TextSTAT 2.7
User's Guide

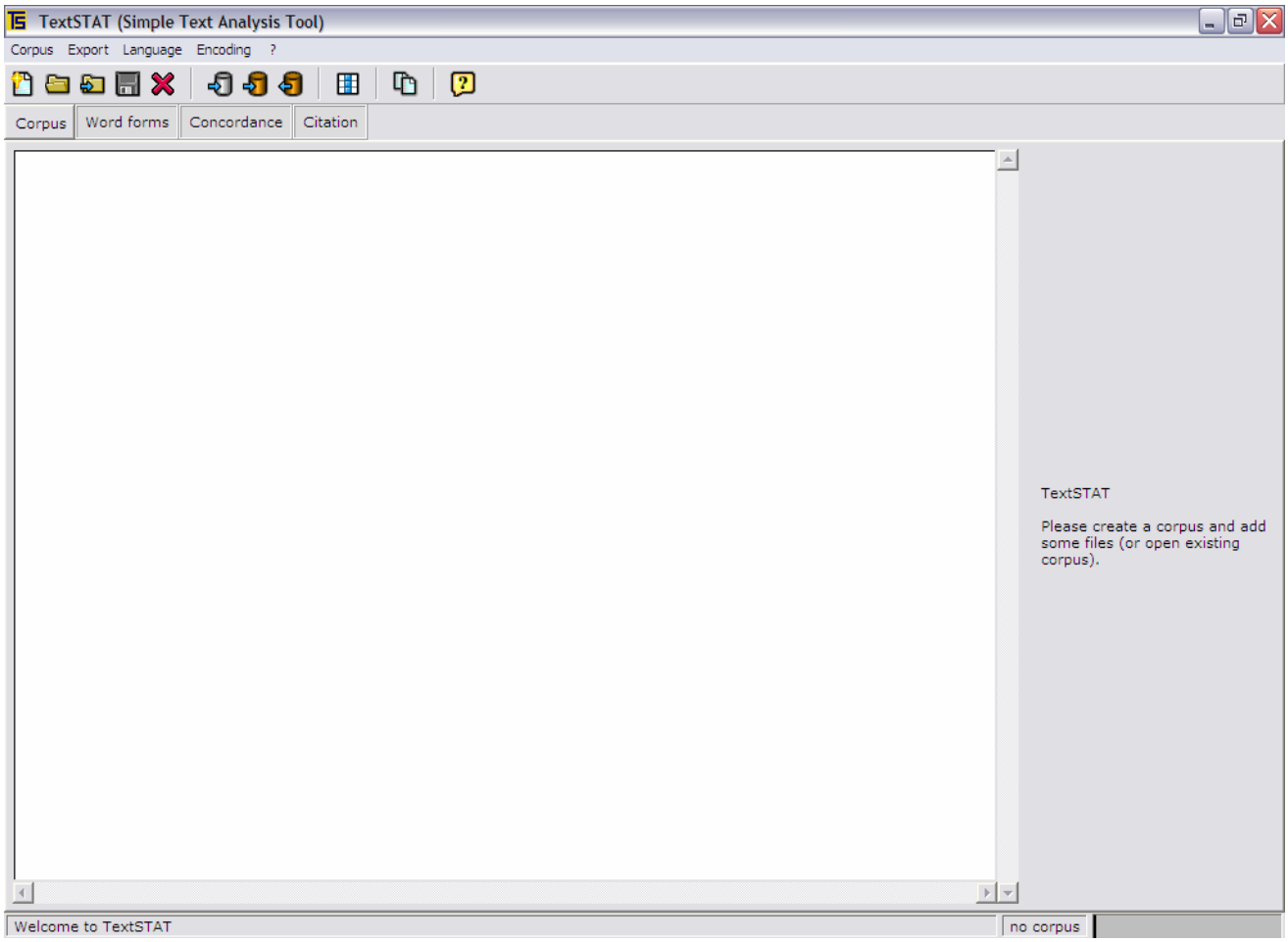
Hüning, M. (2007). TextSTAT (Simple Text Analysis Tool) *2.7*.
http://www.niederlandistik.fu-berlin.de/textstat/software-en.html

Created by Gena Bennett
genabennett@yahoo.com

Table of Contents

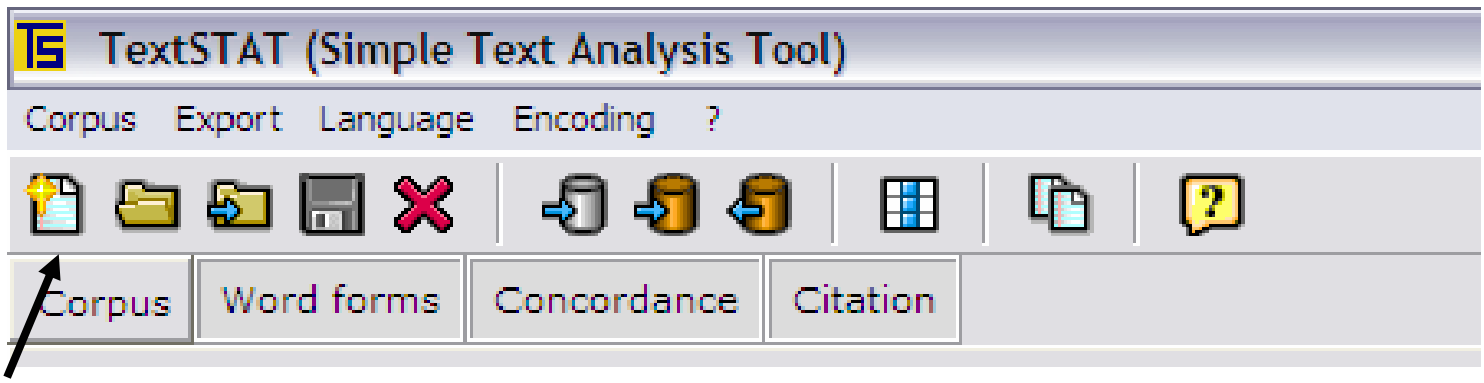
Utilizing TextSTAT…….. 1
Beginning with the corpus……. 1
Creating a new corpus…… 1
Opening an existing corpus…… 2
Adding local files to a corpus…… 2
Adding internet files to a corpus…… 2
The corpus screen…….. 3
Accessing the corpus…….. 3
Word forms…… 3
Frequency list……. 4
Sorting (alphabetically, by frequency, retrograde, by case)…….. 4
Searching for affixes…….. 4
Viewing concordance lines……..... 4
Searching concordance lines…….. 4
Query editor …….. 6
Manipulating concordance lines ……... 7
Citation …….. 7
Viewing extended text …….. 7
Notes ……. 7

When you open TextSTAT, this screen will appear.

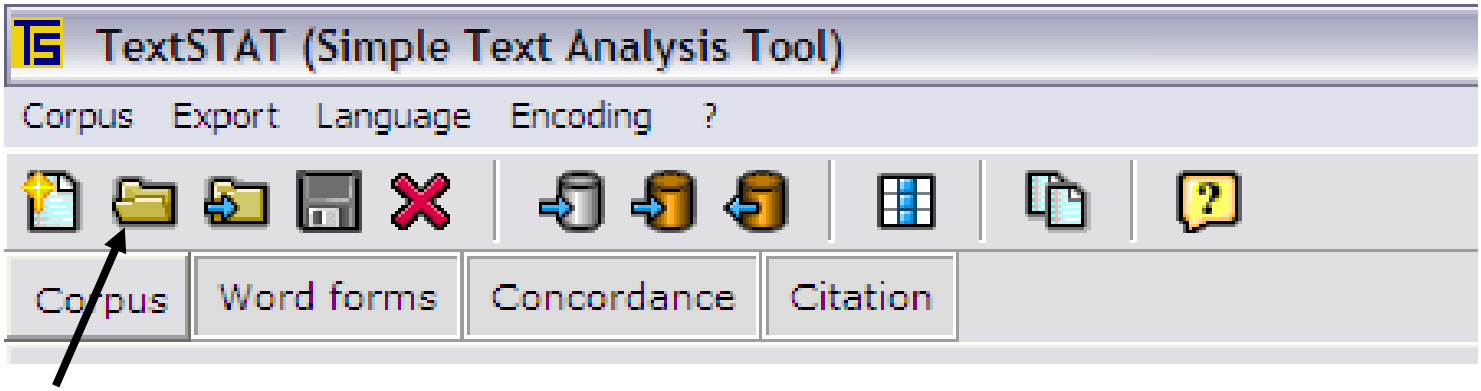


The tool bar along the top of the screen will allow you to utilize the program.

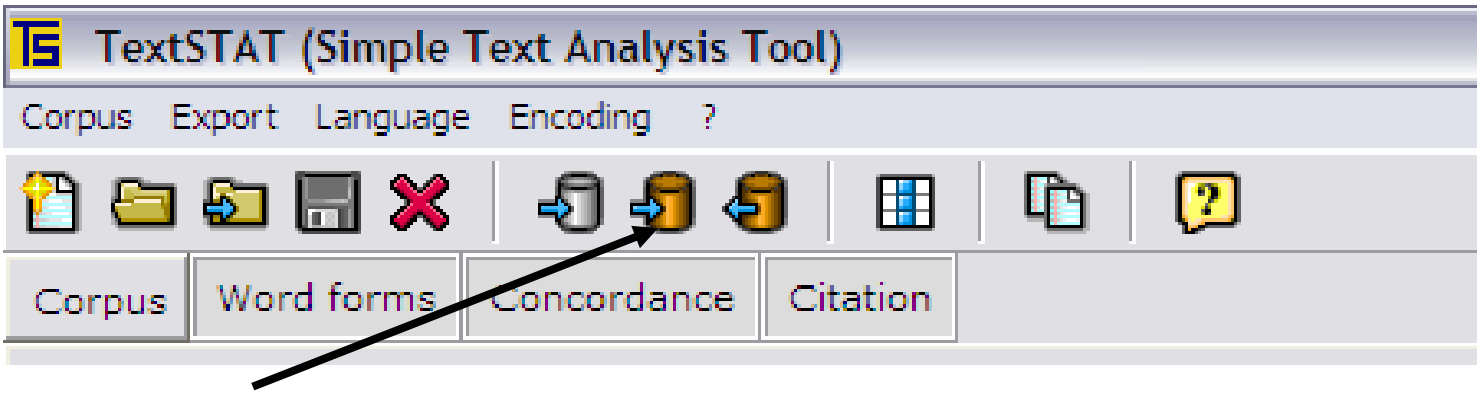
1. Beginning with the Corpus



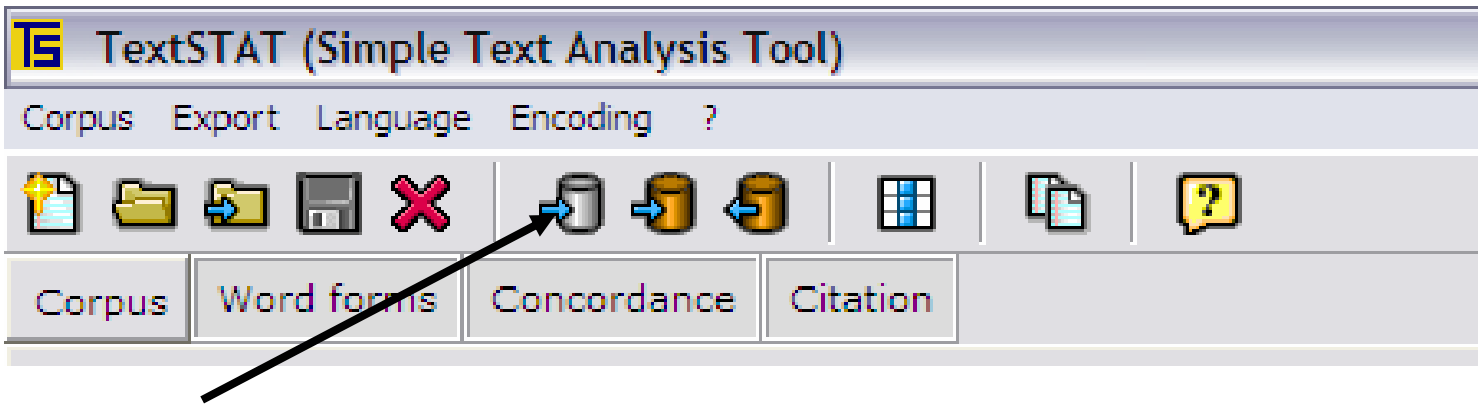
A. To create a new corpus, select the New Corpus button (first from the left). This will open a window allowing you to name and save a new corpus. You must name the corpus before you can add files. Once you name your corpus, a box will appear letting you know you may add files. See steps C and D for adding files.



B. To open an existing corpus, select the Open Corpus button (second from the left). This will allow you to select from your saved corpora to open the corpus of your choice.

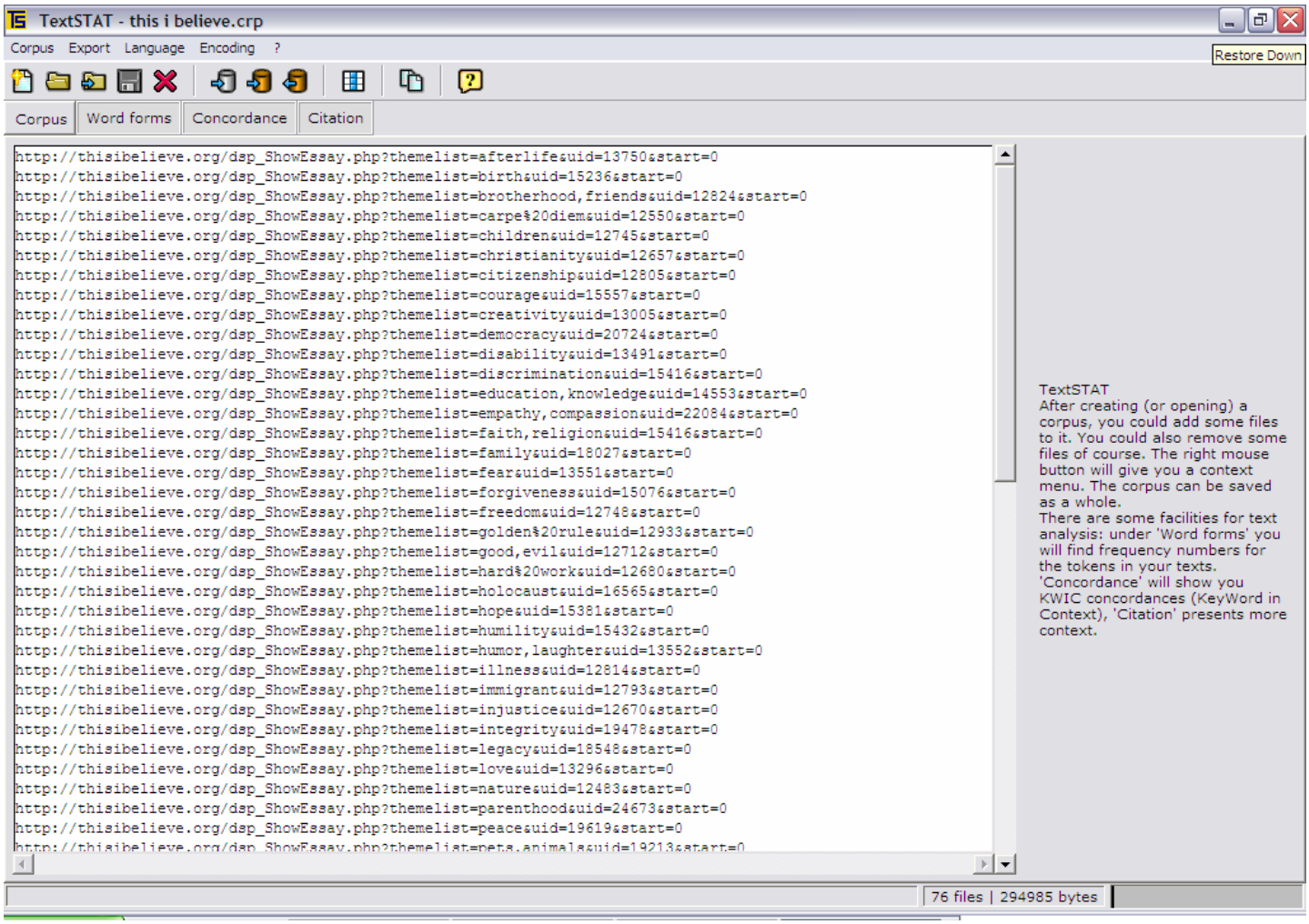


C. To add local files from your computer to a new or existing corpus, select the Add Local Files button (seventh from the left). This will allow you to search your computer to add a .doc file or a .txt file.



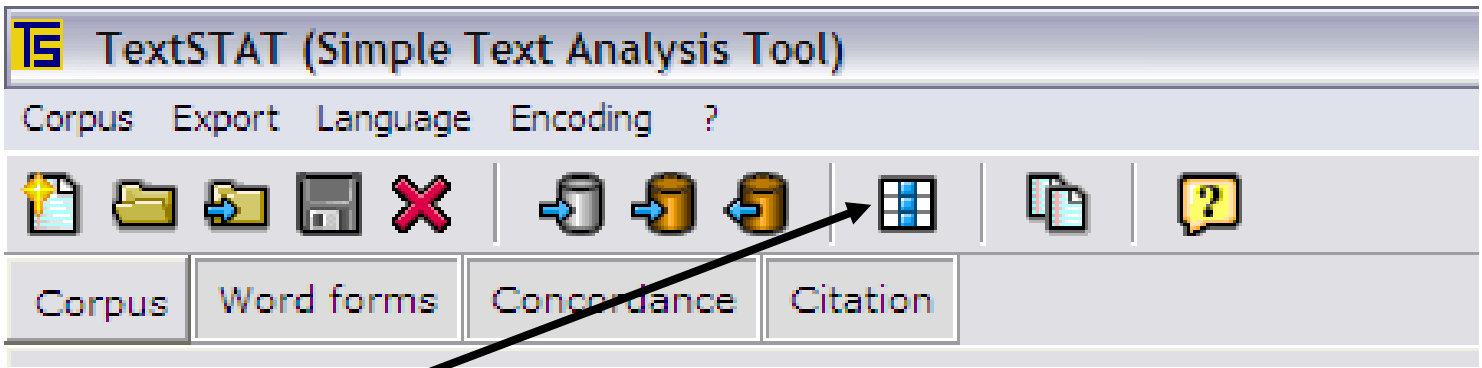
D. To add internet files to a new or existing corpus, select the Add File From Web button (sixth from the left). This will open the Web Spider in order for you to paste in a web address(es).

Once the corpus in complete, the corpus screen may look something like this (displaying each file in the corpus):



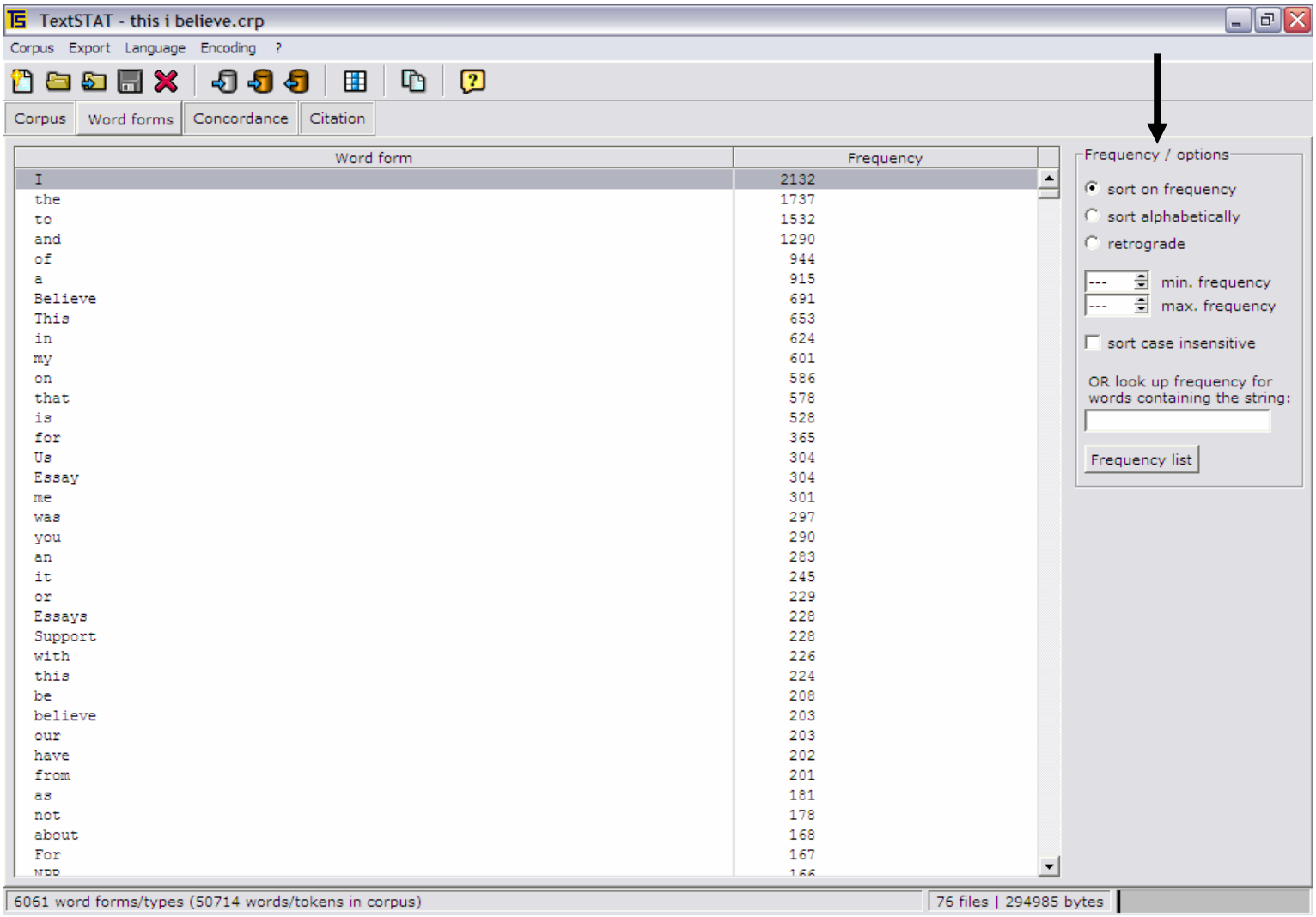
2. Accessing the Corpus

The corpus may be accessed in three ways: *Word Forms*, *Concordance*, and *Citation*. To begin, access the corpus via Word forms.



A. *Word Forms*

To get a word list for the corpus, select the Show Word Frequencies button (third from the right). This will open the *Word forms* screen and display a word list sorted by frequency.



To manipulate the word list, use the options on the right side of the *Word forms* screen.

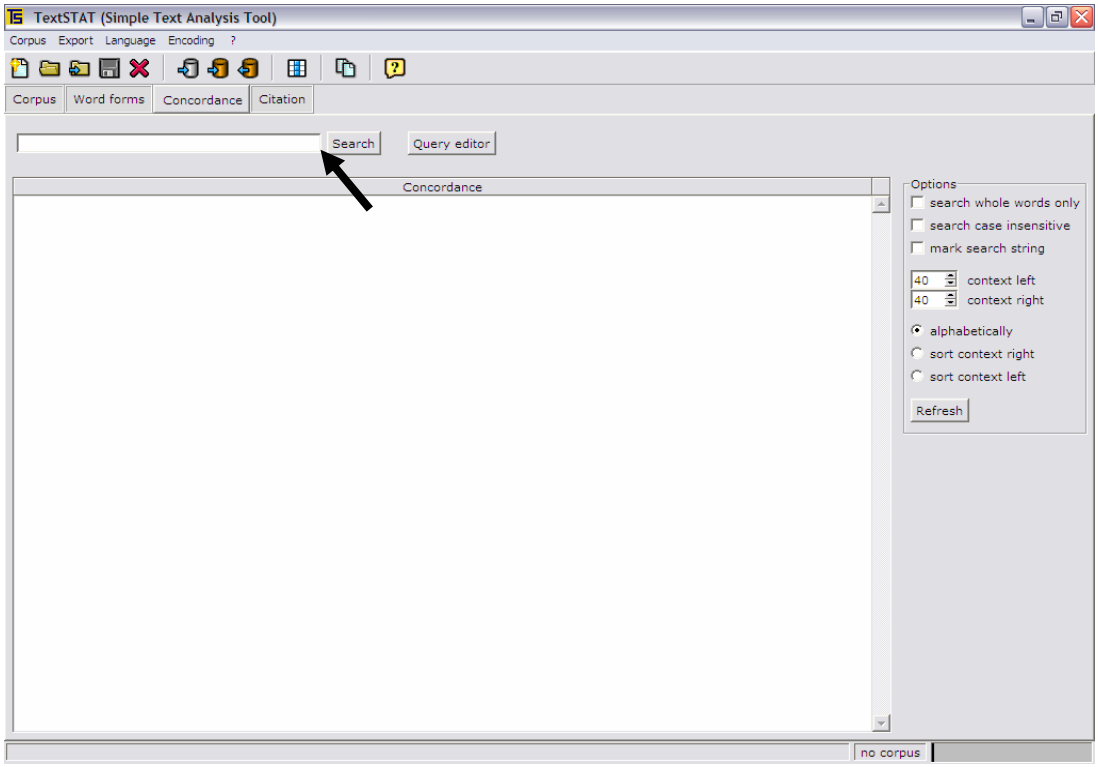
1. Sort the word list according to frequency, alphabetically, or retrograde (alphabetically backwards).
2. Establish a minimum and maximum number of times a word should appear in the corpus in order to show up on the word list.
3. Sort case insensitive so that a word that begins with a capital letter, sentence initial, for example, will be counted with the same word that begins with a lowercase letter.
4. Search for affixes using the search box.

B. *Concordance*

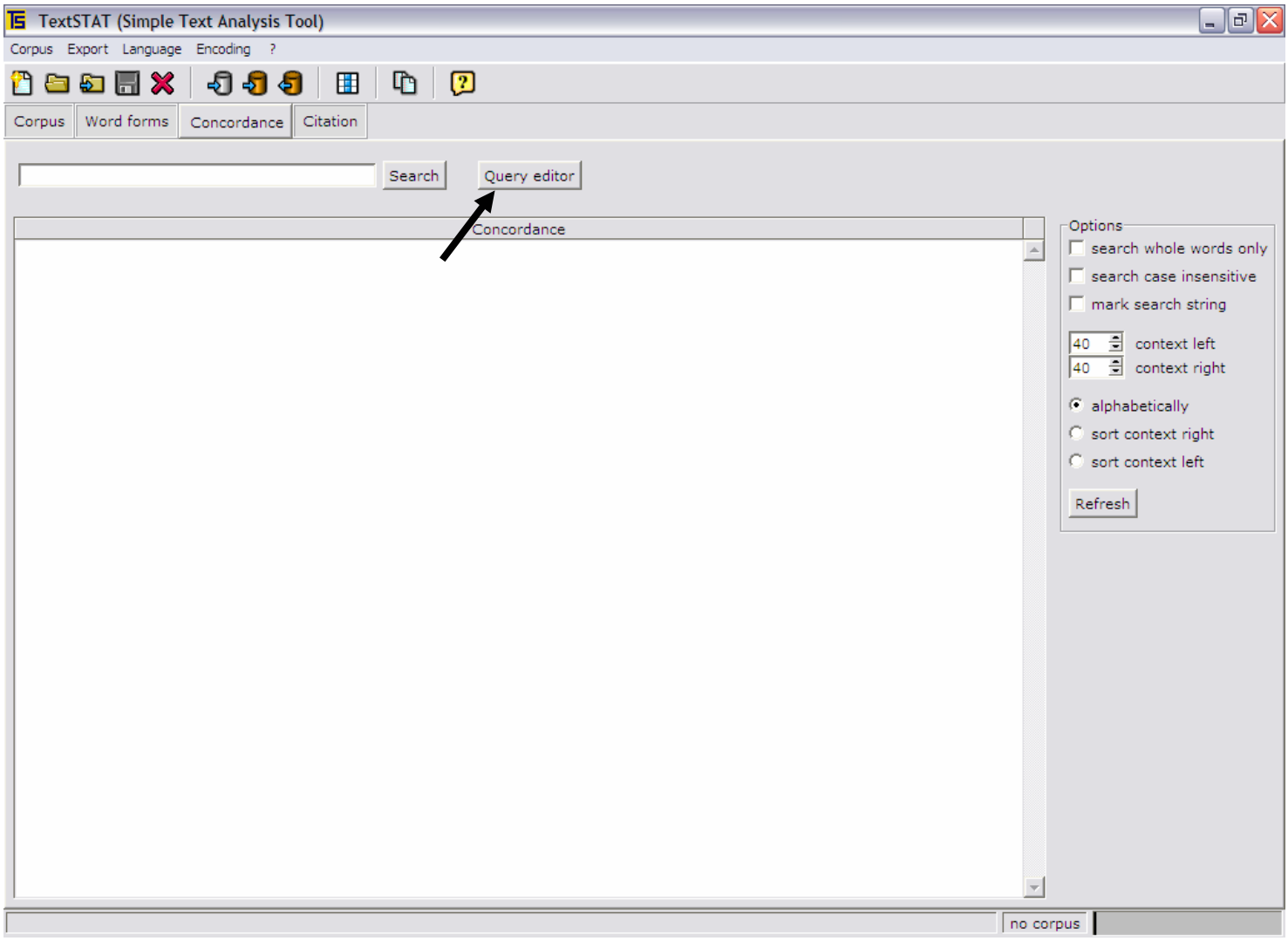
You may view concordance lines from the corpus via two options:

1. Double click on a word from the word list. This will open the *Concordance* screen and show all lines from the corpus that contain the selected word, as well as 40 characters on each side of the word.

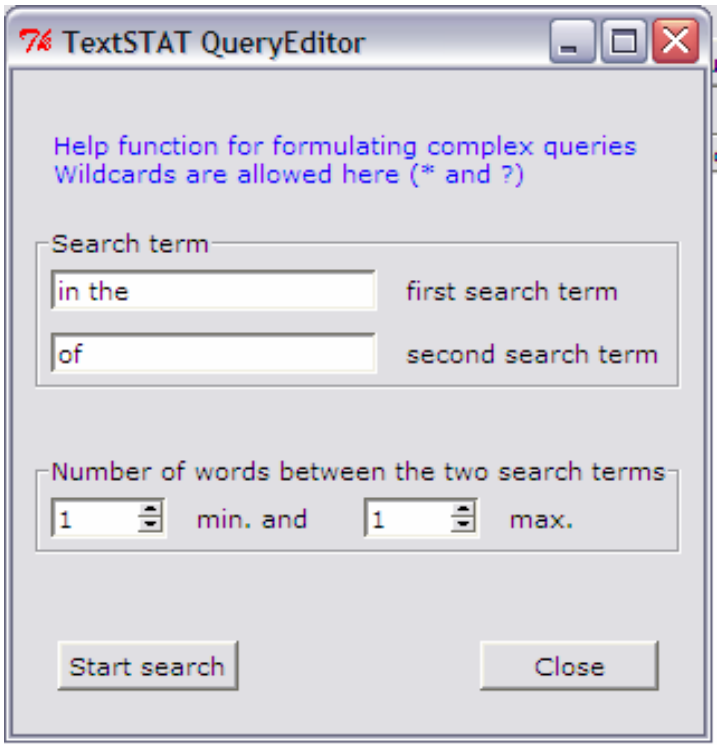
2. Click on the *Concordance* option to open the concordance screen.
   a. Type in your search word/phrase in the search box.



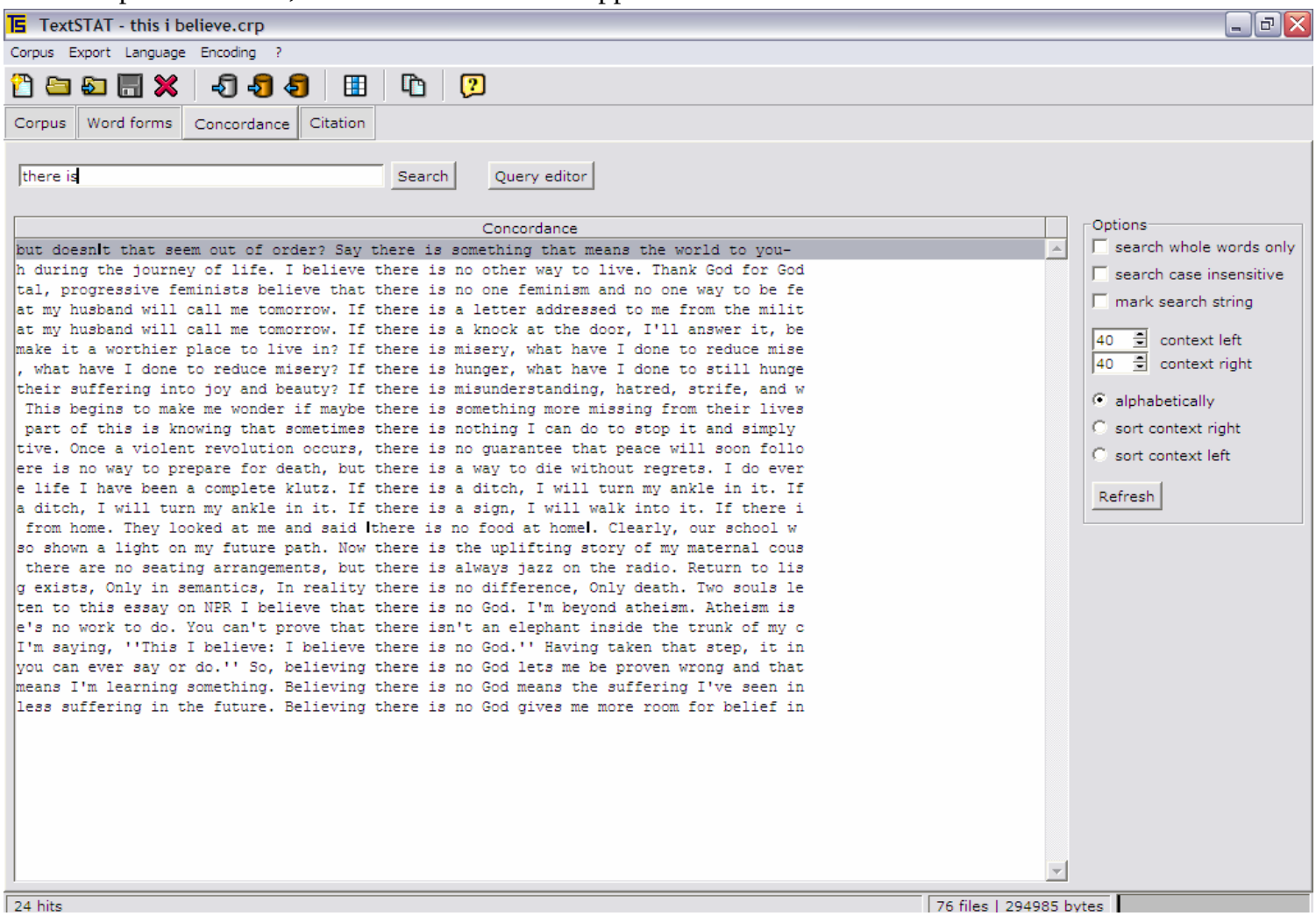
b. To search for a frame (a phrase with a wild card; for example *in the xxxx of*) in the corpus, select the Query Editor button.



Enter the frame, *in the* in the first search term box and *of* in the second search term box; then, select *1* word minimum and maximum between the search terms.



For both options a and b, concordance lines will appear on the screen:



To manipulate the concordance lines, use the options on the right side of the screen.

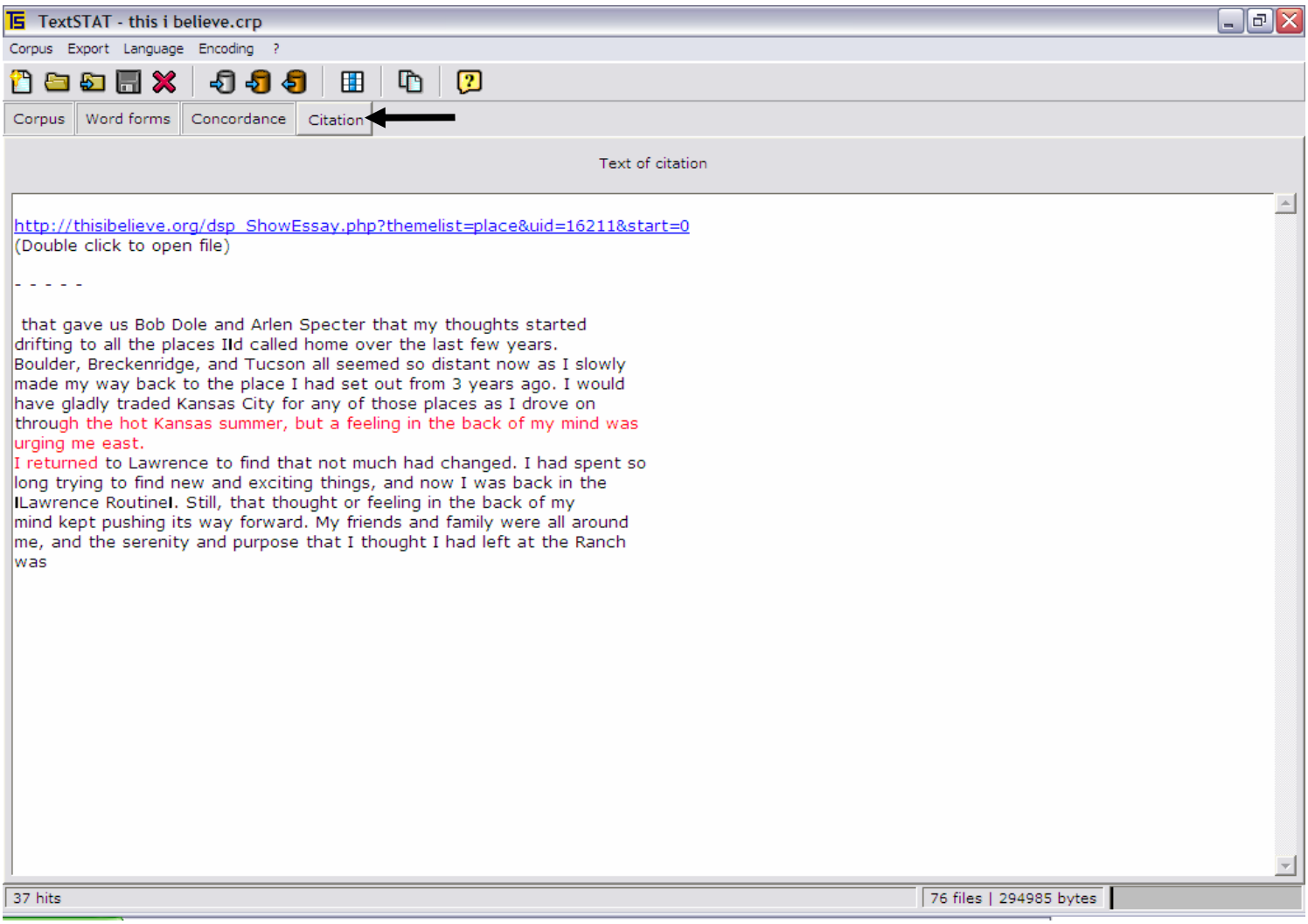
1. The concordance lines may be sorted to the left, i.e. words before the search word/phrase are alphabetized.
2. The concordance lines may be sorted to the right, i.e. words after the search word/phrase are alphabetized.
3. For concordance lines which contain a frame, lines may be sorted alphabetically according to the wildcard.
4. Sort case insensitive so that a word/phrase that begins with a capital letter will be counted with the same word that begins with a lowercase letter.

All of the above options are available on the right hand side of the *Concordance* screen.

C. *Citation*

To see extended text from a word, phrase, or concordance line in the corpus, use the *Citation* screen. The *Citation* screen may be accessed by double clicking on a concordance line. This will open the *Citation* screen to show

1. in blue, a link that contains the original document from which the concordance line came;
2. in red, the original concordance line;
3. in black, 50-75 words on either side of the original concordance line.



NOTES: The *Word forms* and *Concordance* options will only function when a corpus has been opened. The *Citation* option will only function when concordance lines have been selected.