Collocation networks in discourse analysis

- You will be using the **GraphColl** module in **#LancsBox,** a free multi-platform tool for the analysis of language.

- **How to cite the tool?**

  #LancsBox is licensed under BY-NC-ND Creative commons license and is free to use for non-commercial purposes. When you use #LancsBox for presentations or publications, please cite:

  > Brezina, V., McEnery, T. & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, *20*(2), 139–173.

**Step** 1: Download #LancsBox.

1. Go to http://corpora.lancs.ac.uk/lancsbox
2. Extract the contexts of the zip file by right clicking on the file and choosing 'Extract all' from the contextual menu.
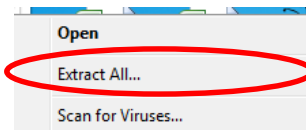
3. Double click on **LancsBox.jar** to run the program.

Step 2: Upload your corpus to #LancsBox following the steps below.

1. Click on '**Browse'**
2. Choose the folder containing the corpus.
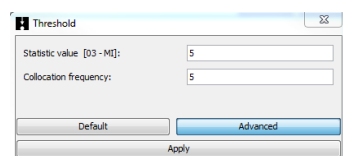3. Select all files (CTRL+A) click on '**Open'.**
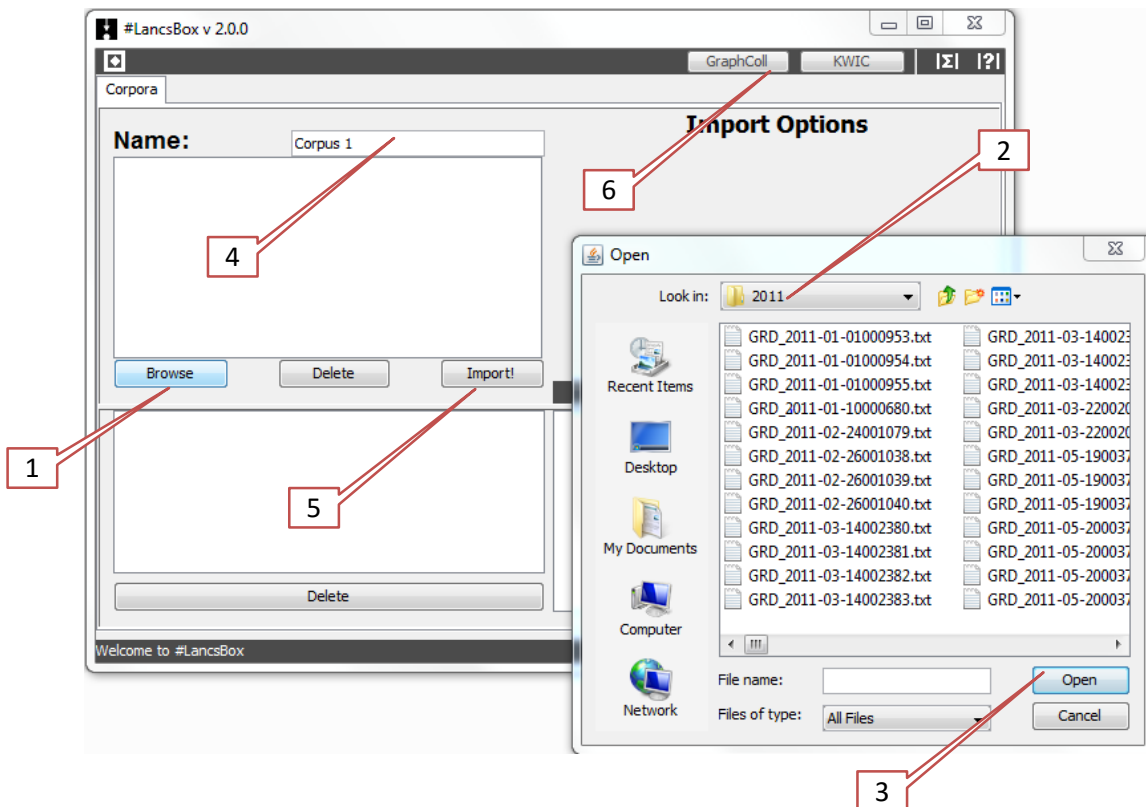
4. **Name** your corpus. If no name is provided, #LancsBox will automatically assign a generic name such as Corpus 1, Corpus 2 etc.
5. Click on '**Import'** to upload the corpus to #LancsBox.
6. Click on 'GraphColl' to start the collocation network module.

Task 1. Create graphs. Work with the LOB corpus.

a) Build the first-order collocation network around the word *time* using MI score and the default settings.

b) How many collocates does the graph display? Are all of them useful?

c) Change the default settings as indicated in the figure below (MI = 5 and above) and search for the node *time* again.

How many results did you get this time?

Task 2. Build collocation networks and explore graphs.

a) Go to the graph you have created in Task 1 c). It should be similar to the graph displayed in the figure below:

b) Find the collocate *spend* in the graph and double click on it. You should get a collocation network similar to the one displayed below:

c) Find the second-order collocate *money* in the graph and double click on it. Comment on the connection between *time* and *money* that you can see in the resulting graph that shows collocates around the node *money*.