ESL student attitudes toward corpus use in L2 writing

Hyunsook Yoon, Alan Hirvela

School of Teaching and Learning, Ohio State University, 333 Arps Hall, 1945 North High Street, Columbus, OH 43210-1172, USA

Abstract

In recent years, there has been growing interest in the use of corpora in L2 writing instruction. Many studies have argued for corpus use from a teacher’s perspective, that is, in terms of how teachers can develop instructional materials and activities involving a corpus-based orientation. In contrast, relatively little attention has been paid to investigations of learners’ actual use of corpora and their attitudes toward such use in the L2 writing classroom. This paper describes a study of corpus use in two ESL academic writing courses. Specifically, the study examined students’ corpus use behavior and their perceptions of the strengths and weaknesses of corpora as a second language writing tool. The study’s qualitative and quantitative data indicate that, overall, the students perceived the corpus approach as beneficial to the development of L2 writing skill and increased confidence toward L2 writing.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Corpus linguistics; Corpora; L2 writing instruction; Academic literacy

1. Introduction

In recent years, there has been growing interest in applying the results of corpus research to second language (L2) pedagogy. Here “corpus research” means research which gathers authentic uses of a target language from a wide variety of sources and then arranges this input in a database, or corpus, in such a way that learners can study specific language uses within the corpus. For instance, a learner of English who wants to use the word “access” in an essay might be confused about which words best “collocate,” or link, with that word. By connecting to a corpus database and then typing in the word “access,” also known as a keyword because it activates a search procedure and something called a “concordance...
"program," the student would eventually see a long list of actual uses of the word with different collocations. These uses occur in the form of portions of sentences, with the keyword positioned in the middle of them. The purpose of the concordance program is to produce such a list. Each example of the keyword in use is a concordance, and the concordancing program provides a series of concordances, as reflected in the following samples for the keyword “access”:

film and tv producers worldwide access to locations available for shooting
a particularly easy and fruitful access to that area of the mind. REP:
their own military contacts to gain access to the arsenals, and within hours
Arab landowner: his bedroom, with access to the boudoirs of his three wives,
seem to be when children have access to wide green spaces in which

The information above is known collectively by the more technical term, “concordance output.” Students can study this output, or information, closely and see how a keyword operates in context with closely related words. They can also obtain a second key type of output known as “collocate output.” Collocate output indicates how often (within the larger corpus) a keyword is collocated, or linked, with specific surrounding words. The popular Collins COBUILD Corpus, for example, provides a list of up to 100 collocates (e.g., prepositions) for each keyword, arranged in order of the frequency of their links with the keyword in the entire corpus. In the case of “access,” for example, students will learn that “to” is its most common collocate in the Collins COBUILD Corpus, and they will encounter additional statistical information about the frequency of this collocation. With this information at hand, students can then compose their own sentences using the keyword with increased knowledge of how to link the word with appropriate collocates.

The general aim of corpus research is to provide a rich collection, or corpus, of such uses covering as wide a range of words and word combinations as possible, and to provide information about the frequency of use of these word combinations.

While the volume of literature on spoken corpora has increased (e.g., Simpson & Swales, 2001), the primary area of interest has been written discourse, with much of the emphasis placed traditionally on analysis of literary texts (Hockey, 2001). With their emphasis on textual analysis and descriptions of written language in use, corpus-based studies are especially amenable to the teaching of reading and writing skills and the development of academic literacy (Flowerdew, 2002). In this vein, Jabbour (2001) points out that “a corpus approach befits teaching second language reading and writing, since both activities are text oriented and make use of words and word combinations, or lexical patterns, within the confines of discourse” (p. 294). In other words, a corpus approach provides meaningful input into the language side of L2 writing instruction. However important rhetorical concerns and issues may be in the teaching and learning of writing in another language, the language domain—syntax, punctuation, style, vocabulary development, etc.—also commands attention. L2 writers, particularly below the advanced level of proficiency, are also language learners eager to develop greater knowledge and command of the linguistic resources necessary for effective writing. A corpus is now seen as a primary contributor of these resources because, as Tao (2001) observes, of its “potential to make explicit the more common patterns of language use” (p. 116).
One of the central principles of the corpus approach to language descriptions is that vocabulary and grammar are interrelated rather than distinctive from each other (Halliday, 1992; Sinclair, 1991). Within this linking of vocabulary and grammar, also known as lexico-grammar, there is an emphasis on the co-occurrence or most frequent combinations of words, i.e., “collocation” (Biber & Conrad, 2001; Conrad, 2000). For instance, the noun “location” is often followed by the prepositions “of” and “for.” This combined focus on lexical input and grammatical function is of considerable value to someone acquiring English as a second or foreign language, as well as to teachers of ESL or EFL. Recent attempts to connect the corpus approach with genre analysis have been particularly beneficial in such domains as English for Academic Purposes (EAP) or English for Specific Purposes (ESP). Genre-based corpus analysis looks for common collocational frameworks within particular genres of communication, and in the process facilitates more effective communication of the kind expected inside specific genres, such as medical or legal English, where specialized word patterns are likely to occur. Given such emphases, the use of corpus data has become increasingly appealing in the context of L2 writing instruction, where the simultaneous focus on vocabulary, grammar, and discourse patterns provides second language writers with the kinds of target language input they especially need to achieve high levels of proficiency as L2 writers.

Along with the lexico-grammatical aspect, corpus-based activity has also been considered beneficial to L2 writing by offering learners a rich experience of real language use (Thurstun & Candlin, 1998). Corpora, which are, as noted earlier, databases of authentic language uses culled from multiple sources, provide learners with large quantities of real-life target language discourse. Exposure to these examples of genuine language use can (a) enrich learners’ understanding of specific uses of target words in a wide variety of contexts and (b) expand their L2 linguistic repertoire. Students’ encounters with these multiple samples of discourse combinations should then contribute to growth as L2 writers (and readers).

In light of such benefits of corpora use in L2 writing, many studies have argued for the inclusion of corpora in teaching materials and classroom activities. This is not a call for a corpus-dominated writing pedagogy, but rather the inclusion of some corpus-based activity with respect to treating the language side of L2 writing. In discussions of corpus work (including this paper), then, references to “corpus pedagogy” or related terminology refer only to those portions of a writing course syllabus devoted to language study incorporating a corpus component.

Also influencing this shift toward a corpus component has been growing concern about the traditional way in which teachers and textbook writers have provided language input for students. Biber (2001) points out that “these decisions have usually been based on the author’s gut-level impressions and anecdotal evidence of how speakers and writers use language” (p. 101). Teachers, too, may depend on their own notions of which language items are most useful for students to learn. Biber goes on to explain that “empirical analyses of representative corpora provide a much more solid foundation for descriptions of language use” (p. 101).

Despite these advantages of a corpus component in L2 writing instruction, a limitation in this field is the fact that the focus has been on corpus use from the perspectives of teachers and materials developers. In the process, little attention has been paid to the use of corpora...
from the perspectives of learners. To date, a meaningful body of empirical research investigating learners’ actual use of corpora and their attitudes toward such use has yet to emerge. Therefore, in addition to discussing how corpora might influence syllabus design and the construction of pedagogical tasks, it is also important, as Stevens (1991) says, “to speculate about what effects such exercises might have on the students’ writing performance” (p. 44). Likewise, compositionists can benefit from knowing more about how students react to such exercises and to what extent they see an impact on writing performance. We need, then, greater understanding of what actually happens when a corpus component is brought into a writing classroom.

The purpose of the study described in this paper, therefore, was to engage in such speculation by looking closely at how ESL students actually use corpora and how they feel about using the corpora in L2 writing instruction so as to gain deeper insight into how learners may benefit from corpus-based writing activity. Since students are expected to be the primary beneficiaries of such an experience, it is essential to know how they respond to a classroom environment in which corpus use plays a major role. With such results in hand, ESL educators can adjust corpus-based techniques so as to create more effective and meaningful learning conditions. Our study of student experiences of and responses to a corpus component in L2 writing courses involved a survey of student attitudes and follow-up interviews with selected students in one intermediate level and one advanced level ESL academic writing course at a large Midwestern research institution.

In this paper, we review the literature that has addressed the advantages and disadvantages of corpus use in the teaching of L2 writing. We then present the results of our study with respect to (a) students’ actual corpus use; (b) their views as to the strengths and weaknesses of the corpus activity; and (c) their overall evaluation of corpus use in L2 academic writing instruction.

2. Corpus use in the teaching of L2 writing

Corpora are believed to be beneficial in L2 pedagogy primarily because of the target language input they provide both learners and teachers, especially information about the frequency with which certain linguistic items are used and the most common word pairings, or collocations. As a number of corpus researchers have pointed out, the linguistic items most often used in authentic discourse are generally the most useful for learners, so “the most common words and their combinations should form the basis of instruction” (Jabbour, 2001, p. 298). A resulting objective in syllabus design, then, is to expose students to the language use that they are most likely to encounter outside the language classroom (Kennedy, 1987). Biber and Reppen (2002) explain that “although frequency information can never be the sole factor used to design materials, it does provide a more solid basis than relying only on intuitions and accepted practice” (p. 201).

As noted earlier, another useful application of corpora is to expose students to multiple authentic texts so as to expand their understanding of the functions of words in particular contexts or genres. By helping students acquire contextualized grammatical knowledge, this corpus-based approach teaches them (indirectly or directly, depending on the explicitness of the instruction) how and where to put words into sentences, which a
dictionary often fails to do (Odlin, 2001). Tribble (2002), in a study of a course employing corpora, argued that corpus-based activity increased “contextual and linguistic awareness raising during an EAP course” (p. 133), and this is the essence of corpora: their capacity for supplying L2 writers with the concrete linguistic input necessary for effective writing.

A third benefit of including a corpus component in a writing syllabus is that it encourages or enhances inductive language learning, an important feature of second language acquisition. Learners are able to make their own inductive discoveries about the target language as they examine multiple examples of a specific linguistic item in use. The result is student-centered discovery learning, which promotes self-confidence and mastery of the learning process (Johns, 1991; Stevens, 1995).

The computer-based nature of a corpus component can also favor ESL/EFL learning because it significantly increases opportunities for exposure to and contact with English language texts. That is, through such computer-based artifacts as the Internet and hypertexts, with their multiple links to other texts, students gain access to an almost limitless supply of target language texts (Conrad, 2000). The computer also makes it easy for students to select texts from whichever discoursal domains they prefer to work within, as opposed to relying on more general texts. A student interested in chemical engineering, for example, can quickly locate and surf chemical engineering-based Internet sites and select those texts of greatest linguistic value relative to her or his needs. Then, too, web-based corpora are becoming increasingly available for language learning and teaching purposes (Sun, 2000). For example, the computer provides access to the vast riches of the Bank of English, a storehouse of tens of millions of samples of authentic English collocations and other linguistic items based at the University of Birmingham in England. Students interested in developing their use of spoken English for academic purposes can explore the Michigan Corpus of Academic Spoken English (MICASE developed at the University of Michigan’s English Language Institute; see Mauranen, 2001; Powell & Simpson, 2001; Swales & Malczewski, 2001). Given this kind of access, students can become independent learners outside class in addition to utilizing such resources in the classroom environment. Hence, Gavioli and Aston (2001) encourage a view of corpora as “resources for learners to use directly, inside and outside the classroom” (p. 244).

While there are positive aspects to the inclusion of a corpus component, some learners may have difficulty in acquiring the skills needed to experience it, and not all students have access to the technology necessary to employ it. Then, too, some concordancing programs are particularly sophisticated and produce language data in formats difficult to interpret. Learners may feel overwhelmed as well as powerless in the face of such complex looking linguistic input. Johns (1994) also mentions the problem of trainability; that is, there is uncertainty as to whether and to what extent students can be trained to use corpus technology. In addition, Cobb (1997) has commented on the amount of time necessary for students to become accustomed to the new technology—time that could well be spent on other, more conventional and time-tested classroom practices.

Furthermore, research has yet to confirm the benefits of corpus-based activities, and only a few empirical studies have investigated student responses to corpus use. Some supplementary findings suggest that students have a positive attitude toward vocabulary learning while using corpora (Cobb, 1997; Thurston & Candlin, 1998). In Thurston and Candlin’s pilot study (1998), for example, participants reacted positively toward this innovation in
vocabulary teaching. However, they also reported some negative reactions, such as, “some students were puzzled by the cut-off sentences of the one-line concordances and daunted by the difficulty of the authentic academic texts” (p. 271). One student complained that “there are [too] many exercises for just three words” (p. 277).

While the previous research did not thoroughly investigate student responses, Sun’s (2000) study was aimed at exploring student reactions to a corpus-based lesson. Using a questionnaire approach, he examined Taiwanese EFL students’ feedback toward web-based concordancing. The majority of the students were positive toward the web-based concordancing, mostly because it allowed them to encounter authentic language use. Among its specific benefits, they felt the approach was most helpful in acquiring knowledge of the actual usage of individual words as well as phrases, and in reading comprehension. The least effective areas, in their view, were writing proficiency, application of grammatical rules, and understanding the meaning of vocabulary. The students also expressed concern about the slow speed of Internet connections and the time involved in conducting an analysis of concordance data. Qualitative data obtained through open-ended questions likewise showed that students perceived data analysis of concordance output as problematic due to the huge amount of data available and difficulties in adjusting to the inductive learning style fostered by corpus pedagogy. As Sun noted, “they seem most comfortable with the traditional role of learning, being in a deductive learning model” (p. 8).

The main value of Sun’s study was his emphasis on student evaluations of corpus-based activities. However, his use of a limited number of open-ended questions failed to provide sufficient qualitative insights into student difficulties in their use of concordancing. More importantly, Sun’s study was somewhat limited in scale. His study was based on student responses to a 3-week lesson designed mainly to teach them about the approach rather than an emphasis on actual use of the approach as a tool in their language classroom. This limits our understanding of how the corpus pedagogy can be incorporated into a conventional classroom framework.

Kennedy and Miceli (2001) conducted perhaps the most intriguing study of classroom teachers using a corpus component in their daily teaching activities. The study concerned teaching Italian as a foreign language in Australia. The researchers compiled their own small corpus that mainly consisted of texts of informal personal writing. Noting that “there is relatively little [research] on how learners actually go about investigations” (p. 77), they aimed to evaluate how independent students became in using the corpus for their own purposes after being trained in the approach. Of particular note here is their use of the “apprenticeship” approach in order to lead the students to successful corpus investigations “in a gradual and guided manner” (p. 77). Based on the results of their study, Kennedy and Miceli highlighted the need to first train students to develop “corpus research” skills necessary for successful corpus investigations prior to immersing them in the corpus activities themselves. Unfortunately, their paper focused only on the evaluation of their apprenticeship approach; reporting of the results of student responses to the corpus approach was consigned to a separate and not yet published paper. The main positive finding was the enhanced ability to provide authentic examples of word use in various contexts and specific expressions in particular types of text. As in Sun’s study, they found that the most serious problem is the time-consuming nature of data analysis.
Thus, Kennedy and Miceli’s work does not provide a fuller understanding of student evaluations of corpus-based pedagogical activity, information that is essential if we are to successfully implement a corpus component in L2 teaching, particularly L2 writing instruction, where such a component is likely to have its greatest value. In order to promote teachers’ and students’ successful pedagogical use of corpora, we need to examine how and in what ways a corpus component is beneficial to development of L2 writing ability, the nature of student difficulties in the use of a corpus, and what needs to be done to help learners overcome those difficulties. We can learn much by studying these issues from the perspectives of student writers.

Another interesting point about the studies by Sun, and Kennedy and Miceli is that both were conducted in foreign language learning contexts, which raises a question as to whether second language learning situations would reveal the same results. The present study was intended to fill the gaps identified above. That is, we used a qualitative format as well as a systematic quantitative methodology to more closely examine ESL student responses to corpus use in a context judged to represent a common ESL writing classroom situation. In particular, we were interested in looking at student use behaviors, their perceptions of the usefulness and disadvantages of the corpus activities, and finally their overall evaluations of the experience with respect to development of L2 academic writing ability.

3. The study

3.1. Research questions

Our study addressed the following four research questions,

(a) In what ways do ESL students think corpus use is beneficial for learning L2 writing?
(b) What difficulties do students have in using a corpus?
(c) How do students feel about using a corpus in writing instruction?
(d) What are students’ overall evaluations of corpus use in L2 academic writing?

3.2. Participants and context of the study

The students involved in the study were enrolled in either an intermediate or an advanced level ESL writing course at a large American university. The intermediate class contained 8 students, meeting every day for 48 min, while the advanced class of 15 students met twice per week for 2 hours per class session. Of the total of 23 students between the two courses, 22 students participated in the survey.

Of these 22 students, 15 were male and 7 female. The overwhelming majority, 82%, was from East Asia, predominantly Chinese (32%) and Korean (23%). Their length of stay in America ranged from 2 months to 2 years and 8 months, with 17 (77%) of these students being in America less than a year. They were from 15 different academic programs, the majority of which were science-related fields, such as engineering, biochemistry, and computer science. Four of the students, who were in the intermediate class, were working for their undergraduate degrees, while the rest were enrolled as graduate students.
While the two classes were aimed at different levels of student writing proficiency, they were taught by the same instructor, a veteran teacher of ESL composition who had gradually integrated a corpus component into his courses during the previous several quarters. This component was meant to enrich the language side of the courses; the courses themselves were designed so that rhetorical issues and topics dominated. The corpus approach he had adopted was somewhat generic in nature, that is, it was not tailored specifically to the level of the courses he taught. Consequently, the teacher used the same core corpus materials and techniques in each of the courses. There was, however, a difference in the amount of time he allotted to corpus use, with more classroom instructional time devoted to corpus use in the intermediate class. In that course the students made a weekly visit (for the first 4 weeks of the course) to a computer lab to conduct corpus searches. In addition, during the first few weeks of the term they engaged in a few classroom-based corpus practice exercises that ran in the neighborhood of 20–25 min per exercise. The instructor also spent up to 20 min per week reviewing the results of students’ corpus activities. With respect to the advanced course, the students performed the same introductory corpus-based practice exercises as the intermediate students in the first few weeks of the course. After that, the instructor devoted perhaps 20 min per week to commentary about the students’ corpus work. Most of the corpus work performed by students in the advanced level course actually took place outside the classroom in the form of supplementary activity.

Both classes worked with the Collins COBUILD Corpus, a large database consisting of millions of entries derived from an even larger database called the Bank of English. The Collins corpus was attractive in part because of its availability to students as a free web sampler, (access to it is available at http://www.titania.cobuild.co.uk) as opposed to a commercial program for which the right of use must be purchased. Because the vast majority of entries in the database come from written sources (e.g., letters, newspaper articles, memos, advertisements), this database appears to be appropriate for writing courses in addition to being highly accessible.

With respect to actual classroom implementation of the corpus component, the approach was as follows. The instructor’s overall philosophy was to lead the students from (early in the course) explanation and demonstration, i.e., guided searches, to (later in the course) independent corpus exploration, particularly in the advanced course. Early in each course, the teacher introduced the Collins COBUILD Corpus, including some framing of its roles in the class activities and the kinds of writing-related knowledge it would provide. He emphasized its value in enhancing students’ knowledge of vocabulary and grammar, and thus its positive effects on the development of writing ability.

Following introduction of the Collins COBUILD Corpus, students in both classes were given instructions on, and demonstrations of, how to conduct a concordance search and how to interpret the results of such a search. This meant teaching them about the concordancing programs described earlier.

The teacher then gave instructions on how to construct prototypes of words, i.e., descriptions or portraits of common usage patterns involving keywords (based on frequency of use information provided by the corpus database). That is, the students were asked to gather the “concordance output” (samples of the keyword and collocates
in use) and “collocate output” (statistical information about frequency of specific collocations in the corpus) of certain key words and then to create what the instructor called “prototype strings.” The prototype strings were, in essence, a synthesis of the students’ findings from the concordance and collocate output produced by the concordancing programs. The instructor’s belief was that these “strings” would better enable students to decide which collocations to use in their own writing. Fig. 1 shows a prototype string for the example word “access” discussed earlier.

To help students learn this process, in the first few weeks of the term the instructor arranged them in pairs and asked them to create, and later present to the class, prototype strings for key words he had provided on a master list. As the students gained familiarity with this process, they were asked to explore the Collins COBUILD database on their own with respect to actual word-related problems they were having in their writing (self-identified problems or concerns, or those supplied by the teacher as he read drafts of their papers and encountered problems worth exploring) and to submit their search results to the instructor by e-mail. The instructor gathered this information and shared it in class via handouts on a weekly basis.

By focusing on genuine problems frequently encountered during the process of writing and revising essays for the course, the instructor presented corpus searches and prototype construction as a problem-solving approach to the language side of L2 writing. To make the searches and subsequent class discussions more meaningful and more practical in value (relative to students’ language-based needs), he emphasized the use of content words, e.g., reporting verbs, as well as often troublesome grammatical features, such as prepositions, and demonstrated how the corpus could assist in such use. As the term unfolded, then, and students generated prototype strings and heard the teachers’ commentary on them in class, they gained valuable practice in the application of corpora to L2 writing and acquired knowledge of word collocations that would serve them in other real-life writing situations (such as composing papers for other courses). Indeed, the students left the two courses with a large collection of prototype strings of keywords, and these could serve as a resource for future writing situations (for a helpful pedagogical illustration of corpus use in a writing class, see Jabbour, 2001).

3.3. Data collection and analysis

In order to gain a more detailed look at corpus use, we conducted a study that covered the entire 10-week term as opposed to focusing on one unit of instructional time and activity of
much shorter duration. In order to achieve a clear understanding of how the corpus pedagogy was implemented, the first author attended class sessions and thus observed the implementation as it unfolded. In addition, the instructor was interviewed informally to gain further insight into the implementation process.

In the 18th week of the course, a questionnaire was administered to each of the two classes during a regular class period. The survey instrument was developed by examination of the literature, e.g., Sun (2000), along with consideration of the specific use of the corpus in the classes serving as research sites for the study. After the researchers had constructed the questionnaire items, the teacher was also asked to review the questionnaire for content validity. The questionnaire consisted of two parts. The first part contained 23 items designed to elicit information about the students’ personal background, computer use, and dictionary use. The second part focused on their responses to the corpus pedagogy, including their perceptions of its strengths and weaknesses. That part consisted of 42 statements to which the respondents were asked to indicate their degree of agreement on a scale of 1–6 (1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree). They also had an alternative option, “no opinion.” For the analysis of this Likert-type data, the mean scores of the students’ responses were computed relative to the six-point scale employed. To enhance presentation of the data, responses were coded into three categories, “agree,” “disagree,” and “no opinion,” by placing all positive answers into the “agree” and all negative answers into the “disagree” categories. Finally, the internal reliability of the instrument was checked by using Cronbach’s Alpha. The reliability was $r = 0.96$, indicating a high level of reliability (see Appendix A for the complete text of the questionnaire).

In addition to the administration of the questionnaire, semi-structured interviews were conducted during the remaining weeks of the term so as to further explore student views as expressed in their own words. In order to explore a wide range of views and provide a more comprehensive portrait of responses to corpus use, students were given the option of including their names on the questionnaire; this enabled us to interview those who gave negative feedback as well as positive feedback. While arranging the interviews, we tried to include the same number of students from both the intermediate and advanced classes. Of the 18 students who gave their names, four were asked and then agreed to participate in the follow-up interviews. Each was interviewed individually for an hour. The interviews were tape-recorded and transcribed for analysis.

4. Results

4.1. Survey results

We present the survey findings in the same order as that of the four research questions posed earlier. We also present the student responses separately by class for two reasons. First, the teacher placed different emphasis on corpus use in each class. He incorporated more corpus-related class activities into the intermediate class, e.g., letting students have hands-on practice, while the advanced class used the corpus rather independently on their
own, with less class time devoted to corpus work. Second, we assumed that the level of the class might have an effect on students’ corpus use and thus also affect their responses to it. In fact, it has been presumed in a number of corpus studies that corpus use is of greater benefit to more advanced learners (Johns, 1991; McCay, 1980; Tribble, 1991; Turnbull & Burston, 1998).

Before looking at student responses to the corpus pedagogy, we will briefly review some demographic information about the participants. With respect to student use of the computer, in general, all of the students responded that they liked to use one. A much larger number of students used a computer several times a day for personal purposes (91%) as opposed to schoolwork (59%). In their computer use for personal purposes, 3 of the 22 participants used English software only, while the rest used either their native language (L1) only or both English and L1. Also, 15 students (68%) reported that more than half of their total computer time was in their L1. Twenty had Internet access at home, with 12 using a cable modem or LAN, as opposed to telephone modem. These figures are important because students’ access to technology impacts their prospects for independent work with corpora. With regard to dictionary use, four students reported not using a dictionary at all. Among the 18 dictionary users, only 2 students used a monolingual dictionary, while 12 used a bilingual dictionary and 4 used both. Interestingly, just 1 student reported using a paper dictionary only, and only 2 students indicated use of both a paper and an electronic dictionary, while the rest used only an electronic dictionary, such as a computer program dictionary or online dictionary. This suggests a preference for using a computer in their language learning, a point of related interest with respect to feelings about corpus-based activity.

In addition to providing this background information about the participants, we need to acknowledge the limited size of the research population, especially in the intermediate course. Clearly, this inhibits generalizability of the results, and the interpretative comments we make about the student responses reflect tendencies only within these research sites. Still, we believe there is some interpretive value in the results beyond these research sites, though of course any larger conclusions can only be extremely tentative in nature and should be used only for the purpose of stimulating further discussion about the potential value of corpus use.

4.1.1. Student assessment of advantages of corpus use

Table 1 shows the student assessment of the usefulness of the corpus and addresses the first research question: “In what ways do ESL students think corpus use is beneficial for learning L2 writing?” The first column is from the intermediate class (hereafter class 1) and the second column from the advanced class (hereafter class 2). Several results are worth noting. While the majority of students in both groups (class 1 = 100%; class 2 = 93%) responded that the corpus use was the most helpful for learning the usage of vocabulary and phrases, a smaller number of students (class 1 = 88%; class 2 = 54%) agreed that acquiring the meaning of vocabulary is an advantage of corpus use. Class 2’s mean score of just 3.46 on the six-point scale is especially interesting in this regard and may be explained by the fact that it was comprised of more advanced students who were presumably equipped with a more extensive vocabulary in English.
In general, the students reported positive feelings about corpus use relative to writing. There was a particularly strong belief in the idea that the corpus use improved their general writing skills (class 1, $M = 5.38$; class 2, $M = 4.71$). There was a slight drop-off in terms of development of academic writing skills (class 1, $M = 5.0$; class 2, $M = 4.57$), but the responses were still rather favorable. As we can see, there was another slight drop-off with regards to increased confidence about L2 writing, but the students clearly felt that use of corpora does build confidence in L2 writing. That there was a stronger belief in the value of corpus activity for actual writing skills development than confidence about writing is interesting, as is the fact that mean scores for the intermediate class were higher in each of these writing measures than for the advanced class. Whether this can be accounted for by the greater in-class use of corpora with the intermediate students is unclear, but this possibility is worth noting.

Also worth noting is the reversal of feelings about the corpus and reading skills, with the corpus component not being seen as especially helpful (class 1, $M = 3.88$; class 2, $M = 2.86$). The rather low mean for the advanced group could simply be due to the possibility that, as advanced students, they felt less of a need for improvement in reading. In an effort to seek elaboration on possible differences between responses to reading and writing, we also asked whether they felt the corpus is more helpful for writing or for reading. Interestingly, 82% of the students responded that the corpus was only helpful for writing. A possible explanation here is that, since the corpus had been presented in class as a tool for writing development, students had little or no reason to link it to reading. Also, since a corpus provides chunks of discourse in the form of portions of individual sentences rather than extended pieces of connected discourse, there may not have been a solid basis for students to link reading to a corpus-based experience.

The relatively low scores with respect to grammar must also be examined. As explained earlier, corpora as pedagogical tools are foregrounded with respect to their grammatical

| Table 1 | Advantages of corpus use ($n = 22$) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Category        | Intermediate class | Advanced class |
| Usage of vocabulary | Helpful (%) | Not helpful (%) | Mean ± S.D. | Helpful (%) | Not helpful (%) | Mean ± S.D. |
| Usage of phrases | 100 | 0 | 5.50 ± 0.76 | 93 | 7 | 5.29 ± 0.91 |
| Meaning of vocabulary | 88 | 12 | 4.62 ± 1.30 | 54 | 46 | 3.46 ± 1.20 |
| Writing skill | 100 | 0 | 5.38 ± 0.74 | 100 | 0 | 4.71 ± 0.73 |
| Academic writing ability | 100 | 0 | 5.00 ± 0.76 | 86 | 14 | 4.57 ± 0.94 |
| Increased confidence about L2 writing | 88 | 12 | 4.88 ± 0.99 | 93 | 7 | 4.42 ± 0.79 |
| Grammar | 88 | 12 | 4.50 ± 1.20 | 79 | 21 | 4.00 ± 1.04 |
| TOEFL | 63 | 37 | 3.88 ± 1.46 | 50 | 50 | 2.91 ± 1.58 |
| Reading | 63 | 37 | 3.88 ± 1.25 | 29 | 71 | 2.86 ± 0.95 |

*a* 1: very unhelpful, 2: unhelpful, 3: somewhat unhelpful, 4: somewhat helpful, 5: helpful, 6: very helpful.
value. As such, the mean scores reported (class 1, \( M = 4.50 \); class 2, \( M = 4.00 \)) are not especially high relative to the emphasis placed on the grammatical usefulness of corpora. Why this occurred is unclear and in need of further investigation. However, one possible explanation is that the students had their own perceptions about grammar, perceptions that did not match those of advocates of the corpus approach. Given that a large majority (82%, as reported earlier) had come from Asian countries, where traditional structural and grammar-translation approaches to language teaching have long dominated, it was likely that this was how they had learned about English grammar, in which case they may not have related the lexical approach of corpora and the emphasis on collocations to grammatical value. In other words, the core notion of lexico-grammar was unfamiliar to them, and so grammar as presented through a corpus orientation did not conform to their notion of learning about grammar.

The continued trend of lower mean scores for advanced students in both the reading and grammar categories also stands out. Throughout Table 1 we see a less favorable response to corpus use among advanced level students, a finding that challenges the assumption in other studies cited earlier that advanced students are more likely to benefit from the use of corpora. This raises important questions about how corpora should be used in courses with such students. Whether there is less need or value in corpus activity with these students, or whether such activity would be more positively perceived if it played a greater role in terms of in-class activity (as it did with intermediate students in this study) than out of class activity, should be explored in future research.

4.1.2. Student problems/difficulties in corpus use

Table 2 provides an intriguing look at some specific areas of possible concern or difficulty with respect to corpus use and addresses the second research question: “What difficulties do students have in using a corpus?” Here, as with Table 1, it is especially interesting to look not just at responses to particular items, but to similarities and differences between the intermediate and advanced level students.

As we can see, for the most part, the mean scores of student reactions to the problems/difficulties in corpus use clustered in the 3.0–4.0 score range, indicating that, on the whole, students found corpora neither very easy nor very difficult to use. Difficulties, or perhaps frustration, with respect to the amount of time necessary to analyze data provided by corpora should be noted, though, especially among the intermediate students (\( M = 4.00 \), with 75% of those students reporting some level of difficulty). When considered in conjunction with the results reported in Table 1, where the intermediate (class 1) students consistently showed a more favorable response to corpus use than the advanced students, these Table 2 figures are particularly interesting. On the one hand, they suggest that, despite the time-consuming nature of the analysis of corpus data, intermediate students still saw benefits in their corpus activity. On the other hand, they raise a concern about students at this level—those who might, in fact, need or benefit most from corpora because of their intermediate level of ability—ultimately turning away from corpus use and its possible benefits out of frustration over what is involved in such use. In contrast, access to computer/Internet was the least concern in their use of the corpus, suggesting that the accessibility issue may not be as serious as some studies have indicated.
Meanwhile, we see that class 2 usually produced lower mean scores than class 1, revealing that they had fewer problems in the use of the corpus. For instance, they did not have much trouble in dealing with real texts and vocabulary in the corpus-supplied information, and data analysis did not take as much time for them, which is probably related to their more advanced level of English. On the other hand, there were a few categories—concordance output analysis (i.e., analysis of concordance samples provided by the corpus), search technique, and the availability of too many sentences in concordance output (i.e., in the lists of concordances supplied by the corpus searches)—in which class 2 presented higher mean scores than class 1, indicating that they had more problems than class 1. While the differences between the two groups are not large, and the mean scores for both groups are relatively low (indicating only modest difficulty in these areas), some speculation is in order. It seems likely that the teacher’s greater emphasis on corpus use in class 1 as opposed to class 2 may have contributed to a greater comfort level among the intermediate students. If so, this is a subtle reminder that writing teachers should not assume that advanced-level students automatically have advanced-level computer skills, and that time spent in class illustrating how to use corpora can be just as fruitful in the advanced class as in the intermediate class.

4.1.3. Student responses to corpus use in writing instruction

Table 3 reports students’ feelings about corpus use in writing instruction and addresses the third research question: “How do students feel about using a corpus in writing instruction?” In general, the students were positive about the corpus use in their classes, though responses among the advanced level students in class 2 were generally not as
favorable as those of the intermediate students in class 1. Here it is once again worth noting that this outcome differs from what others have assumed about the reaction of advanced students to corpus use. As Table 3 shows, in nearly every category they were asked about, the intermediate students had mean responses in the upper 4 to mid 5 range on the six-point scale. For these students, learning how to search for and work with corpora was a relatively easy and helpful experience. While these feelings were echoed, though not quite as strongly, among the advanced students, the relatively low mean scores for “hands-on practice was helpful” (3.69), “using constructed prototypes as a reference in writing” (3.86), and “using the corpus for other courses” (3.54) must be noted, particularly since they refer to actual use of corpora (and not knowledge of corpora).

Looking further at the results in Table 3, the fact that 21 of the 22 participants responded that the prototype construction in the classes was “a useful experience” for language learning is interesting in light of what was discussed earlier with respect to Table 2 results: the students’ lukewarm feelings about the corpus activity with regards to its impact on grammatical improvement. The mean scores here (class 1, $M = 5.50$; class 2, $M = 5.07$) suggest somewhat greater value for the corpus component in this domain than was reflected in Table 2 results.

Perhaps the most intriguing results in Table 3 are those relating to “using of the corpus for other courses” and “using constructed prototypes as a reference in writing.” As can be seen, the majority of class 1 students (88%) used the corpus when writing papers for other courses, too ($M = 4.75$), while a much smaller number of class 2 students (62%) used the corpus for other courses ($M = 3.54$). Similarly, while most of the class 1 students (88%) used the constructed prototypes distributed by the teacher every week as a reference in their writing ($M = 4.87$), a smaller number of class 2 students (71%) used them as a reference ($M = 3.86$). These differences between the intermediate and advanced students are

<table>
<thead>
<tr>
<th>Category</th>
<th>Intermediate class</th>
<th>Advanced class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agree (%)</td>
<td>Disagree (%)</td>
</tr>
<tr>
<td>Search technique was easy to learn</td>
<td>100   0</td>
<td>5.38   0.52</td>
</tr>
<tr>
<td>Hands-on practice was helpful</td>
<td>100   0</td>
<td>5.25   0.89</td>
</tr>
<tr>
<td>Constructing prototype was easy</td>
<td>100   0</td>
<td>4.75   0.71</td>
</tr>
<tr>
<td>Having confidence in constructing prototypes</td>
<td>88   12</td>
<td>4.25   1.49</td>
</tr>
<tr>
<td>Using constructed prototypes as a reference in writing</td>
<td>88   12</td>
<td>4.87   1.25</td>
</tr>
<tr>
<td>Constructing prototype is a useful experience</td>
<td>100   0</td>
<td>5.50   0.76</td>
</tr>
<tr>
<td>Understanding the purpose of corpus use in the class</td>
<td>88   12</td>
<td>4.88   1.25</td>
</tr>
<tr>
<td>Using the corpus by own choice</td>
<td>75   25</td>
<td>4.50   1.41</td>
</tr>
<tr>
<td>Using the corpus for other courses</td>
<td>88   12</td>
<td>4.75   1.39</td>
</tr>
<tr>
<td>Learn more, like more</td>
<td>88   12</td>
<td>5.00   1.07</td>
</tr>
</tbody>
</table>

\(^a\) 1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: somewhat agree, 5: agree, 6: strongly agree.
striking. We once again see the intermediate students responding much more favorably to the corpus activity. Whether this was because of their increased opportunities for in-class use of the corpora or some greater intrinsic appeal of corpora to students with more to learn about the target language, or some entirely different reason, is unclear but deserving of investigation.

4.1.4. Students’ overall evaluations of corpus use in L2 writing

Table 4 results offer insight into techniques for classroom application of corpora by looking at the problem-solving approach employed by the teacher as a way of leading the students to corpus use. They also address the fourth research question: ‘‘What are students’ overall evaluations of corpus use in L2 academic writing?’’ Table 4 reveals that class 1 students were generally satisfied with the information they obtained through corpus searches concerning problems in their writing ($M = 4.75$). In contrast, class 2 students expressed much less satisfaction toward their search results ($M = 4.21$). These results may help explain why, in the previously discussed tables, the advanced students were usually less favorable than the intermediate students in their responses to the corpus activity. That is, if they were not overly satisfied with what their corpus searches produced, i.e., the all-important outcome of corpus use, it was understandably difficult for them to express strong levels of happiness with the pedagogy.

For overall evaluations of corpus use in L2 writing, the students were asked whether they would recommend using the corpus to others. As seen in Table 5, both groups agreed, with some enthusiasm, that corpus use should be introduced to students in their home country as well as in ESL writing courses. Of special interest here, particularly in light of some of the results reported earlier, is that in two of the five response categories the class 2 (advanced) students were more positive than the class 1 students, albeit by small margins. Despite some fairly mild responses to corpus use in previous tables, the advanced students nevertheless felt somewhat strongly about the value of corpora when asked to make recommendations about their use in various settings. Just as this outcome runs against expectations aroused by earlier results, so too does the intermediate students’ relative lack of desire to recommend corpus use in their native countries. Why these results occurred is difficult to say. With respect to the intermediate students, it may be important to remember that they were using corpora in an English-rich environment, one that demanded effective language use in their various courses. Hence, for them the corpus activity had immediate

<table>
<thead>
<tr>
<th>Category</th>
<th>Intermediate class</th>
<th></th>
<th></th>
<th></th>
<th>Advanced class</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agree (%)</td>
<td>Disagree (%)</td>
<td>Mean$^a$</td>
<td>S.D.</td>
<td>Agree (%)</td>
<td>Disagree (%)</td>
<td>Mean$^a$</td>
</tr>
<tr>
<td>Searching corpus for writing problems</td>
<td>75</td>
<td>25</td>
<td>4.75</td>
<td>1.55</td>
<td>71</td>
<td>29</td>
<td>4.21</td>
</tr>
<tr>
<td>Getting the information from the search</td>
<td>88</td>
<td>12</td>
<td>4.88</td>
<td>0.89</td>
<td>79</td>
<td>21</td>
<td>3.86</td>
</tr>
</tbody>
</table>

instrumental value across the curriculum. However, they may have believed that students in their home countries, especially those in EFL settings, would not feel as interested in engaging corpora or undertaking the steps involved in learning how to use them given the lack of instrumental motivation at hand in those settings.

4.2. Follow-up interviews

To provide a more in-depth exploration of student evaluations of corpus use, we collected additional data via follow-up interviews with some of the participants. Here we felt it was especially important to allow students to speak in their own words about the corpus activity. Four students were asked to participate in the interviews based on their responses to the questionnaire, and they are identified by pseudonyms rather than their real names. As noted earlier, in an effort to gain a broad range of student perspectives, we sought balance in terms of their degree of satisfaction toward corpus use and their class level (intermediate and advanced). Consequently, from the intermediate class, we identified one student with a negative attitude (“Sammi”) toward corpus use and another student (“Nick”) with a positive attitude toward it. As for the advanced class, there was one student whose questionnaire feedback was generally negative (“Sachi”) and another with basically positive feedback on corpus use (“Jun”). The interviews explored the reasons underlying their attitudes toward corpora and their self-assessments of the value of corpus use in L2 writing. We also asked for a comparison between corpus use and dictionary use, given that the usefulness of corpora is often juxtaposed against that of dictionaries, a traditional resource for learners, particularly when writing.

Sammi was a 21-year-old male student from the United Arab Emirates majoring in political science. He had been in the United States for 1 year, during which he had taken some intensive, pre-admission English courses at the research site before seeking his undergraduate degree. Because he was the only student in the intermediate class who expressed mainly negative feelings about corpus use, we were especially eager to learn about his reasons for disliking the corpus pedagogy.

### Table 5
Student overall evaluations of corpus use ($n = 22$)

<table>
<thead>
<tr>
<th>Category</th>
<th>Intermediate class</th>
<th>Advanced class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agree (%)</td>
<td>Disagree (%)</td>
</tr>
<tr>
<td>Would recommend corpus use in the same future course</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Would recommend corpus use to other ESL students</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Would recommend corpus use to home country</td>
<td>63</td>
<td>37</td>
</tr>
<tr>
<td>Will use corpus in future writing</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Overall, corpus is a useful resource for writing</td>
<td>88</td>
<td>12</td>
</tr>
</tbody>
</table>

Interestingly, though, Sammi started the interview by indicating that his negative attitude had recently changed to a positive one. He explained that his initial negative attitude was due to difficulty in understanding the corpus approach and in acquiring the techniques for using corpora. He emphasized that it was not easy to learn the approach because it was his first exposure to the corpus. However, the continued use of corpora through the course assignments allowed him to gradually become familiar, and therefore comfortable, with them. Work with them seemed to help his writing, too, as reflected in his observation that “I have got fewer corrections on grammar from the teacher after using the corpus.” He added that an especially helpful class activity was to construct prototypes with classmates, an activity in which students engage in their own modest corpus construction. He went on to comment that all ESL writing courses should add a corpus component, a point of view in sharp contrast to feelings he experienced earlier in his exposure to corpora.

Despite his apparent conversion to corpora, Sammi raised one serious and noteworthy concern about corpus use: that it requires too much time in terms of interpreting the results of word searches in comparison to dictionary use. As he explained:

Interpretation of the corpus results is difficult. It gives me a headache. Look for all the sentences. I don’t know. For dictionary, just put the word, enter, and you find that. Dictionary doesn’t give me a headache. Dictionary doesn’t take long time. But the problem is dictionary give you a lot of words, but I don’t know which one is the right one in the sentence . . . . That is the problem with dictionary. It doesn’t take long time, but it gives me another answer. Doesn’t give me which one fit in this sentence.

Another participant from the intermediate class, Nick, was a 29-year-old male Romanian student pursuing his master’s degree in nuclear engineering. He had been in America for 3 months and was experiencing his first academic term at an American university. Consistent with his favorable reactions to corpus use in the questionnaire, Nick’s comments extolled the virtues of corpora. In particular, he emphasized the idea that the corpus was helpful for learning “actual patterns” of word usage. As he commented, “a dictionary is good for learning a definition, but does not give me the right sentences. I need a definition from dictionary, but find context in the corpora.” Furthermore, and unlike Sammi, time was not an issue for Nick; the interpretation of the corpus search data was not time-consuming. In addition, he did not feel that the amount of corpus information was overwhelming. On the contrary, he felt that there was not enough. Notably, Nick was extremely eager to improve his writing in English, with a particular focus on making his content easily understandable to readers. Having come to regard the corpus as a good resource for his English writing, Nick expressed the wish that he had learned about corpora earlier in his study of English. An interesting side note is that he had already recommended using corpora to his wife in Romania.

Sachi, one of the participants in the advanced level course, was a 25-year-old female student from Japan who was pursuing a master’s degree in Japanese linguistics and pedagogy. She had been in the United States for 2 years and 8 months. Yet, during most of the time, she had taught Japanese in a university foreign language department, so she had been in an academic setting as a student for only 6 months. She had been assigned to the advanced class on the basis of her performance on the university’s writing placement test; previous to this, she had taken a writing course at another university.
Overall, Sachi was negative about corpus use in ESL writing instruction. She reported that while it was interesting to her, as a linguistics student, to look at which words collocated with each other, she did not believe in the benefits of corpora for students from other disciplines, such as those in the sciences. She asserted that “the approach may take them too much time and they may not be interested in learning the too much linguistic stuff.” Perhaps because she had some prior experience with corpora through her linguistics studies, she seemed to conceptualize the corpus approach as a linguistic tool rather than as a pedagogical aid.

For Sachi, a major inconvenience of the corpus was that it was not portable, unlike a paper dictionary. Additionally, she complained that the corpus was not helpful for her academic writing because the corpus was a mixture of speech and writing. Nonetheless, she admitted that the corpus was useful for checking and confirming her writing. She commented that “I use a dictionary as the first choice, and if there is ambiguous, and [I want] to know the word in context, then [I] use the corpus as the second resource.”

The last participant, Jun, was a 30-year-old male Taiwanese who showed a highly positive attitude toward the corpus experience on the questionnaire. He had been in the United States for 7 months and was pursuing his Ph.D. degree in chemistry. He had taken a writing course from the same teacher the previous quarter and so had been introduced to the corpus activity prior to entering the advanced level course. This made him an especially valuable participant because he was able to compare the teacher’s use of corpora in the two different courses. He remarked:

the teacher talked more about the corpus in the last course. I learned it at that time. And now I like it more. Now even when I am writing e-mail, I open corpora. And you know, some phrases, sentences if you are not very confident [whether] it’s right or not, then I check corpora.

Jun reported that he almost always kept the corpus website on his computer screen in order to check his writing against the corpus. He explained that the corpus was particularly advantageous for learning how to put words together: “For dictionary the examples are not so many; sometimes I don’t find what I need but . . . uh . . . using corpora almost every time I can find what I want.” Like Nick, Jun also had a great desire to improve his L2 writing skill. With regard to possible difficulties with corpus use, he reported none. Of particular interest relative to comments and results reported earlier, for him time was not an issue. In fact, he indicated that one benefit of using corpora was how quickly he could find what he needed; by contrast, he felt that searches involving a paper dictionary generally take too long.

While all of the participants agreed that the class activity of prototype construction was a useful learning experience, they usually did not construct prototypes of words they needed in their own writing. Rather, the students looked for the specific usage of words in concordance lists supplied by corpus searches and then applied the relevant information to their writing. Another point of interest is that they generally did not attempt to memorize the results of their searches, which sometimes led them to search for the same words repeatedly. As Jun noted: “Usually I search again for the same problems. But if I search for many times, then finally I remember.”

With regard to the display of concordance information, the participants did not report any problems. Typically, the computer screen will display a long list of cut-off sentences.
using the word or phrase that is the object of the search. As we saw earlier with the “access” sample, rather than displaying complete sentences, the concordancing program will feature, in each line, a portion of the sentence containing the linguistic item in question. Thus, readers in these courses saw as many as 40 partial sentences (as noted previously in our introduction to the Collins COBUILD Corpus), with each containing the word or phrase they were interested in. For three of the four interviewees, the cut-off nature of sentences using the word(s) being investigated was not a concern for them because they were able to focus on the target words in the middle of the concordance output and then understand the context of the sentences. According to Nick, “It is not a problem. Because I focus on the search word, it does not bother me at all.” Only Sammi, the student whose survey responses were the most negative, reported difficulty in analyzing concordance information and complained about too many sentences being present on the screen. Here it is worth noting that Sammi was the youngest and the only undergraduate student participant in the study. He had fewer years of experience in learning English than the other participants. He had studied English for only 4 years, as opposed to the others’ average of 12 years. He reported that he did not learn much English grammar at school in his native country, where classes focused mostly on acquiring basic vocabulary. Perhaps, then, his difficulty in using the corpus was, to some extent, related to the relatively short duration and limited nature of his education in English. This possibility is consistent with the participants’ agreement that students with insufficient grammatical knowledge may have some difficulty in using corpora.

Among the linguistic features of greatest interest to the participants, the usage of prepositions was what the students emphasized most in their corpus searches. A comment from Nick exemplifies this situation best: “My problem in English writing is prepositions. The corpus is especially helpful for solving the problem.” All of them pointed out the effectiveness of corpus use for acquiring the (verb + preposition) construction by looking at the multitude of examples supplied by the corpora.

With respect to their corpus behavior, the participants especially liked using the corpus to monitor their writing and see whether their linguistic choices and phrasing were correct. Sachi, for instance, remarked that “I see the examples, and if my sentence is there, I think that it is right.” This conceptualization of corpora as a hypothesis-testing tool may have been closely linked to their increased confidence about writing arising from corpus use. Of special note here is that all of the interview participants, including those who were relatively negative toward corpus use, acknowledged that its use increased their confidence about L2 writing to some extent. For example, Sammi mentioned that “if I looked at the words in the corpus, I have a confidence because I know it’s what other people use.” Jun noted that “using corpus increased my confidence in English writing because I know many people write in this way, so I have confidence if I follow it.” It is worth remembering, too, the results reported in Table 1 when students were asked about the impact of corpus use on their confidence as writers in English. Mean scores on this item were 4.88 for class 1 and 4.42 for class 2. These scores suggest a relatively strong belief in the power of corpora to strengthen confidence in L2 writing, at least among these students.

In summary, while each interview participant reported different problems in corpus use, all of the participants agreed that the corpus was especially helpful for learning common
usage patterns of words. And here they differentiated between traditional dictionary use and corpus use in the same way. That is, they agreed that a dictionary is useful for acquiring the meaning of words, but a corpus is more useful for learning how and where to put words in context. Perhaps more importantly, they all indicated that the corpus use contributed to increased confidence about L2 writing.

5. Discussion

The results of the survey indicated that the students were generally positive toward the use of corpus activities in L2 writing. They agreed that corpus use was particularly useful for acquiring usage patterns of words and enhancing their writing skill. On the other hand, they did not report serious difficulty in gaining access to computers and the Internet, as opposed to Sun’s (2000) study that identified the speed and stability of Internet connections as major problems in corpus use. In addition, while some studies, e.g., Sun’s (2000) and Thurston and Candlin’s (1998), presented concerns about the new forms of display of concordance information, the students in our study did not report the same level of concern. Both the intermediate and advanced learners generally seemed to know how to deal with words in context, as corpora display them.

On the whole, the participants of the study were positive about corpus use in L2 writing instruction, with the students from the intermediate class usually showing more positive responses than the advanced class students. Several explanations for this, some mentioned briefly earlier, can be offered. First, differences in the amount and type of emphasis by the teacher on the corpus work in each class may have had an effect on students perceptions of corpus use. The intermediate class was engaged in more hands-on and in-class corpus-related class activities than the advanced class, which was encouraged to explore the corpus outside class. Therefore, having received more direct training and practice in corpus use, the intermediate class students may have been better positioned to develop more favorable feelings about corpora. It is also important to note that the intermediate students reported fewer problems in corpus use than the advanced students, suggesting that the former achieved a greater level of mastery of corpus search techniques as a result of their additional training and in-class instruction. In addition, some study participants expressed a preference for print dictionaries and grammar books over corpora due to
their direct presentation of information. However, students in our study recognized the strengths of each resource and found that they could complement each other. This distinction between, as well as combining of, corpus use and the time-honored use of dictionaries may cast some new light on the roles corpora can play in L2 learning and the integration of a new approach with conventional and more familiar methods (Tribble, 1999). Finally, and as we saw earlier, the interview results also suggested that the degree of corpus use may be closely related to, or inspired by, student enthusiasm for improvement in their L2 writing. Two participants favoring corpus use showed a great interest in enhancing English writing. Also worth noting here is a possible cause and effect relationship between these variables and the enthusiasm displayed toward the idea of using corpora for writing in other courses besides ESL courses and for recommending corpus use to others.

While these results generally appear to suggest an endorsement for corpora use, it must be remembered, as noted earlier, that the sample sizes were small and thus prevent generalizability of the study’s findings. Replication of this kind of study with much larger research populations is necessary in order to offer definitive conclusions about corpus use.

As a closing remark, we must also mention that the teacher contributed to students’ positive attitudes toward corpus use. Student responses in the survey and interviews suggest that he successfully incorporated corpus use into different levels of writing classes while effectively adjusting the degree and nature of emphasis on corpus work. In so doing, he made corpus use a learning event that was manageable in each class as well as beneficial and relevant for students’ L2 writing. On the other hand, the generally less enthusiastic, while still positive, responses of the advanced level students suggest that an emphasis on leaving students to their own devices in corpus work may need to be re-examined in the design of future pedagogical applications of corpora.

6. Concluding remarks

In this study, we explored ESL student attitudes toward corpus use so as to shed some very tentative but hopefully useful light on students’ own evaluations of its value in L2 writing. The findings of the study suggest that, overall, the students perceived corpus activity to be beneficial for their English writing, particularly for learning common usage and collocates of words and for building confidence in their writing.

One important point emerging from the study is that the level of student proficiency may be worth special consideration when deciding whether and how to incorporate a corpus component into the teaching of L2 writing. Some students may need more time and practice to become familiar with this new type of activity. It may be pedagogically unwise to simply let students browse through a tremendous amount of sample texts or corpora without proper guidelines. They are perhaps better served by being given some guiding principles to experiment with as they learn this new approach to acquiring lexical and grammatical input. Therefore, at an earlier stage of corpus use, some form of apprenticeship could perhaps be included in course planning and execution. In particular, special attention might be paid to less advanced students so that they can go about using the approach successfully relative to their less extensive knowledge of the target language.
With respect to more advanced learners, encouraging them to conduct corpus searches and analyses independently of classroom applications may well be suitable, but the results of our study suggest that caution should probably be applied. Asking these students about their learning preferences in this regard may be preferable to simply assuming that their advanced status ensures that they will embrace corpora on their own, or with minimal guidance from the instructor.

In conclusion, it should be noted that corpus use as presented here represents one of many possible ways to employ corpora as a tool in L2 writing instruction. Furthermore, it must be remembered that corpus-based work is just one instructional component in a writing course and is not meant to be the dominant pedagogical tool employed. In the courses described in our study, corpus use was, overall, a small part of the larger instructional picture, and this should be construed as a common teaching practice. Consistent with these remarks, an observation by Flowerdew (1996) is especially worth noting:

There is a danger of the enthusiasm for concordancing being inflated to such an extent that concordancing is seen as a sort of language teaching panacea. Carefully conducted evaluative studies will ensure that such an inflated view will not prevail. Instead of concordancing being promoted as a panacea, considered evaluative studies will allow concordancing to be incorporated appropriately into the teacher’s battery of reference teaching resources as the useful additional teaching and learning tool that it undoubtedly is. (p. 112)

Thus, while there may be value in corpus activity, as suggested by the results of our study, in order to ensure more effective learning and teaching, L2 writing instructors need to be prudent in their application of corpus practices.

Acknowledgements

We would like to thank Dr. Jack Rouzer, the course instructor, and his students for their cooperation during this study. We would also like to thank the reviewers of this manuscript and the journal editors for their insightful comments on the paper.

References

Appendix A. Questionnaire about using the Collins COBUILD Corpus in ESL writing

A.1. Background information

1. Name ____________________________ 2. Email address ____________________________

3. Gender: Male _____ Female _____ 4. Program: Undergraduate _____ Graduate _____

5. Major subject (e.g., Business) ____________________________

6. Country you are from ____________________________ 7. Native language ____________________________

8. How long have you been in America? ________ Years ________ Months

9. How long have you been in an American university setting? ________ Years ________ Months

10. In general, do you like to use computer? Yes _______ No _______

11. How often do you use computer for personal purposes (e.g., email)? (Check one)

Several times a day _____ About once a day _____ About 5 times a week _____

About once a week _____ About once a month _____ Seldom _____

Others(list): ____________________________

12. How often do you use computer for school work (e.g., writing a paper)? (Check one)

Several times a day _____ About once a day _____ About 5 times a week _____

About once a week _____ About once a month _____ Seldom _____

Others(list): ____________________________

13. When you use computer for personal purposes, do you use English or your native language?

English _____ Native language _____ Both _____ Others(list): ____________________________

14. How much of your total computer time is in your native language?

Almost none _____ About 25% _____ About 50% _____ About 75% _____ Others(list): ____________

15. Do you have Internet access at home? Yes _______ No _______

16. If yes, do you use the Internet through modem or LAN at home? Modem _______ LAN _______

17. Do you use a dictionary for English writing? Yes _______ No _______

18. If yes, what kind of dictionary do you often use? Please check all that apply.

Bilingual (e.g., English-Chinese) _____ Monolingual (English-English) _____

Paper dictionary _____ On-line dictionary _____ Electronic program dictionary _____

19. Had you heard about corpora before you took this class? Yes _______ No _______

20. Had you used corpora before you took this class? Yes _______ No _______

21. If yes, which corpora did you use? ____________________________

22. Do you use the Collins COBUILD Corpus for your composing and/or revising papers?

Composing _____ Revising _____ Both _____ Others (list): ____________________________

23. Do you use the corpora at home and/or school? Home _____ School _____ Both _____
A.2. Reactions to using the Collins COBUILD Corpus

The following questions are regarding your opinions on using the Collins COBUILD Corpus. Please use the scale below to circle the response that most closely resembles your perspectives.

1: strongly disagree  
2: disagree  
3: somewhat disagree  
4: somewhat agree  
5: agree  
6: strongly agree  
N: no opinion

<table>
<thead>
<tr>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The searching technique was easy to learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. The practice sessions in the computer lab were helpful for learning the technique</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. It is easy to construct prototype strings by use of concordance/collocate output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Constructing prototype is a useful experience for language learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I feel confident in constructing prototype strings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I use the prototype strings that the teacher gave us as a reference when I write</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. The corpus is more helpful than a dictionary for my English writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Using the corpus is helpful for learning the meaning of vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Using the corpus is helpful for learning the usage of vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Using the corpus is helpful for learning the usage of phrases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Using the corpus is helpful for learning grammar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Using the corpus improved my English reading skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Using the corpus improved my English writing skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Using the corpus improved my English academic writing ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. I have some difficulty in using the corpus due to limited access to computer/Internet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. I have some difficulty in using the corpus due to the speed of Internet connection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. I have some difficulty in using the corpus due to time and effort spent on analyzing the data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
18. I have some difficulty in using the corpus due to unfamiliar vocabulary on concordance/collocate output
1 2 3 4 5 6 N

19. I have some difficulty in using the corpus due to cut-off sentences in concordance output
1 2 3 4 5 6 N

20. I have some difficulty in using the corpus due to too many sentences in concordance output
1 2 3 4 5 6 N

21. I have some difficulty in using the corpus due to the limited number of sentences in concordance output
1 2 3 4 5 6 N

22. I have some difficulty in analyzing concordance output
1 2 3 4 5 6 N

23. I have some difficulty in analyzing collocate output
1 2 3 4 5 6 N

24. I have some difficulty in performing the search technique
1 2 3 4 5 6 N

25. The real texts in the corpus are too difficult to understand
1 2 3 4 5 6 N

26. I often use the corpus by my own choice
1 2 3 4 5 6 N

27. I understand the purpose of using the corpus in this course
1 2 3 4 5 6 N

28. I want to use the corpus in my next ESL writing courses too
1 2 3 4 5 6 N

29. When I have problems in English writing, I search for help in the corpus
1 2 3 4 5 6 N

30. When I search for information in the corpus, I usually get the information that I need
1 2 3 4 5 6 N

31. I use the corpus when writing papers for other courses too
1 2 3 4 5 6 N

32. As I have learned more about the corpus, I have come to like them more
1 2 3 4 5 6 N

33. I will use the corpus for my English writing in the future
1 2 3 4 5 6 N

34. Learning about the corpus has increased my confidence about writing in English
1 2 3 4 5 6 N

35. If I had used the corpus earlier, I would have had a better performance on TOEFL
1 2 3 4 5 6 N

36. Overall, the corpus is a very useful resource for my English writing
1 2 3 4 5 6 N

37. Corpus use is more helpful for writing than for reading in English
1 2 3 4 5 6 N

38. Corpus use is more helpful for reading than for writing in English
1 2 3 4 5 6 N

39. The corpus should be introduced in all ESL writing courses
1 2 3 4 5 6 N

40. I recommend using the corpus in the same course in future quarters
1 2 3 4 5 6 N

41. Corpus use should be taught in English classes in my home country
1 2 3 4 5 6 N

42. I will recommend the corpus to other international students at OSU or elsewhere
1 2 3 4 5 6 N