COMPARATIVE REVIEW:
TEXTSTAT 2.5, ANTCONC 3.0, and *COMPLEAT LEXICAL TUTOR 4.0*

Title	*TextSTAT 2.5*	*AntConc 3.0*	*Compleat Lexical Tutor 4.0*
Platform	PC and Macintosh	PC	PC
Access	Downloadable	Downloadable	Used from the Web site
Minimum hardware requirements	Windows (95/98/NT/2000/XP), Linux, and Mac OS-X	Windows (version 3.0) and Linux (version 2.2)	Windows or Linux
Publisher	Mattias Huning http://www.niederlandistik.fu-berlin.de/textstat/software-en.html	Lawrence Anthony http://www.antlab.sci.waseda.ac.jp/antconc_index.html	Tom Cobb http://www.lextutor.ca/
Support offered	No online manual or help; brief explanations and a few screenshots are provided in the software's Web site	Very brief manual is provided in the software's Web site	Directions are provided in each screen; contact: http://www.lextutor.ca/mailer/
Target language	English and Dutch (interface)	English	English and French
Target audience	Beginning to advanced users	Beginning to advanced users	Beginning to advanced users
Price	Free	Free	Free

Review by Luciana Diniz, Georgia State University

INTRODUCTION

With the advance of corpus linguistics research in the past few decades, there has been a proliferation of corpus software. Due to the high cost of most software products, however, having access to minimally sophisticated types of software is sometimes not possible. For this reason, this review will focus on three corpus linguistics programs that are available at no charge over the Internet: *TextSTAT 2.5, AntConc 3.0,* and *Compleat Lexical Tutor 4.0*. It should be noted that *Compleat Lexical Tutor* is not a single program, but a varied collection of tools for students, teachers, and researchers. Only the researcher's section of this application will be reviewed in this article.

COMPARATIVE REVIEW

Apart from being free of charge, the most important feature these three programs share is that they allow users to upload their own corpus. They also contain basic tools to analyze texts. This review will highlight each program's strengths and weaknesses and how they can be useful for both researchers and language teachers.

All three programs have a user-friendly interface. *AntConc* and *Compleat Lexical Tutor* provide explanations on the screen for each feature that the user clicks on. Even though *TextSTAT* does not offer this same type of support, all the options in this program are also straightforward and transparent. Both

TextSTAT and *AntConc* have to be downloaded and installed in a computer, while *Compleat Lexical Tutor* is used directly from a browser with connection to the Internet. Although all three programs allow users to upload their own corpus, *Compleat Lexical Tutor* only permits the users to upload one file at a time. In the other two programs, on the other hand, users can upload numerous files. *Compleat Lexical Tutor,* however, contains several sample corpora that are available within the program, such as the Brown Corpus, the BNC (British National Corpus), and American TV talk, among others, in case users want to verify or compare information by using a general corpus.

Another aspect that differentiates the three programs is their compatibility with multiple extension files. *Compleat Lexical Tutor,* for example, is compatible only with text (.txt) files, while *TextSTAT* also reads Word files (.doc). Both *AntConc* and *TextSTAT* are compatible with Internet files (.html), but *AntConc* only accepts HTML files saved on disk. A useful and innovative feature of *TextSTAT* is that it contains a Web spider that captures the text directly from the Internet. The users can type the Web address and choose the number of pages they want the Web spider to include in the corpus.
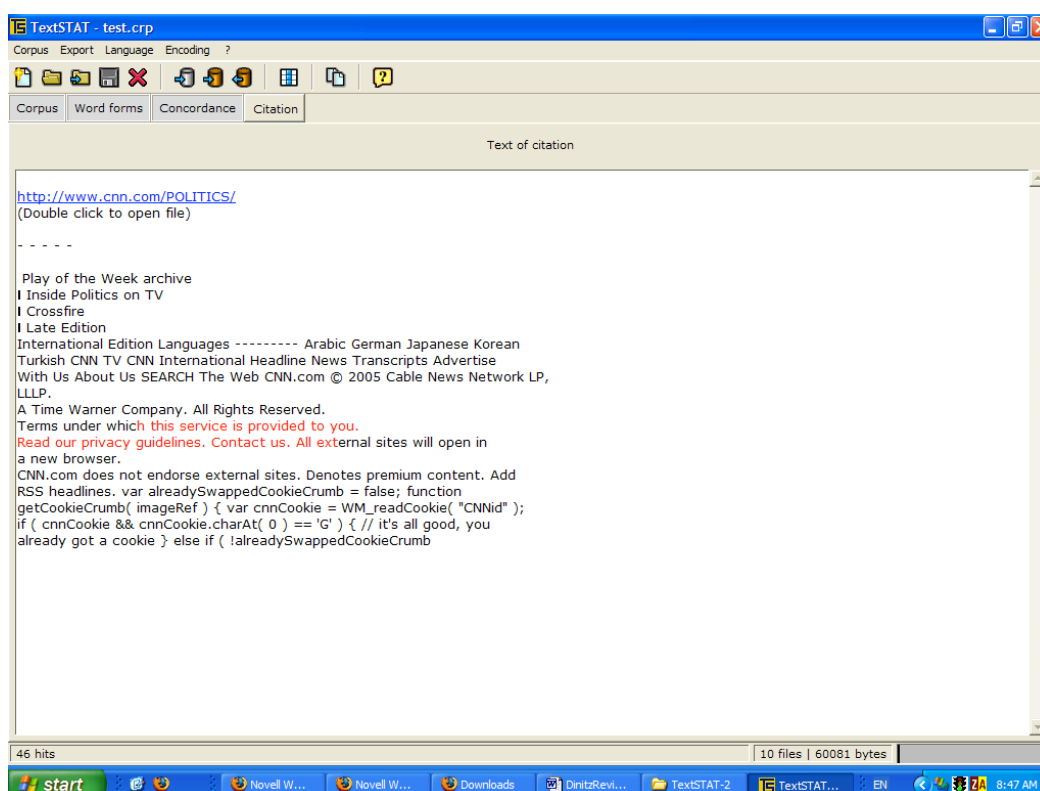

Figure 1. *TextSTAT* screen capturing 10 pages from the CNN Web site and including them in the corpus

At the same time that this element can be helpful, users have to be cautious because the Web spider is not selective on the information it collects from the Web pages, as it retains not only the relevant text, but also advertisements, tables, and any other types of texts. For this reason, it is safer to copy the content and paste it into a text file (or even a Word file if users are working with *TextSTAT*) and then upload this file to the software.

All three programs provide concordance lines (i.e., list of sentences containing a certain word in its contexts) based on the uploaded texts. For instance, both *TextSTAT* and *AntConc* allow users to sort the concordance lines in several forms (e.g., alphabetically by the node word and by the left side and right side of the node word). *TextSTAT* also contains a feature named "query editor," which permits the localization of collocates (i.e., two words that have a high likelihood of occurring together in the same sentence). *TextSTAT* and *AntConc* provide access to the expanded context from the concordance lines. In

TextSTAT, by double clicking on the node word, the users can visualize the expanded context in the Citation Mode, while *AntCont* shows the context on the View File mode.
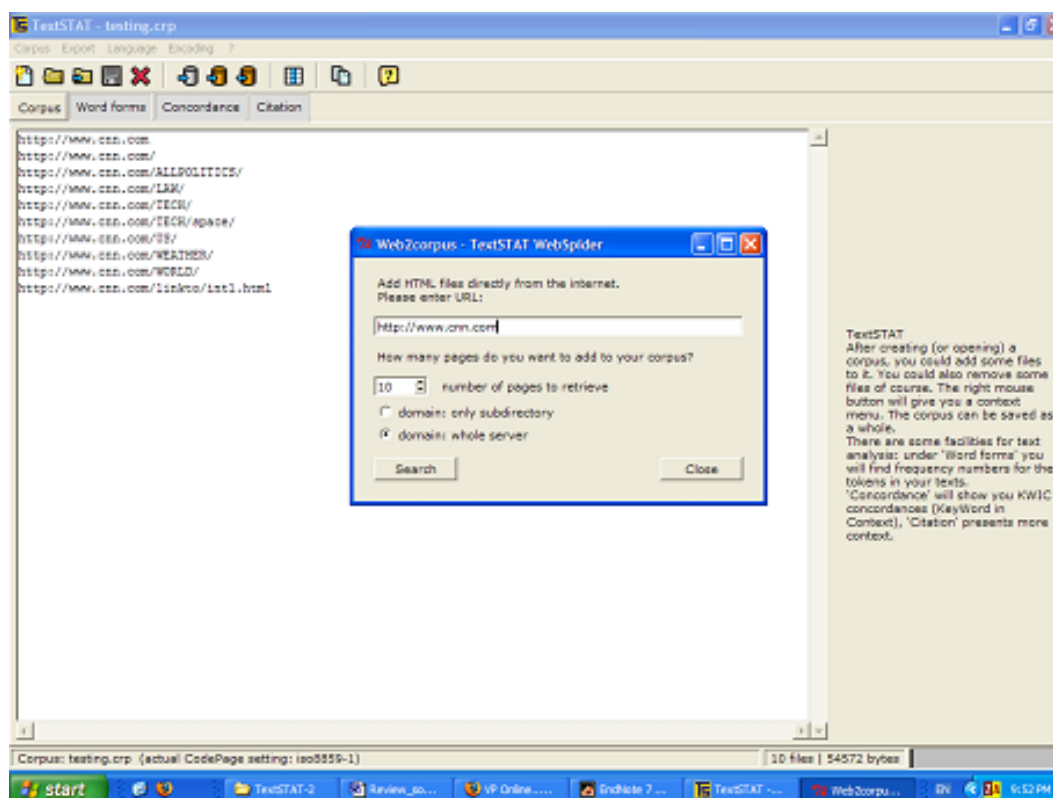

Figure 2. *TextSTAT* screen showing the expanded context of a selected node word

TextSTAT and *AntConc* provide word lists, sorting by frequency, alphabetically, and retrograde. This last feature sorts the word list alphabetically, but backwards. This can be helpful for teachers and students when they are learning poetry, for example, because it provides easy access to rhyming words in a text. Both *TextSTAT* and *AntConc* also run frequency lists either by pre-establishing the minimum and the maximum number of appearances in the corpus or by searching for a specific word.

In both *TextSTAT* and *AntConc,* the users can save the output into separate files. In *TextSTAT,* word lists can be saved in CSV (comma-separated values format) or Excel files and concordances on text or word files. *AntConc* saves all the output in text files only. However, this program contains a helpful feature of allowing individual results to be displayed on separate windows, so that several sources of information can be compared on the screen. In *Compleat Lexical Tutor,* the user has to copy and paste the output into a different file, which sometimes can break the alignment of the results.
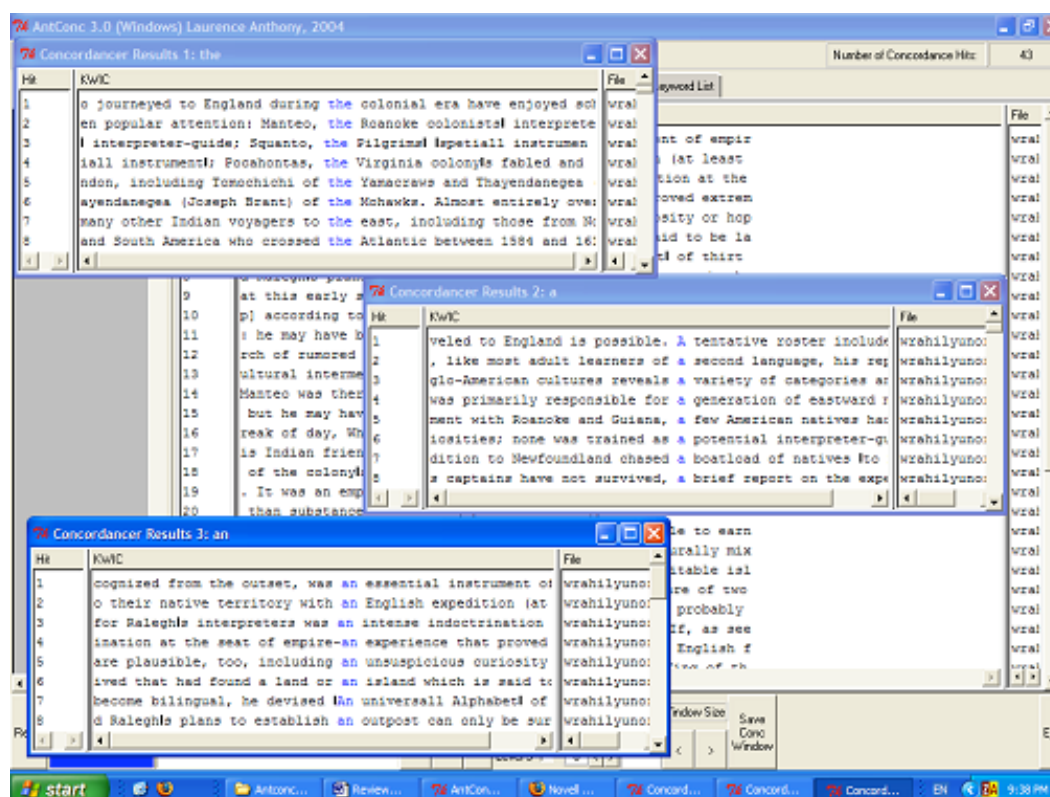
Figure 3. Three mini-screens containing concordance lines are displayed in *AntConc*; the user has access to several types of information on the computer screen

AntConc and *Compleat Lexical Tutors* allow users to look for word clusters. *Compleat Lexical Tutor,* for example, contains a session called *n-gram,* in which users can upload only one small text at a time and search for word chunks. This program also has one feature that allows users to merge a maximum of 10 files into one. However, the user still has to manually copy and paste the output into a text file before using it. *AntConc* apparently holds a great amount of texts. This program permits the search for different sizes and types of clusters.

Even though *Compleat Lexical Tutor* does not offer the same conventional types of text searches as the other two programs, it contains unique features which can be useful for both researchers and teachers. For example, in the Vocabulary Profile feature, users can upload their texts (one at a time) and compare their texts to the most-frequent-words-in-English word list and to the Academic Word List (Coxhead, 2000). This element compares all the words from the uploaded text to those lists and generates an output containing all the words from the uploaded text that are included in the lists. This component can be helpful for teachers to analyze their students' texts and check the variety of words that students are using in their own texts. Moreover, this feature can help teachers and researchers to analyze the complexity of different texts, by looking at how many frequent and academic words the texts contain.
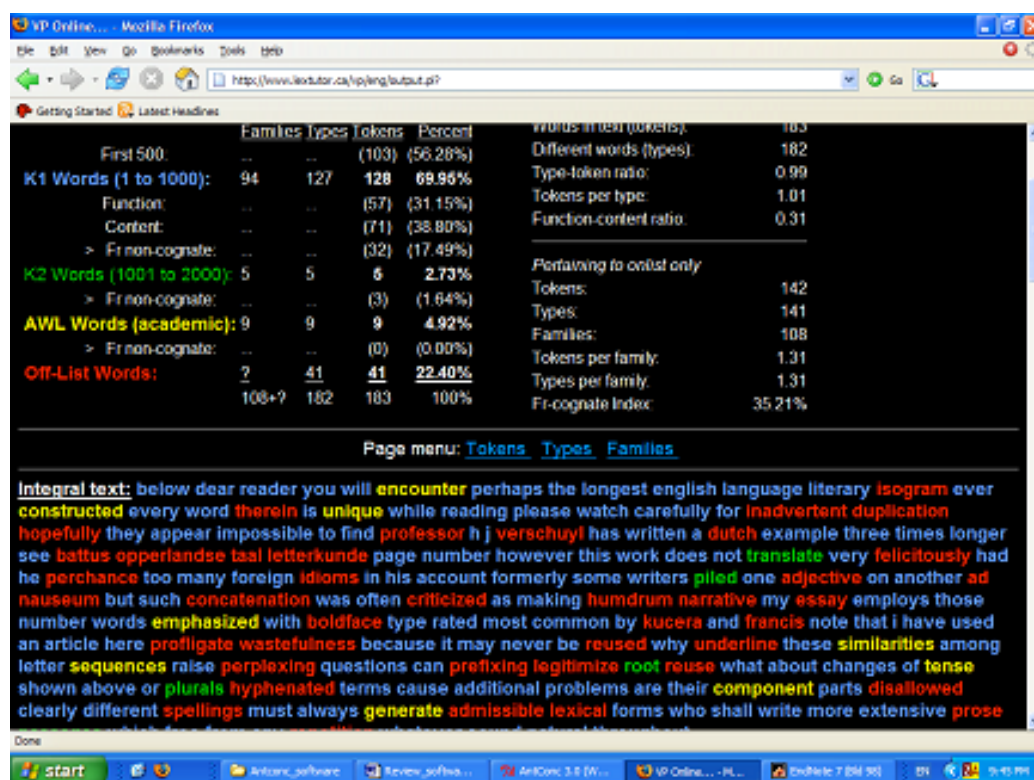
Figure 4. The *Compleat Lexical Tutor* Vocabulary Profile divides the uploaded text into word lists; blue list words are included in the 1000 most frequent words in English, yellow ones are included in the Academic Word List

Along the same lines, *AntConc* has a KeyWord feature that allows users to choose the list of words to which they want to compare their texts. However, *AntConc* does not have embedded word lists as does the *Compleat Lexical Tutor*. Another resource provided by *AntConc* is the possibility of changing the font color, font size, background color, and so forth. This feature can be valuable when teachers are preparing concordance lines to hand out to students. Apart from that, *AntConc* can display the distribution of pre-selected words or word clusters along the text. This feature can provide researchers with insights about the structure of the text.

CONCLUSION

One drawback of all three programs is that they do not provide any type of statistical analysis. *Compleat Lexical Tutor* provides some statistical information, but only about pre-established corpora, not the corpora uploaded by the user. Another drawback is that, apparently, none of the programs works with tagged texts.

As this review has shown, all three programs provide basic text analysis features and can work with uploaded corpora. All three programs are user-friendly and straightforward, but they have limitations, especially for research purposes. The best program to use depends on the researcher's and teacher's needs. Because *TextSTAT, Compleat Lexical Tutor,* and *AntConc* are free of charge, they are accessible to researchers and teachers and can even be used simultaneously, depending on the project's purpose.

ABOUT THE REVIEWER

Luciana Diniz is a PhD student at Georgia State University. She has worked as an EFL/ESL teacher for 6 years.

E-mail: esllsdx@langate.gsu.edu

REFERENCE

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.