WORDSMITH: Text Analysis

By Anthony Bushong

This tutorial will review how to make a batch text file and how to search for keywords within the text files up for analysis. It is based on Linguistech's Wordsmith Tutorial. See this for more detailed tutorials on specific types of queries.

What is Wordsmith?

Wordsmith is an advanced software that documents word frequency and patterns with the ability to sift through a large corpus of documents. This is advantageous in parsing through one person's collection of work or speeches to document common themes or relevant topics.

Getting Started

a. When opening WordSmith, go to settings. In the Settings window, make certain that the radio button for "advanced" is selected in the bottom right hand corner. Then click "OK".

b. Select the "WordList" tab, and then go to File -> New. Select "Choose Texts Now" Move over all the .txt files you will be using. Then select "OK". (pictured right)

c. Next, select "Make a batch now". Make sure you note the location of where the .zip folder is being saved, then click "OK". You have now made a wordlist documenting the frequency of words in each .txt file. (pictured right)

d. Repeat the following steps but this time instead of selecting "Make a batch now," select "Make a wordlist now". Now you have a wordlist documenting the frequency of words in the complete corpus of the .txt files.
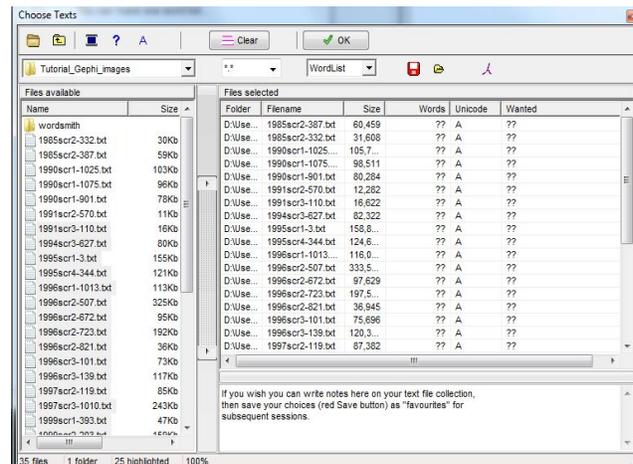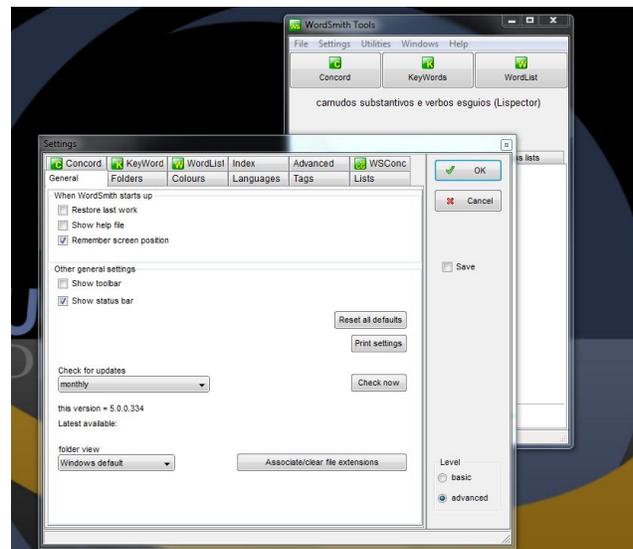
Keywords

When documenting differences between the speeches or works of a specific author, keywords will be especially useful for comparing and juxtaposing what made a specific work different from the rest.

a. To find keywords, go back to the original window. Select the tab for Keywords at the center of the top menu.

b. Then go to File -> New. For the reference corpus wordlist, use the master wordlist that you made.

c. For the keyword list, select any of the individual .txt files. Once you have each selected, click "Make a keyword list now." (pictured right)
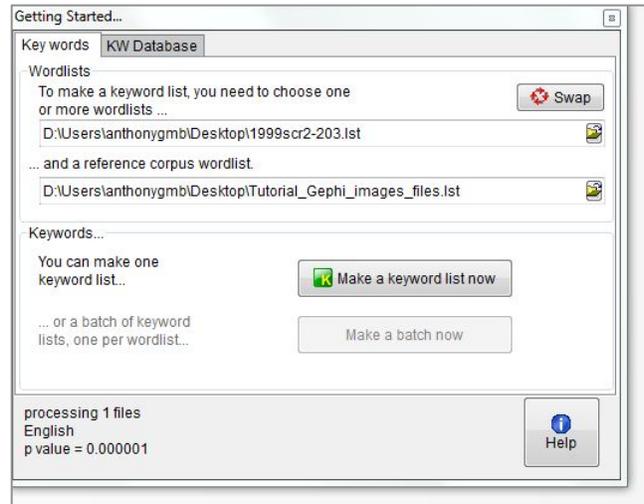
Once you have done this, you will receive the keywords from the specific individual .txt file that were specific to that file when compared against the corpus of files. Congratulations!

You can now use WordSmith.