

Corpus Literacy and Applications in Norwegian Upper Secondary Schools: Teacher and Learner Perspectives

Petter Hagen Karlsen and Marte Monsen, Inland Norway University of Applied Sciences¹

Abstract

This study explores corpus literacy and pedagogical corpus applications among in-service teachers who have corpus linguistics in their educational background, along with their students of English in upper secondary school in Norway. In order to investigate the multifaceted nature of educational practice, the study took a wide perspective encompassing aspects of digitalization in education, general pedagogy, and corpus linguistics. The students (n=154) answered an online questionnaire and four teachers were interviewed. The data are discussed in light of a technology integration framework that takes into account teacher knowledge, resource availability, teacher beliefs, as well as student epistemologies. Findings show little to no corpus literacy among the students, with the exception of one student. The interviews show that the teachers have largely avoided corpora in their teaching practice. One reason why the teachers chose to do so was due to inaccessibility related to paywalls, registration requirements and functionality restrictions. Other potential reasons that were uncovered were tied to their perception of learner competencies, how the affordances of corpora fit their curricular focus, and how their epistemic beliefs about language learning influenced their choices.

Keywords: pedagogical corpus application; English language; corpus literacy

1. Introduction

Corpora may offer a range of pedagogically valuable affordances in the English as a foreign language [EFL] classroom. However, according to Cardona, Didriksen, and Gjesdal (2014: 1), they have not been used to a satisfactory degree in Norwegian schools. Pedagogically oriented corpus linguists claim that applying corpora in the classroom has several benefits. As a resource, corpora grant access to authentic language examples, as opposed to artificial, engineered examples and rules (Boulton & Cobb 2017; Leńko-Szymańska & Boulton 2015), contain information on context, frequency, collocations, and distribution (Boulton 2010), and provide opportunity for the examination of lexical and phraseological

¹ The authors have no competing interests to declare.

patterns, and language variation through register and genre information (Farr 2008). These features have several reported pedagogical affordances. For instance, examining corpora can raise the learners' language awareness by examining real language behavior through authentic language (Leńko-Szymańska & Boulton 2015: 3), it can open up for learners following their own interests in language discovery (Bernardini 2004), and it can promote the acquisition of new learning skills and learner autonomy by focusing on learner-centered corpus interactions with the learner being akin to a discoverer or researcher (see Boulton & Cobb 2017; Cheng, Warren & Xun-feng 2003; Johns 1991; Millar & Lehtinen 2008).

This study examines the role of linguistic corpora in four secondary schools, by interviewing four teachers on their corpus use and classroom experience and by investigating the corpus literacy of the students in seven of their classes by means of a questionnaire. These teachers all had a course in corpus linguistics as part of their teacher education. In an experiment with pre-service teachers, Leńko-Szymańska (2017: 234) concluded that a semester-long corpus course may not be sufficient to give future teachers the technological, corpus linguistic, and pedagogical skills required for the pedagogical application of corpora. Similarly, the four teachers participating in this study had a semester-long corpus linguistics course; however, this study examines in-service teachers, a perspective that has not been explored as much. One study in Norway of in-service teachers' familiarity with corpora found little to no familiarity with the concept among the respondents (Kavanagh in preparation). Our belief is that the perspectives from in-service teachers and their students can offer ecological explanations of why corpora are reportedly underused in English language education in Norway.

We argue that by not only exploring the teachers' perspectives, but also the learners', one gets a better understanding of the dynamics of the classroom, which may reveal opportunities or barriers for pedagogical corpus integration. In order to explore the corpus literacy of the students, we pose the research question: (1) How familiar are upper secondary school students with corpora? In light of the student data, we turn our focus to the teachers to see how their corpus education translated into practice, while also inquiring into their general teaching practice to make visible potential challenges and opportunities for the integration of corpora and corpus tools into the classroom. Thus, we pose a second research question: (2) What beliefs do teachers express about corpora as a pedagogical tool?

We turn first to theoretical perspectives associated with corpora as pedagogical tools. Next, since the research is seen through the lens of technology integration and teacher and student beliefs, we will be discussing these in some detail, before we turn to describing our method in Section 3. In Section 4.1, we first examine the degree to which corpora have affected the participants, before we investigate and discuss the reasons for the teachers' choices in Section 4.2. Ultimately, we conclude in Section 5.

1.1. Corpora as a pedagogical tool

In this study, we focus on the direct application of corpora in the classroom setting, i.e. when teachers or learners interact with them directly (Römer 2011). Such direct applications naturally encroach more noticeably on the everyday experience of both teachers and learners and on educational practice, which is why we mainly examine direct use of corpora. Classroom corpora integration can range from inductive scenarios, where the learners make generalizations from the corpus data, to deductive scenarios, where the learners check some lexico-grammatical rule or pattern against data (Liu & Lei 2017: 31-34). It can also range from learner-autonomous, learner-centered to teacher-led activities (Mukherjee 2006: 12). The most inductive, and learner-centered approach to the pedagogical application of corpora is Data-Driven Learning [DDL] as proposed by Johns (1991), where the learner is conceptualized as a researcher, the computer an informant, and the teacher a director or coordinator (Johns 1991). However, Gilquin and Granger (2010: 359) have defined DDL as 'using the tools and techniques of corpus linguistics for pedagogical purposes', thus broadening the concept considerably. In this study, we take a broad perspective on DDL, as defined by Gilquin and Granger (2010), since we are looking for any sign of direct corpus utilization. Moreover, Callies (2019: 247) defines corpus literacy as "a multicomponential set of complex skills". Drawing on Mukherjee (2004) and Dalton-Puffer (2014), Callies (2019: 247-248) presents four components of teachers' corpus literacy:

1. Understanding basic concepts in corpus linguistics: What is a corpus and what types of corpora are available and how? What can you do—and cannot do [sic]—with a corpus?

2. Searching corpora and analysing corpus data by means of corpus software tools, e.g. concordancers: What is corpus software and how can it be used to search a corpus? How can corpus output be analysed?
3. Interpreting corpus data: How may general trends in language use/change be extrapolated from corpus data?
4. Using corpus output to generate teaching material and activities: How can you make use of corpus material for teaching purposes?

Although this construct has teachers in mind, we adopt some of the points as categories of analysis against which we can examine students' knowledge of corpora. This adoption is possible, we would argue, because students would need a semblance of some of these skills in order to exploit corpus data independently and effectively. A couple of things should be noted, however. First, the points listed above suggest the need for declarative knowledge. We acknowledge that students' exposure to corpora may have been less explicit and less terminology-centered. According to Frankenberg-Garcia (2014), it is unnecessary to teach learners what corpora are or train them to perform linguistic analyses; they simply need to know how to look up their specific language quandaries. We therefore sought to ask questions that could elicit exposure to corpora that has not resulted in declarative knowledge. When we discuss students' familiarity with corpora, we mean in a broad sense any sign of student-corpora interaction. Second, the fourth point is clearly directed at teachers and is therefore not suitable to our current discussion. Third, we emphasize that (a) even if students report to have no knowledge of corpora, it does not necessarily mean that the teachers have avoided it completely, and (b) even if students report having comprehensive knowledge of corpora, it does not necessarily mean that the source of said knowledge is their current English teacher. We only treat these teachers as a likely source of potential corpus knowledge due to their educational background.

Several challenges have been outlined for applying corpora pedagogically. Some have argued the digital medium to be a possible challenge for corpus integration (e.g. Boulton 2010), while others suggest that today's computer access and digital competence among the younger generation may open the door for corpora (e.g. Cardona, Didriksen & Gjesdal 2014; Flowerdew 2009). The latter positions, however, can be tied to Prensky (2001: 1) generationally bounded term 'digital natives', which has been problematized by Bennett, Maton, and Kervin (2008: 783),

whose review of several studies on youths' digital competence concluded that there is little evidence for a distinct generation with sophisticated digital technology skills and learning preferences. Another issue is that many of the major corpora were created for linguistic, not pedagogical, purposes (Braun 2007: 308), and may not be suitable for younger learners. Lastly, Leńko-Szymańska (2014: 261) lists digital and computer skills, usability, access, cost, and lack of knowledge about corpora as other or related problem areas.

2. Theories of Language Learning in DDL

DDL has been linked to three distinct yet connected theories of language learning: the noticing hypothesis, constructivist learning, and sociocultural theory (Flowerdew 2015: 16-19). According to the noticing hypothesis, learners are more likely to acquire language competence if their attention is consciously directed toward linguistic features. Schmidt (2001, 2010) maintains that noticing is a prerequisite for understanding. He also suggests that conscious work to 'notice the gap' between the learners' own output and target language input is necessary to overcome errors in L2 language production (2010: 724). Related to the noticing hypothesis, and as we will argue also relevant in connection to DDL, are theories about linguistic and metalinguistic awareness, for example as put forward by Bialystok (2001). According to for example James and Garrett (1992), Bialystok (2001), Purpura (2004), and Van Essen (2008), the ability to evaluate and explain language practices (metalinguistic awareness) will support language learning, and in order to develop such awareness, students need to pay conscious attention to linguistic features. One could argue that DDL, for example focusing on investigating recurring features in the target language, enhances opportunities for students to focus and reflect on lexico-grammatical features in a way that fosters metalinguistic awareness.

Flowerdew (2015: 25) gives examples of studies that '[illustrate] how a corpus can be mediated to address learner needs in line with principles associated with constructivist learning', by providing opportunities for students to choose among ways of investigating linguistic features. For instance, by choosing or switching between formal grammar guides and corpus searches, the learner is a co-constructer of knowledge about language at different levels.

Flowerdew (2015: 27) also maintains that DDL provides opportunities for peer-to-peer interaction with a view to mentoring and feedback. In this way, fellow students may be able to scaffold each other's learning. According to Vygotskian sociocultural theory, learning occurs when individuals interact with others within their zone of proximal development (Lightbown & Spada 2013: 47). Vygotsky defines the zone of proximal development as 'the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in cooperation with more capable peers' (Vygotsky 1978: 86). In contrast to e.g. processability theory, the emphasis of the model of proximal development is on social interaction itself rather than on the internal processing of information (Cook 2016: 249). According to Vygotskian theory, learners gain control over their cognitive processes through mediation, and knowledge is internalized when they engage in social activity (*ibid.*). As Flowerdew (2015: 27-28) shows, certain ways of working with DDL foster learning in a Vygotskian sense, such as by facilitating peer-to-peer interactions during groupwork where the students also engage with the corpus. In the context of this study, a primary interest is in how students and teachers conceptualize corpus-based approaches, and how it is strongly related to their views of how knowledge production and learning occur.

2.1. DDL as technology integration

Working with corpora in the classroom requires of the teacher technical, pedagogical and corpus linguistics skills (Leńko-Szymańska 2017; Mukherjee 2006). Similar categorization is made in the technological pedagogical content knowledge [TPACK] model, which examines 'the complex web of relationships between users, technologies, practices, and tools' (Koehler & Mishra 2005: 132). The model's predictive power has been questioned due to the lack of evidence for the discreetness of each domain (e.g. Archambault & Barnett 2010; Graham 2011). However, it is beyond the scope of this article to clear up this knowledge construct, and we do not seek to track, test, or explain the interrelations between the teachers' knowledge(s). Instead, we treat knowledge of technology, corpus linguistics, and pedagogy as separate concepts only to operationalize the research, and for descriptive purposes. By technology,

we are particularly interested in the teachers' experience with digital technologies and their influence on classroom practice. In the case of corpus linguistics knowledge, we refer back to Mukherjee's (2006) construct corpus literacy as a basic knowledge of corpora. Finally, pedagogy involves the teachers' choices and reflections on teaching practices, learning, planning, and teacher-student interactions.

However, these abovementioned knowledge domains are not the only aspects said to influence successful integration of technology in the classroom. In the wider framework of technology integration, of which we count pedagogical applications of corpora to be part, teacher knowledge is but one element. According to Kim et al. (2013: 77), teacher knowledge, together with environmental readiness, make up a first-order barrier to technology integration. Environmental readiness can be thought of as the access to and availability of necessary resources and conditions to make use of a novel technology. For instance, integrating corpora in the classroom requires the access to a corpus that is free or not too expensive with an easy-to-use interface, as well as time to plan out its application and implementation, and access to computers or tablets.

Moreover, teacher beliefs pose a second-order barrier to technology integration (Kim et al. 2013: 77). Borg (2011: 371) describes teachers' beliefs as "propositions individuals consider to be true and which are often tacit, have a strong evaluative and affective component, provide a basis for action, and are resistant to change". Teachers' beliefs are theorized to 'facilitate or hinder practice by serving to filter, frame, and guide experience, decisions, and actions' (Fives & Buehl 2012), and thus serve as a potential gatekeeping mechanism. There is no stable, unified definition of the teacher beliefs construct (Pajares 1992; Skott 2015). To operationalize this wide concept, Pajares (1992: 316) suggests discussing educational beliefs about..., e.g. about epistemology, about student performance, etc., which can be thought of as the distinct but related subsystems of one's whole belief system (Fives & Buehl 2012). One can be conscious or unaware of these beliefs, and they can vary in stability. However, beliefs show a degree of plasticity to experience and interaction (Fives & Buehl 2012: 473-475).

In this study, we take a particular interest in the teachers' thinking about the nature of knowledge, about pedagogy, and about their students. We do not seek to track the origins of these beliefs (e.g. Levin & He 2008), nor do we discuss how to transform them (e.g. Fives & Buehl 2014); we

only examine them as possible explanatory terms for the teachers' corpus-related choices. We argue that teachers' beliefs may reveal (dis)congruence between teacher epistemologies and the theoretical underpinnings of DDL. In addition, we also investigate student beliefs—what some researchers refer to as student epistemologies—as potentially influential for corpus use, as we believe student influence to be central to classroom practice. In line with Elby's (2009: 139) broad conceptualization of personal epistemologies, we take students' epistemologies to mean their views about the nature of knowledge and knowing, and about learning.

3. Methods and Materials

3.1. Participants

Several teachers were asked to participate in the study with the selection criterion that they taught English in secondary school in Norway. The teachers who were ultimately selected and agreed to participate were four teachers of upper secondary school, three with a Master's degree in English language education, one with a Master's underway, and all with some formal corpus-related training. Specifically, they all had a similar, semester-long corpus course at Master's level, which focused on both corpus analysis and the critical evaluation of corpus-based studies on the one hand, and the ability to facilitate corpus-based classroom activities and to use corpora as an aid for independent learning on the other. Thus, their courses had both a teaching and learning component, and a corpus linguistics component. They all taught at different schools. Since this selection of participants was done based on the somewhat specific characteristic that they were English teachers with some corpus background, the sampling strategy can be defined as purposive sampling, where '[...] researchers handpick the cases to be included in the sample on the basis of their judgement of their typicality or possession of the particular characteristic(s) being sought' (Cohen, Manion & Morrison 2018: 218). We therefore make no claims about the wider population, but try instead to access individuals with particular knowledge and experience.

The students participating in the survey were those from the teachers' classes. The study includes both general studies and vocational studies programs, but the number of participants is greatly skewed toward general studies, as is the case in the Norwegian upper secondary school system.

Table 1 provides an overview of the schools, the teachers, and the number of participating students in each of the teachers' classes (Class 1, 2 & 3). As shown in the 'percentage of total' column, the number of participating students per teacher is skewed. In the case of Marcus' classes, his second and third class were of a particular vocation that tend to have smaller class sizes than general studies classes. Moreover, not all of their classes were available for participation.

Table 1. Overview of schools, teachers, grades and number of students who completed the questionnaire (Class 1, 2, 3). The teachers' names are pseudonyms.

School	Teacher	Class 1	Class 2	Class 3	Total	Percentage of total
W	Nora	16	21	-	37	24%
X	John	29	30	-	59	39%
Y	Marcus	20	5	6	31	20%
Z	Sarah	8	19	-	27	17%
Total					154	100%

The project was approved by the Norwegian Centre for Research Data [NSD] prior to the commencement of the research, including processes of gathering informants, giving information and receiving consent, in order to ensure ethical conduct that guarantees the participants' anonymity in any subsequent publications. A consent letter with detailed description of the research were signed by everyone involved and all participants were informed about their option to withdraw consent at any point without consequences.

3.2. *Research design*

In this study, we sought both student and teacher perspectives to examine the presence of corpora in classroom practices, and to uncover potentials and challenges for successful corpus implementation. By reaching out to as many parties as possible within the different classes, we sought to obtain democratic validity, which is concerned with whether or not the research is done in collaboration with everyone who has a stake in the research (Newton & Burgess 2008: 26). The students answered an online

questionnaire. Subsequently, the teachers were interviewed based partially on preliminary questionnaire findings.

As discussed in Section 2.1, the domains of technology, pedagogy, and corpus linguistics are important aspects of corpus classroom integration (Leńko-Szymańska 2017; Mukherjee 2006). These constructs formed an overarching framework for both the student questionnaire and teacher interviews. Thus, the following categories emerged: (1) technology-integration, including digital technologies such as computers and websites, (2) pedagogy, relating to general pedagogical approaches, preferences, and language teaching and learning theories as discussed by Flowerdew (2015), and (3) corpus literacy, as described by Mukherjee (2006), concerning familiarity and competence with corpora and corpus linguistics.

The questionnaire was constructed using the program Checkbox (Checkbox.com), and by following recommendations from Cohen, Manion and Morrison (2018) about question length, question order, and item type. It was piloted by a group of students from a different school than the ones involved in the research proper. The questionnaire contained thirty-seven items which were predominantly Likert-type items, but with open and multiple-choice questions interspersed. The students could choose either English or Norwegian versions of the questionnaire. All students chose Norwegian with one exception. The responses to the open-ended questions were translated by the first author. The questionnaire took around twenty minutes to complete, was distributed by the teachers, and was completed in class.

The interviews were semi-structured and sought to ascertain the teachers' perspectives on the use of techniques related to corpora or DDL, and to uncover their perceived challenges and possibilities associated with these concepts. The interview guide contained twenty-three questions starting from why they chose to be English teachers and concluding with specific corpus-related questions. The interviews were conducted in Norwegian, audio recorded, transcribed by one of the researchers, and presented as excerpts translated into English by the same researcher who transcribed them. Each interview lasted a little over an hour, with the exception of Marcus', which lasted almost two hours due to his frequent elaborations. Interviews, transcriptions, and translations were all conducted by the same researcher, the main author of this paper.

3.3. Data analyses

Cohen, Manion and Morrison (2018: 842) suggest frequencies and cross-tabulations for non-parametric data, which is the case for the ordinal and nominal data collected through our questionnaire. Since the sample was neither random nor sufficiently large, no effort was made to generalize statistically to a larger population. The data are presented in frequency tables, as percentages, or illustrated through bar charts. The interview data were subject to a qualitative analysis process of segmenting, coding and reassembling (Boeije 2010: 77-79). The data were first segmented based on emerging themes, and given labels—or codes—to mark these themes. These new units were analyzed both in relation to each other and in light of the questionnaire data, which made up the process of reassembly (see Boeije 2010: 76), a process that relies on the subjective discretion of the researchers.

4. Results & Discussion

In the following we present and discuss the findings of the study. In Section 4.1, we explore whether or not the students have any corpus experience by looking for signs of corpus literacy (as presented in Section 1.1) in the data. We also examine the teacher interviews to see whether or not they have implemented corpora in their practice. Note that students' declarative knowledge of corpora or lack thereof does not necessarily reflect on the teachers' practice but could also be signs of previous experiences, indirect corpus work, or even forgetfulness. Section 4.2 presents and discusses the findings relating to why corpora have not featured to any notable degree in the classes involved. It is subdivided into three parts, namely the corpus dimension, the digital dimension, and the pedagogic dimension. Each of these parts corresponds to Leńko-Szymańska's (2017) claim that teachers need technical, pedagogical, and corpus linguistics skills to integrate corpora in their teaching (see Section 2.1). These categories are not treated as discreet units but are instead used as categories to structure the discussion. Lastly, we also discuss some possibilities for integrating corpora as a language-learning tool for upper secondary students and make some tentative suggestions for ways forward.

4.1. Corpus literacy and application

4.1.1. Students' corpus experience

Figure 1 shows the digital resources the students report using in school based on the open-ended question 'Which digital tools and/or webpages do you use at school?' The item was open-ended to see if any corpus-related tool was mentioned. None of the specified webpages or dictionaries was corpus-related. Keep in mind that these responses might reflect their immediate associations, rather than being exhaustive. Note that the categories shown in Figure 1 are made by the researcher based on students' free responses. For instance, general answers like 'PC' or 'Mac' were categorized as 'hardware'.

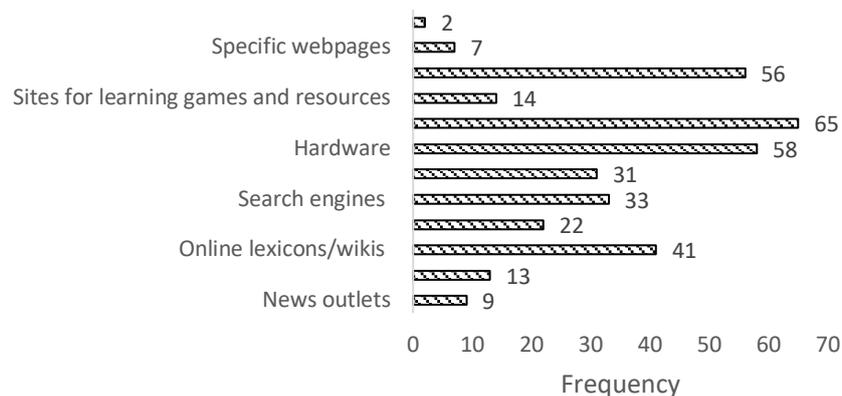


Figure 1. Digital tools the students report using with frequency information

When asked directly in the questionnaire, 94% of the students reported no familiarity with corpora. The remaining 6% who claimed familiarity with corpora were presented with an open-ended question to specify their experience. As can be seen in Table 2, only two of the eight students who had heard of corpora, both taught by Marcus, attempted a description. Student 6 gave a short, concise definition, while student 7 gave a more ambiguous answer. Student 2's answer 'nri' is most likely a typo of the Norwegian word for 'no' (nei).

Table 2. An overview of students' answers to what they have heard about corpora, coupled with each students' teacher

#	Student responses	Teacher
1	Someone has mentioned it, don't know what it is.	Nora
2	Nri	Nora
3	Something about cooperation.	John
4	Have only heard the word before, do not remember the context.	John
5	From our teacher Marcus.	Marcus
6	That it is a tool to find words and expressions that are used in natural contexts among English speaking people. That it is how the language is used in everyday speech by those who have English as their mother tongue.	Marcus
7	It has been mentioned. It helps you find different words for things.	Marcus
8	The teacher took some of us aside and went through what corpus was and whether it was a topic we wanted to learn more about later.	Marcus

Finally, some students responded that they were familiar with a selection of corpus tools in a multiple-choice question (Figure 2). The category 'Other' gave them the opportunity to elaborate freely on what other tools they might have heard of. None of the subsequent responses was corpus-related.

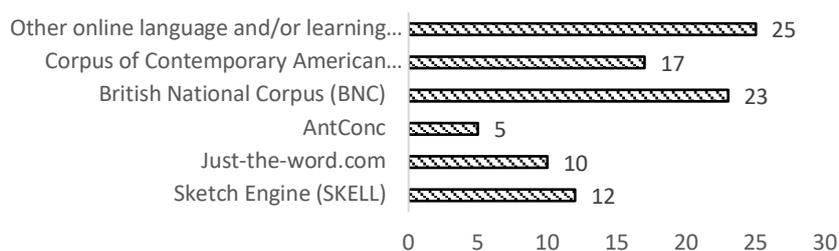


Figure 2. Student answers to the item 'Have you heard about any of the online tools listed below?' with frequency information. The item was multiple-choice, allowing students to select any number of the alternatives given in the chart

4.1.2. The teachers' implementation of corpora

Having examined the students' corpus experiences, we interviewed their English teachers to investigate how corpora had affected their practice. We considered these corpus-trained teachers as the most likely source of the students' experience with corpora. When asked how corpora had influenced her teaching practice, Sarah responded:

(1) Way too little... in an ideal world I would have gladly used it a lot, but the students... they utilizing it themselves is a way too extensive process. I think the interface of these sites are far too difficult and troublesome... it takes too much time to learn [how to utilize] corpora, and I do not think [the students] are in a place language-wise where they would have any benefit from it, either.

She emphasized that corpora were too far removed from the student everyday experience, as they have multiple subjects at the same time, and may not possess the linguistic interest one might wish. When Nora was asked how corpora had influenced her practice, she answered:

(2) I told my students about a corpus assignment, or a corpus study, I did myself ... and told them about the concept, but we have not used it in teaching, we have not looked into it. I have not implemented it in my work yet. In a way, I have the project that I am trying to complete with my Master's thesis, and then there is my job on the other hand. As of now, there is a thin, thin, thin common thread between the two.

Elaborating on the time she told her students about her own research—a study of discourse markers in learner English—she stated,

(3) It was really just a digression, since we were talking about formal and informal language, slang and such, so I suppose it had its source in etymology and language, but they were politely interested [laughter].

John also explained how corpora had not influenced his teaching practice in any significant way, but he thought it had made him more conscious about choosing dictionaries, (4) 'if you search for a word [in a corpus] you get it in use—in actual use—as it is used someplace, as opposed to a invented, made-up example. I find that useful, and I think the students would too, if they can understand what it's all about, that is.'

Notably, the number of students who answered affirmatively on some of the tools seen in Figure 2 does not match the number who claimed familiarity with the term 'corpora'. The teachers found this peculiar, with the exception of Marcus who clarified that he had once taken those he

described as his brightest students aside and introduced them to the very basics of corpora. Since it was a multiple-choice question, the teachers speculated that, although unlikely, perhaps their students had encountered corpora in lower secondary, or, more likely, that they had answered because of an association with the name or initialism (e.g. BNC), rather than skipping the question. Judging from the teachers' skepticism, the item type, and the other student responses about corpora, questions arise about the item's reliability. It is likely that the students responded through some association with the words or acronyms, rather than familiarity with the concept.

These results shed some light on the research question 'How familiar are upper secondary school students with corpora?' None of the students brings up any corpus-associated tools when asked about their digital habits in general, and when asked directly, only a small minority claims any recognition of the terms 'corpus' or 'corpora' in an educational context. Furthermore, those who report familiarity with the phenomenon mostly supply unsubstantial description of the terms, with the exception of two students. Cardona, Didriksen and Gjesdal's (2014: 6) observation that corpora have had little impact in Norwegian secondary schools seems to be reflected in this group despite their English teachers' formal corpus training.

Although the students' reported experiences do not necessarily reflect what the teachers have actually done in the classroom, these findings suggest that corpora have not been featured enough to make an impact on their declarative knowledge, which might further serve as evidence for Leńko-Szymańska's (2017: 272) conclusion, as presented in Section 1.1, that a one-semester corpus course may not be sufficient to make pre-service teachers confident in using corpora pedagogically in their own practice, albeit indirectly. A central premise of corpus literacy, i.e. familiarity with corpora and their uses, is not met by any of the students save one. It is interesting how, despite having discussed corpora with a small group of students, only one of Marcus's students could produce a meaningful definition of corpora. On the whole, however, their lack of knowledge is unsurprising, as the teachers predominantly opted away from using corpora in their practice. A few challenges to DDL are revealed in Sarah's comments about her students' language level, their lukewarm language interest, and the poor user-interfaces of corpora. Next, we further

investigate the reasons behind the teachers' corpus-related choices in their educational practice.

4.2. Why not corpora?

4.2.1. The corpora dimension

Three of the teachers perceived corpora as an academic tool. According to Marcus,

(5) [Integrating corpora in my feedback system] is shown not to be possible, because the corpus world is made one hundred percent for academics, so it becomes sort of 'Do you want access? That will be 3000 dollars a year'. It is not doable for a teacher.

Similarly, John stated that his own corpus education seemed more useful for academic pursuits than for applications in schools. Sarah explained, (6) 'I think that, on a high academic level, it is probably valuable, but [in order to apply it] in schools by having the students use it themselves, for example, it has to be made simpler and more accessible'.

The issue of accessibility is echoed by Nora, as can be seen in this interaction:

(7) Interviewer: How accessible are corpora and corpus resources?

Nora: Inaccessible, are they not? . . . I have experienced it as a threshold, but not extremely so.
I have not been lying awake at night because of it [laughter].

Interviewer: But would it have been worse if every student needed access?

Nora: Yes. If every student had to apply for access, it would be a hassle, and I would lean toward 'no'.

Marcus was the only teacher who explicitly tried to implement corpora in his practice. He has developed a system where he, with the click of a button, added pre-written comments when giving feedback electronically. For instance, if student's writing was unidiomatic, he could press a button that added a comment in the text explaining idiomaticity and providing examples. For language feedback, he attaches language examples from learner dictionaries. He commented that he attempted to link the example function to corpora, but the corpora had paywalls, and no Application Programming Interface (API) that allowed for his needs.

Two interesting perspectives can be gleaned from these statements. First, the accessibility to corpora is an environmental readiness barrier (i.e. inaccessibility, costliness, etc.), as discussed in Section 2.1, since many of the larger corpora have paywalls and are too expensive to introduce to an entire class. It is also a question of teacher knowledge, as they seemed to have mainly interacted with larger corpora such as COCA and the BNC, while being unaware of smaller corpus resources without restrictions such as cost, registration or search limit. On the one hand, this barrier can be explained through the teachers' lack of time, patience, or familiarity to orient themselves in the corpora landscape. On the other hand, it can be seen as a limitation of their education, which seems to have involved a narrow selection of larger, academic corpora.

Second, the teachers' descriptions of corpora as academic research tools reflect Braun's (2007: 308) observation that many of the larger, accessible corpora are designed for linguistics, not pedagogy. If the teachers conceptualize corpora as best suited for the academic level, it is not surprising that they also find it too daunting for their own students. For instance, if their academic experiences with corpora made great demands on their lexico-grammatical knowledge, it can explain why Sarah sees her students' lacking language interest and competence as probable obstacles should they navigate a corpus themselves (see Section 4.1). This belief, which likely is a product of the teachers' own corpus experience, can act as a deterrent for classroom corpus-integration.

4.2.2. The digital dimension

As highlighted in Section 1.1, the digital aspect of corpora has been discussed as a source of both obstacles and possibilities to DDL (e.g. Boulton 2010; Cardona, Didriksen & Gjesdal 2014). Table 3 shows that the majority of the students agree to varying degrees with the statement that they are used to working with digital tools in class (92%). The majority also preferred working with digital tools (77%) as opposed to working without them, with fifteen percent neither agreeing nor disagreeing. Seemingly, the students' beliefs about digital technologies are positively skewed, with mobile phones excluded due to some of the teachers banning them in the classroom. They report familiarity with digital technology in educational contexts, and, more interestingly, they seem to see a relationship between learning and digitalization, but not

quite as strong as the relationship between finding information and digital tools. These beliefs might suggest that they view learning, which can be related to their personal epistemologies (Elby 2009), as positively affected by digitalization and the increased access to information. Their exposure to digital technology appears to be unquestionable and their impressions of digital resources in education are largely positive, so in that sense they are digital natives. There is, however, a clear discrepancy between student beliefs and teacher beliefs, which becomes apparent in the teacher interviews.

The teachers were asked to comment on the results from two of the items in Table 3, namely the students' self-perceived ability to learn new computer programs and digital tools, and their perception that these tools have made learning easier. Nora commented on the relationship between learning and digitalization:

(8) I think it has created winners and losers... it is only a personal, anecdotal hypothesis I have, but I think perhaps the gap between those who are real digital natives and those who do not have a clue is getting wider and wider. I think some of the students have frighteningly bad technological skills... It may well be that they are good with social media, and that they are really good at reading blogs—I don't actually know—but they have difficulty finding documents stored on their Macs, for example, or uploading things to [the school's online learner platform] Canvas.

Sarah described her students' digital skills as (9) 'surprisingly bad, despite having grown up with the Internet'. She further elaborated that they are uncritical of information and will often ask her whether a source is good. She, like Nora, pointed out that they lacked good search strategies: (10) 'When they search for information, they ask questions [to the Google search engine], like sentence-length questions... with question marks and everything.' John calls his students digital competency worryingly low and says that they are quick to ask before really trying themselves. Marcus describes his students' digital competency as 'very, very low', despite having grown up in today's society. He claims they lack fundamental skills, such as cloud-storing, changing languages in Microsoft Word, keyword searches, and awareness of which program to use when opening files.

Table 3. Upper secondary students' responses to their familiarity and preferences concerning digital tools in their education. N=154. All students responded to all items: 100%=154 respondents.

Items	Strongly agree	Partially agree	Neither agree nor disagree	Partially disagree	Strongly disagree
I am used to working with digital tools in English lessons (computer, tablet, the internet).	64 %	28 %	6 %	2 %	0 %
I am used to working with my mobile phone in English lessons.	6 %	7 %	16 %	32 %	39 %
I find it easy to learn new computer programs and digital tools.	37 %	38 %	17 %	7 %	2 %
I learn more in classes where I can use a computer, a tablet or other digital tools.	29 %	35 %	24 %	10 %	1 %
I prefer working with digital aids/tools, as opposed to working without them.	42 %	35 %	15 %	7 %	1 %
I think technology and digital tools have made it easier to find information.	76 %	18 %	6 %	1 %	0 %
I think technology and digital tools have made it easier to learn new things.	49 %	41 %	8 %	2 %	0 %
I wish we would use computers, tablets or mobile phones more often in English class.	16 %	27 %	44 %	10 %	2 %

There is a clear discrepancy between the teachers' negatively skewed impressions of their students' digital competence, and the students' positively skewed impressions of their own abilities. These divergent impressions indicate a potential barrier to classroom corpus-integration. Positive teacher beliefs about technology have been shown to coincide with successful technology implementation in educational practice (Kim et al. 2013). If we keep in mind both the teachers' view of corpora as an academic tool and the lack of user-friendliness of most major corpora, it is unsurprising that the teachers are reluctant to introduce corpora in their teaching repertoire, given their classroom experiences. The technical

aspect of corpora has also been suggested as an obstacle for classroom corpus application; for instance, Farr (2008: 39-40) found that the biggest issues teacher training students had with corpus-based instruction, in which the students themselves utilized corpora, were technology and software related. Learners' frustrations with the technical aspects of corpora have been reported by other researchers as well (e.g. Boulton 2010). Given that both students and teachers in this study express confidence in using digital technology, it may seem the interface and usability of corpora are the greater issue, not technophobia as some suggest (see Boulton 2010: 539). Furthermore, several non-corpus-related applications (apps) are developed with language learning in mind that have user-friendly interfaces and usability metrics directly targeted at younger language learners. Most corpora do not offer the same advantages, as they are often developed for linguists with complex interfaces cluttered with search options and contain technical terminology and codes. There might therefore be a need to design corpus resources with simpler interfaces and fewer options to avoid confusion.

In addition, the teachers said they were surprised by their students' low digital competency, which they expected to be higher for someone growing up in the digital age. This may indicate that the 'digital natives' assumption (see Section 1.1) was part of the teachers' beliefs at the onset of their careers but has since been challenged. As Nora expressed, the digital age has created 'winners and losers', a divide between the technologically competent and the strugglers. Corpora as pedagogical resources introduce not only novel ways of studying language, but also novel ways of utilizing digital technology. In addition, it may potentially widen the gap between the 'winners and losers' of the digital age by building on the varied and unsteady foundation of assumed student digital competency, and consequently alienate the learners who already find simpler digital work difficult.

4.2.3. The pedagogical dimension

Next, we turn to views of pedagogy and how it might explain why the teachers have largely opted away from utilizing corpora in their everyday practice. First, we examine the teachers' descriptions of their English lessons in general. Second, we try to discuss the teachers' epistemological beliefs. These beliefs were explored partly through the teachers'

reflections on students' answers to the questionnaire items 'I think the teacher should be able to answer any question I have about English during English class' (see Figure 3) and 'I do not think the teacher should ask me questions during English class that he/she does not have the answer to himself/herself' (see Figure 4). Both parts endeavored to discover how the teachers' general pedagogical and epistemological beliefs could potentially influence the integration of corpora in teaching.

All four teachers described their lessons as topic-focused (English-speaking countries, global issues, self-chosen topics, etc.), with each topic often spanning several weeks. When inquired about their language-specific teaching, the teachers stated that they had no or limited explicit grammar focus. Marcus taught vocabulary explicitly with his vocational classes through profession-specific word lists, by having the students practice these words in their vocational environment, and through Quizlet (Quizlet Inc. 2019), a website that lets you create flashcards and quizzes for students. The explicit vocabulary teaching for his general studies classes were mainly through feedback on student writing via his system of pre-written comments and attached examples of use described above (see Section 4.2.1). John said his explicit language-focus was through feedback and color-coding of errors. Nora's vocabulary teaching was implicit through working with different topics. If someone needed additional tasks during a lesson, she might send them to Exploringenglish.cappelendamm.no (2012), a site offering English language learning resources to students and teachers, to work on a specific grammatical phenomenon. When discussing corpora and their uses, however, the teachers saw them as tools for etymological research, as reference works for grammar, collocations, and prepositions connected to idiomatic language-use, or as tools for comparative linguistic analyses.

If the teachers' approaches to pedagogy are seen in light of how they described corpora in certain capacities, we start seeing some possible obstacles to corpus integration. Firstly, none of these capacities is directly related to the study of social and cultural themes, as appears to be the pedagogical and curricular focus in these classrooms, but rather directly tied to linguistic inquires. Thus, one could argue that the teachers' beliefs about the usefulness of corpora in education are colored by their experience of corpora as tools for explicit language research, and not as tools for social, cultural or pragmatic study. Their focus on social and cultural themes, which will naturally be influenced by curriculum

demands and number of available teaching hours, does not open for time-consuming, explicit language learning activities such as inductive DDL activities.

Secondly, there seems to be a discrepancy between the teachers' approaches to language study—i.e. through corrective feedback or implicit acquisition—and the explicit language focus theorized in DDL—e.g. inductive learning and the noticing hypothesis (see Section 2). This difference shows a potential mismatch between the teachers' epistemic beliefs, and the type of language study suggested in DDL. In other words, the teachers approach language learning as a case of immersion and feedback, whereas the constructivist approach of DDL puts language at the forefront, favors salience, and requires longer sessions of interpretive work with language data—as its connection to the noticing hypothesis also indicates (see Section 2), DDL requires conscious attention to language. If these disparities in epistemic beliefs are a reality, there are fundamentally different epistemic foundations between DDL and the language pedagogy of the teachers. These differences are a second-order barrier that is hard to grasp, as beliefs about language learning are likely to be central to a language teacher's belief system. We must therefore be sensitive to not only the teachers' resources, knowledge, and needs, but also to their deeply held epistemological beliefs.

On the one hand, we should heed the advice that, '[...] any incremental step or change should be sensitive to the current needs of teachers' (Kim et al. 2013: 83), and we might need to rethink the strong inductive drive in DDL. On the other hand, we should not be afraid to challenge teachers' entrenched epistemological assumptions in manners that may transform their pedagogy to serve student needs in different ways. To achieve this, we suggest scholars interested in pedagogical applications of corpora work closely with teachers and students to examine possible ways to add corpus to their pedagogy incrementally, avoiding invasiveness, and with a mind to complimenting already established practice.

To further investigate the epistemic beliefs of the teachers and students, we included the questionnaire items 'I think the teacher should be able to answer any question I have about English during English class' (see Figure 3), and 'I do not think the teacher should ask me questions during English class that he/she does not have the answer to himself/herself' (see Figure 4.) in the student questionnaire, and asked the teachers to comment on the results. We deemed this information

interesting as DDL is suggested to entail a shift in teacher style from instructor and authority to coordinator and director (Johns 1991: 3). Both questions open up for interpretations from the respondents, but they sought to elicit their views on teachers as experts.

Although mostly skewed toward agreement, both graphs show a high number of students who take no position (33 and 36 respondents respectively). These results may be due to the items' ambiguity, as they can be interpreted in several ways. This ambiguity was pointed out by the teachers when they were asked to comment on the result, which in turn questions the internal validity of the items. It should not be read as an answer to how students regard their teachers as experts, but rather as an indication that most of the students regard the teacher as a knowledge authority in some capacity (frequencies of 95 and 76 respectively)—in what capacity would require more in-depth dialogue with the students.

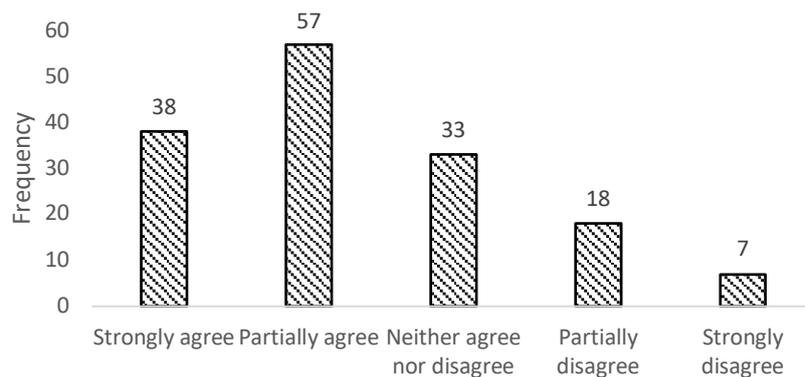


Figure 3. Students' answer to the questionnaire item: 'I think the teacher should be able to answer any question I have about English during English class' by frequency.

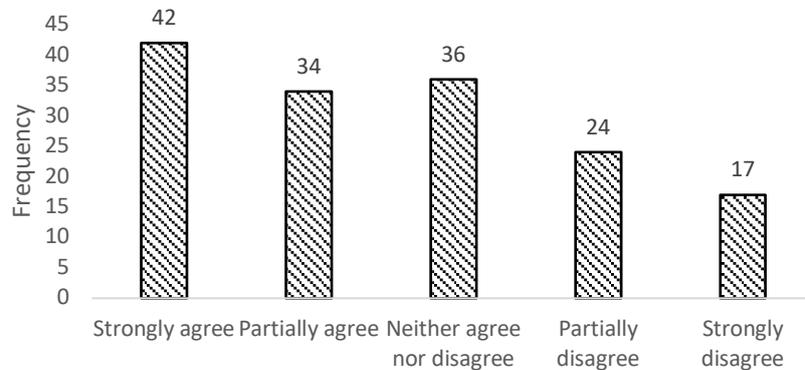


Figure 4. Students' answer to the questionnaire item: 'I do not think the teacher should ask me questions during English class that he/she does not have the answer to himself/herself', by frequency.

Rather than presenting these results in isolation, we used them as prompts in the teacher interviews. The teachers were asked to comment on the graphs in Figures 3 and 4 and give their perspectives on teachers as knowledge experts. Both Nora and John expressed that being a knowledge expert is difficult in the modern information society. Marcus said that a teacher should be someone who knows enough, but still dares to acknowledge when s/he does not have the answer. He did emphasize that a teacher should know the basics, such as concord, and should avoid [L1] interference mistakes. Sarah avoided closed questions with her pupils, and instead focused on the big questions, analyses, critique, and reflection. Nora also mentioned reflection-based schools as more time-appropriate, (11) 'if there is a grammar task, a multiple-choice task, or a sort of right-or-wrong task, yes [there could be a key], but that's a bit "out", isn't it? I am thinking that schools have become more reflection-based'. Similar divisions between language and social topics were made by the other teachers. Sarah said she had more expectations toward her linguistic knowledge, as had her students:

(12) [The students] ask about something there and then, because they have discovered something that does not coincide [with what they know], or they find a pattern that is different, and I am not always able to say 'it is like this because of this' in a simple way without using subject terminology.

There appears to be a division between the teachers' beliefs and expectations of linguistic knowledge (grammar, orthography, etc.) versus knowledge of social and cultural topics (e.g. the #MeToo-movement). They largely approach the English subject similarly, with students using internet resources to explore topics, and they acknowledge the impossibility of being an expert in every such topic. They also seem to promote open exploration, and openness to varying student answers. However, they appear to have a stronger expert identity tied to the linguistic side of their profession, albeit to varying degrees.

Arguably, DDL in its most inductive form promotes a student-centered and explorative approach to language learning not unlike the pedagogies promoted by the teachers when working with varying topics. Here, the teacher's expert, authoritative role shifts toward a facilitative one where language teaching is no longer the teaching of proscriptive grammar rules, and learners become constructors of their own language models (see Section 2). Our findings suggest that the teachers appear more skeptical toward affording their students the same freedom in discovering language, as they are with social and cultural topics, although Sarah's observation that students come to her with unusual language patterns they have discovered seems like prime territory to involve DDL. This skepticism can be interpreted as the teachers' beliefs about their students' language competence, it can be due to their non-linguistic pedagogical focus and priorities, or it can be indicative of their epistemic beliefs about language, i.e. they view language knowledge and thematic socio-cultural knowledge as different domains requiring different pedagogies. Whether one or more of these interpretations are accurate is beyond the scope of the current study to conclude on; however, they provide some potential perspectives for future inquiries.

5. Concluding Remarks

As our findings suggest, the teachers avoided corpora in their teaching practice, and the students reported little to no experience with corpora. First-order barriers, such as paywalls, registration requirements, and functionality restrictions (e.g. APIs), appear to have dissuaded them from integrating corpora in their teaching practice. On one hand, this issue concerns the question of whether corpus resources are available, pedagogically appropriate, and freely accessible. Most corpora are made

for linguistic inquiry in mind, not pedagogic use (Braun 2007), and there are few corpora that are both freely available and appropriate for younger learners. On the other hand, the issue points to the teachers' knowledge about what kinds of corpus resources are out there, as they seem to have been exposed mainly to large, general corpora such as the BNC and COCA.

There are two implications of these findings. First, there is a dearth of free, user-friendly, pedagogically available corpora, which means that new materials need to be developed if corpora are to be of interest to pre-tertiary language teachers. A solution lies in material designers' willingness to create such resources based on teacher and student feedback, and corpus scholars' research on pedagogical corpus use (e.g. Pérez-Paredes 2019; Wicher 2019). Second, the corpus courses offered during teacher training should focus not only on large, general corpora, but also on what resources and possibilities are out there in terms of pedagogically appropriate corpora and how to explore them. This point is further emphasized by the teachers' description of corpora as fundamentally academic and not immediately relevant to their students, a view that is likely linked to their own experiences with corpora. If corpus resources are to find their way into the pre-tertiary classroom, teacher training should provide the type of resources appropriate to the teachers and their students' needs.

Perhaps more difficult to remedy are the discrepancies between teachers' negative view and students' positive view on student digital competency, and the way the teachers see linguistic knowledge in a different way than they see social and cultural knowledge. In case of the former, suggestions have been made to remove the digital component by relying on paper-based concordances (see Boulton 2010); however, the students reported a clear preference toward utilizing computers in their learning process. One direction could be to better understand user-friendliness from a learner perspective when designing corpus resources, so that the burden of explaining and understanding cluttered or advanced interfaces does not fall solely on the teacher. It could also imply that giving students free rein when exploring corpora can lead to confusion, and that the teacher might have to offer instructions. In case of the latter, the way teachers perceive knowledge is difficult to change, if change is even desirable. Instead, future research could explore affordances of corpora

beyond analyzing frequency data and concordance lines and look at how corpora can fit in an educational environment that is topic centered.

This study has revealed some potential barriers to using corpora directly in the classroom by seeking the perspectives of both teachers and learners and by examining classroom practice more holistically. It should be noted that the study only looked at four teachers and that the analyses were predominantly qualitative, which makes generalization difficult. The research was also limited to the upper secondary context and teachers with a common corpus background. Further research is therefore needed for primary and lower secondary school, and with teachers from different educational backgrounds. Although our findings suggest little direct use of corpora, most of the teachers were ultimately positive about trying corpora in their teaching practice. We believe that corpora have a place in the EFL classroom, if we can be more sensitive to teacher and learner needs, and help co-design corpora, corpus resources, and corpus approaches that will benefit and complement the educational process beyond 'what works'.

References

- Archambault, Leanna M., & Joshua H. Barnett. 2010. Revisiting technological pedagogical content knowledge: Exploring the TPACK framework. *Computers & Education*, 55(4), 1656-1662. doi:10.1016/j.compedu.2010.07.009
- Bennett, Sue, Karl Maton, & Lisa Kervin. 2008. The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, 39(5), 775-786. doi:10.1111/j.1467-8535.2007.00793.x
- Bernardini, Silvia. 2004. Corpora in the classroom: An overview and some reflections on future development. In *How to use corpora in language teaching*, edited by John M. Sinclair (pp. 15-36). Philadelphia: J. Benjamins.
- Bialystok, Ellen. 2001. *Bilingualism in development: Language, literacy and cognition*. Cambridge, UK: Cambridge University Press.
- Boeije, Hennie. 2010. *Analysis in qualitative research*. Los Angeles: SAGE.

- Borg, Simon. 2011. The impact of in-service teacher education on language teachers' beliefs. *System*, 39(3), 370-380.
- Boulton, Alex. 2010. Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572.
- , & Tom Cobb. 2017. Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393.
- Braun, Sabine. 2007. Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3), 307-328. doi:10.1017/S0958344007000535
- Callies, Marcus. 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. vi, 267 pages). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Cardona, Margrete D., Anders A. Didriksen, & Anje M. Gjesdal. 2014. Korpusbasert undervisning i fremmedspråkene: La elevenes nysgjerrighet sette dagsorden. *Acta Didactica Norge*, 8(2), 1-26.
- Checkbox.com. 2018. Checkbox (Version 2018Q2). Massachusetts, USA: Checkbox Survey Inc. Retrieved from <https://www.checkbox.com/>
- Cheng, Winnie, Martin Warren, & Xu Xun-feng. 2003. The language learner as language researcher: putting corpus linguistics on the timetable. *System*, 31(2), 173-186. doi:http://dx.doi.org/10.1016/S0346-251X(03)00019-8
- Cohen, Louis, Lawrence Manion, & Keith Morrison. 2018. *Research methods in education* (Eighth edition. ed.). London; New York: Routledge.
- Cook, Vivian. 2016. *Second language learning and language teaching* (Fifth Edition. ed.). New York: Routledge.
- Elby, Andrew. 2009. Defining personal epistemology: A response to Hofer & Pintrich (1997) and Sandoval (2005). *Journal of the Learning Sciences*, 18(1), 138-149.
- Exploringenglish.cappelendamm.no. 2012. Exploring English. Retrieved from <https://exploringenglish.cappelendamm.no/>
- Farr, Fiona. 2008. Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43. doi:10.2167/la414.0

- Fives, Helenrose, & Michelle M. Buehl. 2012. Spring cleaning for the “messy” construct of teachers’ beliefs: What are they? Which have been examined? What can they tell us? In *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors*. (pp. 471-499). Washington, DC, US: American Psychological Association.
- Flowerdew, John. 2009. Corpora in language teaching. In *The handbook of language teaching*, edited by Michael H. Long & Catherine Doughty (pp. 327-350). Chichester, U.K.; Malden, MA: Wiley-Blackwell.
- Flowerdew, Lynne. 2015. Data-driven learning and language learning theories: Whither the twain shall meet. In *Multiple affordances of language corpora for data-driven*, edited by A. Lénko-Szymánska & A. Boulton (pp. 15-36). Amsterdam/Philadelphia: John Benjamins publishing company.
- Frankenberg-Garcia, Ana. 2014. How language learners can benefit from corpora, or not. *Recherches en didactique des langues et des cultures [En ligne]*, 11(1), 1-8.
- Gilquin, Gaëtanelle, & Sylviane Granger. 2010. How can data-driven learning be used in language teaching. In Anne O’Keeffe & Michael McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed., pp. 359-370). London; New York, NY: Routledge.
- Graham, Charles R. 2011. Theoretical considerations for understanding technological pedagogical content knowledge (TPACK). *Computers & Education*, 57(3), 1953-1960. doi:10.1016/j.compedu.2011.04.010
- James, Carl, & Peter Garrett. 1992. *Language awareness in the classroom*. London; New York: Longman.
- Johns, Tim. 1991. “Should you be persuaded”: Two samples of data-driven learning materials. *Classroom Concordancing ELR Journal*, 4, 1-16.
- Kavanagh, Barry. (In preparation). *Bridging the gap from the other side: The uses English teachers have found for corpus methods*. Manuscript submitted for publication.
- Kim, Chanmin, Min K. Kim, Chiajung Lee, J. Michael Spector, & Karen DeMeester. 2013. Teacher beliefs and technology integration.

- Teaching and Teacher Education*, 29, 76-85. doi:10.1016/j.tate.2012.08.005
- Koehler, Matthew J., & Punya Mishra. 2005. What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of Educational Computing Research*, 32(2), 131-152.
- Leńko-Szymańska, Agnieszka. 2017. Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217-214.
- . 2014. Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. 26(2), 260-278. doi:10.1017/S095834401400010X
- , & Alex Boulton. 2015. Introduction: Data-driven learning in language pedagogy. In *Multiple affordances of language corpora for data-driven learning*, edited by Agnieszka Leńko-Szymańska & Alex Boulton (pp. 1-14). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Lightbown, Patsy, & Nina. M. Spada. 2013. *How languages are learned* (Fourth edition. ed.). Oxford: Oxford University Press.
- Liu, Dilin., & Lei Lei. 2017. *Using Corpora for language learning and teaching*. Alexandria, VA: TESOL International Association.
- Millar, Neil, & Ben Lehtinen. 2008. DIY local learner corpora: bridging gaps between theory and practice. *The JALT CALL Journal*, 4(2), 61-72.
- Mukherjee, Joybrato. 2006. Corpus linguistics and language pedagogy: The state of the art – and beyond. In *Corpus technology and language pedagogy: new resources, new tools, new methods*, edited by Sabine Braun, Kurt Kohn, & Joybrato Mukherjee (pp. 5-24). New York: P. Lang.
- Newton, Paul, & David Burgess. 2008. Exploring Types of Educational Action Research: Implications for Research Validity. *International Journal of Qualitative Methods*, 7(4), 18-30.
- Pajares, M. Frank. 1992. Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332. doi:10.3102/00346543062003307
- Pérez-Paredes, P. 2019. The Pedagogical Advantage of Teenage Corpora for Secondary School Learners. In P. Crosthwaite (Ed.), *Data-*

- driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. London: Routledge.
- Prensky, Marc. 2001. Digital natives, digital immigrants. *On the Horizon*, 9(5), 1-6.
- Purpura, James E. 2004. *Assessing grammar*. Cambridge, U.K. ; New York: Cambridge University Press.
- Quizlet Inc. 2019. Quizlet: Learning tools and flashcards, for free. Retrieved from <https://quizlet.com/>
- Römer, Ute. 2011. Corpus research applications in second language teaching. *Annual review of applied linguistics*, 31, 205-225. doi:10.1017/S0267190511000055
- Skott, Jeppe. 2015. The promises, problems, and prospects of research on teachers' beliefs. In *International handbook of research on teachers beliefs* (Vol. [13]), edited by Helenrose Fives & Michele G. Gill. New York: Routledge.
- Van Essen, Arthur. 2008. Language awareness and knowledge about language. In *Knowledge about language* (Vol. 6), edited by Jasone Cenoz & Nancy H. Hornberger. New York: Springer.
- Vygotsky, Lev. S. 1978. *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wicher, Oliver. 2019. Data-Driven Learning in the Secondary Classroom: A critical evaluation from the perspective of of foreign language didactics. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. London: Routledge.