Learning and teaching with corpora: reflections by student teachers
Yvonne Breyer*

European Languages, Macquarie University, Sydney, Australia

(Received 14 April 2008; final version received 15 July 2008)

In recent years, it has become an issue of growing concern that, despite undiminished enthusiasm in the research community, the application of corpus tools and resources in the classroom remains limited. In this paper, I will argue that focusing on the role of the teacher in the process of using corpora in the classroom is an essential step towards popularising this approach. It is vital that future language teachers can discover corpora and concordances as part of their initial training from the perspectives as learner and teacher. To this end, I will present and discuss a case study in which student teachers were introduced to corpus analysis and trained how to teach with corpora. Data on the reflections and opinions provided by the student teachers will highlight the significance and potential of such a course.

Keywords: corpora; classroom concordancing; teacher training

Introduction

The advent of computers has brought about significant changes to the study of language. In fact, the concept of language, the way we study it and what exactly it is we study have changed dramatically since computer technologies have become widely available to the research community. The ability to store language data on computer systems and to gain access to them through a software interface has paved the way for the emergence of modern corpus linguistics. The main subject of inquiry of this approach is language data stored in digital format (i.e. language corpora) and its most powerful tool of analysis is the concordancer. Johns (1986) was among the first to suggest putting this research tool into the hands of foreign language learners. He named this approach ‘data-driven learning’ (DDL), which he defined as ‘the use in the classroom of computer-generated concordances to get students to explore the regularities of patterning in the target language’ (Johns & King, 1991, p. iii). Johns (1988), a strong advocate of this approach, claimed that ‘the concordancer is ... one of the most powerful tools that we can offer the language learner’ (p. 15). Over the past two decades, this form of corpus use for language learning has received much attention alongside countless other studies revolving around corpora and language learning.

In recent years, however, concern has been mounting that the continuing enthusiastic academic endeavour within this field is not reflected in current language
teaching practices. Tribble (2000) notes that ‘not many teachers seem to be using corpora in their classroom’ (p. 31). In the following year, he concludes from a survey he conducted in the UK that ‘IT (not to mention corpus analysis) remains mysterious to most of [his] professional colleagues’ (2001, p. 7). Seidlhofer (2002) discovers that ‘there is very little awareness amongst teachers and students in [her] department of the enormous impact corpus linguistics has on both language description and on the preparation of the very language teaching materials and reference tools they all use’ (p. 216). Referring to the situation in Germany, Mukherjee (2004) remarks that ‘in reality, the influence of applied corpus-linguistic research on the actual practice of language teaching is still relatively limited’ (p. 239). Kaltenböck and Mehlmauer-Larcher (2005) observe that, while there is more progress in tertiary education, ‘in secondary education and general ELT classes, however, computer corpora are still conspicuously absent’ (p. 66). Braun (2005) concludes that ‘corpora, while being the “buzzword” in language research departments, are still far from being part of mainstream teaching practice, if not terra incognita altogether’ (p. 48). Thus, there is a need to address this discrepancy between highly productive research activities on the one hand and lack of application in language classrooms on the other.

There are a number of identifiable factors that contribute to the process of successfully integrating corpus applications in language classrooms. They include the availability and suitability of resources and appropriate training for learners. However, it is important to keep in mind that as the use of corpora remains outside mainstream curricula, the decision to incorporate corpora into language teaching lies ultimately with the teacher. It is the teacher who plays a significant role in the process of introducing corpus work into the classroom but I believe this role has so far been neglected. For this reason, this article focuses on the role of the teacher in the process of teaching and learning with corpora. It is also important to note the key factors that influence the teacher’s decision to implement corpora: motivation, availability of materials and the possession of adequate skills to teach with corpora.

In the first section of this paper I will:

1. explore the role of the teacher in the process of using corpora in language education;
2. address the challenges of that role; and
3. highlight the need to train teachers to use it in their classrooms appropriately.

I will argue that this training is best situated in initial language teacher education (LTE), a view that is in line with recent publications on this topic (cf. Amador Moreno, Chambers & O’Riordan, 2006; Mautanen, 2004; Mukherjee, 2002).

In the second part of this article, I will present and discuss a case study, which reports on a course for student teachers of English as a foreign language (EFL) at the University of Duisburg-Essen, North Rhine-Westphalia, in Germany. In that course, student teachers were introduced to corpus analysis in order to make them aware of the potential of corpus use in language learning and, more importantly, to train them how to teach with corpora. As part of this process the student teachers were encouraged to reflect on their learning experience and analyse it from the perspective of their future role as teachers. In the last phase of their training the student teachers were given the task of creating worksheets for corpus activities as supplements to
standard EFL textbooks. An analysis of selected aspects of this study will shed light on the issues and challenges that student teachers identified regarding teaching with corpora and the implications thereof. Data on the student teachers’ reflections and opinions will highlight the significance and potential of such a course in LTE.

Teaching with corpora
Research on corpora in language learning and teaching has focused primarily on the learner. Some of the topics that previous studies have dealt with are: corpus resources for learners (Braun, 2006; Breyer, 2006b; Vinther, 2004); training in corpus consultation for learners (Gavioli, 2001; Kennedy & Miceli, 2001); and the effectiveness of hands-on concordancing by learners (Chan & Liou, 2005; Cobb, 1997; Cresswell, 2007). However, as Johns (1986) points out early on: ‘it is important that teachers themselves should have experience in using concordance output if they expect their students to make use of it’ (p. 159). It is becoming increasingly clear that this experience needs to be quite substantial; that is, teachers need to have a sound understanding of corpus analysis in order to teach with it. This leads to the question as to when and how teachers should acquire such ‘corpus literacy’ (Mukherjee, 2002, p. 179). Furthermore, in her study on integrating corpora in language study at tertiary level, Chambers (2005) finds that time constraints pose a serious challenge to popularising corpus consultation (p. 113). In view of the complexity of corpus analysis and the interpretation of results, it seems fitting to include corpus training in initial teacher education. In this context, the potential exists to integrate it meaningfully into their training and to devote sufficient time to the subject to allow for in-depth exploration.

Corpora already play an important role in LTE for the purpose of raising language awareness and developing language skills (see e.g. Allan, 1999, 2002; Amador Moreno et al., 2006; Coniam, 1997; Farr, 2008; Hunston, 1995; Tsui, 2004). In this case, corpora are used for the further development of the teachers’ language skills. The teachers experience corpus studies from their role as learners. However, the challenges to the teacher in the process of teaching with corpora, and the implications this has for LTE have so far rarely been addressed.

This lack of focus on the role of the teacher is surprising, given that Johns (1988) observes early on that the direct use of corpora in the classroom involves ‘a shift in the traditional division of roles between student and teacher’ (p. 14). The teacher ‘has to learn to become a director and coordinator of student-initiated research’ (Johns, 1991a, p. 3). That is a change that Johns concedes ‘can be difficult for teachers to come to terms with’ (p. 3). Furthermore, teachers, especially non-native teachers, may not feel competent enough to guide learners through corpus analysis because ‘once the concordancer becomes an important focus of activity in the classroom, many old certainties start to crumble (for example the central position of the syllabus and of the teacher’s key at the back of the textbook)’ (Johns, 1991a, p. 3). Such issues represent major challenges, both pedagogically and linguistically, to the traditional role of the teacher in the classroom, and may quite well be a serious obstacle in the path of introducing corpora in most language classrooms.

The limited availability of ready-to-use corpus teaching materials is another contributing factor that increases the challenge for teachers significantly. Currently, direct corpus applications are neither a standard component of curricula nor textbooks. Indirectly, however, many reference materials such as dictionaries and
grammars are corpus-driven. Furthermore, an ELT textbook series based on the CANCODE corpus was published in the UK (McCarthy, McCarten & Sandiford, 2004). In addition, a number of standalone products comprising corpus-based activities are available. Examples are the Collins COBUILD concordance samplers 1 & 2 (Goodale, 1993, 1995), Classroom concordancing (Tribble & Jones, 1997), the concordancing workbook Exploring academic English (Thurstun & Candlin, 1997), and, for American English, the recently published CorpusLAB series (Barlow & Burdine, 2006). Yet these materials still leave the teacher with a number of challenges to overcome. The teacher has to assess the materials for their appropriateness in the respective learning context; that is language proficiency level, suitable vocabulary, integration into the curriculum and so forth. Apart from these more general difficulties, the teacher needs to possess a degree of ‘corpus literacy’ in order to teach with these materials and integrate them meaningfully into the classroom. The teacher has to guide the learners who will most likely be novice users of corpora. The other option available to teachers is to create their own materials, which in addition to the above mentioned, adds more challenges to the list. These include finding or creating an appropriate corpus, acquiring and learning how to work with concordancing software, creating meaningful exercises, and producing worksheets.

Despite his own enthusiasm for the approach, Johns (1991b) comes to the conclusion that direct applications of concordancing in the classroom ‘represent a considerable challenge to the teacher’s own linguistic sophistication and powers of induction (…) a challenge which has implications for teacher-training which go far beyond the scope and aims of “computer familiarisation”’ (p. 36). He emphasises that this ‘challenge would be even more severe if we expected each classroom teacher to prepare a full range of teaching materials on the basis of concordance output. Clearly, such an expectation would be highly unrealistic’ (p. 36). This highlights the need for more teaching materials with concrete teaching suggestions, and, perhaps more importantly, teaching materials which are integrated into existing textbooks.

This overview has shown that teachers face a number of challenges when wanting to include corpora in their language classrooms. In addition, recent research agrees on the important role teachers play ‘in the process of recontextualising corpora and any useful findings from corpus-based description’ (O’Keefe & Farr, 2003, p. 391), and that ‘mediation by the teacher is a necessary prerequisite for successful application of computer corpora in language teaching and should therefore be given sufficient attention in teacher education courses’ (Kaltenböck & Mehlmauer-Larcher, 2005, p. 81). Gaining a better understanding and appreciation of the challenges teachers face when using corpora in the classroom is a crucial step to a more widespread application of corpora in language teaching.

The stage of initial LTE presents itself as a valuable opportunity for student teachers to explore the use of corpora from the perspective of their role as learner and as teacher. In-depth training can be provided and student teachers can discover the potential of corpora as part of their own studies. If they find their learning experience with corpora to be beneficial, then this is likely to influence positively their decision to use corpora for their own teaching later on. Recognising that there is a significant difference between learning and teaching with corpora, as well as providing student teachers with the required skills, is of great importance. This was the main rationale for conducting the case study, which I will present in the following section.
Training student teachers to teach with corpora: a case study

Research design

The present case study is part of a larger project that explores the integration of corpus linguistics in initial LTE. That project consists of a national survey among student teacher educators on the integration of corpus linguistics in LTE at universities in Germany (Breyer, 2005), the development of a concordancing software for classroom use (Breyer, 2006b) and a course for student teachers described and evaluated below (Breyer, 2006a).

The aim of the course was to create a learning experience for student teachers from two perspectives: as learner and as teacher. On the one hand, the participants were expected to develop a basic understanding of corpus analysis and on the other, to learn about direct applications of corpus-based activities in the classroom. In order to increase their awareness of this learning experience, student teachers were encouraged to reflect on their learning process in relation to their future role as teachers in the form of classroom discussions and reflective writing tasks. In the last part of the course, the student teachers were given the task of creating corpus-based teaching activities as supplemental material to standard EFL textbooks. This required the student teachers to combine their own learning experience with their teaching expertise in order to accomplish the successful transition from the role of learner to teacher.

Current research into language teacher education underlines the significance of creating models of reflected learning experiences in order to ‘allow student teachers to experience themselves the learning process that they are supposed to organise with their EFL students’ (Müller-Hartmann & Schocker-von Ditfurth, 2004, p. 16). This approach is part of a model that the authors propose, which sees ‘language teachers as generators of theories based on a reflection of their own language learning experiences and on an ongoing reflection of their classroom teaching’ (p. 9). In a broader context, such an approach is also in line with well established learning models for adult development such as Kolb’s ‘Experiential learning cycle’ (1984).

One of the central goals of the case study was to gain insight into the student teachers’ perspective on teaching with corpora based on their reflections and feedback. It is hoped that the results of the study will contribute to an improved understanding of the process of using corpora in the language classroom and the ensuing challenges to the role of the teacher. For the purpose of this article, I have focussed on the analysis of a reflective writing task and on the feedback student teachers supplied in a questionnaire regarding the corpus-based learning activity that they produced towards the end of the course.

Set-up and participants of the course

Initial teacher training in North Rhine-Westphalia, Germany, takes between four and five years. As part of their teaching degree in English as a foreign language, student teachers have a choice between a number of courses and lectures from five main areas: linguistics, literature, cultural studies, language practice and teaching methodology (as translation for the German term: Fachdidaktik). Completion of the course described here could be counted as credit towards fulfilment of the requirements of linguistics or teaching methodology courses. In addition, previous participation in the introductory courses for linguistics and teaching methodology and basic IT competence were set as prerequisites for participation in the course.
The course took place at the English department of Duisburg-Essen University, Germany, in Semester 2 of the academic year of 2005. It ran for 11 weeks with two-hour weekly sessions. These sessions took place in the computer lab, which was equipped with individual computer stations connected to the internet. Corpus resources available to the students included corpora from the ICAME CD-ROM, the concordancing software AntConc 3.0.1 (Anthony, 2004), MonoConc 2.2 (Barlow, 2002), ConcApp 4 (Greaves, 2003), and demo versions of both Wordsmith Tools 4.0 (Scott, 2005) and Concordancer 3.2 (Watt, 2004). The choice of resources was limited to that which was made available by the English department and any freely available resources found online.

Eighteen student teachers (14 females and 4 males, aged between 19 and 33) of English as a foreign language participated in the course. The participants were all studying to become teachers of English at secondary educational institutions in Germany. All participants had an advanced level of English language proficiency arising from their secondary school education and their university studies. According to their own assessment as part of a questionnaire, their computer skills ranged from basic (50%), to intermediate (39%) and more advanced skills (11%). Except for one student teacher, the participants had no previous knowledge of corpus linguistics. All participants had previously gained teaching experience (Schulpraktikum) as part of their studies.

Structure of the course

The course aimed to convey the following:

- a basic understanding of corpus-based language analysis;
- the ability to work with concordancing software;
- ways of using corpora in the language classroom;
- the production of teaching materials with corpora and concordances; and
- the integration of these materials in future teaching practices.

The introduction to corpus analysis took place through a series of training units that included presentations followed by related hands-on exercises. These training units formed a major part of the student teachers’ learning experience. They were designed to introduce gradually the subject matter in a combination of theory and practice; an approach which provided the participants with the opportunity to discover facts and theories while actively taking part in the learning process. This put the student teachers into the role of the learner. The purpose of this approach was twofold. Firstly, they were to acquire the necessary knowledge to learn with corpora. Second, the experience was to enable them, as teachers, to reflect on this process when deciding how to best teach with these tools and resources. After completing each training unit a discussion was held in the classroom. Often in the form of brainstorming and collecting ideas on the white board, the student teachers were encouraged to reflect on their learning experience and discuss it from their perspective both as learner and as teacher. In order to provide a relevant framework for these discussions, the Official Teaching Guidelines for Secondary Education in North Rhine-Westphalia\(^1\) (Kernlehrpläne) were introduced early on. As the Teaching Guidelines are binding for all secondary educational institutions in Germany, they were highly relevant for the student teachers’ future professional role. The desired
target competencies for learners of English, as defined in the Guidelines, served as an important frame of reference during the class discussions.

The requisites for successful participation in the course were regular attendance and completion of a number of smaller projects. The projects included writing a reflective essay on one of the training units, reviewing concordancing software and producing a language exercise with concordances for learners. The fulfilment of these requirements leads to a participation credit (Teilnahmenachweis) for the course.

Data collection
The approach adopted for this study is essentially qualitative. Data were collected in the form of a teacher diary, questionnaires, and materials produced by the participants of the course (that is reflective essays, software reviews, and the teaching materials they produced at the end of the course). For the purpose of this article, I will present and analyse findings from two of the data sources obtained throughout the course:

1. the reflective essays on the topic ‘Teaching the use of some and any’; and
2. a questionnaire, which recorded feedback on the student project ‘Creating a language exercise with concordances’.

These sources were chosen for analysis here because they demonstrate the student teachers’ reflections on teaching with corpora from the perspective of their future role as teachers. The analysis of the data serves the purpose of highlighting their concerns and attitudes as well as the challenges and the potential of using corpora.

Findings and discussion
Teaching the use of some and any
The training unit Analysing concordancing output was part of a sequence of units, which were designed to introduce the student teachers to the basics of corpus analysis. In this unit the student teachers were given a paper handout – Worksheet (1) (see Appendix) – with a concordance list of ‘any’ as keyword-in-context (KWIC). This list was created based on a very small ad hoc compiled corpus of materials (approx. 29,000 words) from EFL textbooks commonly used in secondary schools in Germany (Ashford, Aston, & Hellyer-Jones, 1995, 1996). All 19 occurrences of ‘any’ in this corpus were listed on the handout.

To begin with, the group discussed the various features of the layout of a concordance list. The student teachers noted the centred layout, which is inherent to KWIC lists, and the fact that words and sentences appear incomplete on either side. Looking at the language content, the student teachers then went on to list characteristics about the use of ‘any’ according to the KWIC list. They found that the samples were apparently all taken from spoken text; that ‘any’ was almost always directly preceded by a verb, followed mostly by a noun; and that ‘any’ was used either in negative statements or questions. In particular, this last observation piqued their interest as it was confirming the rule they had learned in their secondary education.

The student teachers expressed enthusiasm for this type of ‘language exploration’ but also discovered that they were not always sure about using the correct
grammatical terminology to express their thoughts about the use of ‘any’. This led to a further discussion of possible implications for their future role as teachers, and how this kind of activity in turn might be extremely valuable to do with language learners.

Afterwards, a second handout, Worksheet (2) (see Appendix), with a random selection of 20 occurrences of ‘any’ from the Australian Corpus of English (ACE), was distributed. Upon examination, the student teachers noticed that the use of ‘any’ on this handout was much more varied than on Worksheet (1) and could no longer be defined by the simple rules that had previously applied. In groups they worked through the list and tried to formulate tentative rules for the use of ‘any’ based on Worksheet (2). In the next step, the student teachers compared their results with definitions from the Longman Dictionary of Contemporary English Online (LDOCE). The discrepancy between the simple rules of usage, as they remembered it from the early stages of their secondary education and the use of ‘any’ in the authentic language material from the ACE was of great concern to them. A lively discussion ensued in which the student teachers debated the implications of this for teaching the use of ‘any’ and ‘some’ to foreign language learners. They were particularly concerned about questions such as whether to teach the complete rules as listed in the LDOCE or the simplified rules as provided in many textbooks. As a follow-up task from this training unit, the student teachers had to write an essay (350–500 words) in which they were asked to reflect on their views on ‘Teaching the use of some and any’. In their essays the student teachers made assumptions about their future learners, reflected on their own role as teachers, and dealt with the issue of authentic versus textbook language.

From the 11 essays that were returned, reflections on four main points emerged. These were:

(1) teaching methodology;
(2) language content – teaching ‘real’ English;
(3) using concordance lists; and
(4) definition of the teacher’s role.

A selection of quotes from the essays will be provided and discussed below in order to illustrate these points.

Teaching methodology

During the classroom discussions the student teachers had agreed that learners should be taught the complete rules of use for ‘some’ and ‘any’. In their essays they had to tackle the problem of how to realise this in the classroom. Generally, the student teachers felt that it was of great importance to provide beginners in particular with clearly defined and reliable rules:

I think for beginners it is very important to learn grammatical rules.

I would say that it is important for foreign language learners, especially for beginners, to have strict rules, which they can follow. Language learning is a difficult thing, anyway.

On the one hand pupils perhaps need rules to understand the language, get along with it and use it correctly... I do believe that it is very important for children to have certain rules on which they can rely.
Language learners in general – not just at the beginner level – need some clear and structured rules they can learn, repeat and practice in a first step.

When pupils first encounter new grammatical phenomena, didactic reduction is inevitable. With too many [sic] information at once, we would certainly discourage the pupils.

The assumptions that the participants make about their future students are worth noting here. During the classroom discussion, the student teachers often drew on their past learning experiences in school. Their limited practical experience in the classroom at this point in their career needs to be taken into account as a contributing factor. Relating back to their time as learners, they felt that it was important to provide a ‘safe environment’ for beginners; that is, to teach clearly defined rules, provide easy-to-digest information and not to overwhelm learners with the complexities of language.

This attitude, although not in line with the characteristics that generally define direct corpus use through concordancing, is not uncommon. It is an indication that the student teachers are concerned about successful classroom management and maintaining their authority as teacher. In her study on corpus linguistics, language variation and language teaching, Conrad (2004) also observes that ‘teachers and students seek only definitive answers – such as being able to identify what is grammatical and ungrammatical’ (p. 68). Furthermore, in their desire to present the grammatical rules appropriately for beginners, the student teachers were faced with the problem of how to avoid error fossilisation:

I suppose it could be a problem if they just learned the general rules and then, eventually, are confronted with sentences which do not fit into the system they were taught. I assume this is very confusing and it is hard to look beyond the rules, which one has once learned.

If they do not learn it at an early stage of their language learning, they probably will not learn it at all and will limit their knowledge and language use only to the rules they have learned at school. With this they cannot really become good speakers of English.

This in turn led them to recognise the potential for over-simplification at the cost of teaching authentic language use:

The point is, however, that over-simplification leads to incorrect portraying of authentic language use. . . . Pupils are taught only a part of the rule in a bid to keep it straightforward and simple. What troubles me initially is how something can be taught that is quite obviously wrong. As welcome as it may seem, I am astonished at such a misleading and confusing attempt to make learning easier for EFL students.

In the weeks following this exercise, the issue of authentic texts and vocabulary difficulty was discussed frequently. Various options of dealing with authentic texts in the classroom; for example varying the task rather than the text (Nunan, 1989; Prabhu, 1987) were explored.

Language content – teaching real language

The direct comparison of textbook language samples from Worksheet (1) and authentic language use as represented on Worksheet (2) appears to have emphasised
to the student teachers not only the significance of using authentic language but also the discrepancies between textbook and ‘real’ language:

The material at school is based on rules which are regularly broken by real world’s English.

In the beginning, the student teachers were dealing with the immediate problem at hand – how to teach the use of ‘some’ and ‘any’ – but this quickly opened up to a wider debate on important issues such as what language to teach. This led to an increase in their critical awareness of the textbook materials:

Also with regard to the curriculum it is very important to confront pupils with authentic written and spoken language . . .

On the whole, I think that it is very important that it becomes clear to our pupils that authentic language use very often differs from the rules we learn and teach at school.

Even though the student teachers generally felt that it was important to teach authentic language use, the task made them aware of the difficulties that this might entail. This becomes noticeable in the conflict they display between wanting to teach appropriately for beginners – but not at the cost of teaching ‘real’ English. Should pupils learn English as it is represented by their school books or as it is spoken by native speakers?

Use of concordance lists

After addressing the question of what to teach, the student teachers then focused their attention on how to teach it. Mostly, the student teachers regarded it to be the task of the teacher to introduce the learner to the basic rules governing the use of ‘some’ and ‘any’:

I would suggest that the teacher teaches the simple rules . . . by explaining the rules to the learners.

It could be the easiest and maybe best way to teach pupils that some has to be placed with statements and any with negative sentences and questions.

Once those simplified rules are learned, the concordancer comes into play as a means to help the learners to discover the extended rule set for themselves:

From the third year of English the pupils know many vocabularies [sic] and grammar rules as a basis. Then they can discover rules on their own. One possibility is the concept of data-driven learning.

Another way to teach ‘some’ and ‘any’ is to start with a concordance-exercise, so that the pupils directly learn about any possibility to use ‘some’ and ‘any’. They can formulate the rules themselves . . .

It would perhaps be a nice idea to let the pupils discover these common exceptions on their own, by using a concordancer with chosen texts.

Language learners in general – not just at the beginner level – need some clear and structured rules they can learn, repeat and practice in a first step. But in a second step they have to get prepared to transfer these rules into an authentic context and language
use. DDL offers a wide range for creating such authentic and exploratory tasks and activities for language classes.

The student teachers recognised the value of the corpus and concordancer as tools for the learner to explore the complexities of language and also to lend credibility by allowing the learner to explore authentic texts and discover language use at their own pace.

Since they are able to work with authentic texts, the differences become even more plausible.

If we as EFL teachers are to help our students to develop language awareness, we have to be extremely careful to supply them with adequate teaching material.

This inductive approach to learning or ‘learning by discovery’ lies at the heart of classroom concordancing (Johns, 1988, p. 14). The observations made by the student teachers also reflect Bernardini’s (2002) view of corpora as pedagogical tools that enable learning by discovery and their significance for ‘engaging the learners’ interests, developing autonomous learning strategies, raising their language consciousness, etc.’ (p. 166).

**Definition of the teacher’s role**

Finally, there was an increasing awareness of the shortcomings of available teaching materials and the recognition of the responsibility as a teacher to introduce learners to adequate language content prompted the participants to reflect on their own role as teachers. Just as the student teachers found it challenging to unite the desire to teach authentic language use and finding appropriate ways for teaching beginners, they were torn between their willingness to take risks in the classroom and fear of losing control of the learning process.

One of a teacher’s greatest tasks is to trust in his pupils and also to challenge them sometimes, at least in my view.

The question is, how teachers should handle those exceptions of [sic] rules and if they should teach the rules at all when they are not reliable in class.

Through working with the concordance handouts, the student teachers not only honed their own language skills but also started to reflect on the language content they would be teaching later in their profession. Their learning experiences with the concordance handouts meant that the student teachers turned their attention towards complex and essential issues related to their future teaching practice.

The analysis of the essays on ‘Teaching the use of some and any’ shows that the student teachers saw a strong link between linguistic aspects and pedagogical implications of using corpora in the classroom. The training unit ‘Analysing concordance output’ not only provided the student teachers with a basic introduction to concordancing but also led the student teachers to reflect on several important issues in regards to their role as teachers; that is, to language as content, to teaching methodology, and to the role and potential benefit of using concordance lists as a tool for teaching. The essays show that to them, this tool not only provided opportunities but also poses challenges. Furthermore, the training unit alerted them
to the difficulties of trying to teach authentic language use to beginners in a way suitable for maximum long-term learning outcomes. It turned out that this exercise led not only to an increase of language awareness but teaching awareness as well. This is a significant outcome that illustrates not only the challenges teachers have to overcome in teaching with corpora, but also the very useful role such a course can play in language teacher education.

Creating a language exercise with concordances

Towards the end of the course, the student teachers had been introduced to corpora, concordancing software, and to various applications of corpora in the language classroom. They had discussed and reflected on their learning experiences throughout the course, as well as analysing them in regards to their future role as language teachers. With this in mind, they were asked to put theory into practice and create an exercise with concordances on a topic of their choice. The activity was intended to simulate the situation they might find themselves in as teachers; that is, supplementing existing teaching materials with a hands-on concordancing exercise. Therefore, the task was set leaving the student teachers with the freedom to make decisions on the choice of topic, which corpus to use and so forth. The purpose of this activity was twofold: on the one hand the student teachers gained practical experience in creating such tasks, and on the other their feedback afterwards provided insight into the practical issues arising from the process of creating the teaching materials.

As a starting point the student teachers received EFL textbooks that are commonly used in secondary educational institutions in Germany (e.g. English G). The textbooks were designed for beginners and intermediate learner levels. The student teachers were divided into eight teams. The teams produced ‘data-driven learning’ tasks on the following topics:

- adverbs;
- some and any; have to and negation;
- phrasal verbs;
- concordancer as writing aid for essays;
- comparison of adjectives;
- adjectives and false friends; and
- reading tasks with dictionary work and concordancer.

All teams presented their exercises to the class and afterwards the results were discussed. In the following session, the student teachers participated in a questionnaire that recorded feedback on various aspects of the course. For the purpose of this article, I have focused on the feedback that the student teachers provided on creating the language exercises. Each student was asked to answer two open-ended questions in order to record commentaries on the choice of corpus and task design.

Question 1: During the process of creating the exercise, did you find suitable corpora? Please describe your choice and give reasons.
Question 2: Did you encounter any difficulties during the process of creating the exercise – technical or otherwise? Please report.
Question 1: Choice of corpus

Finding a suitable corpus proved to be the most difficult part of the assignment. The majority of the student teachers (61%) reported that they had been unable to find suitable corpora. Choice of topic and level of lexical difficulty were among the most frequently quoted reasons:

Suitable texts were very hard to find – how can I make sure that the corpus I make includes language that I want to teach?

It was quite hard to find suitable texts which would fit into my exercises.

It was difficult to find corpora for beginners.

I couldn’t really find a suitable text for a corpus. The texts I saw on the internet were too difficult (newspapers!)

When regarding the feedback provided by the student teachers, it is important to keep in mind that they are novice users of corpora. Issues that are significant from the point of view of a corpus linguist, such as the question of representativity, were not at the forefronts of their minds. First and foremost, it was important to them whether or not the task would work in the classroom, whether it was suitable for the topic provided by the textbook in use, and whether it would illustrate the learning target in question. Concern was also expressed regarding the accuracy of the chosen texts:

It was hard to find suitable corpora and took a lot of time searching on the internet. No guarantee of correctness.

I couldn’t find anything for my corpus. The internet is not so reliable and what if the corpus has mistakes?

This issue also frequently came up during classroom discussions. The student teachers believed that it would create a difficult situation for them as teachers if errors in the corpus were discovered during an exercise. They also expressed frustration about a perceived lack of control over the exact contents of the corpus because the corpus was not originally intended as a dedicated learning device for use in the classroom. The feedback provided by the participants further illustrates their concern that unknown vocabulary may pose a serious problem for the learners. This point in particular was frequently the focus during feedback sessions after presentations delivered by the student teachers. They particularly expressed their frustration at not being able to obtain authentic text material that would suit their individual purposes. Based on more traditional approaches to teaching, the student teachers were looking for examples to support their desired learning target rather than language samples that would reflect language use as it naturally occurs. This is in direct opposition to the very core of concordancing and corpus linguistics itself, where the language data come first. This view becomes clearer in the evaluation of Question 2 below.

Question 2: Reported difficulties

Teachers cannot always know the result of DDL exercises. They are also difficult to prepare.

Depending on the corpus, it can be difficult to find suitable examples.
The above comments reveal that the student teachers were very intent on creating ‘closed’ tasks; that is, tasks of which they, as the teacher, already knew the outcome. Hunston (2002) also observes this problem relating to classroom management: ‘[i]f the corpus is consulted and no answer is apparent to student or teacher, or if further difficult questions are raised, the teacher may feel that a loss of expertise has occurred’ (p. 171). It appears that the student teachers viewed the unpredictable nature of concordancing exercises as a source of concern rather than an asset. A feeling of fear of losing control of the teaching process emerged. In addition, the student teachers found it challenging to gauge whether or not the task they were creating was appropriate for the target learner group they had in mind.

I was not sure whether my task was too difficult. It is hard to tell with this kind of exercise.

Perhaps the task would be too difficult for the students.

It is difficult to anticipate what the students know.

A few student teachers experienced technical problems, which were all exclusively related to the handling of the concordancing software:

- technical problems: an online concordancer outcome wasn’t possible to copy
- I couldn’t copy the KWIC lines into Word. The lines lost their formatting.

Despite the reported difficulties, which were the focus of this section, the student teachers presented very interesting learning activities and expressed their enthusiasm for the task. However, as one of the student teachers concludes: ‘a lot more teachers would use DDL to teach grammar/vocabulary if the teaching materials were already prepared. It takes quite some time to create a proper exercise.’

Throughout the course, the student teachers experienced corpus applications in the classroom from their perspective as learners and as teachers – although the transition between the two viewpoints was often seamless. The two components of the course discussed in this paper represent both perspectives. The analysis of the training unit has shown how the student teachers transferred their learning experience with concordances to their role as teacher. During this activity, the student teachers built on their own learning experience, and based on this they identified the strengths and weaknesses of this approach.

The analysis of the data presented for this paper has shown that pedagogical concerns, as expressed by the student teachers, play a major role when teaching with corpora.

**Limitations of the study and implications for further research**

There are some limitations to this study that need to be acknowledged. Due to its exploratory nature, generalisations drawn from the case study presented here must be regarded with caution. In particular, it seems necessary to conduct more longitudinal studies that explore the effect of such a student teacher course on the actual classroom practices of those teachers. More research also needs to be done to determine the extent of corpus that literacy teachers need to possess to successfully teach with corpora. Finally, the course underlying the case study presented here is
situated within the context of initial language teacher education in Germany. Therefore, the results are shaped and influenced by the characteristics of this environment. Furthermore, the participants were training to teach a language that was not their native language. While this is an important consideration when regarding the results of this study, it should also be noted that this is a very common situation in language education. Therefore, the main insights gained from this study should be transferable and applicable in other LTE contexts.

Conclusion

Over the years, the extensive body of research on using corpora in the language classroom has paid tribute to the many opportunities this versatile approach has to offer for language learning and teaching. However, in order for corpus use to be more firmly established in language learning, particularly in secondary education, some challenges have yet to be overcome. As I have argued in this paper, the teacher plays a key role in the popularisation of corpus use in language classrooms. Therefore, training teachers appropriately is essential if they are to be expected to use corpora for their teaching practice. This training is best situated in initial teacher training where student teachers can experience the use of corpora for language learning from two perspectives: as learners and as teachers. If student teachers can discover the potential of corpora for their own learning, then this may foster intrinsic motivation to make use of corpora in their profession as teachers. It also allows student teachers to explore and address the challenges that such a learner-centred approach entails. The analysis of the case study presented in this paper has, furthermore, emphasised that teaching with corpora is very much about teaching and that it is important to consider that teachers are by and large novice corpus users.

In order to enable language teachers to use corpora, it is important for them to gain a thorough understanding of corpus analysis, as their own insecurity may have a negative influence on their decision to apply it in their classroom. From the evaluation of the task that student teachers had to create, it emerged that some of the characteristics generally considered to be positive features of DDL were regarded with caution by the teachers for reasons of classroom management. The difficulties, as reported by the student teachers, regarding the process of creating materials highlight the importance of availability of ready-made and integrated tasks. Beyond this, the case study has also shown that the role of corpora in LTE is potentially more than just about raising language awareness and improving language competence. Working with corpora provides a context that is highly conducive to stimulating discussion and reflection on teaching methodology and raises language as much as teaching awareness. The close relationship between language as content and questions regarding teaching methodology make it an ideal playground for student teachers. Amador Moreno et al. (2006) point out that it is ‘difficult to envisage finding time in the programme of study for training in corpus consultation and analysis’. In light of the results of this study, however, I would argue that such a course on learning and teaching with corpora has much to offer beyond training in corpus literacy. In particular, student teachers’ reflections about their own learning experience and how to make the successful transfer of this into their teaching practice has proved to have the most potential. The student teachers not only gained better language awareness but naturally created a strong connection between the
subject matter of their teaching – that is, language – and how to teach it. The reflective essays particularly displayed their struggle to come to terms with the challenges to their own beliefs and attitudes generated by the work with corpora. This clearly demonstrates the enormous potential corpora bear for language teaching and language teacher training.

Notes
1. The Teaching Guidelines were developed in order to ensure a common educational standard (Bildungsstandards) among the different types of educational institutions in Germany. These binding guidelines define target competencies for the subjects such as English, German and Mathematics.
2. In order to prepare the language exercises despite these difficulties, many student teachers decided to use online concordancers, such as COBUILD Corpus Concordance Sampler or the BNC Simple Search.
3. More detailed feedback on the use of software was collected in the context of the software reviews written by the student teachers, but due to the limited scope of this article I will omit these results here. The issue of suitable concordancing software for language teaching is discussed in Breyer (2006b).

Notes on contributor
Yvonne Breyer is a doctoral candidate at Macquarie University, Sydney, Australia. Her research interests include corpora in language teaching, online distance education and software development for language teaching applications. She has also been working as Online Development Coordinator at the Department of European Languages at Macquarie University since 2006.

References


Appendix

Worksheet (1) KWIC list of any from ‘Klett Schultexte’ corpus.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ing to the shop for her . . . Robert: Do we need bread, Mum? Mrs Croft: Yes, we need some bread, a</td>
<td>any</td>
</tr>
<tr>
<td>2</td>
<td>chocolate biscuits, please. Robert: We haven’t got vegetables, Mum. Mrs Croft: Get some carrots, the</td>
<td>any</td>
</tr>
<tr>
<td>3</td>
<td>beans, too, if you like. Robert: O.K. I won’t get fruit. We’ve still got some bananas and apples. M</td>
<td>any</td>
</tr>
<tr>
<td>4</td>
<td>e bananas and apples. Mrs Croft: No, there aren’t bananas. Look! Robert: Oh, I forgot. David was he milk? Robert: Let’s look. -Oh, there’s some milk</td>
<td>any</td>
</tr>
<tr>
<td>5</td>
<td>we made milk shakes. Mrs Croft: Milk! Have we got milk? Your list is long enough already! Revision</td>
<td>any</td>
</tr>
<tr>
<td>6</td>
<td>ee without cream, please?&quot; &quot;Sorry, we haven’t got cream. Would you like it without milk?&quot; &quot;Doctor have you got a tempe</td>
<td>any</td>
</tr>
<tr>
<td>7</td>
<td>, dear!&quot; Robert said. “I don’t think they’ll want more rabbits.” They went in and asked. “No, sorr</td>
<td>any</td>
</tr>
<tr>
<td>8</td>
<td>‘ve got a terrible stomach-ache. But I don’t want nasty medicine. Doctor: Hmm. Have you got a tempe</td>
<td>any</td>
</tr>
<tr>
<td>9</td>
<td>: Well, don’t look at me, Sarah. I can’t lend you money. I want to buy the new ‘Ghots’ CD today. S</td>
<td>any</td>
</tr>
<tr>
<td>10</td>
<td>was always the only young person there. It wasn’t fun for him. “Mum, it’s so boring,” Rober said.</td>
<td>any</td>
</tr>
<tr>
<td>11</td>
<td>id. “Well,” his mother answered, “he doesn’t know other children. His birthday parties were fun whe</td>
<td>any</td>
</tr>
<tr>
<td>12</td>
<td>ht. “Great Uncle Arthur is lucky. He doesn’t need haircuts because he hasn’t got</td>
<td>any</td>
</tr>
<tr>
<td>13</td>
<td>e? Oh, I see. Yes, they’re great. But I can’t see T-shirts. Sarah: Well, there aren’t</td>
<td>any</td>
</tr>
<tr>
<td>14</td>
<td>can’t see any T-shirts. Sarah: Well, there aren’t in the window, but I’m sure</td>
<td>any</td>
</tr>
<tr>
<td>15</td>
<td>Becky are inside the shop. Becky: Yes, there are some lovely</td>
<td>any</td>
</tr>
<tr>
<td>16</td>
<td>for a hundred years! Shop assistant: Do you need help? Becky: Er – no thanks. We’re just looking</td>
<td>any</td>
</tr>
<tr>
<td>17</td>
<td>d for the milkshakes? Mrs Richards: We don’t need sugar. Only two or three spoons.</td>
<td>any</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Worksheet (2) KWIC list of *any* from Australian Corpus of English (random selection).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>will hang. Plain fabrics look good with almost any heading tape but with patterned fabric you need to think</td>
</tr>
<tr>
<td>2</td>
<td>g painting in adjoining studios than you would in any identifiable philosophies in the City Art Institute and the other six people performing.” Armstrong said they avoi</td>
</tr>
<tr>
<td>3</td>
<td>oot six people on stage and make them look unlike any other six people performing.”</td>
</tr>
<tr>
<td>4</td>
<td>s. “Gillian approached the concert as she would any other film project— that’s what will make the difference.”</td>
</tr>
<tr>
<td>5</td>
<td>before shooting the following night’s concert or of the close-ups. Armstrong covered herself well by sho</td>
</tr>
<tr>
<td>6</td>
<td>of a sudden you wouldn’t be receiving the channel any more. It was critical stuff. John went from Racal</td>
</tr>
<tr>
<td>7</td>
<td>note advis + ing him that if he thought I could be of any assistance to him, to give me a ring. I actually had no int</td>
</tr>
<tr>
<td>8</td>
<td>‘The trick for any customer, no mat + ter what they’re buying, is to put their</td>
</tr>
<tr>
<td>9</td>
<td>rst, ‘that’s why we offer a money-back guarantee on any product we recommend— that we’ve chosen for a custom</td>
</tr>
<tr>
<td>10</td>
<td>prove who could drive up a hill fastest rather than any display to skill or general driving ability. Bearing in m</td>
</tr>
<tr>
<td>11</td>
<td>forethought with the type of conditions expected by any keen four wheel drive enthusiast. Australia was top sc</td>
</tr>
<tr>
<td>12</td>
<td>be meeting you there. He kindly offered to carry any letters or messages. Would you like me to have P</td>
</tr>
<tr>
<td>13</td>
<td>s strong than a curiosity at the complete lack of any trace of these people in my mother’s few enough b</td>
</tr>
<tr>
<td>14</td>
<td>sure you didn’t. There is nothing I much care for any longer. Should I like all this? What did the qu</td>
</tr>
<tr>
<td>15</td>
<td>ation (IVR) model. The second model asserts that any absorption must occur into and eigenstate of the c</td>
</tr>
<tr>
<td>16</td>
<td>munity Libraries. The Committee does not foresee any further school/community libraries being built, e</td>
</tr>
<tr>
<td>17</td>
<td>that at this stage it would be premature to make any judgement about the levels of freight charges beca</td>
</tr>
<tr>
<td>18</td>
<td>hing else. ‘The ten commandments do not contain any creed to which men are answerable for their crime</td>
</tr>
<tr>
<td>19</td>
<td>ere was actually nothing there. No life, not even any trees or grass, just dust. And craters. That’s al</td>
</tr>
<tr>
<td>20</td>
<td>Horton. ‘Your sobriety has tipped the balance Any children on the way?’ he added. ‘We’re hoping</td>
</tr>
</tbody>
</table>