

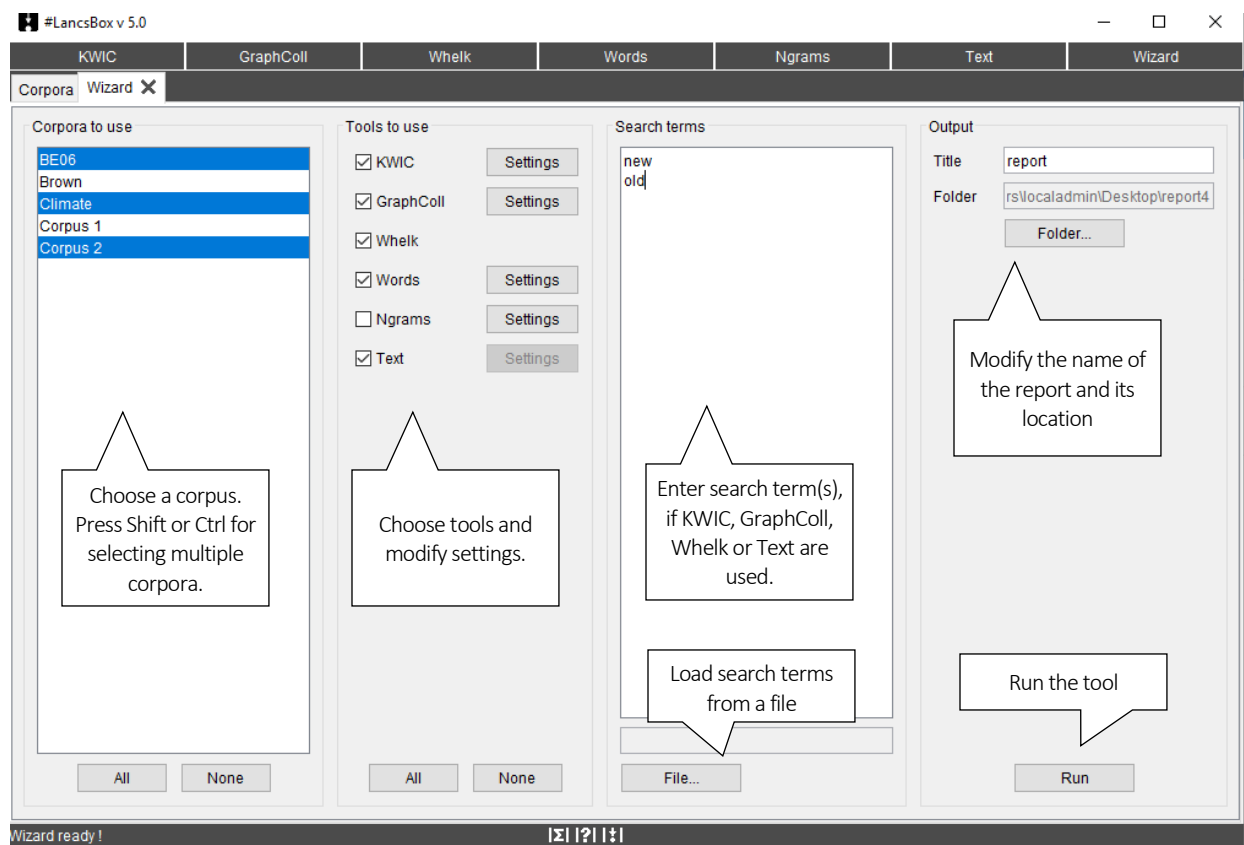
## 10 Wizard

The Wizard tool combines the power of all tools in #LancsBox, searches corpora and produces research reports for print (docx) and web (html).

It can be used, for example, to:

- Carry out simple or complex research.
- Produce a draft report.
- Download all relevant data.

### 10.1 Visual summary



## 10.2 Selecting settings and running Wizard

---

Wizard produces research reports automatically. All you need to do is select the corpus/corpora and procedures to use.

---





1. Select corpora in the 'Corpora to use' panel (left). For multiple adjacent corpora, hold Shift on the keyboard while selecting; for multiple corpora that are not next to each other, hold Ctrl (or Control on mac).
2. If the corpus you want does not appear on the list, go to 'Corpora' tab and add it.
3. Choose the tools you want to employ. The choice of the tools depends on the type of analysis you want to perform.
  - KWIC: analysis of concordance lines.
  - GraphColl: collocation analysis.
  - Whelk: analysis of frequencies in individual texts.
  - Words: analysis of frequencies and dispersions of individual lexical items and keyword analysis.
  - Ngrams: analysis of frequencies and dispersions of ngrams.
  - Text: analysis of broader contexts.
4. Adjust the tool settings by clicking on the 'Settings' button next to the tool.
5. Enter search term(s), if, KWIC, GraphColl, Whelk or Text are used. Alternatively, load search terms from a text (txt) file.
6. Choose the location where the report and extracted data will be saved or leave default (Desktop).
7. Press 'Run'

## 10.3 Data analysis

---

Wizard produces the data analysis in the background and informs the user about the progress. The complete data set is saved together with a report.

---

1. The data set includes the following folders.
  -  csv
  -  images
  -  tsv
  -  xml
2. csv (comma separated files; open in Excel or Calc) include data for Lancaster Stats Tools online.
3. images (png; open in a standard graphics app) include graphs and other graphical output.
4. tsv (tab separated files; open in Excel or Calc) include complete data output from the individual tools.
5. xml (extensible markup language; opens in a text editor) includes a complete Wizard data set, which can be used for further computational processing.

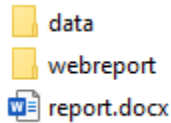
## 10.4 Research report

---

Wizard produces a structured data report in two formats: docx and html

---

1. The .docx report can be easily edited in Word, Writer or a similar word processor.



2. The length and contents of the report depend on the number of corpora and tools that were selected.
3. The report follows the structure of an academic research report.

Created by #LancsBox Wizard

June 9, 2020 - 22:07

# Comparison of British and American English

---

## 1 Introduction

This research report was automatically produced by #LancsBox (Brezina et al. 2020), a corpus analysis tool developed at Lancaster University. It uses cutting-edge technology and statistical sophistication (Brezina 2018) to analyze and visualize corpus data. For more information and tips on research report writing see the [Research Report Guide](#).

## 2 Method

### 2.1 Data

The study analyzed the following corpora:

Table 1. Corpora used

Name	Language	Texts	Tokens	Additional information
Brown	English	15	1,014,361	Types: 49,686 Lemmas: 44,622
L-O-B	English	15	1,007,677	Types: 48,349 Lemmas: 43,920

In the study, 2 corpora were used of the total size of 2,022,038 running words (tokens) in 30 texts. A full description of the corpora is available in [data\tsv\corpora](#).

### 2.2 Procedure

#LancsBox (Brezina et al. 2020) software package was employed to analyse the data. The following tool from the package was used: KWIC. The KWIC tool generates a list of all instances of a search term in a corpus in the form of a concordance. The following search terms were used: "new", "old" and "some".

## 3 Results