interfaces for designing lesson plans, with her preference being COCA due to cost, similar to Makenzie. Like Luka, Aubrey learned to design activities that required her students to become more autonomous. While Luka asked his students to select and write down collocations which interested them, Aubrey thought more broadly, often asking her students to look up the meaning of a word using COCA's Word tool, instead of checking with her. Braun (2007) also found corpora supported more autonomy in the classroom. Out of the classroom, Aubrey stated that she "double-checks" words that she thinks are unfamiliar to her in COCA when marking assignments.

Makenzie, who arguably experienced the most development in the training programme, reflected positively about her corpus literacy development, stating that the sessions were "so useful" that she planned to share what she had learned with her colleagues. She mentioned how corpora have become a tool that she can switch to when she needs to change her teaching routine, something that she notes as especially important when teaching young learners. Makenzie commented on her frustrations with performing searches in COCA. Given her inexperience with corpora before the training programme began, an idea of separating the more basic skills with learning to use corpora from planning activities with corpus data has been suggested (also see Ma et al., 2021). This is a possibility for future training sessions, perhaps having one training group for those with little to no prior corpus experience, and another group with trainees who have prior exposure to corpora. This could allow the trainer to focus on providing more guidance to the former group.

All teachers reported that corpora brought authenticity to their lessons, which aligns with previous research findings (e.g., Basal et al., 2024; Farr, 2008). For example, Luka commented that corpora brought authenticity to the lesson by asking his students to look at authentic sentences in the corpus, which could explain, without him needing to explain in words, where contrastive linkers could be placed within the sentence. Aubrey commented that the untidiness of corpus sentences reminded her and her students that they were using authentic language. However, this authenticity could be a challenge too at times, when it distracts the students or conflicts with the grammar she needs to teach for standardised tests. Makenzie commented on the slight tension of liking the authenticity that corpora bring to her lessons, but also felt the need to censor profanity when having her young students search through a corpus. The idea of selecting corpus sentences for lesson planning was also a preference of teachers in previous research (e.g., Xodabande & Nazari, 2022). Yet Makenzie also acknowledged that finding potentially inappropriate words is the reality of using authentic language.

Despite the affordances and insights that were gained from instituting this training, the training and the study design have some limitations. In particular, teachers were not provided pedagogical resources for designing corpus-based activities, such as design principles as suggested in previous research (e.g., Ma et al., 2021). Secondly, although teachers participated in training sessions together, little interaction occurred between participants. This was likely to minimise the length of the online meeting, and since some of the experienced in-service teachers were confident in designing activities on their own. A final limitation was that many reflective journals were submitted with only partial completion, suggesting that this data instrument may not be suitable for private in-service teachers with high teaching demands. Rather, video reflections may be a more suitable solution.

Conclusion

Teachers in the three case studies highlight the benefits that corpora brought to their classroom, including authenticity and learner autonomy. They also experienced challenges related to the time needed to become familiar with corpora/corpus tools. All teachers demonstrated an understanding of essential corpus literacy skills, such as understanding the tools within a corpus interface, and how language can be selected and utilised in building activities for language learning. Teachers also reported continued use of

corpora three months after the training programme concluded and expressed an interest in learning more about corpora.

Though this paper focuses on detailing the experiences of three teachers, all seven experienced being on a learning journey. Indeed, further research into tailoring training programmes to teachers' existing corpus experience is an area worth more investigation to close the gap between corpus research and classroom practice.

Notes

1. <https://www.sketchengine.eu/skell/>: a minimalistic corpus query interface.

References

- Abdel-Latif, M. M. A. (2020). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practises. *ReCALL*, 33(1), 34–48. <https://doi.org/10.1017/S0958344020000129>
- Allwright, D. (2003). Exploratory Practise: rethinking practitioner research in language teaching. *Language Teaching Research*, 7(2), 113–141. <https://doi.org/10.1191/1362168803lr118oa>
- Basal, A., Akbaş, E., & Bal-Gezegin, B. (2024). Corpus literacy training for in-service English language teachers. *Profile Issues in Teachers Professional Development*, 26(1), 65–80. <https://doi.org/10.15446/profile.v26n1.106854>
- Bennett, C. (2024). *Investigating the experiences of in-service English language teachers in the use of language corpora for teaching purposes: An international action research study*. [Unpublished doctoral dissertation]. Trinity College Dublin.
- Bennett, C., & Uí Dhonnchadha, E. (2023). Becoming corpus literate: An in-service EFL teacher education framework for integrating corpora into EFL teaching. *Journal of Applied Corpus Linguistics*, 3(1), Article 100048. <https://doi.org/10.1016/j.acorp.2023.100048>
- Braun, S. (2004). Designing and exploiting small multimedia corpora for autonomous learning and teaching. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom: Selected papers from the sixth international conference on language corpora (TaLC 6)* (pp. 29–46). https://doi.org/10.1163/9789401203906_004
- Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3), 307–328. <https://doi.org/10.1017/S0958344007000535>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage.
- Breyer, Y. (2009). Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153–171. <https://doi.org/10.1080/09588220902778328>
- Ebrahimi, A., & Faghih, E. (2017). Integrating corpus linguistics into online language teacher education programs. *ReCALL*, 29(1), 120–135. <https://doi.org/10.1017/S0958344016000070>
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25–43. <https://doi.org/10.2167/la414.0>
- Farr, F., & Leńko-Szymańska, A. (2024). Corpora in English language teacher education: Research, integration, and resources. *TESOL Quarterly*, 58(3), 1181–1192. <https://doi.org/10.1002/tesq.3281>
- Farrell, T. S. C. (2022). *Reflective practice in language teaching*. Cambridge University Press.

- Farrell, T. S. (2012). Reflecting on reflective practice: (Re)visiting Dewey and Schon. *TESOL Journal*, 3(1), 7–16. <https://doi.org/10.1002/tesj.10>
- Frankenberg-Garcia, A. (2012). Integrating corpora with everyday language teaching. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching language corpora* (p. 33–51). Masaryk University Press.
- Hanks, J. (2017). Integrating research and pedagogy: An exploratory practice approach. *System*, 68, 38–49. <https://doi.org/10.1016/j.system.2017.06.012>
- Kavanagh, B. (2021a). Bridging the gap from the other side: How corpora are used by English teachers in Norwegian schools. *Nordic Journal of English Studies*, 20(1), 1–35. <https://publicera.kb.se/njes/article/view/24331>
- Kavanagh, B. (2021b). Norwegian in-service teachers' perspectives on language corpora in teaching English. *Nordic Journal of Language Teaching and Learning*, 9(2), 1–17. <https://doi.org/10.46364/njltl.v9i2.933>
- Leńko-Szymańska, A. (2022). Training teachers and learners to use corpora. In R. R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 509–524). Routledge. <https://doi.org/10.4324/9781003002901>
- Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. In A. Turla, B. Mikołajewska, & D. Stanulewicz (Eds.), *Insights into technology-enhanced language pedagogy* (pp. 129–144). Peter Lang.
- 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>
- Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.
- 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., & Yang, J. (2022). Teacher paths for developing corpus-based language pedagogy: A case study. *Computer Assisted Language Learning*, 37(3), 461–492. <https://doi.org/10.1080/09588221.2022.2040537>
- Mercer, S. (2020). The wellbeing of language teachers in the private sector: An ecological perspective. *Language Teaching Research*, 27(5) 1054–1077. <https://doi.org/10.1177/1362168820973510>
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Poole, R. (2020). Corpus can be tricky”: revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641. <https://doi.org/10.1080/09588221.2020.1825095>
- Stanley, P. (2016). Economy class? Lived experiences and career trajectories of private-sector English-language-school teachers in Australia. In P. Haworth & C. Craig (Eds.), *The career trajectories of English language teachers* (pp. 174–185). Symposium Books.
- Xodabande, I., & Nazari, M. (2022). Impacts of a corpus linguistics course on in-service EFL teachers’ corpus literacy. *CALL-EJ*, 23(1), 318–346. <https://callej.org/index.php/journal/article/view/383>
- Zareva, A. (2016). Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1), 69–79. <https://doi.org/10.1093/elt/ccw045>

Appendix A. Learner Needs Per Teacher

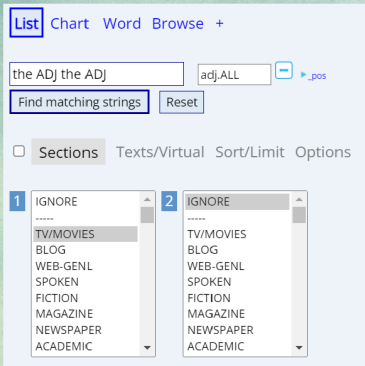
Participant	Learner needs
Luka	<ul style="list-style-type: none"> - Continuous and perfect tenses - Linking words - Conjunctions – making longer sentences
Aubrey	<ul style="list-style-type: none"> - Conditionals – all types, especially 3rd conditionals and mixed - I wish + verbs - Comparatives: far/much bigger than ... / a bit friendlier - Concession linking devices: despite/in spite of, although, even though, though
Makenzie	<ul style="list-style-type: none"> - Continuous forms of the present and past perfect - Comparative and superlative language: not as X as, the X the X, one of the X - Expressing degrees of similarity / difference ; similar to vs different from - Conditionals - Disaster management

Appendix B. Sample Corpus Training Materials

Superlatives

B1 activity: Using the theme of the lesson or photos from students/the internet, Ss use the superlative structure (one of the *) to describe what they see.

B2 activity: Investigate frequently used adjectives in the following structure. →



The screenshot shows a search interface with the following elements:

- Search bar: "the ADJ the ADJ"
- Filter: "adj,ALL" with a dropdown arrow and "pos" label.
- Buttons: "Find matching strings" and "Reset".
- Options: "Sections", "Texts/Virtual", "Sort/Limit", "Options".
- Two dropdown menus labeled "1" and "2", both set to "IGNORE".
- Dropdown menu options: "TV/MOVIES", "BLOG", "WEB-GENL", "SPOKEN", "FICTION", "MAGAZINE", "NEWSPAPER", "ACADEMIC".

Appendix C. Follow-Up Survey Questions

Note: Drop-down responses were taken from the original scoping questionnaire for participants' convenience in completing the follow-up survey.

1. What did you find most effective in the training sessions? [drop-down menu responses]
 - PowerPoint
 - Recordings of training sessions
 - I felt comfortable that the trainer was a language teacher like myself.
 - Other
2. What did you find more effective outside of the training sessions? [drop-down menu responses]
 - Using a needs analysis to learn to use corpora
 - Performing searches
 - Reflecting on planning and teaching with corpora in my class
 - Other
3. Which online resources do you use while teaching in the classroom? (write-in response)
4. How comfortable are you with using technology in the classroom? (Please select 1-5, with 1 being 'very comfortable' and 5 being 'not comfortable at all'.)
5. Have you used a corpus website for language teaching and planning? (Yes/No)
6. Which corpus websites have you used? (write-in response)
7. Do you currently use a corpus for language teaching in your classroom? (Yes/No)
8. Have you printed parts of a corpus for language teaching in class? (Yes/No)
9. Have you used a corpus to design materials for class? (Yes/No)
10. How confident do you feel in designing activities with a corpus for the classroom? (Please select 1-5, with 1 being 'very confident' and 5 being 'not confident at all'.)
11. How often do you use a corpus in class? (write-in response)
12. Which language skills do you use corpora for? (reading, colligations, student autonomy, etc.) (write-in response)
13. Which student levels have you used corpora with? (write-in response)
14. Would you be interested in learning to use corpora more for teaching purposes? (Yes/No)
15. What specifically would you like to learn about when using corpora for language teaching? (write-in response)

About the Author

Cathryn Bennett received her PhD in Applied Linguistics from Trinity College Dublin. Her doctoral research focused on corpus literacy training for in-service ESL/EFL practitioners teaching in the private sector. Recently, her research interests have extended to multimodal analysis and teaching pragmatics to pre-service teachers in Intercultural Virtual Exchange/Telecollaboration.

E-mail: cabennet@tcd.ie

ORCID: <https://orcid.org/0000-0002-0074-9594>