

Quickstart Guide to AntConc

THE MCGRAW CENTER FOR TEACHING AND LEARNING

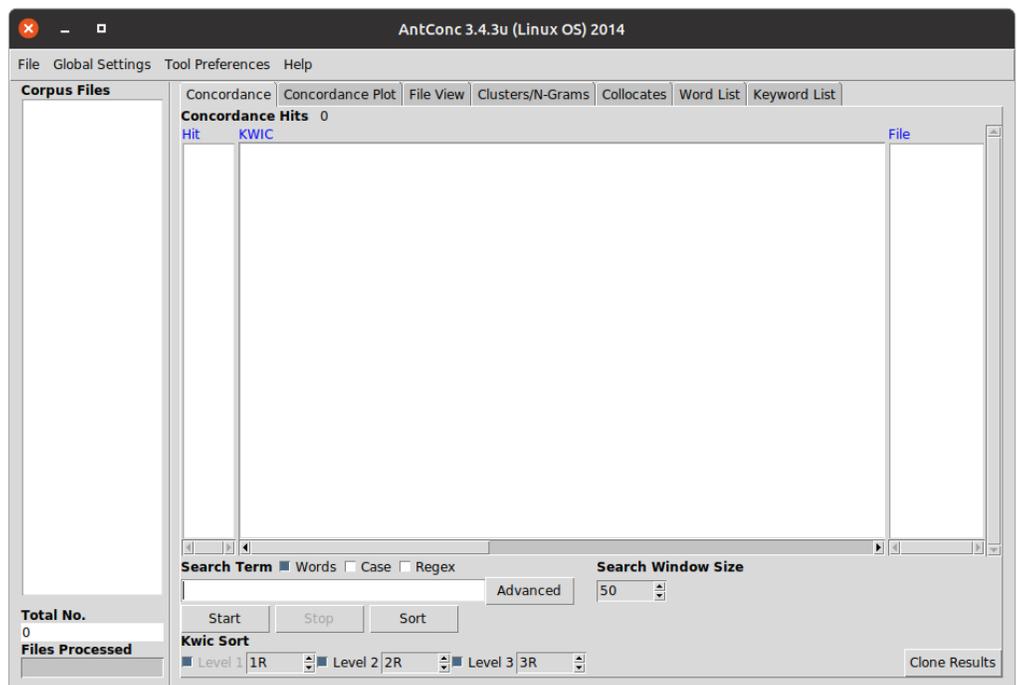


[AntConc](#) is a freely-available desktop corpus search and concordancing program available for Windows, Mac, and Linux. AntConc is easy to use but offers many powerful functions for analyzing corpora, including concordancing, collocation search, and word frequency, keyword, and n-grams lists. It also, importantly, supports the use of regular expressions in search queries. Regular expressions enable advanced searching beyond simple term matching.

The interface

After you have downloaded, installed and opened the software, you should see an interface similar to the one pictured to the right. The exact appearance of the interface may differ from platform to platform, but the basic elements are the same.

The left-hand column displays a list of the text files included in your corpus and some information about that corpus. Note that AntConc only works with plain text files.



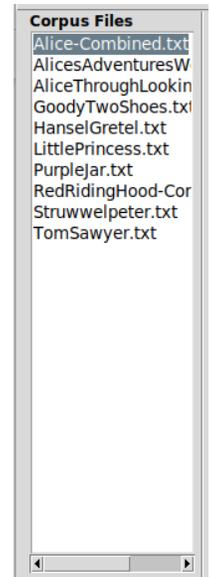
The larger, main area can display results of the various actions that can be taken on text as represented by the tabs across the top of this column: Concordance, Concordance Plot, File View, Clusters/N-Grams, Collocates, Word List, and Keyword List.

The area just below this main window includes tools for performing searches and for configuring how the results of those searches will be displayed.

Adding documents and performing searches

The corpus you intend to work with **must be a collection of plain text documents** (.txt). Place your documents in a folder on your computer. In AntConc, open the **File** menu and select **Open Dir..** This will allow you to navigate to that folder and ingest all the text documents contained within. You can also add files individually by selecting **Open File(s)...**

After files are added to AntConc, they should appear in a list on the left-hand side of the application. You can click once on any of the files in this list and select the **File View** tab to see the contents of the file. Keep in mind that the formatting of these files does not matter, but be sure to remove any copyright statements or other materials that you wish to exclude from your searches.



Performing a concordance

A concordance lists the occurrences of certain words in the corpus ordered by how frequently those words are used as well as the context in which those terms appear. Select the **Concordance** tab, then enter a search term in the search box near the bottom. Click the **Start** button to generate the concordance. Clicking on the term within these results will take you the File View to see the occurrence in a broader context.

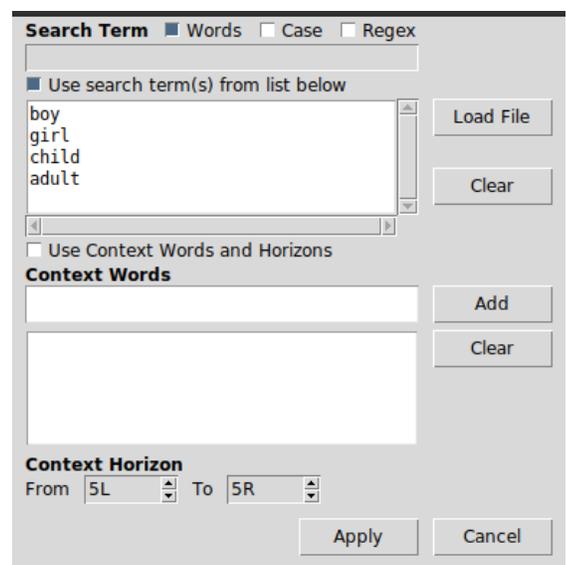
The **Concordance Plot** tab provides a graphical view of the same results giving you a sense for where in each document the term appears.

Searching for several terms

The previous search found occurrences of an exact search term. You may be interested, however, in also finding variants of that search term within the same search. For example, if your first search was the term 'children', you might also like to include the terms 'child', 'childish', or 'childhood'. You can include a **wildcard** (*) in your searches to account for many of these variations. Try searching for something like 'child*'. This type of search will require that the search term begins with 'child' but that it can end with anything else (or nothing). Note that this wildcard can also appear at the beginning of the search term or even in the middle of the search term.

You can also enter multiple search terms in the search box by separating each term with a vertical line (|), also called a *pipe*. So, for example, you could search for: boy|girl|son|daughter.

You can find a list of these special wildcard tokens by going to the **Global Settings** menu in the main menu and selecting **Wildcards**.



Another, quite powerful, way to search for multiple terms is to use the Advanced Search feature. This allows you to enter multiple terms that you would like to include in the search. Because you are manually entering the list of terms, you can include any terms you would like rather than simply variants of the same root word. Click the **Advanced** button to open a window with options for advanced searching. Enter several terms in the upper window, one on each line, and be sure to check the box labeled **Use search term(s) from list below**. Click **Apply** to close this window and the **Start** to perform the search.

Notice also that the Advanced search window allows you to also include a list of words that should appear **within the context** (the words on either side) of the occurrences. The **Context horizon** setting will allow you to vary how inclusive this context is.

Clusters and Collocates

The **Clusters/N-Grams** and **Collocates** tabs serve somewhat similar functions: they give you a sense for which words tend to occur next to or close to your search term and sort those results by frequency. Experiment with adjusting the size of the cluster and whether the search term should appear on the left or right.

A word of caution: Searches such as these are sorted by frequency and can give the appearance that results that appear at the top of the list are more frequent than they actually are. Pay attention to the column title **Freq** to see the actual number of occurrences.

Word Lists and Keyword Lists

The Word List tab does not require a search term at all. This function provides a list of the most commonly occurring terms across the entire corpus.

The Keyword List function compares keywords from two text sources (a reference text and an analysis text). This can be extremely useful in identifying terms that are characteristic or unique to a corpus.

References

https://www.mtholyoke.edu/courses/rschwartz/DigitalHistory/AntConc_readme.pdf

<https://lib.manhattan.edu/c.php?g=728252&p=5752364> (includes an excellent sample corpus of Gothic fiction works that you can use)