Helping Language Learners Get Started with Concordancing

Stephen Jeaco*
Xi'an Jiaotong-Liverpool University, China

Abstract

While studies exploring the overall effectiveness of Data Driven Learning activities have been positive, learner participants often seem to report difficulties in deciding what to look up, and how to formulate appropriate queries for a search (Gabel, 2001; Sun, 2003; Yeh, Liou, & Li, 2007). *The Prime Machine* (Jeaco, 2015) was developed as a concordancing tool to be used specifically for looking up, comparing and exploring vocabulary and language patterns for English language teaching and self-tutoring. The design of this concordancer took a pedagogical perspective on the corpus techniques and methods to be used, focusing on English for Academic Purposes and including important software design principles from Computer Aided Language Learning. The software includes a range of search support and display features which try to make the comparison process for exploring specific words and collocations easier. This paper reports on student use of this concordancer, drawing on log data records from mouse clicks and software features as well as questionnaire responses from the participants. Twenty-three undergraduate students from a Sino-British university in China participated in the evaluation. Results from logs of search support features and general use of the software are compared with questionnaire responses from before and after the session. It is believed that *The Prime Machine* can be a very useful corpus tool which, while simple to operate, provides a wealth of information for language learning.

Key words: Concordancer, Data Driven Learning, Lexical Priming, Corpus linguistics.

Introduction

This paper presents the results of an evaluation of a concordancing program which the author developed as part of his doctoral studies (Jeaco, 2015). After presenting a brief introduction to why the software was developed, some of the theories and studies which had an influence on this work will be discussed. Then the basic design of the software will be introduced and the evaluation itself will be presented.

The Prime Machine for Language Learning

The desire to develop *The Prime Machine* as a new corpus tool grew out of professional experience as an English language teacher and manager of language teachers in China. At the time when I began work on this project, I had been interested in corpus linguistics for several years, but I had had limited success in passing on this enthusiasm to my students or colleagues. Part of the problem was being able to find ways to systematically present convincing examples from corpora which learners could understand and appreciate. Another aspect of the problem was finding ways to introduce the functions of corpus software tools without needing to explain complicated procedures or difficult to grasp background information about the corpus linguistic theories underpinning the results.

** Tel: + 86 51288161301; E-mail: Steve.Jeaco@xjtlu.edu.cn; HS431, Xi'an Jiaotong-Liverpool University, 111 Ren'Ai Lu, Suzhou Industrial Park, Suzhou, P. R. China*

Given the limited time available in class and a deep sense of the need to help my Chinese learners of English develop skills to explore language themselves, one of the main reasons for developing the concordancing tool was so that it could be an additional language resource to which my students could turn in order for them to check the meaning and use of words as they were composing, to consider alternative wordings as they were proof-reading and editing their own work or the work of a peer, and to explore in their own time some of the vocabulary which they had encountered briefly in a class session and the different contexts and environments in which it typically occurs.

Corpora and Language Teaching

Developments in corpus linguistics over the last few decades have had a great impact on the understanding of how language operates and how it is used, providing tools for lexicography, research and language study, and allowing users of these tools to draw on evidence which can be found in the patterning of language choices in texts. Learners and teachers (whether they know it or not) are using more materials based on patterns from corpora, and the language learning dictionaries and textbooks of major publishers include patterns such as collocations and draw on authentic examples, often with the corpora used or corpus-derived wordlists prominently displayed in the blurb. However, the impact of corpus technology on self-study and in the classroom has not been as great as the shift in the academic research or publishing fields (Timmis, 2003). Indeed, it would seem that of the vast numbers of language teachers working around the world, only a relatively small number attempt to motivate learners to use concordancers, often finding that learning to navigate the user interfaces requires a deep understanding of linguistic jargon and that learners only experience a limited amount of success in being able to process snippets from authentic sentences which have been decontextualised.

Nevertheless, hands-on use of corpora with language learners has been successful in a number of different teaching situations; for a review of the use of corpora with learners see Yoon (2008) and Kennedy and Miceli (2010). There are several reasons highlighted in the literature which explain why concordancing software can be especially useful for learners. Data Driven Learning (DDL) is the main way that corpus linguistics tools have been implemented in the classroom. DDL can assist learners and teachers in deciding what should be learned, and can provide new meaning-focussed approaches to problem areas such as prepositions (Johns, 2002). The common patterns of syntax associated with particular items of vocabulary are not typically available in dictionaries, but can be explored through corpora (Sinclair, 1991). The concordancer can create an "ideal" space where language learners can test their hypotheses about language use (Kettemann, 1995; cited in Meyer, 2002). If an approach is taken where the learner is seen as a "traveller" rather than a "researcher", Bernardini (2004) argues that concordancing tasks can be used as a means of meeting a variety of language teaching goals. Concordancing skills can also be seen as supporting life-long learning (Mills, 1994; Kennedy, 1998).

There have been many studies into the use of corpora specifically as a means for vocabulary building. Some of the earliest tasks in the Data Driven Learning classroom were centred around comparing words with a similar meaning or comparing different word forms. The idea of looking up two words and exploring the results has been a mainstay in articles introducing classroom concordancing (Johns, 1991; Tsui, 2004); as well as in the methodologies of various studies on the use of concordancers; and advice for teachers or teaching training (Coniam, 1997). Thurstun (1996) created materials for learners using lists of concordance lines, with a view to enabling them to recognise the common syntax of selected academic vocabulary and then use the terms for specific writing functions. Cobb (1999) used a concordancer as a means for students to develop their own personalised dictionaries, suggesting that new examples from a corpus could help students strengthen their knowledge of these words.

Although practitioners and participants in such studies have reported the comparisons as being rewarding, feedback from previous studies on the use of concordancers has also shown that learners can find formulating suitable queries quite challenging and knowing what to look up can be a hurdle (Gabel, 2001; Sun, 2003; Yeh, et al., 2007). From the students' perspective, exploration using carefully selected concordance lines may seem to take too long (Thurstun, 1996). In a recent study by Luo and Liao (2015), corpora were shown to be more

effective than online dictionaries as reference resources in error correction in writing, but participants also showed strong attitudes regarding difficulties related to the time needed, unknown words in the concordance lines, rule induction, having cut-off sentences and having too many examples. In addition to these issues, as Anthony (2004) argues as he presents his classroom concordancer (*AntConc*), software for concordance exploration is not usually designed specifically with learners in mind. It is true that *AntConc* goes some way towards simplifying the interface of a concordancer, but there are still many obstacles to getting started and knowing enough about the tools and functions in order to use them. It has been argued that effort should be put into trying to make concordancing software better in terms of its user-friendliness and its suitability for language learners (Horst, Cobb, & Nicolae, 2005; Krishnamurthy & Kosem, 2007).

The Prime Machine aims to make insights about language based on Hoey's theory of Lexical Priming (2005) accessible and rewarding. The software has been designed to provide a multitude of examples from corpus texts and additional information about typical contextual environments. Hoey argues that priming is "the result of a speaker encountering evidence and generalising from it" (2005, p. 185), and also considers some of the challenges that learners of a foreign language face due to limited opportunities to encounter language data naturally, and also due to the severe limitations of wordlists and isolated grammar rules. *The Prime Machine* was developed following key principles from Second Language Acquisition. First and foremost, the concordancer and concordancing activities are a means of leading language learners to read multiple examples from authentic texts. The SLA principle of exposing language learners to target language in use (Krashen, 1989; Nation, 1995-6) provides a basis for this. Another fundamental principle from SLA is that of focussed attention and noticing (Doughty, 1991). Schmidt claims that "intake is what learners consciously notice" (1990, p. 149). A link between concordancing activities and Laufer and Hulstijn's *involvement load hypothesis* (Laufer & Hulstijn, 2001) has also been made clear by Lee et al. (2015). Tomlinson argues the positive effects of noticing language features within authentic texts, and the learners' recognition of a gap in their own language use can be strengthened if the discovery process can be one in which the language learners uncover features for themselves (Bolitho et al., 2003; Tomlinson, 1994, 2008). It is hoped that *The Prime Machine* goes some way to providing a platform for these kinds of discovery as it has been designed specifically to facilitate noticing of patterns and tendencies (Jeaco, 2017).

It is possible to evaluate a piece of software like *The Prime Machine* by carrying out a series of system evaluations or by conducting a user evaluation. A user evaluation considers how well the system meets the expectations of its users, and how performance and accuracy affect the attitudes and actions of the users, and these can be measured through both feedback mechanisms such as questionnaires, interviews or focus groups, and through looking at the preferences expressed in records of users' interactions with the software. Following a user evaluation, priorities for further development become clear as software engineers can focus on ways to build on the more positively viewed aspects of the software, or they can look at which parts of the system were underappreciated or neglected and use system evaluation techniques to focus on these in isolation and attempt to improve them. As the software was designed for language learning and teaching, it is important to consider how principles from Computer Aided Language Learning (CALL) could be applied for the evaluation. Chapelle (2001) makes suggestions for the judgemental analysis of CALL software (p53-4), the appropriateness of task (p59) and the empirical evaluation of tasks (p68). She provides a list of six qualities as follows:

- Language learning potential
- Learner fit
- Meaning focus
- Authenticity
- Impact
- Practicality

Each of these qualities should be considered when evaluating the effectiveness of a concordancing tool for language learning. However, as Krishnamurthy and Kosem (2007) point out, it is also important for software designers to get feedback from teachers in a pilot scheme in order to ensure teachers will want to use it. Scott's own reflections on perceptions of the user-friendliness of *WordSmith Tools* include an important point that

teachers need to have confidence in their own abilities to use software, and what it should be used for, otherwise their fears for loss of face can be an inhibiting factor (Scott, 2008).

Overview of the Software

Like many other software applications, one of the main visual components of *The Prime Machine* is a set of tabs which can be used to switch between different functions and different pages of results. Figure 1 shows the tabs which appear at the top of the screen. A range of corpora are available, including the British National Corpus (BNC, 2007), with sub-corpora from the BNC based on the main groupings provided by Lee (2001), corpora constructed from the academic journals of Hindawi (2013), and other newspaper and specialist corpora.

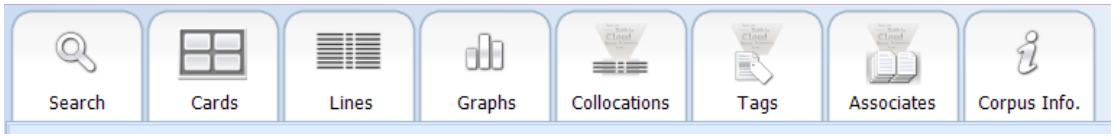

Figure 1. The Tabs Across the Top of the Screen in *The Prime Machine* Concordancer.

Search Tab

The usual starting point for language learners and teachers using the software is a specific word or collocation. The search tab provides two boxes where words or phrases can be entered. As users start to type, the corpus which is currently selected is accessed, bringing up lists of words and collocations for complete words. If the word or phrase entered into the system is not found in the current corpus, the user can seek additional spelling support, or click to check whether the word or phrase exists in any of the other corpora which are loaded into the system. The software was designed to make comparisons between two words, two word forms from the same family, words with similar meanings, and related collocations easy to make by providing search suggestions based on words entered, and by presenting results for two searches side-by-side on screen [1]. The search tab also allows for comparisons of the same item across two corpora, and some other tools more tailored to corpus linguistic research.

Cards and Lines Tabs

When language learners are first presented with concordance lines in the normal manner for corpus linguistics, namely Key Word in Context (KWIC), it can be quite hard for them to understand exactly what each horizontal line of disjointed text extract represents, and how they should go about trying to *understand* and *learn* from each example. Once they get used to the KWIC display, there are of course many advantages including the way in which lexical and grammatical patterning can be made more obvious (particularly through different sorting mechanisms) and the way in which many examples can be viewed together. However, in *The Prime Machine*, information about the source of each concordance line is made available to the user and the Lines Tab and the Cards Tab provide different layouts of the concordance data, with the aim of making different aspects of the contextual environment more noticeable.

One of the main differences in the presentation of concordance lines in *The Prime Machine* is the Cards Tab and the card for the currently selected KWIC line. For several years, the possibility of presenting more context to learners in a concordancer had been part of a vision I had had for helping learners become more confident and more familiar with corpus data. In the literature, there have been many reports of students finding the KWIC display difficult, at least at first. While some writers have played down the importance of this, and others have suggested it could be a benefit (Stevens, 1991), since my concordancer was being built from scratch, it seemed sensible to try to find an alternative way to display the information. As can be seen in Figure 2, the Lines display is similar to the KWIC display of other concordancers, but the card for the currently selected line shows

complete sentences above and below the sentence containing the node, with gentle highlighting of the line of text which contains the node. At the top of each card, the caption shows strong collocations within the nearby context of the node and the source type and citation is also prominently shown. The Cards Tab presents the list of concordance lines in the form of cards, but obviously compared with the Lines tab, fewer concordance lines are visible.

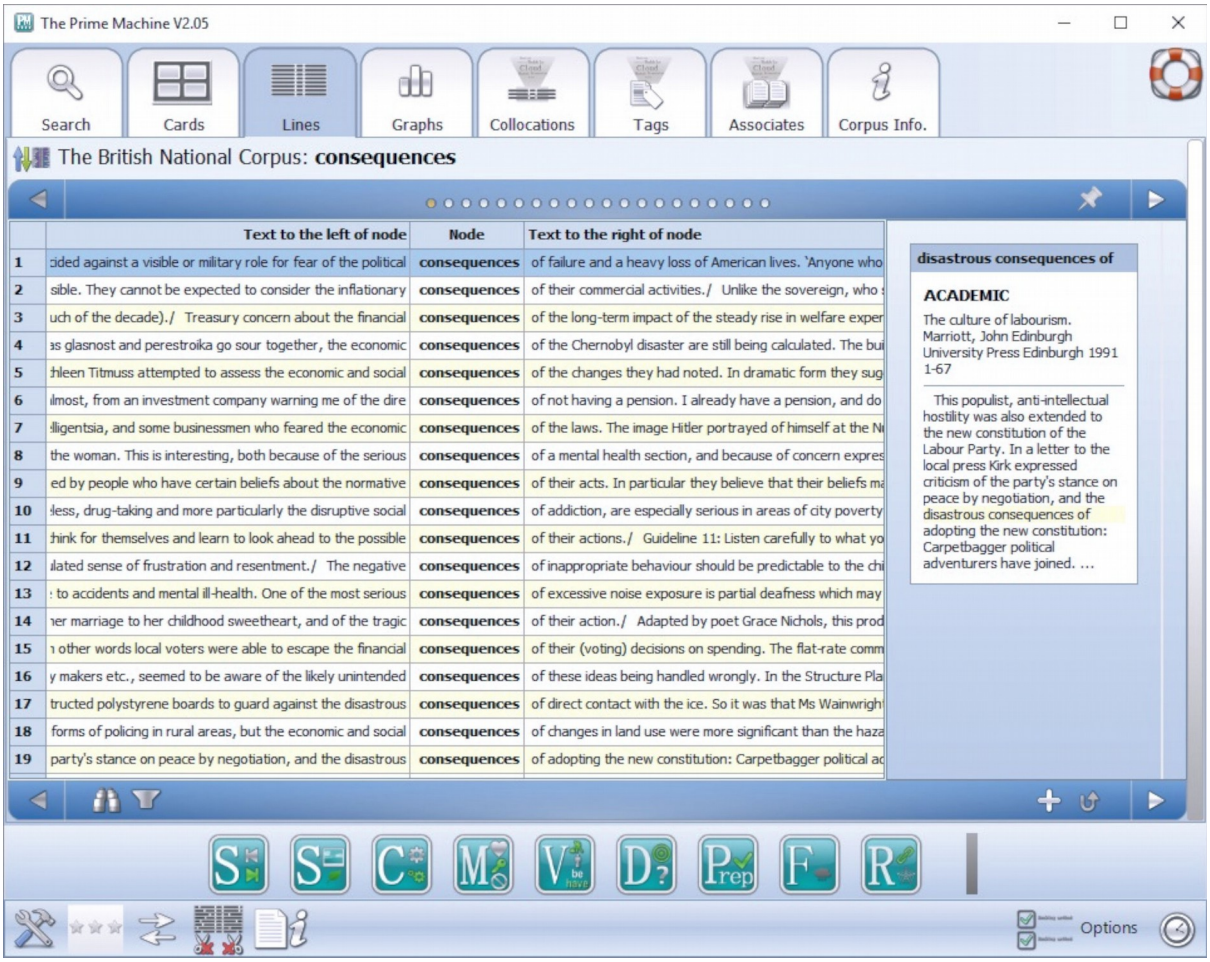

Figure 2. Example of the Lines Tab Showing the Card for the Currently Selected Concordance Line. (Incidental Data from a Query for the Word consequences in *The British National Corpus*)

The Collocations Tab

Unlike some of the other features of language which have been uncovered through the approaches of corpus linguistics, collocation is a term with which language teachers are certainly expected to be familiar, and from the widespread use of the term in section headings and dictionary panels it is clear that students are being encouraged to gain an understanding of it too. Other concordancers typically show collocations as lists of words rather than complete phrases, but learners may need to see the words together for these visual representations of the collocations to have an impact.

The default measure for collocations in *The Prime Machine* is based on specific ordering and proximity of the collocates, so it is possible to present each as a complete collocation rather than isolated words. In this way, the items in the clouds or tables on the Collocations Tab should provide a stronger impression and provide

learners with the opportunity to experience the phenomena introduced in one of Firth's ([1951]1957) memorable assertions: "A word in a usual collocation stares you in the face just as it is" (p. 182).

Other Tabs

Additional information about the typical environments in which the search query may be found in the corpus are shown on the other tabs. When the user looks up a specific vocabulary item, icons indicating strong tendencies draw attention to different aspects of its typical context. The Graphs Tab shows the proportion of concordance lines within specific contexts, and should draw learners' attention to a selection of features that will resonate with language teachers and will help learners engage with the data in the concordance lines more easily, including the use of articles and prepositions, passive voice and modal verbs. Pre-calculated summaries for words and collocations are also provided covering a range of features from the theory of Lexical Priming. Information on the other tabs also makes it possible for language learners and teachers using *The Prime Machine* to explore the patterns of words or collocations occurring in texts or sections labelled with a wide range of metadata, and as they occur with other words and collocations in different text categories. Finally, the Corpus Info. Tab provides information about the currently viewed corpus and its division into text categories.

Research Questions

This paper follows a user evaluation and reports on attitudes of language learners who used the software in a language learning activity. The following research questions are considered:

1. Can the students find examples which they consider helpful?
2. Which kinds of information do they look at most? How many results do they look at?
3. Which of the search support features are used most frequently?
4. How do they feel about the software? Would they want to use it in the future?

Methodology

Participants

Volunteers from an English-medium university in Eastern China were invited to participate in the project through short announcements before lectures and through the student email system. None of the students were currently studying modules taught by the researcher. Three sessions were scheduled for the same day, and these face-to-face sessions took place on a Saturday to avoid any conflict with class teaching. Students were able to indicate a preferred slot through the university's virtual learning environment system (VLE), (*Moodle* version 1.9), and an information sheet was also provided for them to review before the first session.

Materials

The materials for the evaluation included two questionnaires, a set of instructions demonstrating various aspects of the software, a brief user manual for the software and a set of essay question prompts. The first questionnaire included demographic questions as well as questions relating to the students' own views on their use of a range of language learning reference tools such as dictionaries, electronic dictionaries and search engines, etc. Therefore prior to using the new software, participants were presented with a broad range of relevant study resources available as choices in the early part of the first questionnaire, and for the questions relating to student habits and their attitudes regarding the best resource for several specific language learning issues, the option of concordance lines was not in any way foregrounded. The first questionnaire also included questions about peer review and more general attitudes towards language study.

The second questionnaire explicitly picked up on one of the questions from the first questionnaire and asked students whether their view of the importance of examples had changed as a result of taking part in the project. There were also questions about how much they used several of the main features of the software and how useful they perceived them to be. There were also a range of questions designed to gather their views on appropriate future uses of the software and any suggestions for improvements.

Both of these questionnaires were delivered electronically through the VLE. Examples of resources were provided on a printed A3 sheet, so that students would not need to flip between screens. This had examples of dictionary entries, popular search engines or mobile phone apps and a picture of concordance lines.

Printed instructions were given to the participants, providing step by step guidance on the overall procedure from answering the first questionnaire, downloading the software, working through the examples, writing the essay, and performing the follow up tasks later. In order to make the writing task relevant to students from a wide range of university programmes, prompts were written on a range of topics related to contentious but non-threatening issues which had been discussed in the news, following the style of popular language proficiency examinations.

Procedure

Participants volunteering to take part in the project were required to attend a face-to-face session in one of the university computer labs. At the beginning of each session, the information sheets and consent forms were distributed and then students were invited to complete the first questionnaire on the VLE. After completing the questionnaire, the students were free to start working through the instruction sheet, download the software and look through the user manual. When the questionnaires had been completed, the researcher worked through all the examples using a computer attached to a data projector. The participants were free to just watch or to try using the software themselves. At the end of the presentation, blank lined sheets were distributed to students who preferred writing essays by hand, while others loaded Microsoft Word and started to work on their essays on the computers. The students were then given one hour to write their essays. During this time, they were free to consult any other resources and to make use of the software. Formal examination conditions were not enforced.

Once students had submitted their essay to the researcher, they were free to leave. Within the next two days, individual feedback on each essay was sent to each participant. The template used by the researcher for this feedback included some comments based on each of the four criterion from the public band descriptors for IELTS (www.ielts.org). The feedback also included three screen shots showing sets of concordance lines related to three words or phrases used in the essay, as well as two *Microsoft Excel* spreadsheet attachments showing up to 100 more of the lines for these. A table of other single items or pairs of items to compare was also given. This feedback was then sent to each participant and he or she was invited to complete the second questionnaire online once he or she had reviewed the feedback, making use of the software again if he or she wished.

Four students participated in a pilot study several days before the main sessions took place, and a few minor changes were made to the procedure, the wording of some items, and some small aspects of the software's operation.

Logs

For research into the use of corpus tools with language learners, Pérez-Paredes, Sanchez-Tornel et al. (2011) argue that tracking of user actions through logs is essential in order to determine actual use rather than reported use. *The Prime Machine* was designed to include the capability of collecting logs of various actions triggered by mouse or keyboard movements during the evaluation.

Table shows a summary of the kinds of actions which are logged. During formal evaluations where participants have consented to the collection of this kind of data, logs are sent when the application is not busy retrieving data from the server or when the application closes.

Table 1
User Actions Which Can Be Automatically Logged by the Software

Action Category	Examples	Details
Search Support	• Auto-complete for single words; • Auto-complete for collocations; • Suggestions for words with similar meanings; • Spelling support request; • Request for a word or collocation to be checked in other corpora; • Alternative corpus selected after other corpora have been checked for a word or collocation not found in the current corpus. • Use of other navigation buttons ("Back", "Forward", "Home" or "Swap").	Words / collocation clicked
Query Blocked	• Rules for query syntax not followed; • Too few or too many words entered in a single query; • Word or collocation not found in the currently selected corpus; • Combination of words not found in the currently selected corpus.	Search string
Query	• Single search; • Compare mode search for two different queries; • Compare mode search for two different corpora; • Requests for more lines or collocation data.	Search string
Tab	• Cards Tab; • Lines Tab; • Collocation Tab; • Graphs Tab; • Tags Tab; • Associates Tab.	Number of seconds viewed
Other	A variety of other actions including the use of filters, access to help screens, changes to options, changes of the main corpus and use of various visual elements including the "Priming Dock". Also details such as the number of lines/cards viewed is stored.	

As can be seen, a range of categories have been created, allowing the grouping of log data in terms of search support features, actual queries, viewing of results and other features such as changes to options and access to help.

Findings

A total of 25 students attended one of the face-to-face sessions, completing the questionnaire and submitting an essay. All 25 participants were Chinese and came from Mainland China. The vast majority of the participants were female, with just 3 male participants. In terms of the academic programmes from which the students came, the most common was Financial Mathematics with 14 students, and this was followed by English and Finance (5 students), and 3 from engineering or computer science programmes, 2 from Chemistry and 1 from Economics. The ages of the participants ranged from 18 to 22, with 3 students from Year 1, 7 students from Year 2 and 15 students from Year 3. Given the programmes represented, the gender balance and the home provinces of the

participants broadly reflected the whole student population from which they were drawn. The participants reported that they had studied English for between 7 and 15 years, with 19 out of 25 students having studied English for 10 years or more.

Following the demographic questions, the first set of questions in the questionnaire was related to the students' reported use of reference tools to help them with their English. As can be seen from Figure 3, by far the most popular choice was mobile phone dictionary apps, with 21 students claiming to use these very often, and 3 students selecting 4 out of 5 for this item. Just one student reported a lower score (2/5) tending towards never. Interestingly, this student was the same student who indicated very often for concordance lines and one of the four students who indicated 5/5 for English-English dictionary with Chinese translations. Following mobile phone or electronic dictionaries, the next most popular choice was search engines. It is also clear that paper dictionaries are disfavoured, and electronic means through mobile phone apps or search engines are clearly favoured. As expected, the other clear finding was that for the majority of students concordance lines are not at all regularly used, with 72% of respondents claiming never to use them at all, and a further 20% choosing the second lowest rating. Three of the 5 students who chose 2/5 for concordance lines did not rate any of the resources below 2. The student who rated concordance lines 3/5, also selected neutral scores for half of the resources and did not select 1 or 5 for anything.

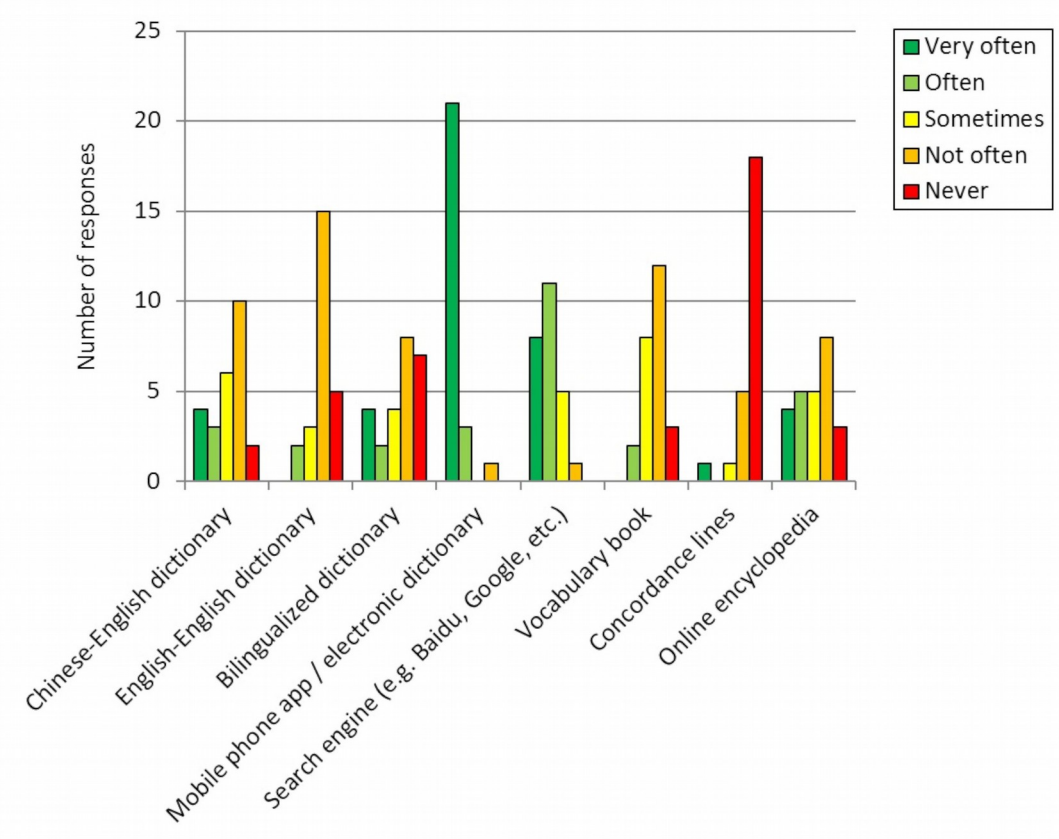

Figure 3. Reported Use of Different Resources.

The next set of questions was related to which resource listed on the handout students thought would be the most useful for five specific kinds of language problems. Figure 4 shows the number of students who selected each of these.

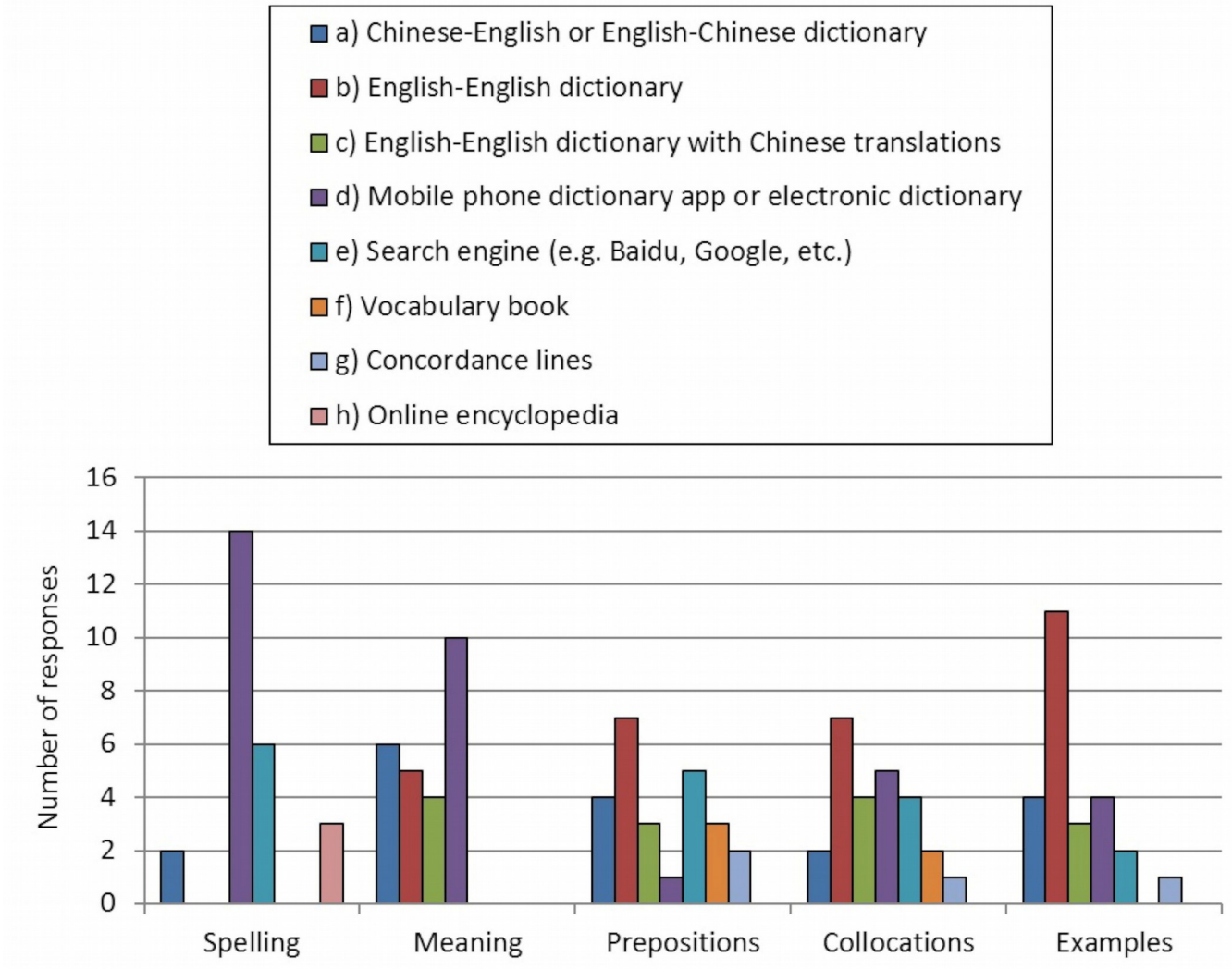

Figure 4. Judgements Given by Participants on the Best Resource for a Variety of Language Issues.

It is clear that mobile phone or electronic dictionaries were perceived to be the best choice for spelling and meaning, while English-English dictionaries were considered best to check prepositions, collocations or to find examples. Interestingly, search engines were not considered the best choice by any students when checking the meaning of words and were less popular than all three paper dictionary types and mobile phone dictionaries as a source for examples. The only three areas where search engines were considered the first choice by 16% or more of the students were for spelling (24%), prepositions (20%) and collocations (16%). This would suggest that search engines are used for language purposes by the students to check spelling and co-text rather than to provide information about meaning or examples.

Again, it is evident that concordance lines were not considered the best resource for any of these problems by the vast majority of students. There was also an interesting mismatch between the answers to the previous question about reported frequency of use and the resources which were considered most useful. Only three students chose concordance lines for any of the problems, and all three of these students had reported actual use of concordance lines as being 1 (never) or 2. The student who had rated concordance line usage so highly in the earlier question chose the option for "Chinese-English or English-Chinese dictionary" and the option for "Mobile phone or electronic dictionary" for all of the problems. This suggests that the student who had reported

using concordance lines very frequently was perhaps using them for other work or considered them to be a supplementary resource rather than a key one.

Another obvious conclusion which can be drawn from these data is that the vast majority of students (16 out of 25) consider translation dictionaries or mobile phone and electronic dictionaries to be suitable resources to check meanings. The wording of this question was "Checking a word which has several different meanings" and it is surprising that students place confidence in dictionaries which often only have a limited range of translations.

As explained earlier, after submitting the essay, students left the first session and were sent individual feedback within the next two days. They were then invited to complete the second questionnaire. Although 25 students took part in the face-to-face session, two students did not complete the second questionnaire.

Finding Useful Examples

In terms of reported use during different stages of the session, the results were fairly evenly spread. Figure 5 shows that the "Writing", "Checking/Editing" and "Reviewing feedback" stages were all rated as "Often" or "Very often" by at least 13 students. The "Planning" stage, however received fewer positive responses, with only 6 students selecting "Often" or "Very often" and this was the only stage where any students reported never making use of the software.

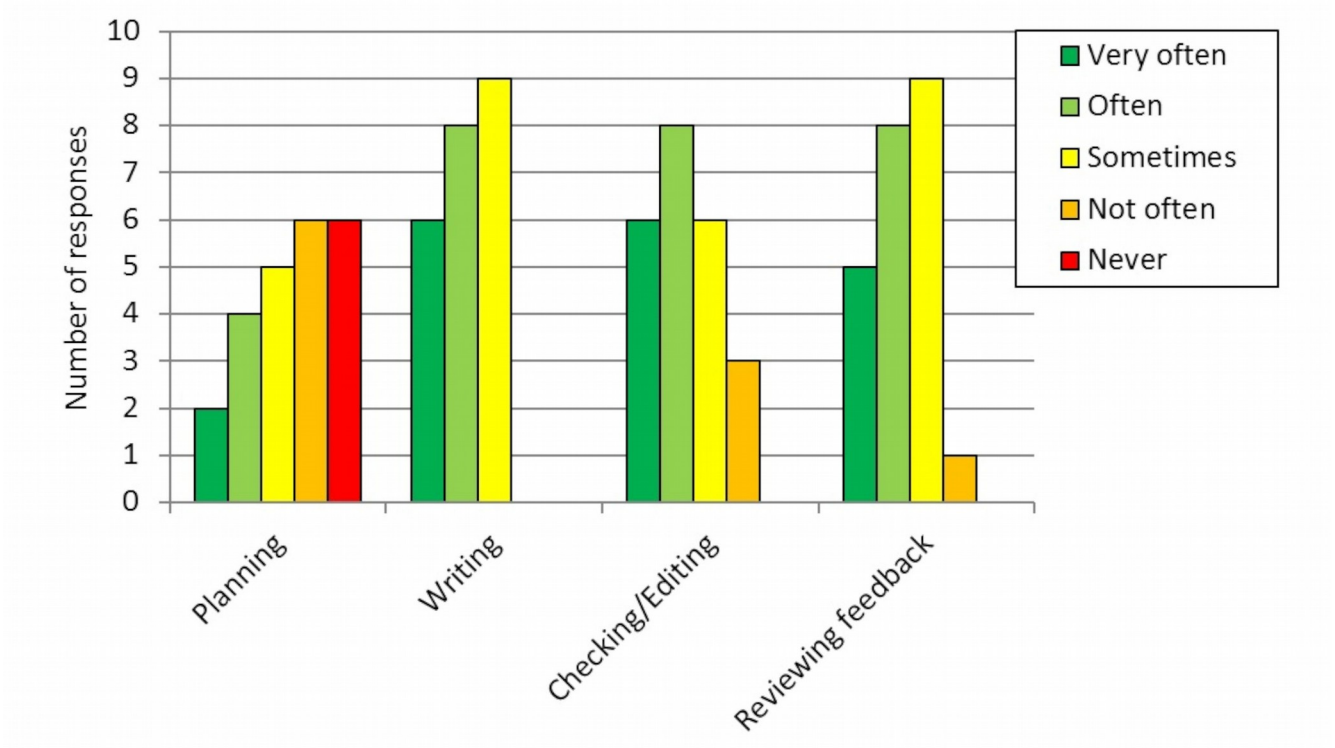

Figure 5. Reported Frequency of Use During Different Stages of the Writing Task.

Average ratings were 2.57 for planning, 3.87 for writing, 3.74 for checking or editing before submission and 3.74 for reviewing feedback from the teacher. The similar average scores for Writing, Checking/Editing and Reviewing (Wilcoxon Signed Ranks Tests: Checking/Editing-Writing z=-.408, p=.683, effect size r=-.060; Reviewing-Writing z=-.408, p=.683, effect size r=-.060; Reviewing-Checking/Editing z=-.0.37, p=.971, effect size r=-.005) mask individual differences, however, as different students reported use of the software at different levels. Only three students rated these three areas equally.

However, it is hard to find evidence of actual use of the software in the logs, which suggests that students were either exaggerating their use of the software or reporting attitudes rather than actual use. The strength of the results is somewhat weakened if the question is interpreted as being representative of attitudes, but the varied results do suggest that different students feel that the software would be useful for different stages of the writing process.

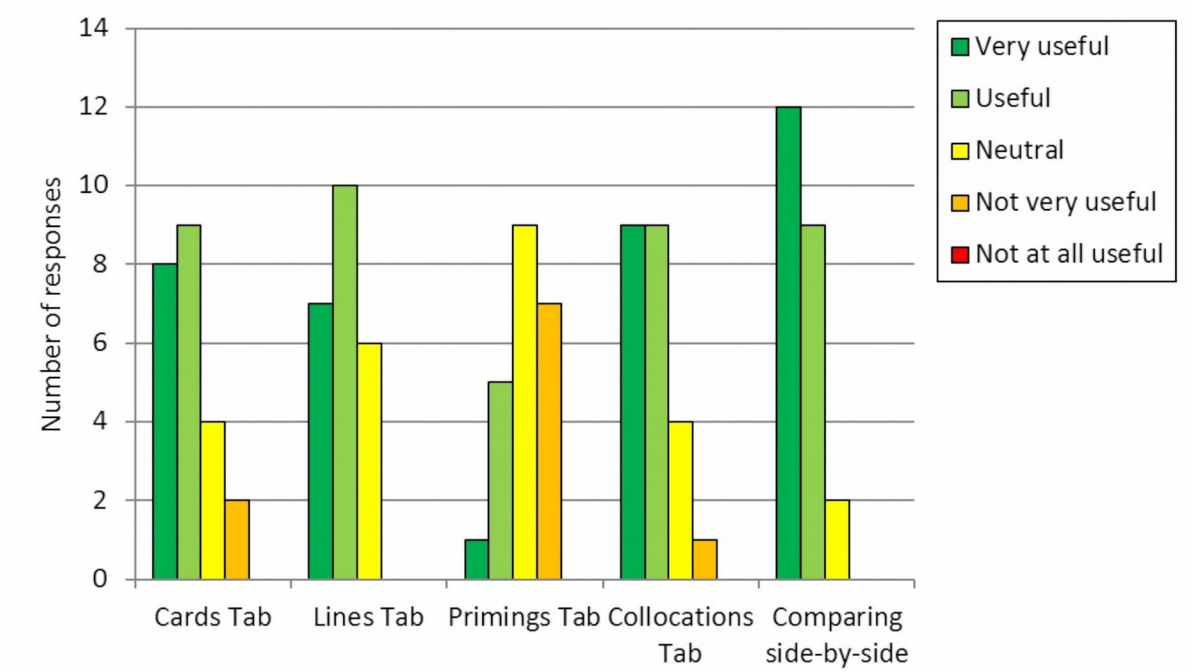

Figure 6. Evaluation of the Usefulness of Some of the Main Features of the Software

From the graph in Figure 6, it is clear that students rated both the cards and lines tabs quite positively, with approximately 74% of those who answered the second questionnaire choosing Useful or Very Useful. It is worth noting that although the Cards Tab seems more mixed with 2 students reporting it was not very useful, 6 of the 23 students (26%) rated the Cards Tab above the Lines Tab. Having both ways of viewing the data may cater for different learner preferences and different uses.

The Graphs Tab[2] received the least positive feedback, with a much lower average rating (Wilcoxon Signed Ranks Tests: Graphs-Cards z=-3.456, p=.001, effect size r=-.510; Graphs-Lines z=-3.337, p=.001, effect size r=-.492; Collocations-Graphs z=-3.072, p=.002, effect size r=-.453), however it is worth noting that 6 out of the 23 students (26%) rated it as very useful or useful. The student who rated the Graphs Tab as "Very useful" had lower ratings for all the other features except the Cards Tab.

The Collocations Tab was generally very positive. The student who rated the Collocations Tab at 2 also rated the Cards Tab and Graphs Tab as 2, but rated the Lines Tab as 4 (useful). Clearly, this student preferred looking at the information in the KWIC view, but from the logs it seems that he or she did not view the tables for collocations.

By far the most striking result from Figure 6, however, is that being able to compare results side-by-side was rated very highly indeed.

The results of the questionnaire questions related to the frequency of use during different stages of the task and the students' evaluation of the usefulness of some of the main features provide evidence that the first research question has been positively answered: the learners reported that they could find examples which they considered to be helpful.

Types of Information Viewed

Table 2
Logs Showing the Number of Views and Time Spent on Different Tabs in the Software

	Number views	Total time	Average number of seconds
Cards Tab	160	6485	40.5
Lines Tab	113	9328	82.5
Graphs Tab	53	2479	46.8
Collocations Tab	70	4325	61.8
Tags Tab	35	813	23.2
Associates Tab	48	6615	137.8

Table shows that the logs seem to support the views regarding the usefulness of different tabs, with Cards and Lines having much higher event counts and generally more time being spent on Cards, Lines and Collocations. When looking at these figures, however, it is worth noting that the Cards Tab was set as the default results tab for all users, so this will have received a log for every search which was completed. However, looking at the number of cards viewed for each event, the logs show that an average of 15.1 cards were viewed with a range between 1 and 65. Only 17 out of the 160 events had fewer than 10 cards marked as having been viewed. Since only a few cards are visible unless the user scrolls down, this seems to confirm that some users viewed quite a few results on the Cards Tab.

It is worth bearing in mind, however, that the vast majority of the events were from the sessions on Saturday, and the time in Table should be treated with caution since it is likely that students may have left a tab visible when stopping to listen to another part of the demonstration. The times are calculated for the whole time that the application is "active" (in the sense of being the window with the current focus), so this kind of data is more reliable when students are completing a task in another window rather than switching attention to a data projector during a demonstration or working on a paper-based activity.

From the logs, only 4 students seem to have made use of the software after Saturday, and figures for use across different tabs for later use are shown in Table 3.

Table 3
Number of Views, Time Spent and the Number of Different Users for the Results Tabs After the Main Input Session.

	Number views	Total time	Users
Cards Tab	10	186	4
Lines Tab	9	1679	4
Graphs Tab	4	91	3
Collocations Tab	7	74	4
Tags Tab	4	22	2
Associates Tab	3	26	2

Again, it is clear that most time was spent on the Lines Tab. Although, figures for the Graphs Tab may seem a little disappointing, it is worth noting that there were a total of 188 clicks on the priming icons on the dock and 18 users made use of this feature to switch to the Graphs Tab.

In terms of use of the ability to compare results side by side, the logs show that a fair proportion of searches were made like this. Of the 281 logs from 22 users, 56% of searches were for one term only, while 44% were made in compare mode. Three users did not appear to make any queries. Using the logs for the right-

hand retrieval only, 85% of the compare mode searches were to compare different queries across the same corpus, while 15% were comparing the same query across two different corpora.

The summary of the log data which has been provided here addresses the second research question, which was concerned the kinds of information viewed and the number of results. It is clear that overall the students spent most time on the Lines Tab, followed by the Cards and Collocations tabs. The logs also showed some engagement of the students with the different kinds of information and the number of results, measured by the number and range of events logged and the number of concordance cards viewed.

Support Features

As well as being able to compare results easily, another set of important design features were related to search support. The third research question was to ascertain which of these search support features would be used most frequently. A total of 54 queries from 16 users were logged as having been blocked by the software. Six of these were related to spelling errors, 1 was because a Chinese word had been entered. Nine blocked queries contained collocations where the incorrect format had been given (lack of spaces or additional full stops, etc.), and 20 blocked queries were because the phrase was not stored as a collocation in the system. Four queries were blocked because it seems nothing had been entered in the search box. A further 14 queries were blocked but information is not provided in the logs.

As well as preventing users from making queries and waiting only to discover that no results are found, the software also included other features such as auto-complete, collocation suggestions, synonyms, other word forms and spelling support. From the logs, auto-complete for words was used 12 times, and 9 of these were for words or word forms which did not form part of the demonstration. Collocations were selected from the drop-down box 9 times, 8 of which were for collocations not part of the demonstration. Spelling support was requested 5 times, but from the logs it does not seem to be the case that the student made a subsequent search using the correct spelling. This suggests that either the spelling component was too slow or did not provide useful suggestions, or perhaps that students were trying it out rather than actually wanting to use it to assist with their spelling.

A quarter of the all the search queries in the concordancer were made for words or word forms not part of the worksheet, and these were made by 13 different users. In the second questionnaire, students were asked to report on whether or not they had looked up words or phrases not connected with their task. Eleven students reported that they had, and 7 said that the search was useful and 4 interesting/fun, including one student who chose both useful and interesting/fun. Just 1 student said that this was a waste of time, but it is worth noting that overall this student was highly positive in his/her responses to the questions about the usefulness of each tab, having rated everything 5/5 except the Graphs Tab which was still rated positively at 4/5. These results might suggest that overall the software is likely to have potential for the kind of serendipitous learning which has been reported in DDL and "discovery learning" activities (e.g. Bernardini, 2004).

These results provide an answer to the third research question, demonstrating that the most frequently used search support features seem to be those which can be found on the main search screen such as the spelling support and the auto-complete features for words and collocations.

Interest and Future Expectations

Another set of questions on the second questionnaire related to whether or not students thought corpus examples, collocation information and the software itself would be useful for students like themselves. These questions were framed to be Yes/No questions with a required comment box to explain their reasons. Only 2 out of the 23 students who completed the second questionnaire responded negatively to the question of whether corpus examples were important. From their comments, it seems that both of these students were unsure of the relevance of the corpus examples to their own language production, with one stating "we do not use those examples very often", and the other stating that he/she did not think it was useful for academic writing. However, of the vast majority of students who responded positively, 6 mentioned examples, 8 mentioned usage, 4 mentioned collocations, and 2 mentioned reliability. Encouragingly, one student wrote, "the examples helped me

to think differently and get some information", and another mentioned that corpus examples were useful because students have little opportunity to see how native speakers express themselves.

The second question in this group related to the importance of understanding collocations. All 23 students responded positively to this question. In the comments, 9 students mentioned the need for this kind of information to avoid making errors or to improve accuracy, and 8 students mentioned the importance of knowing how to use words.

The last question in this group asked students whether the software tool was useful. Out of the twenty-three students, all but one responded positively. The student who selected "no" was one of the two students who used the software most after the Saturday session. However, the actual comment made by this student is still positive about the software's usefulness; as is clear from the full response, his/her reservation is due to his/her belief that other software packages may be able to provide similar information in a more convenient way:

"It has many many tools and looks useful, but some important usage can be replace[sic] by other APP."

Overall, it seems that the software was received very positively, especially considering that from the results of the first questionnaire it is very clear that very few students had used concordancers before. All but one of the students responded positively to the question about the usefulness of the software, and even the student who responded negatively did so in a highly positive manner. As explained earlier, two students chose not to complete the second questionnaire and their reasons for dropping out are not known. Neither student withdrew formally from the project and it is likely that other pressures such as coursework deadlines and mounting pressure for the final examinations may have influenced their choice not to complete the second questionnaire. Nevertheless, even if the non-participation of these students is interpreted as being lukewarm or negative towards the usefulness of the software overall, the proportion of positive responses as a total of all 25 participants is still 88%. Students who completed the second questionnaire gave a variety of reasons why they thought it was useful, with 4 mentioning being able to compare or see differences between words. Two mentioned the resources specifically. One student simply stated "It help [sic] students like a teacher". Another student demonstrated a good understanding of how different resources will be suitable for different occasions:

> *"This software may not be my first choice when I look up a word, because [an] electronic dictionary is much more convenient. However, [the] function of the software is complete and I would like to use it as the complement of my first choice."*

One other student mentioned that it was not so "convenient" to use; however, 4 other students commented favourably on the "convenience" of the software. Another student focused specifically on the way in which the software can help students discover semantic associations of words writing:

"I think, it can tell us whether a word is positive or negative. This is interesting and useful!"

Other comments included positive evaluations of the software in terms of helping students to learn effectively (1), the amount of detail (3), and its potential in helping with academic writing (3). One student also said that it was useful for students from different "levels".

The positive response is also evident in all of the responses to the question "In future, do you think you would like to use software like this again?" 10 out of 23 students chose "Yes, definitely", and the remaining 13 chose "probably". None of the students chose "Not sure", "Probably not", or "Definitely not". When asked to select from three situations when the software should be used, 7 chose "In class with a teacher", 16 chose "In class for pairwork activities", and 14 chose "Outside class independently". Given that almost 70% of the students thought the software was suitable for pairwork, and 2 of these students had reported that they did not think peer activities were useful in the first questionnaire, it seems that the software may have potential to as a teaching tool to enhance pairwork tasks.

The positive responses to the questions about corpus examples, collocation information and the software itself, coupled with these highly positive responses to questions about possible future uses of the software go some way to addressing the fourth research question. However, one factor which needs to be considered in relation to these largely positive responses is that in China there is a cultural desire to please. It is hoped that the influence of this on the questionnaire responses was reduced through the precaution of not revealing who had created the

software until the debrief message was sent. Nevertheless, the results should be considered in the light of these cultural influences.

Discussion

Taking these results as a starting point for the evaluation of *The Prime Machine*, this section will return to the 6 qualities of CALL software which were presented from Chapelle (2001).

The first quality, "Language Learning Potential" when applied to this project might include a judgemental analysis of the level of interactivity and the suitability of the range of target forms the software can provide. It would seem fair to award the software highly in this area since its very design encourages students to look up vocabulary themselves and to interact with the different tabs of data which are presented, and it also supports a wide range of comparisons between words and collocations or between corpora. It is also clear that the software has great potential for providing students access to a very wide range of target forms, both in terms of the level of analysis from individual word types, to similar words and collocations, and in terms of the range of text types from different disciplines and genres which are contained in the corpora which have so far been used. The question of whether target forms are acquired and retained, as has been mentioned above, is still one which needs to be explored, but the responses to the second questionnaire as presented here suggest that students were able to identify the importance of the software in supporting language use and accuracy and as a means of obtaining information about language.

In terms of the second quality, "Learner fit", the software would also seem to stand up very well. As a tool for exploring words and phrases the software provides a great amount of control. The questionnaire responses indicating how students viewed exploration of words or phrases not directly related to their essay writing also provides evidence that the software has potential for incidental or less directed learning. To facilitate autonomy and unsupervised exploration, one of the main aims for the design of the software was to provide more adequate support, hints and guidance to learners, as compared with other leading concordancers. Within the context of higher education, the software seems to have been very well received by students of different levels. The evidence from the questionnaire on how students reported using the software, the variation in their preference for different tabs of information and also the different views on how it could be used in future suggest that it might cater well for different learners with different learner styles. Since students were overwhelmingly positive, but positive about different aspects, it could be claimed that there is some empirical evidence that the software has succeeded in this respect. Based on the positive responses from students, it would seem that the innovations in the design of *The Prime Machine* alleviate some of the difficulties reported in previous DDL studies using other software. The difficulties or frustration in formulating and performing search queries which was observed in previous studies (Gabel, 2001; Sun, 2003; Yeh, et al., 2007) may have been alleviated by the search support features. The availability of the Card view and being able to compare results side-by-side, could also explain why there appeared to be fewer of the kinds of difficulties related to time or the presentation of results reported in other studies (Luo & Liao, 2015; Thurstun, 1996). However, clearly longer-term attitudes and measurements of change in performance over time would need to be considered. Nevertheless, designers of other concordancing interfaces could consider adding features like these if they wish to make their software more learner-friendly.

A focus on meaning also seems to be evident both from a judgement of the software and task, as well as empirical evidence in the form of questionnaire responses. The high rating of the compare feature suggests that students were interested in understanding how different words were used. The reported use of the software as part of a writing task also provides some evidence that students could see how the software could be used to help communicate their meaning effectively in writing, although as was mentioned earlier the logs suggest that these attitudes were probably based on their ideas about how the software could be used, rather than based on their actual experience using the software. Clearly, a longer study with log data matching reported views would be desirable.

In terms of "authenticity", the task design was highly relevant given the number of students who go on to take language tests such as IELTS as well as tests for their EAP modules, but it lacked the authenticity of being

actually part of the degree programme itself. However, the learners clearly demonstrated a belief that the software would be useful in the classroom or for self-study, and the overwhelmingly positive indication that they would definitely or at least probably want to make use of the software again in the future is good evidence that the software has to some extent met its aims as being a tool suitable for classroom or home use.

The "impact" of the software could be measured in terms of the comprehensiveness of feedback and software logs. While the log data was a little disappointing in terms of quantity, the evaluation has demonstrated that the level of detail which can be provided about different actions made by users of this system does have great potential. It is certainly clear that students rated the experience of using the software as a positive experience and in this respect the evaluation so far has been highly successful. The limited evidence of actual use of the software, especially after the main face-to-face part of the evaluation, points to a need for further research in order to ensure that the positive impact in terms of the perceptions of the students would also follow through to a positive impact on longer-term use. One of the main limitations of the evaluation in terms of its face validity was that although the participants were completing a writing task suited to their learning context, the essay was not part of their formal studies and was administered towards the end of the semester when other pressures such as assessed coursework and upcoming exams may have meant they were less inclined to put the usual amount of care and attention into it. In order to encourage greater use of the software so that attitudes would be based on more direct and prolonged exposure to the interface and results, participants could be given opportunities to access it over a longer period. The software needs to be made available so students can access it as and when they encounter language learning needs. Even in a shorter term study, if permission could be gained for students to bring with them early drafts of assignments or materials from their classes, participants would be much more likely to look up more words and phrases than when writing for an additional essay which may not have any long or short term benefits beyond general improvement of their language abilities.

Regarding the use of concordancers by language learners, the results of the first questionnaire were consistent with Timmis (2003) in that the participants had had very little prior exposure to direct use of concordancing software. Given the learning background of learners in China, it would be unrealistic to expect a sudden shift in their understanding of effective language learning processes, but the highly positive response to the software suggests that providing students with a new way of looking at language can be very effective, especially when supported by the kind of evidence which *The Prime Machine* can readily provide. Of course, a very important consideration with any kind of teaching software is whether or not teachers will be interested and willing to make use of it and to recommend it to their students. The design of the software was made by drawing on my own extensive experience as a language teacher and as a manager of language teachers. However, the importance of getting teacher input on software design (Krishnamurthy & Kosem, 2007) and responding to teachers' fears (Scott, 2008) should not be overlooked. Clearly, further exploration of the perceptions of teachers and input from them will be a key to making *The Prime Machine* a well-used tool as well as a useful tool for language learning.

The last quality is that of "practicality". The fact that the evaluation ran smoothly with a single server which was actually a desktop machine purchased in 2011 and was located outside the university local area network suggests that the minimum requirements are reasonable. *The Prime Machine* has now been running on a central server at the university for more than a year, and in the near future, this server will be accessible from outside its host institution. For further details see www.theprimemachine.com.

Conclusion

This paper has focussed on one aspect of the evaluation of *The Prime Machine*. It has considered the results of the small scale evaluation which took place over a short period of a few days, and it has also considered the scope of this evaluation within a wider framework. Despite being somewhat limited in size and duration, the questionnaire-based study has provided interesting insights into the acceptability of the software, face validity and student attitudes before and after and has also provided some concrete areas for future development. While the remaining ground drawing on frameworks from Computer Aided Language Learning for detailed evaluation

of the software as a learning and teaching tool is wide, this initial evaluation has served to demonstrate confidence that the project meets its overall aims. While there is also much scope for detailed evaluation of specific features and mark-up processes, as well as opportunities for performance enhancement of the computer processes behind the software, the participants' enthusiasm suggests that the software is providing some meaningful data and provides at least face validity for the hidden processes.

Through this small evaluation involving undergraduate students, the software has been shown to have considerable potential as a tool for the writing process. Since this evaluation was carried out, *The Prime Machine* has been developed further now includes additional tools for exploring vocabulary in terms of semantic tags and other features. As it continues to be developed, it is believed that *The Prime Machine* will be a very useful corpus tool which, while simple to operate, provides a wealth of information for English language teaching and self-tutoring.

Notes

[1] For a fuller explanation of the way these features work, for more details about the other features of the software and for the pedagogical reasons behind the design see Jeaco (2015) and Jeaco (2017).

[2] At the time of the evaluation, the label on this tab was "Primings Tab", and the questionnaire asked respondents to comment on it using this name. However, the label was subsequently changed to "Graphs Tab" as this better matches the purpose and scope of the tab.

References

Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. Paper presented at the Interactive Workshop on Language e-Learning, Waseda University, Tokyo.

Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. M. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 15-36). Amsterdam: John Benjamins.

BNC. (2007). The British National Corpus (Version 3 BNC XML ed.): Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/.

Bolitho, R., Carter, R., Hughes, R., Ivanič, R., Masuhara, H., & Tomlinson, B. (2003). Ten questions about Language Awareness. *ELT Journal, 57*(3), 251-259.

Chapelle, C. (2001). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research*. Cambridge: Cambridge University Press.

Cobb, T. (1999). Giving learners something to do with concordance output. Paper presented at the ITMELT '99 Conference, Hong Kong.

Coniam, D. (1997). A practical introduction to corpora in a teacher training language awareness programme. *Language Awareness, 6*(4), 199-207.

Doughty, C. (1991). Second Language Instruction Does Make a Difference. *Studies in Second Language Acquisition*, 13(4), 431.

Firth, J. R. ([1951]1957). A synopsis of lingistic theory, 1930-1955. In F. R. Palmer (Ed.), *Selected Papers of J R Firth 1952-59* (pp. 168-205). London: Longman.

Gabel, S. (2001). Over-indulgence and under-representation in interlanguage: Reflections on the utilization of concordancers in self-directed foreign language learning. *Computer Assisted Language Learning, 14*(3-4), 269-288.

Hindawi. (2013). Hindawi's open access full-text corpus for text mining research. Retrieved 6 November, 2013, from http://www.hindawi.com/corpus/

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive on-line database. *Language Learning & Technology, 9*(2), 90-110.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Jeaco, S. (2015). The Prime Machine: a user-friendly corpus tool for English language teaching and self-tutoring based on the Lexical Priming theory of language. Unpublished Ph.D. dissertation, University of Liverpool. Retrieved from https://livrepository.liverpool.ac.uk/2014579/

Jeaco, S. (in press). Concordancing lexical primings. In M. Pace-Sigge & K. J. Patterson (Eds.), *Lexical Priming: Applications and advances* (pp. 273-296). Amsterdam: John Benjamins.

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In T. Johns & P. King (Eds.), *Classroom Concordancing* (Vol. 4, pp. 1-13). Birmingham: Centre for English Language Studies, University of Birmingham.

Johns, T. (2002). Data-driven Learning: The perpetual change. In B. Kettemann, G. Marko & T. McEnery (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 107-117). Amsterdam: Rodopi.

Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology, 14*(1), 28-44.

Kennedy, G. D. (1998). *An Introduction to Corpus Linguistics*. London: Longman.

Kettemann, B. (1995). On the use of concordancing in ELT. *TELL&CALL, 4*, 4-15.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional evidence for the Input Hypothesis. *The Modern Language Journal, 73*(iv), 440-464.

Krishnamurthy, R., & Kosem, I. (2007). Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes, 6*(4), 356-373.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26

Lee, D. Y. W. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology*, 5(3), 37-72.

Lee, J. H., Lee, H., & Sert, C. (2015). A corpus approach for autonomous teachers and learners: Implementing an on-line concordancer on teachers' laptops. *Language Learning & Technology*, 19(2), 1-15.

Luo, Q., & Liao, Y. (2015). Using Corpora for Error Correction in EFL Learners' Writing. *Journal of Language Teaching & Research, 6*(6), 1333-1342.

Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction.* Cambridge: Cambridge University Press.

Mills, J. (1994). Learner autonomy through the use of a concordancer. Paper presented at the Meeting of EUROCALL, Karlsruhe, Germany.

Nation, I. S. P. (1995-6). Best practice in vocabulary teaching and learning. *EA Journal*, 3(2), 7-15.

Pérez-Paredes, P., Sanchez-Tornel, M., Alcaraz Calero, J. M., & Jimenez, P. A. (2011). Tracking learners' actual uses of corpora: guided vs non-guided corpus consultation. *Computer Assisted Language Learning, 24*(3), 233-253.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Scott, M. (2008). Developing WordSmith. *International Journal of English Studies, 8*(1), 95-106.

Sinclair, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stevens, V. (1991). Classroom concordancing: Vocabulary materials derived from relevant, authentic text. *English for Specific Purposes*, 10(1), 35-46.

Sun, Y.-C. (2003). Learning process, strategies and web-based concordancers: a case study. *British Journal of Educational Technology, 34*, 601-613.

Thurstun, J. (1996). Teaching the vocabulary of academic English via concordances. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, Chicago.

Timmis, I. (2003). Corpora and Materials: Towards a Working Relationship. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 461-474). London: Continuum.

Tomlinson, B. (1994). Pragmatic awareness activities. *Language Awareness, 3*(3-4), 119-129.

Tomlinson, B. (2008). Language acquisition and language learning materials. In B. Tomlinson (Ed.), *English Language Learning Materials: A Critical Review* (pp. 3-13). London: Bloomsbury Publishing.

Tsui, A. B. M. (2004). What teachers have always wanted to know - and how corpora can help. In J. M. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 39-61). Amsterdam: John Benjamins.

Varley, S. (2009). I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning, 22*(2), 133-152.

www.ielts.org. IELTS | Researchers - Band descriptors, reporting and interpretation. Retrieved 16 January, 2014, from http://www.ielts.org/researchers/score_processing_and_reporting.aspx

Yeh, Y., Liou, H.-C., & Li, Y.-H. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning, 20*(2), 131-152.

Yoon, H. (2008). More than a linguistic reference: The infuence of corpus technology on L2 academic writing. *Language Learning & Technology, 12*(2), 31-48.

About the Author

Dr. Stephen Jeaco is an Associate Professor at Xi'an Jiaotong-Liverpool University. He has worked in China since 1999 in the fields of EAP, linguistics and TESOL. His PhD was supervised by Professor Michael Hoey and focused on developing a user-friendly corpus tool based on the theory of Lexical Priming.