Corpus consultation is gaining in prominence as a language learning tool. This approach to language analysis has made its way into the language classroom where its presence ranges from the presentation of printed concordance data with accompanying tasks to the direct use of concordancing software by learners themselves to carry out analyses of self-selected language features. Activities of the latter kind place concordancers fairly and squarely alongside dictionaries and grammar books as significant tools in the language learner’s kit. Recent studies have indicated that research is needed to provide support for the integration of corpus consultation into the language learning environment. Here, the response of second year undergraduate EAL students was examined to a course assignment that required them to investigate language features characteristic of a range of genres using a popular concordancing software program, Wordsmith Tools. Results showed that students generally had a positive response to corpus consultation and were able to identify benefits clearly, particularly in the areas of vocabulary acquisition and increased awareness of syntactic patterns. Most of the students indicated they are likely to use concordancers in the future and this interest is strongest amongst those students who have clear goals for their language learning. Course assignments produced by the students demonstrated an increased awareness of lexico-grammatical usage, particularly with regard to vocabulary use, phrases and colligational patterns. A number of obstacles to greater uptake of concordancing are identified and suggestions are made to overcome those obstacles.

Keywords: corpus consultation; concordancer; concordancing; genre; corpus linguistics; corpora; Wordsmith Tools

Introduction

Corpus consultation in general – and concordancing in particular – have been making slow but steady progress into the language learning classroom for some 30 years now. Initially corpora were consulted almost exclusively by language teachers but there is now increasing emphasis on the benefits of providing language learners with access to corpus data.

Much that has been published in this area to date has focused on the types of classroom activities that students can do with corpus data, either mediated by...
teacher selection or through direct access by the learners themselves using concordancing software such as Wordsmith Tools (Scott, 1999). Many studies point to the advantages of this approach but few have researched the extent to which learners actually benefit, particularly from exposure to extended and large-scale incorporation of corpus consultation into the language syllabus.

The study reported here involved EAL students (n = 19) who were taking an undergraduate second year BA course, the aim of which is to enable advanced students of English to develop their language proficiency by consulting corpora using concordancing software. A key outcome is that the students on the course will become ‘language researchers’ who are able to use a concordancer to gain linguistic insights into how language features are used, both in large general corpora or in smaller specialised corpora that represent genres of English. The idea is that they will use this knowledge to supplement that gained from dictionaries, grammar textbooks or tutors.

The study sought to investigate the extent to which this is actually the case and focused on issues such as: how do the students perceive the concordancing approach alongside more traditional tools such as a dictionary? Which would they prefer and when and why? What strategies do they use when working with concordancers? What advantages do they think the concordancer brings – and what may be the obstacles to them adopting it as a tool in their future language learning?

Whilst it was a small-scale local project it could also assist language teachers elsewhere to identify the benefits of concordancing as well as to identify which strategies to develop amongst students who are being introduced to this approach. The English language studies course appears to be one of very few courses internationally that has adopted corpus consultation as a major tool for language learning and it therefore provides fertile ground for research in this field.

The following questions were addressed:

(1) How can classroom concordancing complement traditional language learning resources such as a dictionary and grammar textbook?

(a) What benefits do students perceive that concordancing brings to language learning?

(b) How does concordance knowledge assist with vocabulary acquisition? (i.e. what does it add to dictionary knowledge?)

(c) How does concordance knowledge assist with acquisition of grammatical features? (i.e. what does it add to knowledge accessible from a grammar textbook?)

Background reading

Although there is evidence of the use of corpus data in language teaching from as early as 1969 (McEnery & Wilson, 1997) interest in small-scale corpus concordancing was not seriously awakened until the mid-1980s, most notably with the work of Johns (1986 & 1988, cited in Chambers, 2005). Since then, computerised text analysis has been adopted by teachers, course designers, materials developers and, most recently, by learners. This latter use is designed to promote what Johns (1997) has described as data-driven learning, or DDL. In DDL, ‘the learners’ own discovery of grammar based on evidence from authentic language use becomes central to the learning process’ (Stevens, 1995, p. 3), i.e. the learner as language detective.
Sinclair (1997) examined the potential impact of computer-processed language data on language teaching and indicated why language teachers should pay attention to developments in corpus linguistics. Firstly, a corpus approach supports the use of examples of real language in the classroom (as opposed to invented ones) and corpus data can provide language teachers and learners with illuminating (and often counter-intuitive) guidance as to frequent collocations and other language patterns. Form-meaning links can be taught in order to minimise the learning load and the language learner can use corpus evidence to help develop individual creativity in language use.

In the literature numerous activities have been suggested for classroom concordancing. Tribble and Jones (1997) provide a bank of approaches and tasks that has become a classic resource for the concordance-minded teacher. Stevens (1995), Wichmann et al. (1997), Aston (2001), and Henry and Roseberry (2001) identify the wide range of learning contexts that concordancing now finds an increasingly frequent place in – from academic reading courses to investigating idiom in newspaper texts. Even course books are now written using a concordancing approach, such as *Exploring academic English: A workbook for student essay writing* (Thurstun & Candlin, 1997) and, more recently, *Phrasal verbs in American English* (Barlow & Burdine, 2006). Corpus evidence is therefore increasingly likely to be presented to language students in some form or other.

Amongst the emerging uses of corpus consultation, one with clear potential for advanced EAL learners is in the area of genre analysis, in which learners use concordancing in order to gain knowledge of the characteristic features of particular types of English. Whilst the term genre has been defined in numerous ways in relation to linguistic and other fields (e.g. Swales, 1990; Bhatia, 1993; Biber, Conrad, & Reppen, 1998), it can also be said that it has a reasonably clear application with regard to corpora, in its sense of discourse type (Kennedy, 1998) or ‘discourse occurring in a particular context’ (Hunston, 2002, p. 160). The Brown Corpus, published in 1964, set a standard for later corpora with its sampling of texts from a large number of genres or text categories. This approach continues today and, in addition, there has been a recent trend to create small micro corpora which can be used to investigate the language features, or register, of particular genres (Ghadessy, Henry, & Roseberry, 2001). With regard to applications of this type of genre analysis in language teaching, Cobb (1997) reports on a study in which Omani students created their own dictionaries using evidence gained by concordancing a corpus of commercial English. In addition, Lin (1999) used concordancing to raise students’ awareness of the linguistic features of the report genre, Fuentes uses corpora to teach ESP to undergraduate computing and business students (2006) whilst Gavioli (2001) introduced learners to corpus consultation by getting them to investigate the language features of a corpus of lonely hearts advertisements. As Paltridge (2001) affirms: ‘Corpus studies clearly have much to offer descriptions of genres and their application in the language learning classroom’.

Chambers (2007) reviews progress in facilitating direct corpus consultation amongst language learners, noting that such activities appear to be on the increase. In an earlier 2005 study she reports on a second year undergraduate module on language and technology at the University of Limerick. The module aims to encourage undergraduate EAL students to use corpora as a resource in their language learning alongside other resources such as a dictionary or course book. She reports positive responses from students but claims that despite her study and other
publications that emphasise the benefits of corpus consultation for language learners (e.g. Bernardini, 2002), there is little evidence so far to suggest that direct corpus consultation is coming to be seen as a complement or alternative to consultation of a dictionary, course book, or grammar by the majority of learners. She points to the need for research to underpin the integration of corpora and concordancing into the language-learning environment. The study reported here sought to carry out such research.

The setting
The research involved 19 students who were enrolled in the English language studies course in the Bachelor of Arts in English as an additional language (BA EAL) at a public tertiary institution in New Zealand. Eighteen were female and one was male. Seventeen were from China, one was from Iran and one was from Morocco.

The English language studies course is compulsory for BA EAL major students. However, it can also be taken as part of a minor sequence by students of one of the other four BA language majors, as an elective by students in the Graduate Certificate in English as an additional language (GCertEAL) or by students from other undergraduate programmes in the institution e.g. the Bachelor of Business (BBus). It can also be studied as a stand-alone course by part-time students (Certificate of Proficiency: COP).

The students have two hours of classroom-based and two hours of computer lab-based study per week for 14 weeks, followed by a study week and an exam week. In the classroom-based theory sessions, the syllabus includes an introduction to corpus analysis and the notions of corpus and genre, discussion of a selection of corpus findings that have relevance for language learners, selected readings on the benefits of concordancing and a programme of concordance-focused tasks that are designed to illuminate key lexico-grammatical aspects of English. In order to develop awareness of different genres, samples of different text types are analysed with reference to a list of lexico-grammatical features that may be more or less frequent in each type. In the weekly computer lab-based sessions students are progressively taught how to use the three applications in the Wordsmith Tools suite: Wordlist, Concord and Keyword, which, according to their author, bring to concordancing the versatility exemplified by the Swiss army knife (Scott, 2001). The students learn to create frequency lists from different corpora (and sub-genres within them); to produce concordances and to access and assess data relating to collocations and clusters; and to produce Keyword lists that identify words that are unusually frequent in a small corpus in comparison with a large reference corpus. The language focus of concordancing tasks early in the course is quite prescriptive and guided, but there is increasing opportunity for students to focus on aspects of language use that are problematic for them and for which a traditional grammar book/dictionary approach may not have been successful. A group project early in the course requires the students to work together to analyse selected lexico-grammatical features as exemplified in a large general corpus. A subsequent individual project requires each student to focus on a particular genre of English (chosen for reasons of study, employment or personal interest), to identify and investigate language features that appear to be typical of the chosen genre and to write a report to present their findings. The features are identified through a combination of prediction, reading of text samples and use of the Wordlist and Keyword functions of Wordsmith Tools.
Key project findings are later summarised in a five-minute oral presentation to the whole class. It is this second assignment that provides the focus of this study.

Wordsmith Tools Version 4 is installed in all five computer labs in the School of Language Studies. Students on the BA programme can access the program in class time or during scheduled drop-in times. In addition to Wordsmith Tools, students on the English language studies course are also introduced to online concordancers available via such sites as the Virtual Language Centre (www.edict.com.hk/vlc) or Tom Cobb’s Compleat Lexical Tutor site (www.lextutor.ca). The school has a growing collection of corpora for students to access. The main corpora used for training purposes and for introductory studies are the Wellington Corpus of Spoken New Zealand English and the Wellington Corpus of Written New Zealand English (one million words each) but students can also access such corpora as MICASE (Michigan Corpus of Spoken Academic English), ACE (Australian Corpus of English), the Brown Corpus, the LOB (Lancaster-Oslo/Bergen) Corpus and the LLC (London-Lund) Corpus. This range enables exploration of regional variation as well as variation between spoken and written English. Students are also taught how to create their own small specialist corpus and several do choose to create one for their individual assignment.

Methodology

Studies into the use and effects of corpus consultation with language learners have been both quantitative and qualitative, although Chambers (2007), in surveying the types of methodologies used to date, notes that the majority have so far been qualitative. Cobb (1997) suggests a number of reasons why quantitative studies have so far been problematic in investigations of corpus consultation, including the lack of user protocols in the software and the length of time taken to become familiar with the software. Chambers (2007) claims that the lack of quantitative studies is no weakness and that, at this early stage of the development of corpus consultation, the qualitative studies currently being carried out are very useful for researchers and practitioners keen to discover what is being done in the field, what is working and what problems have been encountered. This study follows the current trend in that it is largely qualitative but also includes quantitative information regarding students’ concordancing usage.

An introductory questionnaire (Appendix 2) was administered to gain information about the students’ language learning preferences, particularly with regard to the use of dictionaries and grammar books and the use of computer assisted language learning (CALL). The questionnaire was piloted with a group of students from the BA EAL course but not from the class that is the subject of the research.

The students carried out an individual project between Weeks nine and 13 of the course. The project required them to select a genre of spoken or written English and, using concordancing software, to investigate, two to four lexical or grammatical features that they identified as being characteristic of their chosen genre. They were asked to complete a reflective log (Appendix 3) as they went through the various stages of the project. The log focused their attention on how they were approaching the use of corpus data for the analysis of language. It asked them to consider the strategies they were using and also asked them to comment on the knowledge they felt the concordancer gave them and to compare their findings to a dictionary or grammar text.
On completion of the project five students were selected to join a semi-structured
discussion (Appendix 4) in order to explore further some of the comments they made
in their reflective logs. The students were selected on the basis of the type of
comments they made in their logs, in an attempt to obtain a range of attitudes
towards a corpus-based approach to language learning. They are referred to as
Students 1, 2, 3, 4 and 5 in the discussion below.

The individual projects completed by the students were analysed to identify
insights gained through the concordancing process and also to determine to what
extent these insights correlated with their earlier comments about the benefits of
cordancing.

Findings and discussion

Findings are presented according to each of the three sub-questions followed by
results as they relate to the main question. Some background data regarding the
students, gained from the initial questionnaire, are presented first.

Background information gained from the questionnaire

All except four students had previously studied English language programmes in
New Zealand. The length of time they had been in New Zealand varied from one
year and eight months to 12 years. Twelve were aged between 25 and 34, three were
under 25 and three were over 35 (one age was not given).

The majority of students felt they were quite competent with regard to PC use,
five indicating very competent and only one indicating not very competent. Sixteen
of the 19 stated they had had no knowledge of corpus consultation prior to the
English language studies course. At the time of completing the questionnaire four
found the course very interesting, four quite interesting, eight interesting and three
not very interesting.

Findings relating to Question 1: What benefits do students perceive that concordancing
brings to language learning?

Students were twice asked if they would continue to use concordancing after the
course finished and what they thought the advantages were. They were first asked in
the initial questionnaire that was administered prior to the individual project and
again in the final section of the reflective log. Tables 1 and 2 compare the responses
the students made each time and show that the benefits identified by students became
much clearer and more specific when they were answering the second time.

Table 1. Intentions to use concordancing after the course.

<table>
<thead>
<tr>
<th>Intention to use software after the course had finished</th>
<th>Questionnaire (n = 19)</th>
<th>Reflective log (n = 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes definitely</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Possibly</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>No probably not</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
The proportion of students indicating that they would definitely use concordancing in the future fell, probably indicating a more realistic awareness of the actual benefits of concordancing, as well as the limitations. Four students remained positive that they would use the approach, however, and the proportion of those who said they would probably or possibly use concordancing actually increased. In the interviews, Students 1 and 4 reported that they were already using a web concordancer to assist with assignment writing. They had the concordancer open at the same time as Microsoft Word and said they could move ‘quickly and easily’ from one to another. ‘While I am doing my assignment I open the Brown corpus’ Student 1 commented.

In Table 2 the earlier questionnaire responses (left-hand column) focus on quite general comments regarding assistance with improving English through increased knowledge of rules and word use, whereas the later responses (right-hand column) indicate a raised awareness of particular ways that concordance evidence could supplement other knowledge available to the students. In their logs each student identified a number of benefits (whereas in the questionnaire it was one benefit per student) and these centred on the evidence of language usage that concordance lines provide. As Student 1 later comments, the concordancer ‘provides heaps and heaps of examples’ of language usage. Collocation and cluster data were clearly identified as being beneficial. Student 4 said that she could check collocations as she wrote by using a web concordancer and commented that it was ‘faster than using a dictionary . . . . I just write and check the website at the same time’.

Factors affecting students’ attitudes to concordancing

A prediction that seemed likely to be borne out was that those students following the three year EAL major programme would have a greater interest in, and motivation towards, language study and therefore may respond more positively to a course designed to assist them with this. However, those students who reported that English language studies was a very interesting course were not from the EAL major but from other programmes of study. The EAL major students were either interested or not very interested (see Table 3).

Of the nine students who said in the questionnaire that they would definitely use concordancing in the future, only two were from the EAL major whereas the four students from the Bachelor of Business course all indicated they would definitely use this approach (see Table 4).

As can be seen in Table 5, intentions had changed to some extent by the time students completed the reflective log, probably reflecting a more realistic view of the

<table>
<thead>
<tr>
<th>Questionnaire response (n = 19)</th>
<th>Reflective log response (n = 14)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Useful to improve English (6)</td>
<td>Evidence/examples of language usage (9)</td>
</tr>
<tr>
<td>Find grammatical rules (4)</td>
<td>Useful to analyse syntactic patterns (6)</td>
</tr>
<tr>
<td>Learn how to use words (4)</td>
<td>Information re collocations and clusters (4)</td>
</tr>
<tr>
<td>Will help if dictionary/grammar can’t (3)</td>
<td>Can identify differences between genres/styles (2)</td>
</tr>
<tr>
<td>Depends on need (1)</td>
<td></td>
</tr>
<tr>
<td>Can help with other study (1)</td>
<td>Faster than a dictionary (1)</td>
</tr>
</tbody>
</table>

Note: *some students indicated more than one benefit.
potential use of concordancing for these students, as well as the limitations. Four out of seven students indicating ‘yes, definitely’ in the questionnaire data changed to ‘probably’ or ‘possibly’ in the reflective log. However, a ‘possibly’ changed to ‘yes, definitely’.

Other factors
With regard to other factors that might influence attitudes towards concordancing, age does not seem to be important. Those indicating that they would definitely use concordancing in the future are represented in all age groups except for the over 45 group. Length of time in New Zealand also did not seem to be significant, as those

Table 3. How interesting the course is.

<table>
<thead>
<tr>
<th>Programme</th>
<th>Very interesting</th>
<th>Quite interesting</th>
<th>Interesting</th>
<th>Not very interesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAL major</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Jap major</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BBus</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GCertEAL</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>EAL BBus conjoint</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COP</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4. Intention to use concordancing by programme (questionnaire).

<table>
<thead>
<tr>
<th>Main programme</th>
<th>Yes definitely</th>
<th>Probably</th>
<th>Possibly</th>
<th>Probably not</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAL major</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Jap major</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>BBus</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>GCertEAL</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>EAL BBus conjoint</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>COP</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Totals</td>
<td>9</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 5. Intention to use concordancing by programme (reflective log).

<table>
<thead>
<tr>
<th>Main programme</th>
<th>Yes definitely</th>
<th>Probably</th>
<th>Possibly</th>
<th>Probably not</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAL major</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Jap major</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BBus</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>GCertEAL</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>EAL BBus conjoint</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COP</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>14</td>
</tr>
</tbody>
</table>
indicating an intention to use the concordancing approach ranged from students who had been in New Zealand for around a year to those who had been here three or more years. The nature of previous language study in New Zealand (private provider or tertiary institute, IELTS focused or more general) was also not found to be a predictor of attitude towards concordancing.

Findings relating to Question 2: How does concordance knowledge assist with vocabulary acquisition? i.e. what does it add to dictionary knowledge?

In the initial questionnaire, students indicated that they valued dictionary consultation highly. Thirteen out of 19 felt it was very useful and five thought it quite useful. The frequency with which they consulted dictionaries was not quite as strong but 13 students stated they consulted a dictionary very often or often.

When compared with students’ intentions to use concordancing in the future it was interesting to note that those who placed the strongest value on dictionary consultation were often the ones who said they were less likely to use concordancing in the future. This becomes a stronger pattern in the reflective log responses. This could indicate that students who are less traditional in their approaches to language learning may be more open to concordancing as a language learning tool.

In the reflective log and in the interviews that followed, students commented positively on the evidence of word usage that concordances gave them. Several students commented in particular on the collocation and cluster information that they were able to access. Student 4 said that for collocations ‘the dictionary doesn’t really help’ and Student 5 stated that concordance evidence can help to know ‘how to use particular words’. Two students commented on the authentic nature of the evidence presented by the concordancer.

Findings relating to Question 3: How does concordance knowledge assist with acquisition of grammatical features? i.e. what does it add to knowledge accessible from a grammar textbook?

In the initial questionnaire, students indicated that they did not value grammar textbook consultation as highly as dictionary consultation. Eight out of 19 felt that it was very useful and nine thought it quite useful. The frequency with which they consulted grammar texts was much lower than for dictionaries. The majority (13) consulted a grammar text sometimes.

Table 7 compares the students’ intentions to use concordancing with their attitudes towards grammar book consultation.
Grammar consultation was valued quite highly although, as noted with regard to dictionary usage in Question 2 above, those students who stated that they thought grammar textbook consultation was very useful were not the ones who said they were most likely to use concordancing in the future.

In the reflective log and in the interviews that followed, students commented positively on the information regarding language patterns and structures that concordances gave them. Six students commented on how concordances enabled them to notice patterns clearly and to ‘be critical’ with language. It is interesting to note that Student 3 indicated that she was pleased that concordancing had ‘pushed her to learn grammar’ in a new way. She said that ‘to be honest, I hate grammar . . . I have done a lot . . . and I am still confused’. This appears to fit with the notion that advanced level EAL students who may have studied grammar for many years could respond positively to concordancing as a fresh approach to developing their knowledge of syntactic patterns.

Findings relating to the main question: How can classroom concordancing complement traditional language learning resources such as a dictionary and grammar textbook?

Figure 1 summarises the views of the students who completed the reflective logs regarding the value of concordancing for language learning.

The data presented above in response to the three sub-questions have identified the reasons why these students agreed that corpus consultation was helpful for their learning. Students point out that concordancing presents them with examples of authentic language in such a way that they can easily analyse structural patterns and can find evidence of language usage, particularly relating to collocations and clusters. As Student 2 notes, concordancing ‘provides very clear evidence’ and Student 4 ‘found something new . . . about the use of words’.

In the semi-structured discussion, students distinguished the different functions of dictionary, grammar book and concordancer as follows: dictionary for meaning, grammar book for rules and concordancer for evidence and patterns. This indicates that they clearly saw the three tools as having complementary functions. So why do many of the students indicate that they will either possibly or probably not use a concordancer in future? Table 8 shows the disadvantages of using a concordancer that were identified by those students who completed the reflective logs and these were explored in the discussion.

Availability of the software is clearly an issue, although two of the students who were interviewed stated that web-based concordancers were easy to use alongside...
writing tasks that they were engaged in. The sheer volume of data that can be presented by the concordancer can also be daunting for many, as is the presentation of truncated sentences in concordance lines (although the full context can be easily displayed if required). In the interviews the issue of trusting the corpus or concordancer was mentioned a number of times. Student 1 wondered if the data were contradictory: the dictionary/grammar or the corpus/concordancer? Student 3 was unsure what to make of evidence (from a corpus of spoken English) that was clearly ungrammatical as far as she was concerned.

**Findings from the students’ assignments**

The genres and language features selected by the students ranged widely and often adventurously. Seven chose spoken genres (including phone conversations, broadcast interviews, teacher monologue and radio talkback) whilst the majority of the students selected written genres (including fiction, academic writing and financial news reports). Most students opted to investigate sections of the corpora to which they had been introduced on the course. One student, for example, focused on religious texts, Section D of the Wellington Corpus of Written New Zealand English, whilst another investigated the texts in the telephone conversation section of the Wellington Corpus of Spoken New Zealand English. Six students decided to create small corpora (approx 50,000 words) to investigate a genre of interest to them. These

<table>
<thead>
<tr>
<th>Disadvantages of using a concordancer</th>
<th>No. of students</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited availability – can’t use at home</td>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>Presents too much data/too many examples</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Concordance lines are incomplete sentences</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Don’t trust the corpus</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Context – lack of or too varied</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Limitations of the corpora themselves</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Still need previous knowledge e.g. re idioms</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Takes time</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Wastes paper</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>
self-created corpora included movie scripts, online financial news, horoscopes and employment advertisements. Reasons given for the choice of genre varied but were chiefly related to likely usefulness for future study or work and/or personal interest. On examining the assignments completed by the students, it is clear that all students developed a greater understanding of their selected language features. However, what did they appear to learn that derived from corpus consultation, as opposed to dictionary or grammar book? The following discussion focuses on three of the interviewed students.

Student 3 investigated features identified in a 50,000-word corpus of astrology texts (horoscopes) that she had collected from the world wide web. Her analysis included modal will, the lexical verb get and the adverb well, all features that the Keyword list had shown to be relatively more frequent in her corpus than in English generally. She found will to be more four times more frequent in her corpus than in the Wellington Corpus of Written New Zealand English. The instances were largely the predictive meaning of will, not surprisingly, but occurring in a range of sentence and clause types, of which she noted that the first conditional structure was but one of many. In examining the concordance of will the student also noted a number of frequent colligational patterns as well as regularly occurring phrasal verbs: ‘... others will tend to see you as ...’, ‘... others will open up to you when they need ...’, ‘You can bet a Goat will know when the fillet is on sale.’, ‘Aquarius will spend countless hours mulling ...’. Student 3 noticed the use of will as a noun: ‘Those who bear its will’ and will in the sense of wish in ‘... consider him top dog, if you will’.

This student also analysed the verb get. Whilst she had no problem finding concordance examples of common meanings of get such as receive, reach and move to, she clearly found it helpful to study the verb as it is used in phrasal verbs in the context of relationships (which she thought typical of her selected genre). She noted concordance lines which featured get along with, get in the way, get away with, get back at/to, get carried away with; for example: ‘Taurus is easy to get along with’, ‘Nobody is going to get away with bending the rules’. This student said in her interview that ‘phrasal verb meanings are sometimes hard to see directly ... we have to consider the meaning of the whole sentence’.

A third feature for this student was the adverb well. A discovery for her was the frequency in her corpus of adjectives featuring well: well advised, well armed, well documented and well groomed: items that would be listed separately in a dictionary. She also noted the frequent use of well in verb phrases such as ‘might well be’ and ‘could well be’, a usage she had not noticed previously and one which she noted was the last of the meanings given for well in the dictionary she had consulted. As noted earlier, Student 3 said in her interview that concordancing had helped to overcome her antipathy to grammar, and it is clear from her assignment that she had relished this opportunity to explore the language of a genre of English that interested her.

Student 4 investigated a corpus of academic writing (Section J of the Wellington Corpus of Written New Zealand English). In her interview she commented that she had ‘predicted a lot but still found something new about use of words’ in her project. Unlike many of the students, she had gone straight to the concordances without consulting a grammar or dictionary first as she felt she had sufficient background knowledge. Despite exhortations from the lecturer, similar to those given by Chambers (2005), to be wary of extremely frequent words such as prepositions, Student 4 chose to investigate of. She had justification in that it was top of the Keyword list that she produced. She was faced with 6725 occurrences but used the
left and right sorting functions of Concord to identify regular patterns. She noted the frequency of *amount + of*: *the greater part of the mountain, one of the central concerns, many of the species*. She noticed the significance in this type of text of the defining use of *of* in a sequence of nouns: *the stages of cultural shock, the assessment of net gains and losses, the sequence of results*. She also noted the regularity of the phrases *instead of* and *in spite of* but did not remark on other usages such as *way of* + *gerund*, a pattern which was very frequent in her concordance.

Student 4 also investigated *may* in academic discourse. With 239 occurrences, it was the most frequent modal verb in her Keyword list. The most common meaning she identified was that of ‘indicating possibility that something is true’ (e.g. *Drowned trees may provide habitats for some insects*) but she also noted that it has a meaning of ‘implying doubt’, a reference to the ‘tentative’ usage of language such as *may, appear, seem* etc. An example she gives from her concordance is ‘To some extent it may be due to the fact that . . .’. This student also pointed out that she had been previously confused by the pattern *may have + part participle* but that the numerous concordance examples of this pattern had helped to clarify it for her. Interestingly, Student 3 had noted the pattern *could well be/might well be* but Student 4 did not remark on *may well be* which was just as frequent in her corpus. She noted but did not comment on the use of *may* in statements that are then contrasted with something more important: *Teenagers may become more fluent in grammar and vocabulary, but they will almost always keep some trace of their foreign accent*. A problem here is that the *but* clause may not feature in the truncated concordance line, which would escape a novice in concordancing.

Student 5 analysed features she identified as unusually frequent in telephone conversations (a section of the Wellington Corpus of Spoken New Zealand English). In her interview she said she had found the concordancer very easy to use and that it was helpful for learning phrases and word usage ‘in order to speak English like a native person’. Two of the features she analysed were *bit* and *anyway*, both examples of language that O’Keefe, McCarthy and Carter (2007) describe as *relational* rather than *transactional* and profitably explored via a corpus. The concordance of *bit* clearly revealed for Student 5 the very frequent patterns of *bit of + noun* (*a bit of money*) and *bit + adjective* (*a bit different*). The student was also pleased to note the use of *little + bit* as in *a little bit of mine*, as well as the occurrence of phrases such as *it was a bit like, wait a bit, in just a bit, bit by bit*. Not surprisingly, her analysis lacked awareness of discourse categories such as ‘hedging’ or ‘vagueness’ but she had nevertheless identified the examples as being significant and could see how they were used in conversation. Her analysis of *anyway* was equally satisfying for her, in that she found native speaker usages of *anyway* as a call-closing strategy (*. . . various bit and pieces, yeah, anyway, thank you very much!*), as a way to say that something is relevant despite other things that have been said (*. . . anyway, the thing is, I’m pleased they are done . . .*) and as a way to change conversation topic (*. . . but, anyway, getting back to what we were talking about . . .*). Student 5 commented that the dictionary had not given enough examples of her chosen features for her to understand how to use them in context, but that the concordancer had done this.

**Conclusions**

Clearly a majority of the students in this study see that corpus consultation has benefits for them as language learners. In the reflective logs, 10 out of 14 students say
they will probably or definitely use a concordancer in their future language development. Cheng, Warren, and Xun-feng’s study (2003) confirmed that it was ‘both possible and worthwhile’ to include a corpus-focused course in an undergraduate English major programme, and the results of the current study would seem to similarly validate the inclusion of the English language studies course in the BA EAL and GCertEAL.

There is evidence too that more advanced level learners may be looking for ways of learning that are different to more traditional methods of grammar and dictionary consultation that they have used for quite some time. As noted also by Chambers in her study (2005), the discovery approach to language knowledge that is a feature of data-driven learning appears to have a strong appeal to some learners at this advanced level.

The analyses of language features that students carried out in their individual projects confirm that concordancing can assist learners to identify and explore aspects of language use that may not have been clarified via the consultation of dictionaries or grammar books, and that it can do this in the context of an approach that is genre-based, enabling a focus on texts that are relevant for learners’ future goals, be they study or employment related. In addition, those students who were enrolled in programmes that had a clear goal beyond English language study were often the ones who showed the most enthusiasm for corpus consultation, seeming to indicate that the need to engage with a particular genre or type of English for study or employment purposes was a motivator for this approach to language learning.

This study also seems to suggest that those students who may be deemed to be more ‘traditional’ in their language learning habits (i.e. valuing and frequently consulting grammar texts and dictionaries) are the ones less likely to be open to concordancing as an approach to language learning. If concordancing is to be introduced to students such as these there is therefore a need to emphasise early in a programme of study how the concordancer can complement dictionary/grammar book information rather than replace it.

Obstacles to concordancer use that are highlighted by the students include the sometimes overwhelming amount of corpus data that they encounter as well as limited access to the concordancing software. Tutors of courses such as the English language studies course featured in this study therefore need to reconsider how they introduce students to corpus consultation. The amount of data with which students are confronted needs to be ‘screened’ in some way, at least initially, and the introduction of corpus analysis theory, whilst required in an undergraduate course, should perhaps be delayed until the students have experienced what the concordancer can do. The suspicion or doubt that some students have about the accuracy or reliability of corpus data also needs to be addressed if they are to see the concordancer on a par with dictionaries and grammar books. Analysis of language from a corpus of spoken English will provide an opportunity to examine the grammar of speech as compared to the grammar of the written language – an aspect that is rarely covered with language learners.

Easily accessible web-based concordancers are increasingly available. Several students in this study are clearly making regular use of them, and to some extent have found them to be superior in comparison to the overly (for their purposes) sophisticated Wordsmith Tools program. The image conjured up by two of the students in this study of the EAL student writing in Word on her PC and switching
back and forth from her online concordancer as she checks collocations and syntactic patterns is an appealing and exciting one.

Limitations
All 19 students completed the initial questionnaire but 14 completed the reflective log. Towards the end of the course it seemed that the pressure of the project itself meant that students were not able to give sufficient time to reflection (although five students were very keen to participate in the discussion). This is a drawback of focusing a research project around a task that is a major summative assessment for students (30% of final course grade).

Further research
The present study indicates an awareness of, and an interest in, the benefits of concordancing amongst advanced learners of English in a tertiary context. There is some anecdotal evidence that several of the students surveyed are continuing to use a concordancer in subsequent courses. A longer term study would assist in identifying the factors that might influence the development of regular corpus consultation as a language learning tool.

Those students who chose to explore spoken genres discovered features of language that are not often presented in traditional courses or grammar texts. There would seem to be considerable benefits in using data from corpora of spoken English in language classrooms, either ‘raw’ for those more proficient students, or ‘partially digested’ for those at a lower level.

Despite the extra time that was involved, six students were sufficiently motivated to create their own corpora for their individual project. Further investigation is recommended into similar activities by students with a focus on the factors that prompt students to undertake corpus creation and on particular insights into language use that they gain as a result. All the students on the English language studies course were encouraged to create a corpus of their academic texts, effectively a learner corpus, and to use this as a comparison corpus. The usefulness for students of consulting corpus data about their own language use fell outside the current study but would be worthy of further investigation.

Notes on contributor
Steve Varley is a senior lecturer at Unitec Institute of Technology in Auckland, New Zealand. He teaches and coordinates the BA English as an additional language programme, which includes a second year course devoted to developing corpus consultation skills. He is interested in investigating how a corpus-focused approach in general and concordancing in particular can benefit advanced learners and users of English as an additional language.

References


Appendix 1

Terms

Advanced level EAL learners are speakers of English as an additional language who have an English proficiency level that has been assessed at IELTS 6 (or equivalent) or above and who are engaged in formal study of English.

A concordance is a collection of all the occurrences of a word in a selected corpus, presented so that the search word (or node word) is in the centre of the computer screen.

Concordancing (or corpus consultation) involves using lexical analysis software (e.g. Wordsmith Tools (Scott, 1999)) to create and analyse word lists, keyword lists and concordances of the occurrence of selected search words. Information about collocations and clusters is also available.

A corpus (pl corpora) is a structured collection of texts, usually stored in electronic form.

Data driven learning (DDL) is a term coined by Tim Johns (1997) to describe the learning that takes place when language students use contextual evidence such as corpus data to deduce meaning and to discover grammatical patterns

Appendix 2

Integrating corpus consultation into the language learning environment

QUESTIONNAIRE

Please complete the following questionnaire.

1. Personal information

Country of origin: _________________  □ Male  □ Female

My main programme of study is

□ BA EAL major  □ BA EAL minor  □ GCertEAL  □ BA/BBus conjoint
□ BBus  □ COP  □ other please specify: ______________

Age group: □ below 25  □ 25 to 34  □ 35 to 44  □ 45 or over

How long have you been in New Zealand _________ yrs _______ mths

Have you studied an English course in New Zealand before enrolling in this BA programme?  □ Yes  □ No
Please specify: __________________________________________
________________________________________________________________________
________________________________________________________________________

2. Language learning preferences

Please comment on how you prefer to use the following resources to help you to learn English.  Tick (✓) one box.

I think looking up the meanings of words in a dictionary is
□ very useful  □ quite useful  □ a little useful  □ not useful

I look up the meanings of words in a dictionary
□ very often  □ often  □ sometimes  □ rarely/never

I think consulting a grammar book is
□ very useful  □ quite useful  □ a little useful  □ not useful
I consult a grammar book
☐ very often ☐ often ☐ sometimes ☐ rarely/never

I think that using computer assisted language learning (CALL) software is
☐ very useful ☐ quite useful ☐ a little useful ☐ not very useful

I use CALL software outside the classroom (at home or in drop-in lab)
☐ very often ☐ often ☐ sometimes ☐ rarely/never

3. Using computers
How would you rate your ability to use a computer?
☐ very competent user ☐ quite competent user ☐ not very competent user

Do you have a PC at home that you can access easily? ☐ yes ☐ no

How often do you use the drop-in time in the computer lab?
☐ more than twice a week ☐ twice a week ☐ once a week ☐ rarely/never

4. Corpora and concordancing
How much did you know about corpora and concordancing before you enrolled in the English language studies course?
☐ a lot ☐ something ☐ a little ☐ nothing

Have you found the English language studies course so far
☐ very interesting ☐ quite interesting ☐ interesting ☐ not very interesting?

Have you downloaded the free version of Wordsmith Tools onto your home computer?
☐ yes ☐ no

If no, can you say why? __________________________________________

Have you accessed any of the web concordancers outside of the classroom?
☐ yes, often ☐ yes, sometimes ☐ no, never

If no, can you say why? __________________________________________

Do you think that you will use concordancing after this course has finished?
☐ yes, definitely ☐ probably ☐ possibly ☐ no, probably not

Why/why not? ____________________________________________________
________________________________________________________________

Thank you for completing this questionnaire.

Appendix 3

Integrating corpus consultation into the language learning environment

Reflective log

Log number: ______

As you carry out your individual project please make a note of the following.
This log is not part of your assessment. It will NOT affect your final grade in any way.

First stages
1. What genre have you chosen to examine? _______________
2. Why did you choose this genre?
☐ an interest/hobby  ☐ previous/future study  ☐ previous/future employment
☐ other reason: please specify ___________________________________

3. What language features do you predict you will find in the genre?
________________________________________________________________________________
________________________________________________________________________________

4. How did you make this prediction?
☐ guessed  ☐ read some extracts first  ☐ previous knowledge
☐ other method: please specify ______

5. What corpus have you chosen to use as your main corpus?
_________________________________
Why? __________________________________________________________________________
_______________________________________________________________________________

Collecting your data
Please complete this section as you work on your project. I would like to know the order in which you carry out the project tasks e.g. making wordlists, looking features up in a dictionary or grammar book, making a keyword list etc.

Note the sequence in which you collected your data to analyse: use a separate sheet of paper if you need to and attach it to this log.

1. ________________________________
2. ________________________________
3. ________________________________
4. ________________________________
5. ________________________________
6. ________________________________

How easy was it to decide how to collect the corpus data?
☐ very easy  ☐ quite easy  ☐ easy  ☐ not easy

Who did you get assistance at any stage from
☐ your teacher(s)?  ☐ classmates?  ☐ no-one?  ☐ someone else?  Please specify who:
______________________________

Analysing your data
Please comment on how you analysed the data you collected.

I analysed the concordances first and then checked a grammar/dictionary
☐ yes  ☐ no

I looked at a grammar/dictionary first then analysed the concordances
☐ yes  ☐ no

Evaluation of your use of concordancing in this project
Please answer the questions in this section when you have completed (or nearly completed) your individual report.

1. Corpus consultation is helpful for advanced language learners like myself
☐ I strongly agree  ☐ I agree  ☐ I agree to some extent  ☐ I disagree
2. I think the following are **advantages** of using corpus consultation for language learning:

   ___________________________________________________________

   ___________________________________________________________

3. I think the following are **disadvantages** of using corpus consultation for language learning:

   ___________________________________________________________

4. Do you think that you will use concordancing software after this course has finished?  
   □ yes, definitely  □ probably  □ possibly  □ no, probably not

   Please comment on why or why not:

   ___________________________________________________________

   ___________________________________________________________

**Appendix 4**

**Focus group**

I will arrange to meet 4 to 5 of the students who completed the reflective logs in order to discuss the information they provided on the log: both to clarify and to expand on it. The discussion will be focused around the following guiding questions. I will ask supplementary questions as appropriate. I will record and then transcribe key sections of the discussion.

Questions identified are:

1. How did you choose a genre to analyse? Was it an easy choice?

2. How did you predict the characteristic features? Were you right?

3. How did you approach the task? Was it easy to decide what to do first, next, next...? Who did you get assistance from?

4. How do you think that concordancing can help you to learn English?

5. How does it compare with a dictionary and a grammar book?

6. Will you use this approach in the future do you think? Why/why not?

7. Any other comments you would like to make?