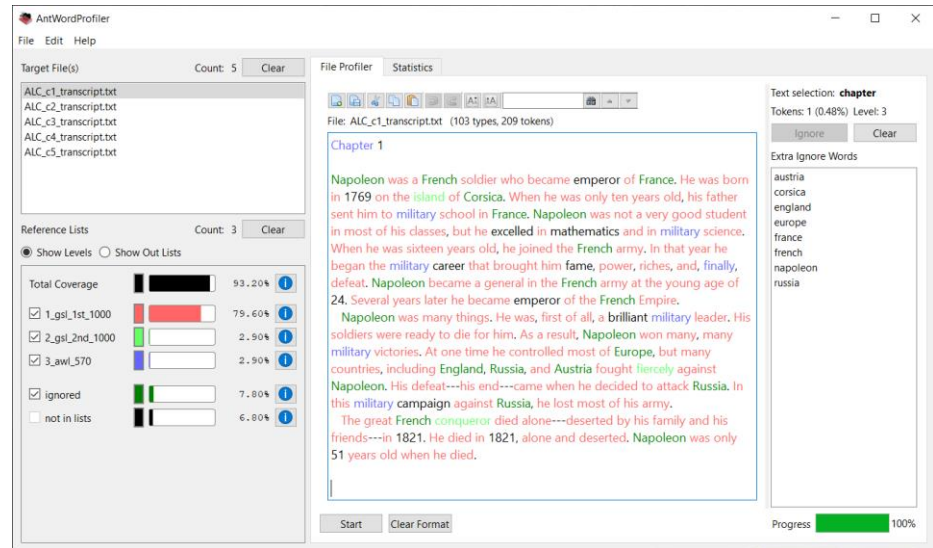# AntWordProfiler (Windows, MacOS, Linux)
Build 2.2.0

Laurence Anthony, Ph.D.
Center for English Language Education in Science and Engineering, School of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.
May 25, 2024

Introduction

AntWordProfiler is a freeware vocabulary profiling and editing tool. It can process TEXT files (.txt) (**UTF-8 encoded**), Microsoft WORD (.docx) files, and PDF (.pdf) files, generating a color-coded view of the vocabulary levels for individual files or statistics for across a set of files. Users can load in row-based or tree-based 'level lists' and expand these lists adding extra forms using a 'master list'. Users can also load in an 'ignore list' or

add ignored words dynamically while processing texts. *AntWordProfiler* runs on any computer running Microsoft Windows (tested on Windows 10), Apple MacOS (tested on OSX 10.15 Catalina), or Linux (tested on Linux Mint 17). *AntWordProfiler* is developed in *Python* and *Qt* using the *PyInstaller* compiler to generate executables for the different operating systems.

Getting Started

Windows - Installer
Double click the *AntWordProfiler.exe* file and follow the instructions to install the application into your Programs folder. You can delete the .exe file when you are finished. You can start the application via the Start Menu.

Windows - Portable
Unzip the *AntWordProfiler.zip* file into a folder of your choice. In the *AntWordProfiler* folder, double click the *AntWordProfiler.exe* file to launch the program.

Macintosh OS X
Double click the *AntWordProfiler.dmg* file to create a AntWordProfiler disk image on your desktop. Open the disk image and drag and drop the AntWordProfiler app onto the Applications folder (or into another location if you desire). You can then launch the app by double clicking on the icon in the Applications folder or the Launchpad.
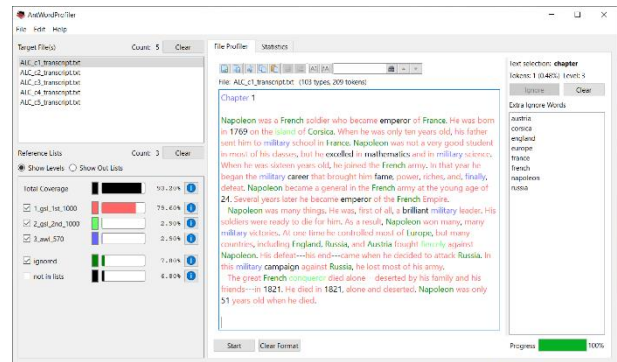
Linux
Decompress the *AntWordProfiler.tar.gz* file into a folder of your choice. In the *AntWordProfiler* folder, double click the *AntWordProfiler.sh* file to launch the software. On the command line, type ./AntWordProfiler.sh to launch the software.

Loading Target Files

Step 1: From the "File" menu, select "Open Input File(s)" or "Open Input Dir".

Step 2: Choose the files you want to process in the Dialog box. You can load TEXT files (.txt) (**UTF-8 encoded**), Microsoft WORD (.docx) files, and PDF (.pdf) files.

Loading Reference Lists

Step 1: From the "File" menu, select "Open Reference List File(s)..." or " Open Reference List Dir...".

Step 2: Choose the list you want to use in the Dialog box

Profiling and Editing Individual Files

Step 1 Double click on the file name in the Target File(s) list. The file will appear in the main "File Profiler" window.

Step 2 View the vocabulary profile in the left window frame. The profile can be switched to highlight words in the chosen text or words *not* in the chosen text. Individual reference files can be activated or deactivated to see their effect on the total vocabulary coverage.

Step 3 Click the information button (🛈) to see the words that appear in each level.

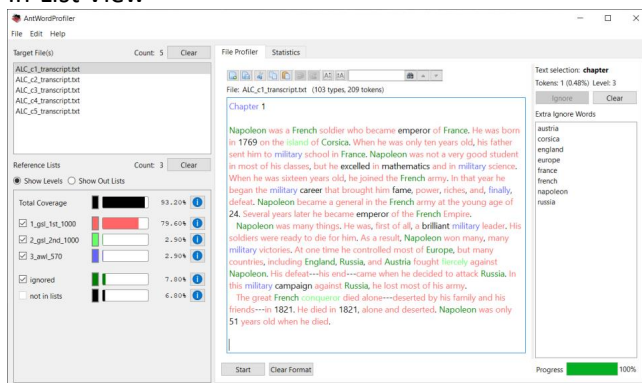**Step 4** Edit the chosen file as necessary (e.g., to adjust the vocabulary level).

Step 5 Click the Start button in the File Profiling window to update the profile coverage statistics.

Step 6 To add words in the chosen file to the ignore list, select the word and then click the "Ignore" button in the File Profiling window. Only words that are not in the reference lists can be added. Alternatively, after selecting a word, right click, and choose the action from the pop-up dialog.
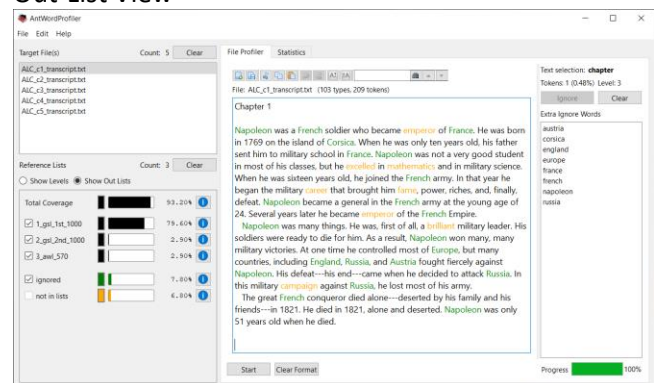
Profile Views

In-List View

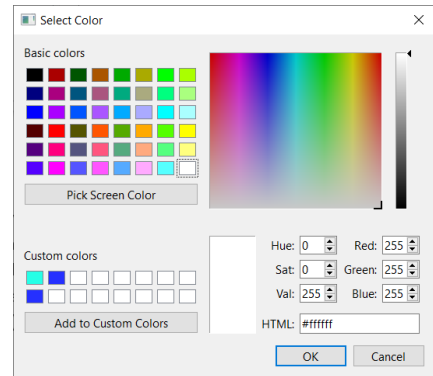Out-List View

Useful hints:

Hint 1 If you select and copy one or more words from the chosen file, you can then paste the selection directly into the "Extra Ignore Words" list. All words that are currently not in the reference lists will be automatically added. This can be useful if you want to quickly identify all "out-of-list" words in a text. Once the words are added to the "Extra Ignore Words" list, you can remove any specific words in the list by selecting them and clicking "Clear".

Hint 2 If you click on the "not in lists" information button (🛈) you can immediately see all the words in the

file that are currently not in the reference lists or ignore list (so called "out-of-list" words). Again, you can copy and paste these into the "Extra Ignore Words" list and edit the list as necessary.

Hint 3   The "File Profiler" window contains various useful text editing tools. Use these tools to load/save files, cut/copy/paste words, undo/redo text edits, increase/decrease the display font size, and search for words.

Hint 4   To change any of the profiling display colors, click on the color indicator box to the left of the main coverage meter and select an alternative color.
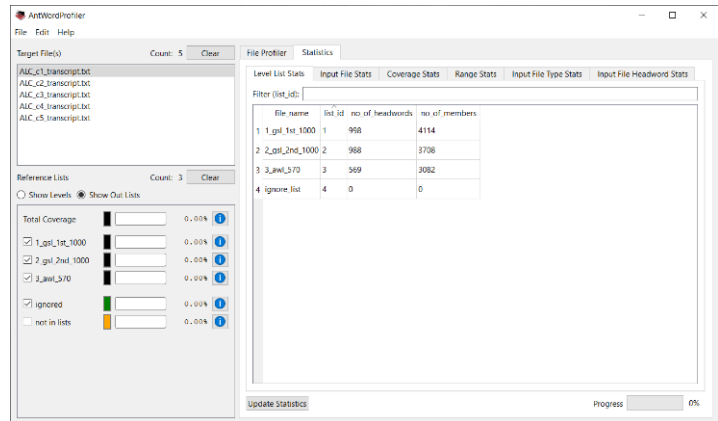
Generating Statistics for All Files

Step 1   Double click on the "Statistics" Target File(s) list and click the "Update Statistics" button at the bottom of the main window.

Step 2   Click on the sub-tabs to see the various statistics tables

Step 3   Use the "Filter" option at the top of each statistics table to filter the results.

Step 4   If you make any changes to the target files or reference lists, click the "Update Statistics" button again to refresh the results.
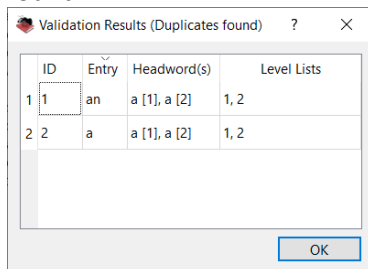
Useful hints:

Hint 1   The results in each statistics table be sorted according to the column values by clicking on the header for that column. Clicking the header again will reverse the order.

Hint 2   The "Filter" option can be used in combination with the sort options to quickly identify various language patterns (e.g., the most frequent words in a particular level).
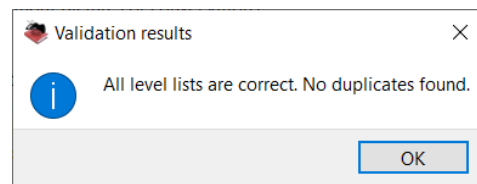
List Validation

Pre-loaded level lists can be checked for duplicates via the File menu->Validate Level Lists option. On selecting the option, one of two dialog boxes is shown depending on if duplicates are found or not. If duplicates exist, the first entry will be considered during profiling.

Duplicates found   No duplicates found

Additional Features

The following shortcuts are available as is standard on the operating system:

Windows

CTRL-A	⇨ Select All	CTRL-X	⇨ Undo
CTRL-C	⇨ Copy	CTRL-Y	⇨ Redo
CTRL-V	⇨ Paste	F5	⇨ Refresh/Update the profile
		CTRL-F	⇨ Find Next
		CTRL+SHIFT-F	⇨ Find Previous

Macintosh

Command-A	⇨ Select All	Command-X	⇨ Undo
Command-C	⇨ Copy	Command-Y	⇨ Redo
Command-V	⇨ Paste	Command-SHIFT-Z	⇨ Refresh/Update the profile
		Command-F	⇨ Find Next
		Command +SHIFT-F	⇨ Find Previous

CITING/REFERENCING *AntWordProfiler*

Use the following method to cite/reference *AntWordProfiler* according to the APA style guide:

Anthony, L. (YEAR OF RELEASE). *AntWordProfiler* (Version VERSION NUMBER) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

For example, if you download *AntWordProfiler 2.0.0*, which was released in 2022, you cite/reference it as follows: Anthony, L. (2022). *AntWordProfiler* (Version 2.0.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.antlab.sci.waseda.ac.jp/

KNOWN ISSUES

None at present.