AntWordProfiler (Windows, MacOS, Linux)

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

VERSION HISTORY

2.2.0

New features:
1) Updated the software to work with QT5, making it easier to build the app for Linux.
2) Revised the coverage viewer display to make the coverage values easier to see.

Bug features:
1) Fixed a few issues when processing user files vs the inline text editor that sometimes caused the stats page to not update correctly.
2) Fixed a problem that causes the radiobuttons in the global settings to not act independently.

2.1.0

New features:
1) The backend database has been completely rewritten to allow for much faster processing of target files.
2) The latest version of Qt (the graphical engine) and the main user interface has been redesigned to hopefully look better on MacOS.
3) *AntWordProfiler* can now run as a native app on both MacOS-Intel and MacOS-Silicon.

2.0.1

Bug features:
1) In the previous version, clicking "Start" in the "File Profiler" tool would cause all the information for all files in the "Statistics" tab to be updated before showing the profile for the selected text. This could slow the performance so now the "Statistics" page will not be updated. To update the "Statistics" tab, you will need to click the "Start" button in that tab.
2) In the previous version, clicking "Start" in the "File Profiler" tool would cause a temporary file to be created and displayed if the file profiler was not showing the contents of a currently selected file. Now, if no file contents is being shown and the "Start" button is clicked when a file is selected, the contents of that file will be shown.
3) Progress during the "File Profiler" updating is now shown in steps.
4) The PDF to text converter engine has been updated and some settings have been changed to allow it to deal with a wider variety of PDF oddities.

2.0.0

New features:
1) *AntWordProfiler* has now been completely re-written in Python. Allowing for a much-improved user experience. All existing functionality has been ported to the new version except for thesaurus-based file simplification. If this featured is desired by users, I will consider adding it to a future release.
2) The build process to compile the app for Windows has been updated. Now the application is released with an installer or as a portable application. The app is also digitally signed by Microsoft so fewer security errors will appear.
3) The build process to compile the app for Mac OS has been updated. Now the application is released as a dmg file that can be installed into a normal location (e.g., the Applications folder). The app is also digitally signed and notarized by Apple so no security errors will appear.

4) The Linux build has also been updated. Now the app can be run by clicking on the AntWordProfiler.sh file in the main folder.

1.5.1
This is a minor upgrade to fix a bug that caused the help file to not be shown correctly in the help menu.
New Features
1) Help file updated.

1.5.0
This is a minor upgrade to update the icon, and allow the app to run more smoothly Big Sur.
New Features
1) Icon updated.
2) The left and right panes of the file view profiling tool can now be resized. This is necessary for Mac.
3) The help file is now split into help, license, and version history and each file is now accessible independently in the help menu.
4) The various help files are now updated to reflect new changes

1.4.1
This is a minor upgrade addressing a refresh issue that was noted on Macintosh OS X 10.9 (Mavericks) operating systems.
Bug fixes:
1. Due to changes introduced to the graphics rendering on Macintosh OS X 10.9 (Mavericks), AntWordProfiler did not refresh correctly after certain actions making the screen look to freeze. The earlier workaround was click outside of the AntWordProfiler window and back into the window to force a refresh. This new version of AntWordProfiler fixes the problem at the code level and so no refresh issues should be seen in the future. The workaround has involved creating a standard build of the Perl programming language based on Citrus Perl (5.16.0) and the latest Tcl graphics engine. The introduction of the latest versions of these tools may introduce other very small changes in results due to the use of the latest Unicode standard, but these are unlikely to affect normal users.

1.4.0
This is a major update that adds a few new features.
New Features
5) Sorting of results can now be made at two independent levels chosen by the user from the choices of (frequency, range, or word).
6) Users can now opt to use their own token definition instead of the default one.
7) The help menu option now launches this PDF file instead of a basic text file.
8) Warnings are now issued before closing the file viewer window (this might become annoying!)
Changes
1) Some of the widgets in the interface have been renamed to make their actions easier to understand.
2) The ordering of global settings categories is now alphabetical.
3) Highlight colors have now been adjusted to be more pleasant on the eye.
4) The regular expression associated with the default token definition is now shown.
5) The thesaurus option is now turned OFF by default.
6) This readme file is now formatted as a PDF file.
Bug fixes
1) The Stop button has been removed from the main window because it never 'stopped' the program properly before. To force close the program, just click on the button at the top of the window.
2) The progress bar now in the main window now updates properly during a long operation.

1.3.1
This is a minor update addressing one small bug and adding one or two new features
New Features
1) Added an option in the thesaurus settings menu to allow users to select to automatically update the file view after editing it with the thesaurus tool.
2) Massively improved (>10 times) the speed of the file viewer tool when highlighting texts.
Bug fixes
1) Corrected a bug that caused the local frequency count in the file viewer to report values for characters as well as words.
2) Tidied the interface a little so large frequency values do not become hidden in the file viewer tool
3) Fixed a bug that caused the program to hang if the Profiler Tool start button was clicked when no files or lists were imported.

1.3.0
This is a major update adding new features and correcting various bugs in the 1.2x versions.
New Features
1) Revised the names and cleaned contents of the internal baseword lists (approved by Paul Nation)
2) Added an option to view a complete frequency list of the target files in the profile
3) Added an option to view words not in the target files in the profile
4) Added new sort options (word, range, freq)
5) Added a batch process option (main window only)
6) Revised this readme file
7) Added a new thesaurus viewer (activated via a menu option)
8) Revised the profile output to include cumulative frequencies
9) Improved the User File Viewer tool to display all level lists in a scrolled pane.
10) Added new editing functions (Previous, Next) to the User File Viewer tool
11) Added a new edit function to simplify the target text via a level-marked thesaurus.
Bug fixes
1) Corrected an unintuitive feature that duplicated the listing of internal baseword lists after importing a settings file.
2) Corrected a bug that caused the windows to not be centered when first displayed.
3) Corrected a bug that prevented more than eight baseword lists to function correctly.

1.2.1
This is a minor update to correct a few bugs and minor issues in the 1.200 version.
New Features
1) Added a new icon to the software.
Bug fixes
1) Fixed an issue that caused the global settings options to not appear on Macintosh OS X.
Minor issues
1) The icon for AntWordProfiler has been update to a graphic that matches other *AntLab* tools. I hope you like it!
2) The height of widgets has now been fixed to make them fluid and thus allow them to take their full height on each operating system.

1.200
This version includes many changes and so deserves its 1.2 status. A list of new features and changes are below.
New Features

1) Redesigned the software engine so that the program now works smoothly with Unicode (UTF-8) encoded files. This means the program will work with any language (including Asian languages) provided that the target files, and baseword lists are saved in the UTF-8 encoding.
2) Redesigned the settings files to be easier to edit and work with (for creating user generated settings files). The settings file is no longer an XML file. Rather, it is a special file with a very simple format.
3) The main window title now shows if the program is using a user-settings file or not.
4) Any errors do not show appear on the main window anymore. This makes debugging a little more difficult, but improves the user experience.

1.104
Bug fixes
1) Fixed code to correctly update the level list word database when analyzing new texts. Previous versions would produce strange results if the User File view was used without generating a main set of results first.

1.103
New Features
1) Recoded to allow Macintosh OSX version to be compiled easily.
Bug fixes
None

1.102
Bug fixes
1) Deleted two file menu options that were redundant (Choose Main Word List, Choose User Word List)
2) Added widget names to interface that were included in the settings file but not reflected.
3) Repositioned widgets to allow non-English names to be used successfully.
4) Revised the internal level lists to exactly match those in Paul Nation's original Range program (not Range BNC).The previous versions of AntWordProfiler used the first three Range BNC lists. Note that either set of lists can be imported and used as is. Note also that the placement of some entries in the Range lists is questionable (see Bug fix 1 in version 1.101).

1.101
Bug fixes
1) Corrected the placement of "M" in the Nation Baseword List 1. Originally this had its own entry, but now it is correctly placed as a family member of "BE"
2) Corrected the trimming of non-token characters from the beginning of file lines. These caused the not-in-list results to include a stray entry count.
3) Updated and redefined the settings of the compiler for AntWordProfiler. This may resolve the problems reported by some users in Asia who cannot launch the program.
4) Added XML header to user settings file

1.100
New Features
1) Introduced a database back engine, allowing much larger files to be processed.
2) Hugely improved the speed of processing. Now, 7 million words can be processed in around 1 minute.
3) Introduced a tag ignore feature allowing tagged files to be processed without change.
4) Greatly simplified the program (and interface) by adopting Paul Nation's Range format for level list files.
5) Importing and exporting of the settings is now possible via an XML file.
6) Simplified the save format options. The Excel save format may be reintroduced later.
7) Changed the name of the program to more accurately represent what it does.
8) Re-wrote this readme file to reflect the new changes.

Bug fixes
1) Several mistakes in the language of the interface have been corrected.
2) Errors reports that occur when pressing Cancel in some dialog boxes have been fixed.

1.01
Bug fixes
1) Corrected how to read in main word lists and level lists so that blank lines and lines starting with non-token characters are ignored. In version 1.00, blank lines would generate very strange results.
2) Added character encodings for Chinese, Taiwanese, and Korean (Japanese encodings were already included in version 1.00)

1.00 This is the first version of the program.