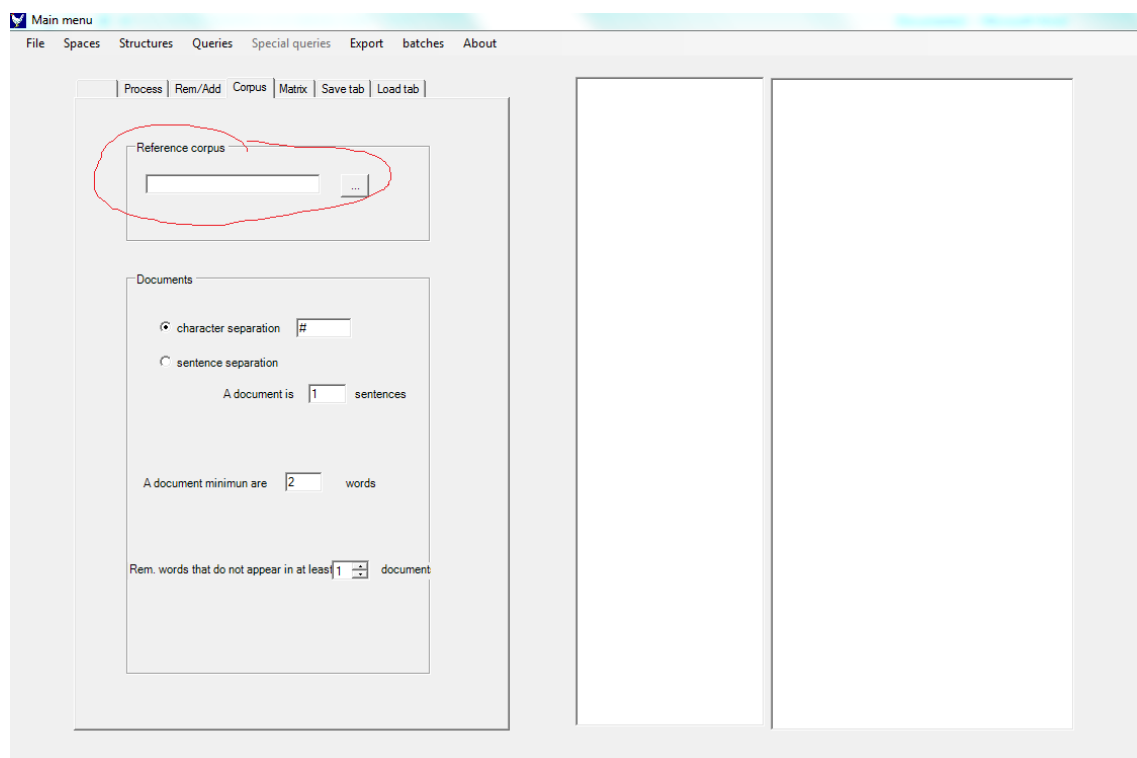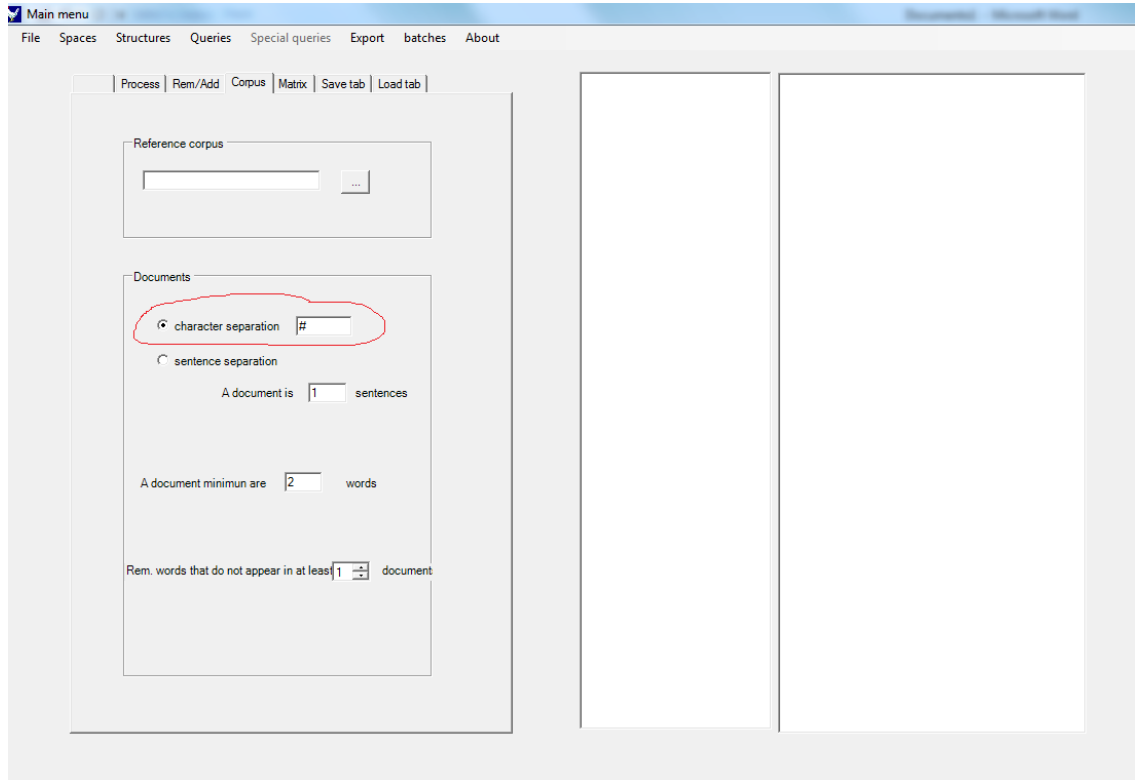# Gallito 2.0 Quick Start

I.   Training a corpus

(Before you start, don't forget to give writing permissions to the directory so that it can install the program in C:\Program Files\elsemantico.com\Gallito_2.0 64Bits)
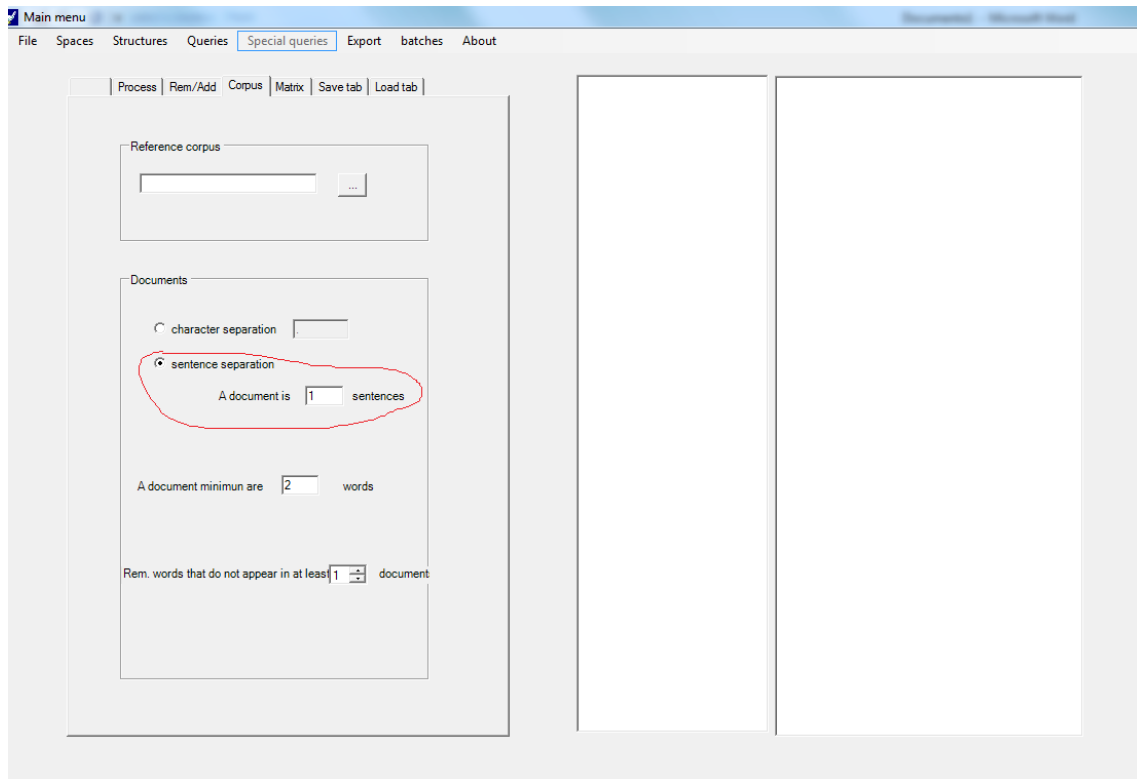
Step 1: Select a corpus file. For instance, the corpus_test.txt file in the distribution package. The File must be in TEXT or ASCII format. The Gallito 2.0 trainer process cannot read WordPerfect, Word, or other word processing files. Notice that Word processing programs usually permit you to convert your files into ASCII or TEXT format.
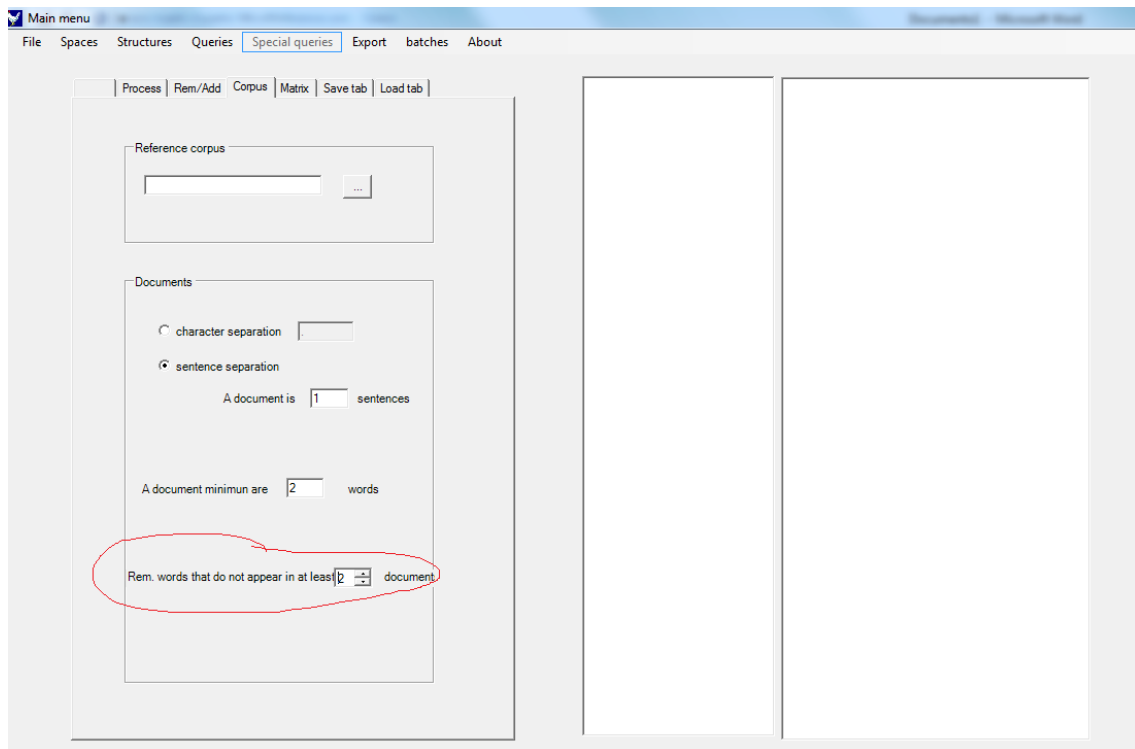
Step 2: Select the context window. You can choose sentences, paragraphs, or sections. In the case of paragraphs or sections, they must be separated by a character (usually "#"). You must specify this in the "character separation" textbox. For paragraphs, you can change the "carriage return" into the chosen character using a text processing program.
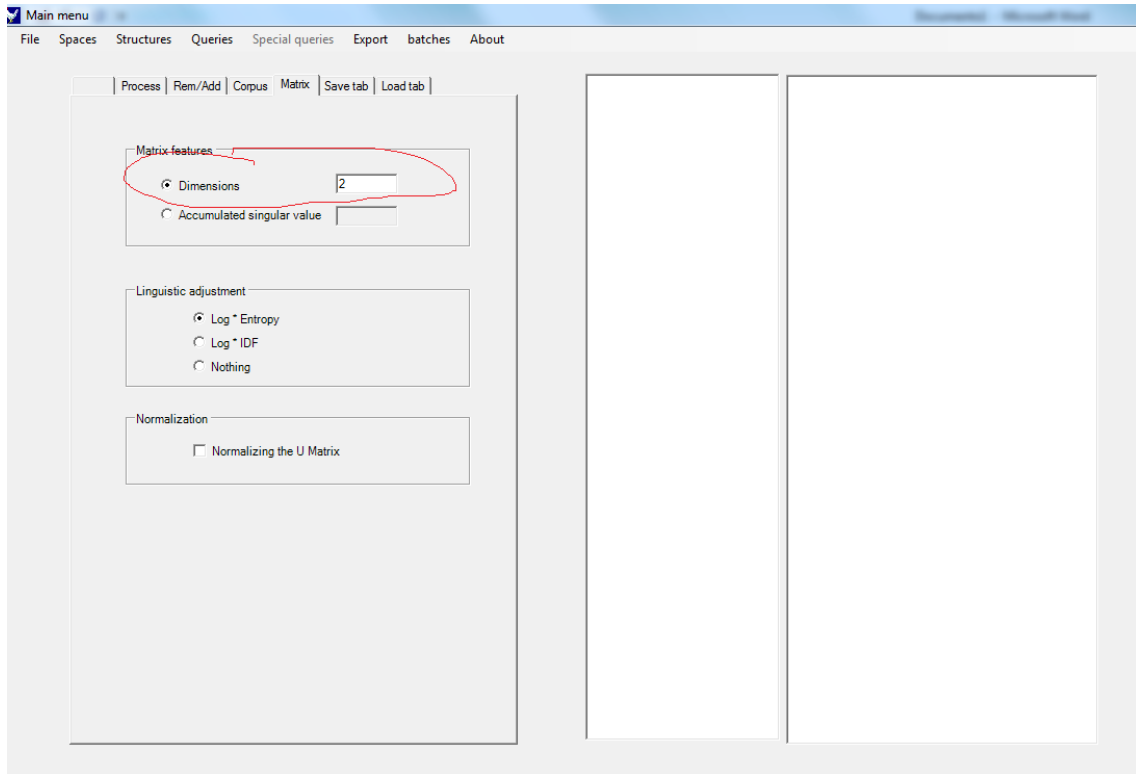
In the case of sentences, check the "Sentence Separation" box and specify that the document is 1 sentence long. (If the number is higher, you will be specifying 2 or more sentences for the contextual window, which is only advisable for specific reasons). The periods in the text will serve as separators.
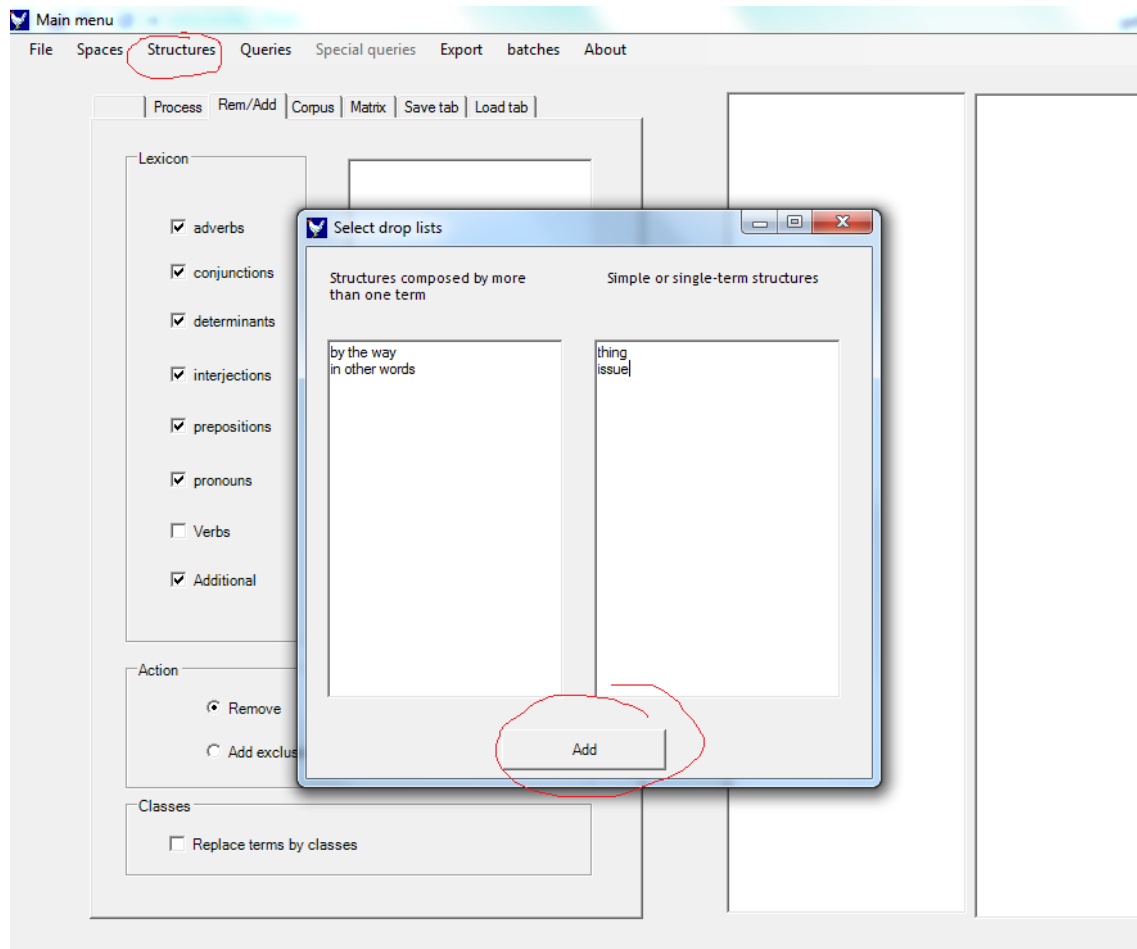
Step 3: Remove hermit terms (words must appear at least twice to be processed)
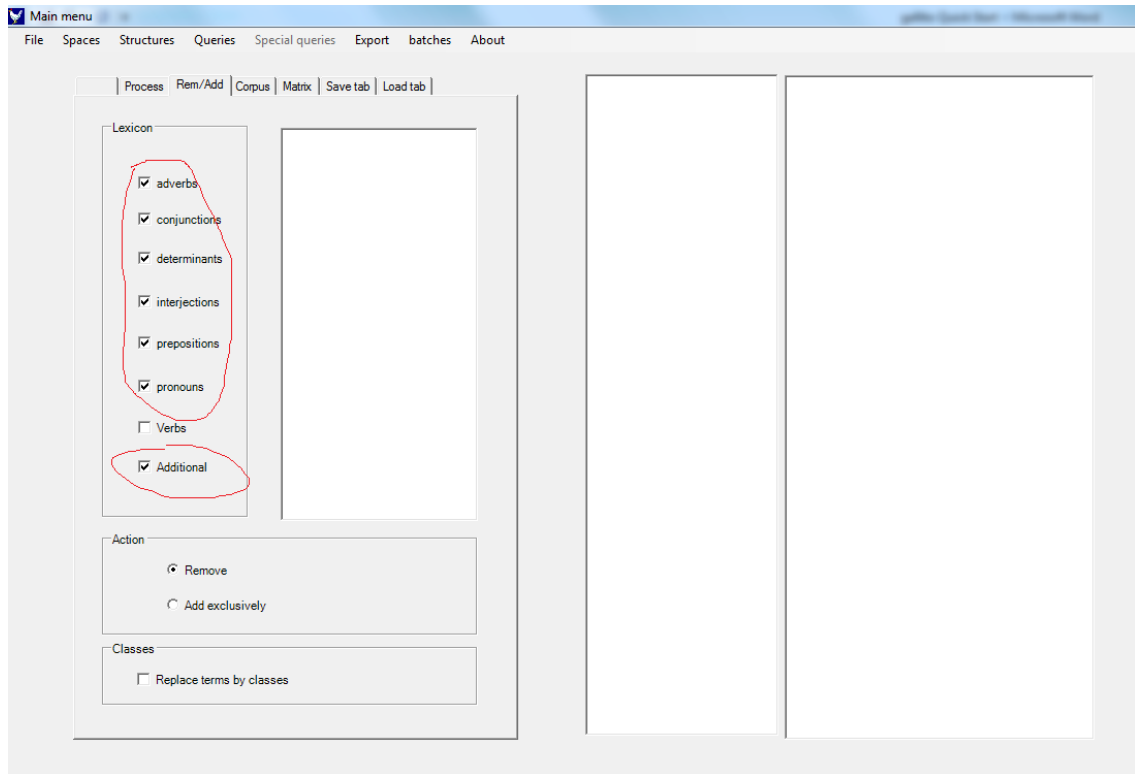
Step 4: Set the final dimensions. Dimension selection depends on size and topic. Take a look at the empirical literature.
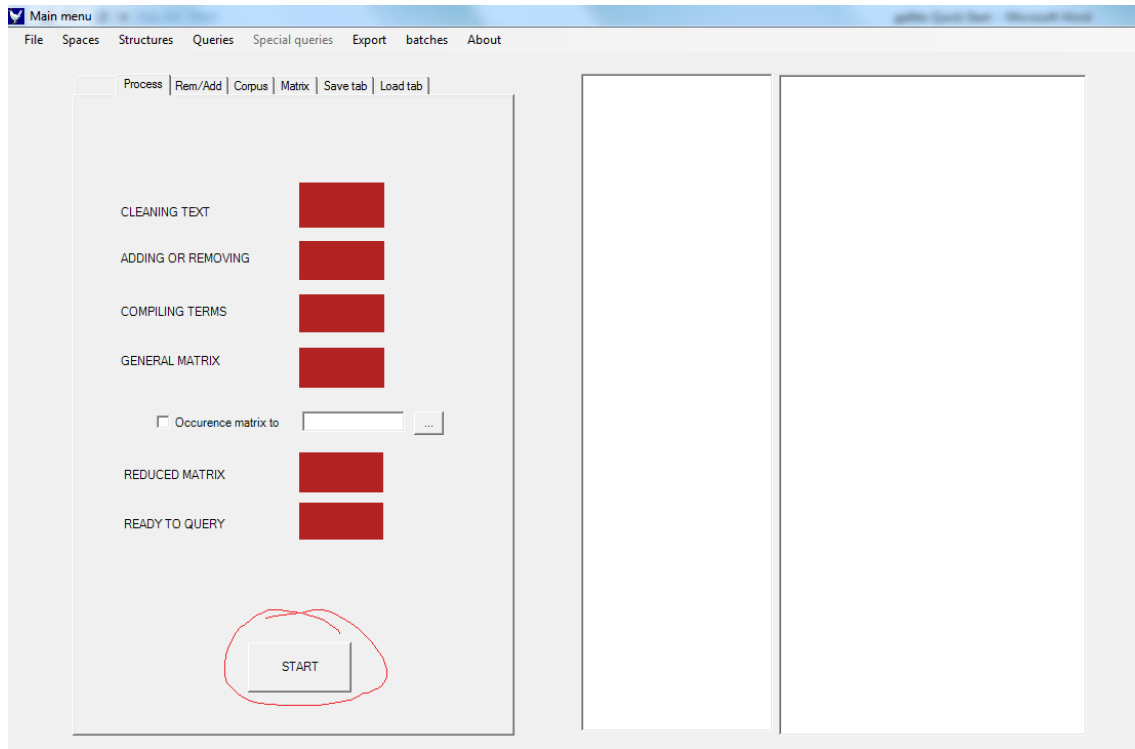
Step 5: Make a stop list.

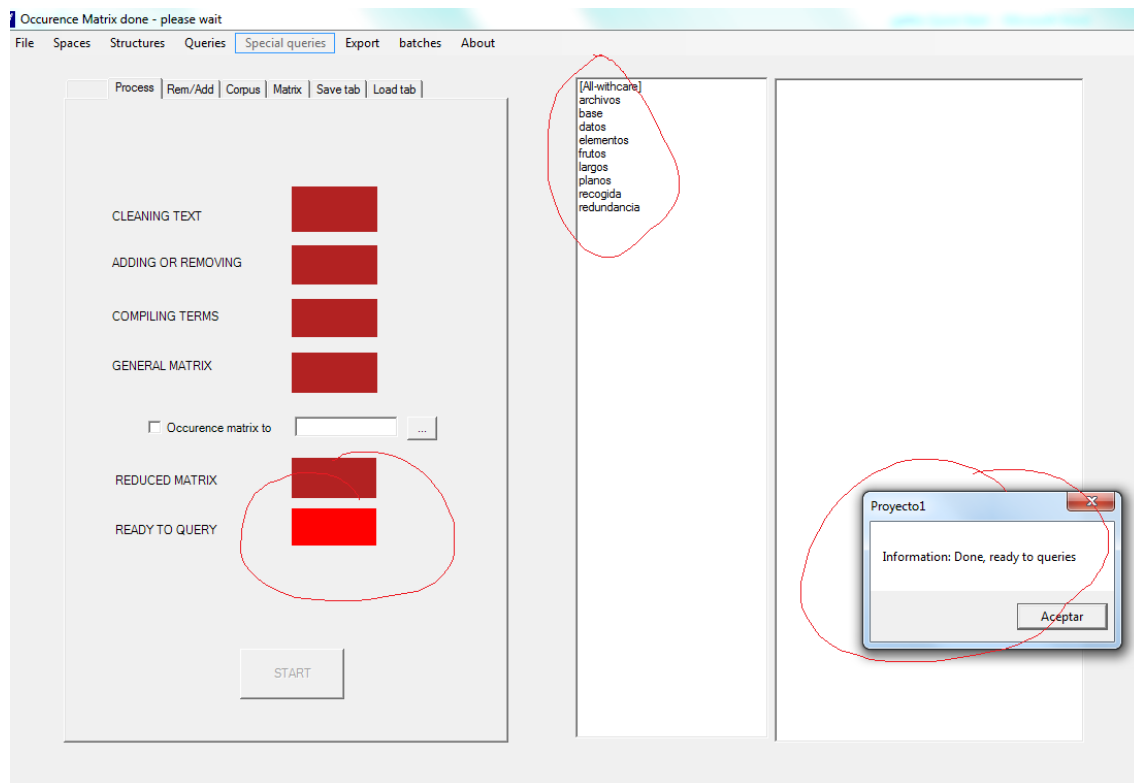**Step 6:** Select function particles and stop lists (to remove).

Step 7: Click on the Start button.

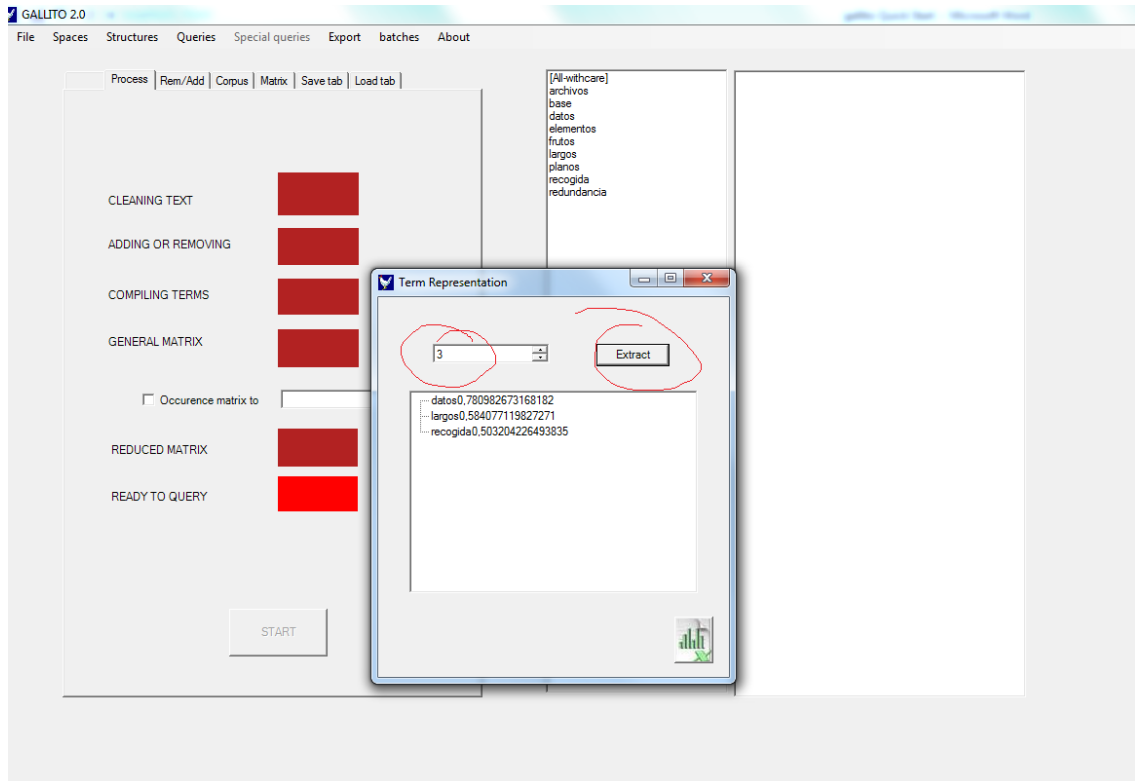**Step 8:** Now you are ready for queries!

II. Some examples of easy queries

Step 1: Word neighbors

Step 2: The most representative words.

III. Saving a space

Just one Step: Select the directories and names for all of the 7 space files (you can give them any names you like, but remember which is which). Then click Save. It is advisable to save the files in binary format.