Annual Review of Applied Linguistics (2011), **31**, 205–225.

doi: 10.1017/S0267190511000055

Corpus Research Applications in Second Language Teaching

Ute Römer

Over the past few decades, corpora have not only revolutionized linguistic research but have also had an impact on second language learning and teaching. In the field of applied linguistics, more and more researchers and practitioners treasure what corpus linguistics has to offer to language pedagogy. Still, corpora and corpus tools have yet to be widely implemented in pedagogical contexts. The aim of this article is to provide an overview of pedagogical corpus applications and to review recent publications in the area of corpus linguistics and language teaching. It covers indirect corpus applications, such as in syllabus or materials design, as well as direct applications of corpora in the second language classroom. The article aims to illustrate how both general and specialized language corpora can be used in these applications and discusses directions for future research in applied corpus linguistics.

Over the past few decades, corpora, corpus tools, and corpus evidence have not only revolutionized linguistic research but have also had an impact on second language learning and teaching—probably a use that "the compilers [of corpora] may not have foreseen" (Johansson, 2007, p. 17). John Sinclair's work with COBUILD (Collins Birmingham University International Language Database),[1] Tim Johns's data-driven learning (DDL), and Dieter Mindt's empirical grammar research can be considered particularly groundbreaking developments in the context of pedagogically oriented English corpus linguistics in the 1980s (see Johns, 1986, 1991; Mindt, 1981, 1987; Sinclair, 1987, 1991). In the field of applied linguistics, more and more researchers and practitioners value what corpus linguistics has to offer to language pedagogy, and the number of publications on related topics has increased steadily. (See, e.g., Ädel, 2006; Aijmer, 2009; Aston, 2001; Braun, Kohn, & Mukherjee, 2006; Campoy, Belles-Fortuño, & Gea-Valor, 2010; Gavioli, 2006; Granger, Hung, & Petch-Tyson, 2002; Kettemann & Marko, 2006; Lombardo, 2009; Reppen, 2010; Römer, 2005; Scott & Tribble, 2006; Sinclair, 2004a; and the proceedings of the first seven events in the Teaching and Language Corpora [TaLC] conference series: Aston, Bernardini, &

Stewart, 2004; Botley, Glass, McEnery, & Wilson, 1996; Burnard & McEnery, 2000; Frankenberg-Garcia, Flowerdew, & Aston, 2011; Hidalgo, Quereda, & Santana, 2007; Kettemann & Marko, 2002; Wichmann, Fligelstone, McEnery, & Knowles, 1997).[2] Also, there is now a wide range of fully corpus-based reference works (including dictionaries and grammars) available to learners, and a number of dedicated researchers and teachers have made concrete suggestions on how concordances and exercises directly derived from corpora could be used in the second language (L2) classroom.

I would, however, still be hesitant to say that corpora and corpus tools have been fully implemented in pedagogical contexts and would argue that much work still remains to be done in bridging the gap between research and practice. The practice of English language teaching (ELT) to date, at least, seems to be only marginally affected by the advances of corpus research, and comparatively few teachers and learners know about the availability of useful resources and get their hands on corpus computers or concordances themselves (see Mukherjee, 2004). In addition, current language-teaching materials still differ considerably from actual language use as captured in corpora (e.g., Römer, 2005). Despite the obvious and recognized strengths of corpus use in a pedagogical context, for example that corpora highlight what lexical items and collocations are typical in the language, and that they provide us with large amounts of natural language examples (see Hunston, 2002; Römer, 2005), it seems that there is still a lack of awareness of corpora and, in some cases, resistance toward corpora from students, teachers, and materials writers.

The aim of this article is to provide an overview of pedagogical corpus applications and to review recent publications in the area of corpus linguistics and language teaching. It aims to illustrate how both general language corpora, such as the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA), and specialized corpora, such as the Michigan Corpus of Academic Spoken English (MICASE) or the International Corpus of Learner English (ICLE), can be used in these applications. This article also discusses how much has been achieved so far in the field and points toward future research in applied corpus linguistics.

When we refer to applications of corpora in L2 teaching, this includes both the use of corpus *tools,* that is, the actual text collections and software packages for corpus access, and corpus *methods,* that is, the analytic techniques that are used when we work with corpus data. In classifying pedagogical corpus applications, that is, the use of corpus tools and methods in a language teaching and language learning context, a useful distinction (going back to Leech, 1997) can be made between direct and indirect applications (see Figure 1). This means that, indirectly, corpora can help with decisions about *what* to teach and *when* to teach it. Indirect corpus applications thus have an effect on the teaching syllabus and the design of teaching materials. Corpora can also be accessed directly by learners and teachers in the L2 classroom, and so "assist in the teaching process" (Fligelstone, 1993, p. 98). Direct applications mainly affect *how* something is taught and learned. They actively involve the learner and teacher in the process of working with corpora and concordances. The following sections will feature a few important lines of research and developments in both areas

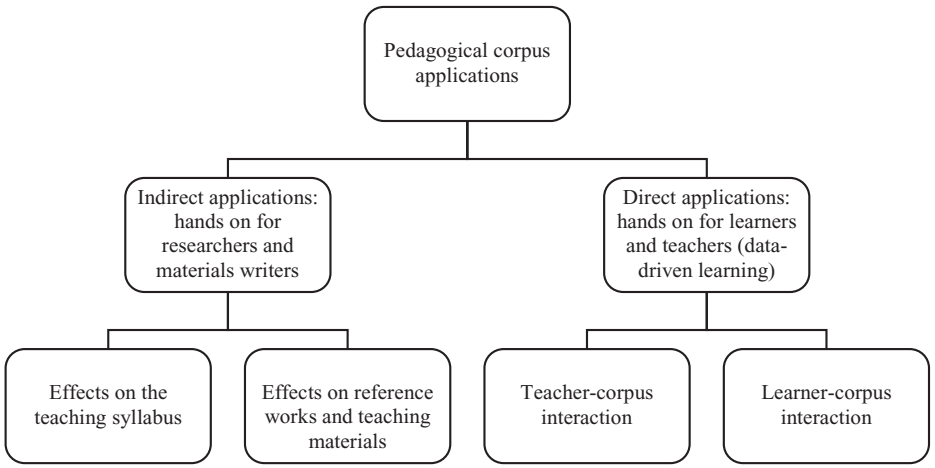

Fig. 1. The use of corpora in second language learning and teaching.

as presented in Figure 1. As the figure shows, we can identify different types of direct and indirect applications, depending on whom or what is affected by the use of corpus methods and tools. In the discussions below, I will consider these distinctions and refer throughout to the pedagogical uses of general corpora as well as specialized corpora. The focus will be on applications of English language corpora with occasional reference to studies which use corpora of other languages. This reflects the current state of the art where considerably more corpora of English than other languages are publically available and used in pedagogical contexts.

INDIRECT APPLICATIONS OF CORPORA AND CORPUS RESEARCH IN SECOND LANGUAGE TEACHING

As Barlow (1996) observed, "[t]he results of a corpus-based investigation can serve as a firm basis for both linguistic description and, on the applied side, as input for language learning" (p. 32). This implies that corpora and the evidence derived from them can greatly affect course design and the content of teaching materials. Existing pedagogical descriptions are evaluated in the light of "new evidence" (Sinclair, 2004c, p. 271), and new decisions are made about the selection of language phenomena, the progression in the course, and the presentation of the selected items and structures (cf. Mindt, 1981; Römer, 2005).

Corpus Research and Selecting What to Teach

Corpora have proven to be an invaluable resource in the design of language teaching syllabi that emphasize communicative competence (see Hymes, 1972, 1992) and give prominence to those items that occur frequently in the language

and that learners are hence most likely to encounter in real-life communicative situations. In the context of computer corpus-informed ELT syllabi, the first and probably most groundbreaking development was the design of the *Collins COBUILD English Course* (CCEC) (Willis & Willis, 1989), an offshoot of the pioneering COBUILD project in pedagogically oriented lexicography (see Sinclair, 1987). The contents of this new, corpus-driven "lexical syllabus" are "the commonest words and phrases in English and their meanings" (Willis, 1990, p. 124). With its focus on lexis and lexical patterns, the CCEC responds to some of the most central findings of corpus research, namely, that language is highly patterned in that it consists of an immense degree of repeated word combinations, and that lexis and grammar are inseparably linked. Also worth mentioning is a much earlier attempt to improve the teaching of English vocabulary that was made long before the advent of computers and electronic corpora. In 1934, Michael West organized a conference "to discuss the part played by corpus-based word lists in the teaching of English as a foreign language" (Kennedy, 1992, p. 327). About 20 years later, West's (1953) *General Service List of English Words* (GSL) was published and has since then exerted great influence on curriculum design (cf. Kennedy, 1992; Willis, 1990). As the title indicates, West's GSL suggests a syllabus that is based on words rather than on grammatical structures. It is also based on frequently occurring words rather than on rare ones. Of course, frequency of occurrence is not the only criterion that should influence decisions about the inclusion of items in the teaching syllabus; there are other relevant criteria, such as "range, availability, coverage, and learnability (Mackey, 1965, p. 188)" (Kennedy, 1992, p. 340; see also Nation, 1990), but it is certainly an immensely important one (see also Leech, 1997). It can be safely assumed that learners will find it easier to develop both their receptive and productive skills when confronted with the most common lexical items of a language (and the patterns and meanings with which these items typically occur) than when the language-teaching input they get gives high priority to infrequent words and structures that the learners rarely encounter in real-life situations.

Another current strand in applied corpus research, which aims to inform the teaching syllabus and stresses the importance of frequency of occurrence, examines language items in actual language use and compares the distributions and patterns found in general reference corpora (of speech and/or writing) with the presentations of the same items in teaching materials (course books, grammars, and usage handbooks). The starting point for these kinds of studies usually includes language features that are known to cause perpetual problems for learners—for example, for German, discourse particles (Jones, 1997), modal verbs (Jones, 2000), the passive voice (Jones, 2000), and prepositions (Jones, 1997); and, for English, future time expressions (Mindt, 1987, 1997), *if*-clauses (Römer, 2004a), irregular verbs (Grabowski & Mindt, 1995), linking adverbials (Conrad, 2004), modal verbs (Mindt, 1995; Römer, 2004b), the present perfect (Lorenz, 2002; Schlüter, 2002), progressive verb forms (Römer, 2005, 2007), and reflexives (Barlow, 1996). For all these phenomena, researchers have found considerable mismatches between naturally occurring German and English and the type of German or English that is put forward as a model in the

examined teaching materials. They have, as a consequence, called for corpus-inspired adjustments in the language-teaching syllabus (particularly as far as selection and progression are concerned) and for revised pedagogical descriptions that present a more adequate picture of the language as it is actually used. A case in point here is the misrepresentation of the functions and contextual patterns of English progressive forms in EFL (English as a foreign language) teaching materials used in German schools. Progressives that refer to repeated actions or events, for example, are considerably more frequent in authentic English than in textbook English, where the common function of repeatedness is neglected and the focus is on single continuous events (cf. Römer, 2005, pp. 261–263).

Specialized corpora, for example, those capturing a particular language for special or specific purposes (LSP), can also influence syllabus design for LSP courses. As Gavioli (2006) pointed out, in English for specific purposes (ESP), "working out basic items to be dealt with is a key teaching problem" (p. 23). A keyword analysis based on a corpus that contains the specific text or discourse type in question (e.g., English business letters, medical research articles, or newspaper editorials) can help solve this problem and assist teachers in "focus[ing] their efforts in terms of selection of language contents" (Pérez-Paredes, 2003, p. 1). An important issue for ESP teachers (who may or, what is more likely, may not be experts in the specific discourse they teach) is that they should give priority to teaching those words and expressions that their learners will need later on to be able to handle texts in their subject area. For instance, having access to a corpus of biology readings and lectures, to give just one example of a science course in English described by Flowerdew (1993), can enable teachers to successfully address this issue and make informed decisions about item and text selection for their course. Another example of how corpora can impact general EAP (English for academic purposes) teaching is Coxhead's (2000) Academic Word List (AWL). The AWL, based on a corpus of academic writing, contains the vocabulary items that are most relevant and useful to EAP learners. Simpson-Vlach and Ellis (2010) have taken Coxhead's idea from word to phrase level and devised an Academic Formulas List (AFL), which consists of word combinations that occur significantly more often in academic than in nonacademic speech and writing. Both studies (AWL and AFL) take an indirect approach to using specialized corpora in language teaching and contribute to improving the teaching of EAP through informing syllabus design, for example, by helping with the selection of items to be included in the course.

Studies on learner corpora, that is, systematic computerized collections of the language produced by language learners, are also highly relevant for syllabus design (cf. Aston, 2000; Granger, 2002, 2009) since they provide insights on "the needs of specific learner populations" (Meunier, 2002, p. 125) and help to test teachers' intuitions about whether a particular phenomenon is difficult or not (Granger, 2002, p. 22). Nesselhauf (2005), for example, identified unidiomatic verb-noun collocations in German learner English (e.g., *reach an aim, drive a bike*), indicating that combinations such as *achieve an aim* or *ride a bike* may deserve particular attention in teaching materials designed for learners with first language (L1) German.

Corpus Research and Designing Teaching Materials

The results of the comparisons between corpora and course books mentioned earlier not only inform the language-teaching curriculum but also help with decisions about the presentation of items and structures in reference works and teaching materials. Corpora research has exerted a huge influence on reference publishing and has led to a new generation of dictionaries and grammar books. Nowadays, "people who have never heard of a corpus are using the products of corpus research" (McEnery, Xiao, & Tono, 2006, p. 97). In the context of ELT, the publications in the Collins COBUILD series constitute a major achievement. Based on real English and compiled with the needs of the language learner in mind, the COBUILD dictionaries, grammars, usage guides, and concordance samplers (see Capel, 1993; Carpenter, 1993; Goodale, 1995; Sinclair, 1990, 1992, 2001) offer teachers and learners more reliable information about the English language than any of the more traditional reference grammars or older non-corpus-based dictionaries. Two major advantages of the COBUILD and other corpus-based reference works for learners (e.g., Biber & Conrad, 2009; Biber, Leech, & Conrad, 2002; Bullon, 2006; Carter & McCarthy, 2006; Hornby, 2005; Peters, 2004; Rundell, 2007) are that they incorporate corpus-derived findings on frequency distribution and register variation, and that they contain genuine instead of invented examples. Particularly worth mentioning here is the student version of the entirely corpus-based *Longman Grammar of Spoken and Written English* (Biber et al., 2002). The importance of presenting learners with authentic language examples has been stressed in a number of publications (e.g., Fox, 1987; Kennedy, 1992; Römer, 2004a, 2005; Sinclair, 1991, 1997). Kennedy (1992), for instance, cautioned that "invented examples can present a distorted version of typicality or an over-tidy picture of the system" (p. 366), and Sinclair (1991) called it an "absurd notion that invented examples can actually represent the language better than real ones" (p. 5). Thanks to the corpus revolution, language learners can today choose from a range of reference works that are thoroughly corpus-based and that offer improved representations of the language they want to study. While course books and other materials used in the language teaching classroom have long been lagging behind this development and have been rather unaffected by advances in corpus linguistics (at least as far as the English as a foreign language market is concerned), the first attempts are now being made to produce textbooks that draw on corpus research and are fully based on real-life data, that is, on language that has in fact occurred (e.g., Barlow & Burdine, 2006; Carter, Hughes, & McCarthy, 2000; McCarthy, McCarten, & Sandiford, 2005). Some particularly innovative developments are currently taking place in Japan, where Yukio Tono of the Tokyo University of Foreign Studies hosts the world's first corpus-based English conversation program on television (watched by a million Japanese language learners) and publishes a series of accompanying textbooks that have become best sellers in Japan (Tono, 2011). Each unit of the program focuses on one of 100 keywords (selected from a BNC frequency list) and its most common collocation patterns. Model skits and exercises illustrate the use of these collocations.

Another branch of general corpora research that has exerted some influence on the design of reference works and, to a lesser extent, teaching materials is the area of phraseology and collocation studies. Scholars like Biber, Johansson, Leech, Conrad, and Finegan (1999); Hunston and Francis (2000); Kjellmer (1984); Lewis (1993, 1997, 2000); Meunier and Gouverneur (2007); Nattinger (1980); Pawley and Syder (1983); and Sinclair and Renouf (1988) have emphasized the importance of recurring word combinations and prefabricated strings in a pedagogical context because of their great potential in fostering fluency, accuracy, and idiomaticity. Although corpus-based collocation dictionaries (e.g., Benson, Benson, & Ilson, 2010; Hill & Lewis, 1997; Lea, 2002) are available, and although information on phraseology (i.e., about the combinations that individual words favor) is implicitly included in learner dictionaries in the word definitions and the selected corpus examples—and sometimes even explicitly described, for example, in the grammar column in the COBUILD dictionaries and in the *COBUILD Grammar Patterns* reference books (e.g., the "V *over* n" pattern for the verb *argue*; Francis, Hunston, & Manning, 1996, 1998)—such information and exercises on typical collocations are as yet largely missing from language-teaching course books (or they are inadequate; cf. Meunier & Gouverneur, 2007). Like Hunston and Francis (2000), I see a necessity in and "look forward to [more] information about patterns being incorporated in language teaching materials." (p. 272)

It has also been demonstrated in recent publications (e.g., Aijmer, 2002; Altenberg & Granger, 2001; Herriman & Boström Aronsson, 2009; Leńko-Szymańska, 2007; Nesselhauf, 2005) how the findings of studies based on learner corpora such as the ICLE can "enrich usage notes" in learners' dictionaries (Granger, 2002, p. 24), or how they "can provide useful insights into which collocational, pragmatic or discourse features should be addressed in materials design" (Flowerdew, 2001, pp. 376–377).

DIRECT APPLICATIONS OF CORPORA AND CORPUS RESEARCH IN SECOND LANGUAGE TEACHING

While the indirect approach centers on the impact of corpus evidence on syllabus design or teaching materials and is concerned with corpus access by researchers and—although to a lesser extent—materials designers, the direct approach is more teacher- and learner-focused. Instead of having to rely on the researcher as mediator and provider of corpus-based materials, language learners and teachers get their hands on corpora and concordance tools themselves and find out about language patterning and the behavior of words and phrases in an "autonomous" way (Bernardini, 2002, p. 165). Tim Johns, who pioneered direct corpus applications in grammar and vocabulary classes at the University of Birmingham (UK) in the 1980s, suggested to "confront the learner as directly as possible with the data, and to make the learner a linguistic researcher" (Johns, 2002, p. 108). Johns (1997) also referred to the learner as a "language detective" and formulated the motto "Every student a Sherlock Holmes!" (p. 101). This method, in which there is either an interaction between the learner and the

corpus or, in a more controlled way, between the teacher and the corpus, is now widely known under the label *data-driven learning* (DDL; see Johns, 1986; 1994).

Examples of Direct Pedagogical Corpus Applications

Following Johns's example, a number of researchers have discussed ways in which corpora and concordances can be used by language learners. Bernardini (2002), for example, described the positive effects of what she calls corpus-aided discovery learning with the BNC, and described corpora as "rich sources of autonomous learning activities of a serendipitous kind" (p. 165; see also Bernardini, 2000b; 2004a). She envisioned the learner in the role of a "*traveller* instead of a *researcher*" (Bernardini, 2000a, p. 131; italics in original), and was less "interested in the starting or end point of a learning experience" than in what the learner experiences in between, on her or his journey (Bernardini, 2000a, p. 142). Kettemann (1995), too, stressed the exploratory aspect of DDL and considered concordancing in the ELT classroom "motivating and highly experiential" for the learner (p. 30).

To give a concrete example of a possible DDL task, learners could be asked to compile concordances of a pair of near-synonyms (such as *speak* and *talk;* see concordance samples in Figure 2) and work out the differences in the collocational and phraseological behavior of these words. Further examples of DDL activities are described in Aston (1997, 2001, 2009); Bennett (2010); Fligelstone (1993); Gavioli (2001, 2006); Hadley (2001); Johns (1991, 2002); Kaszubski (in press); O'Keeffe, McCarthy, and Carter (2007); Oksefjell-Ebeling (2009); Reppen (2010); Sripicharn (2004); and especially in Tribble and Jones (1997). Also worth mentioning are some materials and tools for the creation of DDL exercises that are available online, for example, exercises based on MICASE (see http://micase.elicorpora.info/, links to "ESL/EAP Teaching Materials" and "ESL Self-Study Activities"), Tom Cobb's "Compleat Lexical Tutor" (http://www.lextutor.ca/), a tool collection "[f]or data-driven language learning on the web" (Web site blurb), and a module on "Using concordance programs in the modern foreign languages classroom" of the Information and Communications Technology for Language Teachers, http://www.ict4lt.org/en/en_mod2–4.htm.

The DDL method of using learner-centered activities with the teacher as a facilitator of these activities has, up to now, mainly been discussed with reference to ELT and English language corpora. Some of the few studies that deal with other languages include Whistle (1999) and Chambers (2005), who report on introducing DDL activities to the teaching of French in order to supplement other CALL (computer-assisted language learning) tasks; Dodd (1997) and Jones (1997), who show how corpora of written and spoken German can be exploited "to give students a richer language-learning experience in the foreign language environment" (Dodd, 1997, p. 131); and Kennedy and Miceli (2001, 2002, 2010) who suggest corpus consultation for learners of Italian.

```
1       ery? Can I think.  .. Well I 'd like to speak about the gallery I like to speak for m
2       d morning.   Hello! Yeah,  I 'd like to speak about the the squeeze on the benefits.
3       ught people might be less willing to to speak at the meeting if they knew it was bein
4       d   pointing. And er when you get up to speak at the conference,  you have to give yo
5       re 's not quite as bad as when I had to speak for Amnesty on Radio Essex last year an
6       ke to speak about the gallery I like to speak for myself and er just the visual arts
7       't got a word  . you 're not allowed to speak for the rest of the week. he 's hiding
8       Well football fans? well David I ca n't speak on behalf of Hibs,  all I can say Mm. i
9       a member of the public I 'm not here to speak on behalf of the theatre at all. I migh
10      rhaps people could as actually come and speak to me afterwards,  if they, if they thi
11      e will if you 'd like to come along and speak to him individually afterwards he will
12      hat, then.   When do you start again? I speak to Stella now Do you speak, do you star
13      e appropriate way of doing it.  Shall I speak to Paula about that then?  Yeah. Paula,
14      er Okay.   I 'll speak to Simon.  I 'll speak to Simon erm about borrowing his P C at
15      ot to go     home and do? Are you sure? Speak to me Yes Okay,  right, one person from
16      rt is to go back one step,  not just to speak to people who are experiencing hurt, bu
17      ikey.  Speak up loud, you 've got ta be speak up loud and clear.  No. Uniform. Unifor
18       I think that they may be frightened to speak up and that they 're scared that if the
19       of talking about.  Excuse me could you speak up just a little bit?  Yes yes er Thank
20       first floor.   Oh I 'm sorry!  Can you speak up then?  Oh sorry!  Have I Mm mm. Have


1       ney. All of that being said and Mr will talk a bit more about the figures when he com
2        have an M P here on the phone line and talk about this er proposed pay rise. They ca
3       ? Mm.  . Erm I just want to go back and talk about a few things we just touched on ea
4       are?   Erm Without drawing it,  try and talk about a square.  What would happen? Erm
5        Mm.   Like we had you know,  who could talk about experiences you know,  in a, in a
6       ey include Gerry Addams and we we could talk about the other side,  we could talk abo
7       ot  , you know, I 'm not er gon na even talk about the disaster at airport things, yo
8       Mm.   And basically,  again can we just talk about what we 're trying to achieve? My
9        on, and Mr   I know in a minute we 'll talk about the number of people who attend co
10       Becky. No, Becky. Becky  .  And we 'll talk about this in committee and let you   kn
11      ople or all about that I do n't want to talk about that I want to talk I want to talk
12       get them around the table and begin to talk about those sort of things.  Show a bit
13       bit, we just have done.  We started to talk about the solar system.  How far have we
14      them   anyway. Right.  What else did we talk about and we need to know, we have n't g
15      ber   of possible pieces there,  Did we talk about this?  We did.  Ah, right, yes. So
16      .   They look tired and worn out.  They talk quietly amongst themselves sometimes fin
17      s boss, the professor to come along and talk to us,  and let,  some of their time is
18      than then, because as a kid they do n't talk to you about them things do they like, y
19       going to talk to somebody who will not talk to you,  who will not s possibly even sm
20      how nice the man was when he started to talk to Jason oh I 'll go and help you how fr
```

Fig. 2. Concordance samples of "speak" and "talk," based on the spoken part of the BNC.

The Effectiveness of Direct Pedagogical Corpus Applications

Advantages of corpora work with learners have been suggested by scholars like Sinclair (1997), who noted that, for the learner, "[c]orpora will clarify, give priorities, reduce exceptions and liberate the creative spirit" (p. 38). Likewise, many researchers and teachers in the TaLC tradition are convinced that DDL can empower learners to find out things for themselves, that a data-driven approach raises the learners' language awareness, and that corpora have great pedagogic potential. The effectiveness of DDL has actually been indicated in studies on the teaching and learning of grammar and vocabulary by Boulton (2009); Cobb (1997); Cresswell (2007); Granath (2009); Johns, Hsingchin, and Lixun (2008); and Yoon (2008). As Yoon (2008) observed in a study on the influence of corpora use on L2 academic writing, "students assumed more responsibility for their writing and became more independent writers" (p. 31). On a similar note, Johns et al. (2008) found that "CALL activities [including DDL] did have a positive impact on students' English learning" (p. 503), and Boulton (2009) noted that in his empirical study "corpus samples led to more successful results than traditional pedagogical resources" such as bilingual dictionaries or usage manuals (p. 50). Concordancing has not only been shown to be a useful way "to mimic the effects of natural contextual learning" (Cobb, 1997, p. 314),

but researchers have recently also highlighted its use and usefulness in translation teaching (Bernardini, 2004b; Kübler, 2011) and for error correction in foreign or second language writing (cf. Bernardini, 2004a; Chambers, 2005, 2007; Gaskell & Cobb, 2004). These descriptive empirical studies demonstrate that corpora nicely complement existing reference works, and that they may provide information that a dictionary or grammar book may not provide. They are based on experiments in which groups of students use corpora, corpus-derived exercises, and/or concordances in the classroom, while control groups of students use traditional teaching materials.

Although according to Granath (2009), "advanced students definitely benefit from working with corpora" (p. 59), it is not clear whether this is also the case for beginning learners of a foreign language. Dealing with concordances based on large collections of native-speaker conversations, newspaper texts, research articles, or novels that contain many unknown words and complex sentence constructions may be too challenging a task for beginners, and not all learners may feel comfortable working with computers in language learning. A further limitation may be that it initially takes time for teachers to familiarize themselves and/or their students with online corpora or offline corpus tools. However, once they have been given basic training in corpus use, teachers should be able to create simple concordance exercises tailored to their learners' needs and access information about collocation and language patterning. Teachers could, for example, access the COCA interface (http://www.americancorpus.org/) to search for a word or phrase and retrieve concordance samples and lists of collocations for this word or phrase. As a recent survey on teachers' needs has shown (see Römer, 2009a), teachers of a language that is not their L1 often require native-speaker advice on language points. Computer corpora that have been described as "tireless native-speaker informant[s], with rather greater potential knowledge of the language than the average native speaker" (Barnbrook, 1996, p. 140), can offer help in such situations.

UNRESOLVED ISSUES AND DIRECTIONS FOR FUTURE DEVELOPMENT

Despite the progress that has unquestionably been made in the field of applied corpus linguistics, there is still considerable room for development. A number of tasks can be formulated to enhance pedagogical corpus applications and help corpora make a stronger contribution to improving language learning and teaching. The tasks and future developments I envisage are grouped under three topics and discussed in turn in the following paragraphs: (a) focusing on learner and teacher needs, (b) fostering indirect uses of corpora in L2 teaching, and (c) fostering direct uses of corpora in L2 teaching.

While many corpus researchers (including myself) claim that corpus linguistics has an immense potential to help improve language pedagogy, I would argue that they do not always make sufficient efforts to reach practitioners with what could be called the corpus mission and to find out about what teachers actually want and need. My suggestion would therefore be to focus our attention more on language teachers and their needs and see how we could support them in

their work. A survey of 78 practicing English language teachers (Römer, 2009a) brought to light that a number of wishes and everyday problems of German EFL teachers could actually be addressed by applied corpus linguists. Among the things the teachers who participated in this survey called for were, for example, better teaching materials, support in creating materials, and native-speaker advice. One possible response to these wishes would be to introduce more teachers to corpus resources that are already freely available online. If teachers received basic training in working with corpora and had access to computers with a good Internet connection and hence to online corpora (e.g., COCA, MICASE, and MICUSP [Michigan Corpus of Upper-Level Student Papers]),[3] they could design the required materials themselves whenever they needed them—for example, a worksheet on the most frequent nouns in the newspaper subsection of COCA or an exercise around a concordance of the common phrase *on the other hand* in the academic student papers included in MICUSP. They would also see that corpora, as large collections of native or expert speaker/writer output, can replace, for example, the "always available native speaker informant" they asked for, and that questions like "What prepositions go with that verb?" can easily be answered by looking at a right-sorted concordance of the verb in question.

Another task on our list should be to pay more attention to the needs of learners and consider which groups of learners may profit most from which type of materials. Related to this issue are questions centering around the learners' willingness and ability to deal with computer corpora, online search interfaces, and concordance exercises prepared by their teachers. DDL may work well with the computer-savvy student who is ready to explore larger amounts of language data, but it may not be the best solution for the techno-phobic student who prefers a teacher-centered, controlled type of instruction. For the latter type of learner, a good compromise solution may be an approach to DDL without computers. Boulton (2010), for example, described the benefits of using paper-based DDL materials and called for more DDL worksheets that are readily downloadable for teachers from Internet sources. The needs of learners will probably also vary considerably not only by learner type but also by learner level, course type, and learner objectives. As discussed earlier, working with large corpora that capture complex material with long sentences and low-frequency vocabulary items might intimidate beginning or intermediate learners who have a limited vocabulary. Two ways of addressing this problem could be to use a corpus of children's literature (see Johns et al., 2008) or a corpus consisting of "graded reader texts which contain a limited number of headwords" (Allan, 2009, p. 25). Depending on whether we are dealing with intermediate or more advanced learners, for instance, our focus in designing corpus-derived materials may shift to more specialized vocabulary and its preferred patterns of usage. Similarly, participants in a business English class or international students of mechanical engineering will probably profit most from working with materials that are tailored to their specific needs and discourse in their field of study. That means that when making decisions on what to teach and how to teach it, it is important to consider the learners' language background and what discourse community they eventually want to be a member of and be able to communicate with.

In terms of fostering indirect uses of corpora in language learning and teaching, more work probably has to be put into the creation of reliable corpus-based language descriptions for learners and teachers, especially descriptions of specialized discourses, such as academic English or business English. This implies that there is a need for more large specialized corpora that can be used as bases for creating dictionaries, usage guides, and grammars tailored to the needs of different groups of learners. Also, most available corpora consist predominantly of written language material—which does not come as a surprise given the amount of work that goes into transcribing and encoding speech data for computer-based analysis. In terms of helping learners develop their communicative competence, spoken corpora may, however, be the better basis of pedagogical language descriptions. Generally, I see more scope for research activities that are inspired by the needs of learners and teachers and that take the learners' communication needs and common learning problems into account. For example, language points that tend to be particularly difficult for learners could be identified through comparative analyses of corpora capturing learner (or novice) and native-speaker (or expert) performance data, or through contrastive linguistic analyses based on parallel corpora of the learners' native and target languages (e.g., Granger, 2004). Further comparative studies of lexical-grammatical features in corpora and course books (as already described) could also provide valuable insights into mismatches between real language and school language that need to be remedied. In selecting language points worth analyzing, I would suggest a shift in focus from individual words or empty grammatical structures to phraseological items (e.g., formulas, n-grams, phrase-frames, lexical bundles). Recent research in corpus linguistics has provided massive evidence for the inseparability of lexis and grammar (see, e.g., Biber, Conrad, & Cortes, 2004; Hoey, 2005; Hunston & Francis, 2000; Meunier & Granger, 2008; Römer, 2009b, 2010; Römer & Schulze, 2009; Sinclair, 1991, 2004d; Stubbs, 2001), so it appears only reasonable to provide learners with a more integrated perspective and offer them a view on the language that highlights frequently used phrases that function as important meaning-carrying units in the discourse.

In terms of fostering direct uses of corpora in language learning and teaching, corpus researchers would do well to help create more DDL exercises and corpus-derived teaching materials in general. In the future, I would hope to see more publications (similar to Tribble & Jones [1997] or Barlow & Burdine [2006]) that contain ready-made exercises based on authentic speech and writing from different text types and language varieties and focused on language items that are of central importance and/or troublesome for learners. Another important step we need to accomplish if we want DDL to gain more ground is create a DDL-friendly environment that encourages learner and teacher involvement. Teachers and learners have to be provided with access to corpora that are available on the Internet or to offline corpora and easy-to-use software packages for corpus analysis. A popularization of corpora and their pedagogical use also requires some basic training in accessing corpora and in working with concordances or collocation lists. Such training is crucial because concordance output, at first glance, may seem hard to handle, and because "a corpus is not a simple object, and it is just as easy to derive nonsensical conclusions from the evidence

as insightful ones" (Sinclair, 2004b, p. 2). The good news is that, in most cases, only a few hours of orientation are required before learners and teachers can "enter the world of the corpus" and start exploring language patterns (Sinclair, 2004c, p. 297).

CONCLUSION

This article has focused on the relationship between corpus research and L2 teaching. It has discussed a range of developments in the emerging field of applied corpus linguistics and sketched avenues for future research activities that may positively impact L2 teaching. I hope to have shown that corpus resources and methods have great potential to improve pedagogical practice and that corpora can be used in a number of ways, indirectly to inform teaching materials and reference works and directly as language learning tools and repositories for the design of data-intensive teaching activities. I have also tried to make clear that a lot still remains to be done in research and practice before corpora will eventually arrive in the classroom. Communication between corpus researchers and practitioners has to be improved considerably so that teachers and learners get the support they need and deserve.

In a questionnaire he sent out to language teachers, teacher educators, and applied linguistic researchers in 2008, Christopher Tribble referred to the (direct) use of corpora in language teaching as "a minority sport."[4] One purpose of this article was to describe the potential of pedagogical corpus applications and to provide ideas on what can be done to foster direct and indirect corpus use in L2 teaching and to bring corpora and corpus tools to a larger group of learners and teachers. If we take up some of the ideas mentioned here, and if we are successful in improving communication among researchers, teachers, and materials designers, we can get more people involved in DDL and related activities and thus perhaps make applied corpus linguistics more of a majority sport.

NOTES

1 COBUILD is a project in pedagogically oriented lexical computing started in the 1980s.

2 TaLC conferences take place every other year in different European countries. TaLC conferences 1 through 9 were held in Lancaster (in 1994 and 1996), Oxford (in 1998), Graz (in 2000), Bertinoro (in 2002), Granada (in 2004), Paris (in 2006), Lisbon (in 2008) and Brno (in 2010). Since 1999, conferences with a similar focus to TaLC, organized by the American Association for Applied Corpus Linguistics (now the American Association for Corpus Linguistics [AACL]), have been held in North America (Ann Arbor, MI, in 1999 and 2005; Flagstaff, AZ, in 2000 and 2006; Boston, MA, in 2001; Indianapolis, IN, in 2002; Montclair, NJ, in 2004; Provo, UT, in 2008; and Edmonton, Alberta, Canada, in 2009).

3 MICUSP is freely accessible through a user-friendly search and browse interface at http://search-micusp.elicorpora.info/. The URLs for COCA and MICASE are http://www.americancorpus.org/ and http://quod.lib.umich.edu/cgi/c/corpus/corpus, respectively.

4 A copy of the survey is available at http://www.surveyconsole.com/console/TakeSurvey?id=463263.

ANNOTATED BIBLIOGRAPHY

Sinclair, J. M. (Ed.). (2004a). *How to use corpora in language teaching*. Amsterdam, The Netherlands: John Benjamins.

This collection offers a useful state-of-the art overview of central research topics and practical applications in applied corpus linguistics. It is a book for teachers and applied linguists who are looking for guidance in how to include corpora and corpus evidence in their teaching and research. The book brings together 13 valuable contributions on topics including corpora use in the classroom, corpora as support tools for teachers, the availability of learner corpora, contrastive analyses of corpus data and textbooks, and central steps in corpus analysis. In his opening and closing chapters, John Sinclair, arguably the father of modern corpus linguistics, sets the scene and takes a look into the future of pedagogical applications of corpora.

Aijmer, K. (Ed.). (2009). *Corpora and language teaching*. Amsterdam, The Netherlands: John Benjamins.

This edited volume can be seen as a follow-up publication to Sinclair's (2004a) collection in that it provides an account of developments in pedagogical corpus applications 5 years on. It contains 12 chapters, offering mainly European scholars' perspectives on topics such as the relationship between corpus analysis and L2 language acquisition, applications of learner corpora in language teaching, DDL, and the compilation and use of new types of pedagogical corpora (consisting of textbook materials and spoken learner language).

Reppen, R. (2010). *Using corpora in the language classroom.* Cambridge, UK: Cambridge University Press.

Reppen's book provides a brief and practical introduction to the direct use of corpora in language teaching. The book focuses on creating corpus-derived teaching materials and on using corpus Internet resources in language teaching. Written in a very accessible style and filled with concrete examples and hands-on activities, it is an ideal resource for teachers who wish to bring corpora into their classrooms.

Simpson-Vlach, R. C., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*, 487–512.

This article discusses the development of a corpus-derived, psycholinguistically and pedagogically validated list of formulaic sequences in academic speech and writing. This list, the Academic Formulas List (AFL), consists of sequences of words (e.g., *in terms of, on the other hand, due to the fact that*) that are frequent across genres and disciplines in spoken and/or written academic language, occur more commonly in academic than in nonacademic texts, and express a range of functions that are central in academic discourse. The article constitutes a wonderful example of how pedagogically inspired corpus work can result in a resource that is bound to positively affect the teaching and learning of (academic) English.

REFERENCES

Ädel, A. (2006). *Metadiscourse in L1 and L2 English*. Amsterdam, The Netherlands: John Benjamins.

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language*

acquisition and foreign language teaching (pp. 55–76). Amsterdam, The Netherlands: John Benjamins.
Allan, R. (2009). Can a graded reader corpus provide "authentic" input? *ELT Journal, 63*, 23–32.
Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics, 22*, 173–195.
Aston, G. (1997). Enriching the learning environment: Corpora in ELT. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 51–64). London, UK: Longman.
Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 7–17). Frankfurt, Germany: Peter Lang.
Aston, G. (2001). *Learning with corpora*. Houston, TX: Athelstan.
Aston, G. (2009). Using BNC-XML in the classroom. In L. Lombardo (Ed.), *Using corpora to learn about language and discourse* (pp. 163–198). Frankfurt, Germany: Peter Lang.
Aston, G., Bernardini, S., & Stewart, D. (Eds.). (2004). *Corpora and language learners*. Amsterdam, The Netherlands: John Benjamins.
Barlow, M. (1996). Corpora for theory and practice. *International Journal of Corpus Linguistics, 1*, 1–37.
Barlow, M., & Burdine, S. (2006). *American phrasal verbs (CorpusLAB Series)*. Houston, TX: Athelstan.
Barnbrook, G. (1996). *Language and computers. A practical introduction to the computer analysis of language*. Edinburgh, UK: Edinburgh University Press.
Bennett, G. R. (2010). *Using corpora in the language learning classroom.* Ann Arbor: University of Michigan Press.
Benson, M., Benson, I., & Ilson, R. F. (2010). *The BBI combinatory dictionary of English*. Amsterdam, The Netherlands: John Benjamins.
Bernardini, S. (2000a). *Competence, capacity, corpora. A study in corpus-aided language learning*. Bologna, Italy: CLUEB.
Bernardini, S. (2000b). Systematising serendipity: Proposals for concordancing large corpora with language learners. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 225–234). Frankfurt, Germany: Peter Lang.
Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165–182). Amsterdam, The Netherlands: Rodopi.
Bernardini, S. (2004a). Corpora in the classroom: An overview and some reflections on future developments. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15–36). Amsterdam, The Netherlands: John Benjamins.
Bernardini, S. (2004b). Corpus-aided language pedagogy for translator education. In K. Malmkjaer (Ed.), *Translation in undergraduate degree programmes* (pp. 97–112). Amsterdam, The Netherlands: John Benjamins.
Biber, D., & Conrad, S. (2009). *Real grammar. A corpus-based approach to English*. London, UK: Pearson Longman.
Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at. . .:* Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*, 371–405.
Biber, D., Leech, G., & Conrad, S. (2002). *Longman student grammar of spoken and written English*. London, UK: Longman.
Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London, UK: Longman.
Botley, S., Glass, J., McEnery, T., & Wilson, A. (Eds.). (1996). *Proceedings of Teaching and Language Corpora 1996.* Lancaster, UK: University Centre for Computer Corpus Research on Language.
Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*, 37–54.
Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*, 534–572.

Braun, S., Kohn, K., & Mukherjee, J. (Eds.). (2006). *Corpus technology and language pedagogy*. Frankfurt, Germany: Peter Lang.
Bullon, S. (Ed.). (2006). *Longman dictionary of contemporary English* (4th ed.). London, UK: Longman.
Burnard, L., & McEnery, T. (Eds.). (2000). *Rethinking language pedagogy from a corpus perspective*. Frankfurt, Germany: Peter Lang.
Campoy, M. C., Belles-Fortuño, B., & Gea-Valor, M. L. (Eds.). (2010). *Corpus-based approaches to English language teaching*. London, UK: Continuum.
Capel, A. (1993). *Collins COBUILD concordance samplers 1: Prepositions*. London: HarperCollins.
Carpenter, E. (1993). *Collins COBUILD English guides 4: Confusable words*. London: HarperCollins.
Carter, R., Hughes, R., & McCarthy, M. (2000). *Exploring grammar in context. Grammar reference and practice.* Cambridge, UK: Cambridge University Press.
Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English.* Cambridge, UK: Cambridge University Press.
Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning and Technology, 9*, 111–125.
Chambers, A. (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 3–16). Amsterdam, The Netherlands: Rodopi.
Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System, 25*, 301–315.
Conrad, S. (2004). Corpus linguistics, language variation, and language teaching. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 67–85). Amsterdam, The Netherlands: John Benjamins.
Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213–238.
Cresswell, A. (2007). Getting to "know" connectors? Evaluating data-driven learning in a writing skills course. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 267–288). Amsterdam, The Netherlands: Rodopi.
Dodd, B. (1997). Exploiting a corpus of written German for advanced language learning. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 131–145). London, UK: Longman.
Fligelstone, S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME Journal, 17*, 97–109.
Flowerdew, J. (1993). Concordancing as a tool in course design. *System, 21*, 231–244.
Flowerdew, L. (2001). The exploitation of small learner corpora in EAP materials design. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small corpus studies and ELT. Theory and practice* (pp. 363–380). Amsterdam, The Netherlands: John Benjamins.
Fox, G. (1987). The case for examples. In J. M. Sinclair (Ed.), *Looking up: An account of the COBUILD project in lexical computing* (pp. 137–149). London, UK: Collins ELT.
Francis, G., Hunston, S., & Manning, E. (1996). *Collins COBUILD grammar patterns 1: Verbs*. London, UK: HarperCollins.
Francis, G., Hunston, S., & Manning, E. (1998). *Collins COBUILD grammar patterns 2: Nouns and adjectives*. London, UK: HarperCollins.
Frankenberg-Garcia, A., Flowerdew, L., & Aston, G. (Eds.). (2011). *New trends in corpora and language learning*. London, UK: Continuum.
Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*, 301–319.
Gavioli, L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In G. Aston (Ed.), *Learning with corpora* (pp. 108–137). Houston, TX: Athelstan.
Gavioli, L. (2006). *Exploring corpora for ESP learning*. Amsterdam, The Netherlands: John Benjamins.
Goodale, M. (1995). *Collins COBUILD concordance samplers 4: Tenses*. London: HarperCollins.

Grabowski, E., & Mindt, D. (1995). A corpus-based learning list of irregular verbs in English. *ICAME Journal, 19*, 5–22.

Granath, S. (2009). Who benefits from learning how to use corpora? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 47–65). Amsterdam, The Netherlands: John Benjamins.

Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3–33). Amsterdam, The Netherlands: John Benjamins.

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In E. Connor & T. A. Upton (Eds.), *Applied corpus linguistics. A multi-dimensional perspective* (pp. 123–145). Amsterdam, The Netherlands: Rodopi.

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. A critical evaluation. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 13–32). Amsterdam, The Netherlands: John Benjamins.

Granger, S., Hung, J., & Petch-Tyson, S. (Eds.). (2002). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam, The Netherlands: John Benjamins.

Hadley, G. (2001). *Concordancing in Japanese TEFL: Unlocking the power of data-driven learning*. Retrieved from http://www.nuis.ac.jp/~hadley/publication/jlearner/jlearner.htm

Herriman, J., & Boström Aronsson, M. (2009). Themes in Swedish advanced learners' writing in English. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 101–120). Amsterdam, The Netherlands: John Benjamins.

Hidalgo, E., Quereda, L., & Santana, J. (Eds.). (2007). *Corpora in the foreign language classroom*. Amsterdam, The Netherlands: Rodopi.

Hill, J., & Lewis, M. (1997). *LTP dictionary of selected collocations*. Hove, UK: Language Teaching.

Hoey, M. P. (2005). *Lexical priming. A new theory of words and language.* London, UK: Routledge.

Hornby, A. S. (Ed.). (2005). *Oxford advanced learner's dictionary* (7th ed.). Oxford, UK: Oxford University Press.

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge, UK: Cambridge University Press.

Hunston, S., & Francis, G. (2000). *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam, The Netherlands: John Benjamins.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Harmondsworth, UK: Penguin.

Hymes, D. (1992). The concept of communicative competence revisited. In M. Pütz. (Ed.), *Thirty years of linguistic evolution. Studies in honour of René Dirven on the occasion of his sixtieth birthday* (pp. 31–57). Amsterdam, The Netherlands: John Benjamins.

Johansson, S. (2007). Using corpora: From learning to research. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 17–30). Amsterdam, The Netherlands: Rodopi.

Johns, T. F. (1986). Microconcord: A language-learner's research tool. *System, 14*, 151–162.

Johns, T. F. (1991). Should you be persuaded—Two samples of data-driven learning materials. In T. F. Johns & P. King (Eds.), *Classroom concordancing. ELR Journal, 4*, 1–16.

Johns, T. F. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 27–45). Cambridge, UK: Cambridge University Press.

Johns, T. F. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115). London, UK: Longman.

Johns, T. F. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107–117). Amsterdam, The Netherlands: Rodopi.

Johns, T. F., Hsingchin, L., & Lixun, W. (2008). Integrating corpus-based CALL programs in teaching English through children's literature. *Computer Assisted Language Learning, 21*, 483–506.

Jones, R. L. (1997). Creating and using a corpus of spoken German. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 146–156). London, UK: Longman.

Jones, R. L. (2000). Textbook German and authentic spoken German: A corpus-based comparison. In B. Lewandowska-Tomaszczyk, P. J. Melia (Eds.), *PALC'99: Practical Applications in Language Corpora* (pp. 501–516). Frankfurt, Germany: Peter Lang.

Kaszubski, P. (2011). IFAConc - a pedagogic tool for online concordancing with EFL/EAP learners. In A. Frankenberg-Garcia, L. FLowerdew, & G. Aston (Eds.), *New trends in corpora and language learning* (pp. 81–104). London,UK: Continuum.

Kennedy, G. (1992). Preferred ways of putting things with implications for language teaching. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4–8 August 1991* (pp. 335–373). Berlin, Germany: Mouton de Gruyter.

Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology, 5*, 77–90. Retrieved from http://llt.msu.edu/vol5num3/kennedy/

Kennedy, C., & Miceli, T. (2002). The CWIC project: Developing and using a corpus for intermediate Italian students. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 183–192). Amsterdam, The Netherlands: Rodopi.

Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology, 14*, 28–44. Retrieved from http://llt.msu.edu/vol14num1/kennedymiceli.pdf

Kettemann, B. (1995). On the use of concordancing in ELT. *Arbeiten aus Anglistik und Amerikanistik, 20*, 29–41.

Kettemann, B., & Marko, G. (Eds.). (2002). *Teaching and learning by doing corpus analysis.* Amsterdam, The Netherlands: Rodopi.

Kettemann, B., & Marko, G. (Eds.). (2006). *Planing, gluing and painting corpora. Inside the applied corpus linguist's workshop*. Frankfurt, Germany: Peter Lang.

Kjellmer, G. (1984). Some thoughts on collocational distinctiveness. In J. Aarts & W. Meijs (Eds.), *Corpus linguistics. Recent developments in the use of computer corpora in English language research* (pp. 163–171). Amsterdam, The Netherlands: Rodopi.

Kübler, N. (2011). Working with corpora for translation teaching in a French-speaking setting. In A. Frankenberg-Garcia, L. Flowerdew, & G. Aston (Eds.), *New trends in corpora and language learning* (pp. 62–80). London, UK: Continuum.

Lea, D. (2002). *Oxford collocations dictionary for students of English.* Oxford, UK: Oxford University Press.

Leech, G. N. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1–23). London, UK: Longman.

Leńko-Szymańska, A. (2007). Past progressive or simple past? The acquisition of progressive aspect by Polish advanced learners of English. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 253–266). Amsterdam, The Netherlands: Rodopi.

Lewis, M. (1993). *The lexical approach*. Hove, UK: Language Teaching.

Lewis, M. (1997). *Implementing the lexical approach*. Hove, UK: Language Teaching.

Lewis, M. (2000). *Teaching collocation. Further developments in the lexical approach*. Hove, UK: Language Teaching.

Lombardo, L. (Ed.). (2009). *Using corpora to learn about language and discourse*. Frankfurt, Germany: Peter Lang.

Lorenz, G. (2002). Language corpora rock the base: On standard English grammar, perfective aspect and seemingly adverse corpus evidence. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 131–145). Amsterdam, The Netherlands: Rodopi.

McCarthy, M., McCarten, J., & Sandiford, H. (2005). *Touchstone student's book 1.* Cambridge, UK: Cambridge University Press.
McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies. An advanced resource book.* London, UK: Routledge.
Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 119–141). Amsterdam, The Netherlands: John Benjamins.
Meunier, F., & Gouverneur, C. (2007). The treatment of phraseology in ELT textbooks. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 119–140). Amsterdam, The Netherlands: Rodopi.
Meunier, F., & Granger, S. (Eds.). (2008). *Phraseology in foreign language learning and teaching*. Amsterdam, The Netherlands: John Benjamins.
Mindt, D. (1981). Angewandte Linguistik und Grammatik für den Englischunterricht. In P. Kunsmann & O. Kuhn (Eds.), *Weltsprache Englisch in Forschung und Lehre: Festschrift für Kurt Wächtler* (pp. 175–186). Berlin, Germany: Schmidt.
Mindt, D. (1987). *Sprache—Grammatik—Unterrichtsgrammatik. Futurischer Zeitbezug im Englischen I.* Frankfurt, Germany: Diesterweg.
Mindt, D. (1995). *An empirical grammar of the English verb. Modal verbs*. Berlin, Germany: Cornelsen.
Mindt, D. (1997). Corpora and the teaching of English in Germany. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 40–50). London, UK: Longman.
Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In E. Connor & T. A. Upton (Eds.), *Applied corpus linguistics. A multi-dimensional perspective* (pp. 239–250). Amsterdam, The Netherlands: Rodopi.
Nation, P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle.
Nattinger, J. R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly, 14*, 337–344.
Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, The Netherlands: John Benjamins.
O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching.* Cambridge, UK: Cambridge University Press.
Oksefjell-Ebeling, S. (2009). *Oslo interactive English:* Corpus-driven exercises on the Web. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 67–82). Amsterdam, The Netherlands: John Benjamins.
Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 191–226). London, UK: Longman.
Pérez-Paredes, P. (2003). Small corpora as assisting tools in the teaching of English news language: A preliminary tokens-based examination of Michael Swan's *Practical English Usage* news language wordlist. *ESP World, 6.* Retrieved from http://www.esp-world.info/articles_6/pascual.htm
Peters, P. (2004). *The Cambridge guide to English usage.* Cambridge, UK: Cambridge University Press.
Reppen, R. (2010). *Using corpora in the language classroom.* Cambridge, UK: Cambridge University Press.
Römer, U. (2004a). Comparing real and ideal language learner input. The use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 151–168). Amsterdam, The Netherlands: John Benjamins.
Römer, U. (2004b). A corpus-driven approach to modal auxiliaries and their didactics. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 185–199). Amsterdam, The Netherlands: John Benjamins.

Römer, U. (2005). *Progressives, patterns, pedagogy. A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam, The Netherlands: John Benjamins.

Römer, U. (2007). Looking at *looking:* Functions and contexts of progressives in spoken English and "school" English. In W. Teubert & R. Krishnamurthy (Eds.), *Corpus linguistics. Critical concepts in linguistics* (Vol. *4*, pp. 3–14). London, UK: Routledge. (Reprinted from *The changing face of corpus linguistics* by A. Renouf & A. Kehoe, Eds., 2006, Amsterdam, The Netherlands: Rodopi)

Römer, U. (2009a). Corpus research and practice: What help do teachers need and what can we offer? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 83–98). Amsterdam, The Netherlands: John Benjamins.

Römer, U. (2009b). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics, 7*, 140–162.

Römer, U. (2010). Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews. *English text construction, 3*, 95–119.

Römer, U., & Schulze, R. (Eds.). (2009). *Exploring the lexis-grammar interface.* Amsterdam, The Netherlands: John Benjamins.

Rundell, M. (Ed.). (2007). *Macmillan English dictionary for advanced learners* (2nd ed.). Oxford, UK: Macmillan.

Schlüter, N. (2002). *Present Perfect. Eine korpuslinguistische Analyse des englischen Perfekts mit Vermittlungsvorschlägen für den Sprachunterricht*. Tübingen, Germany: Narr.

Scott, M., & Tribble, C. (2006). *Textual patterns. Key words and corpus analysis in language education*. Amsterdam, The Netherlands: John Benjamins.

Simpson-Vlach, R. C., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31*, 487–512.

Sinclair, J. M. (Ed.). (1987). *Looking up: An account of the COBUILD project in lexical computing*. London, UK: Collins ELT.

Sinclair, J. M. (Ed.). (1990). *Collins COBUILD English grammar*. London: HarperCollins.

Sinclair, J. M. (1991). *Corpus concordance collocation.* Oxford, UK: Oxford University Press.

Sinclair, J. M. (Ed.). (1992). *Collins COBUILD English usage*. London: HarperCollins.

Sinclair, J. M. (1997). Corpus evidence in language description. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 27–39). London, UK: Longman.

Sinclair, J. M. (Ed.). (2001). *Collins COBUILD English dictionary for advanced learners*. London: HarperCollins.

Sinclair, J. M. (Ed.). (2004a). *How to use corpora in language teaching*. Amsterdam, The Netherlands: John Benjamins.

Sinclair, J. M. (2004b). Introduction. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 1–10). Amsterdam, The Netherlands: John Benjamins.

Sinclair, J. M. (2004c). New evidence, new priorities, new attitudes. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 271–299). Amsterdam, The Netherlands: John Benjamins.

Sinclair, J. M. (2004d). *Trust the text. Language, corpus and discourse.* London, UK: Routledge.

Sinclair, J. M., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter & M. McCarthy (Eds.), *Vocabulary in language teaching* (pp. 140–158). London, UK: Longman.

Sripicharn, P. (2004). Examining native speakers' and learners' investigation of the same concordance data and its implications for classroom concordancing with EFL learners. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 233–245). Amsterdam, The Netherlands: John Benjamins.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics.* Oxford, UK: Blackwell.

Tono, Y. (2011). TALC in action: Recent innovations in corpus-based English language teaching in Japan. In A. Frankenberg-Garcia, L. Flowerdew, & G. Aston (Eds.), *New trends in corpora and language learning* (pp. 3–25). London,UK: Continuum.

Tribble, C., & Jones, G. (1997). *Concordances in the classroom. A resource book for teachers*. Houston, TX: Athelstan.
West, M. (1953). *A general service list of English words*. London, UK: Longman.
Whistle, J. (1999). Concordancing with students using an "off-the-Web" corpus. *ReCALL, 11*, 74–80.
Wichmann, A., Fligelstone, S., McEnery, T., & Knowles, G. (Eds.). (1997). *Teaching and language corpora.* London, UK: Longman.
Willis, D. (1990). *The lexical syllabus. A new approach to language teaching*. London, UK: HarperCollins.
Willis, D., & Willis, J. (1989). *Collins COBUILD English course*. London, UK: HarperCollins.
Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning & Technology, 12*, 31–48. Retrieved from http://llt.msu.edu/vol12num2/yoon.pdf