

## Historical Thesaurus of English

The database can be used either to look at words in particular areas of meaning or to collect words from particular periods to use in searching texts.

Log on: <http://libra.englant.arts.gla.ac.uk/historicalthesaurus/>

1. Click on **Sections Available** to see the range of material on offer. You can search particular sections or all of them.

2. Go to **Search Menu**. Click on **Browse**. Enter 'popish' as your search word. The return shows that 'popish' occurs in two categories, **Roman Catholicism** and **Pope**.

Click on **Roman Catholicism** and scroll down to find 'popish' highlighted in red. Look along the columns for further information. The first date shows that the word was first recorded in 1528, while the dash on the right of the screen shows that it continues in use. 'de' is a style label meaning 'derogatory'.

Use the back button to return to your browse headings and select **Pope**. Usage of the word in the meaning of 'papal' is recorded only between 1540 and 1567.

By skimming through the categories, you can get an idea of how words in these areas peaked at certain (predictable) historical periods.

### Currently available sections:

Anger  
Art  
Atmosphere (weather)  
Cinema  
Clothing  
Dancing  
Emotion (general)  
Excitement  
Gratitude  
Health and disease  
Healing/cure  
Humility  
Jealousy  
Kinship  
Mental pain/suffering  
Music  
Navigation  
Performing arts  
Photography  
Pity  
Pleasure/enjoyment  
Pride  
Religion  
War – Armed hostility

3. Return to the **Search Menu** and choose **Label**, then option C, **Chronological Order**. Enter **Religion** as the section and **ct** as the two letter code for 'contemptuous'. This will return 47 words applied in this way. A search of **de** 'derogatory' will produce a further 26 words.

4. Return to the **Search Menu** and choose **Dates**, then **Current in Period**, section **Religion**, **between 1500 and 1600**. This will give you an idea of the range of words available (or not) to someone writing about religion at this time. Note that these words can have first or last dates outside the period but still be current within it.

## **WordSmith Tools**

Developed by Mike Scott.

**Version 3:** <http://www.lexically.net/wordsmith/>

**Version 4:** Free download of software in evaluation mode; purchase registration code to unlock and use all features - <http://www.oup.com/elt/catalogue/isbn/6890?cc=global>

**Version 5:** in beta testing - <http://www.lexically.net/wordsmith/version5/index.html>

### **A. WordList**

WordList creates lists of all the word-forms in a text or corpus, and displays these alphabetically and by frequency, as well as producing basic statistics.

1. From the **Text Retrieval Programs** folder on the desktop, open WordSmith Tools, and select **WordList** from the **Tools** menu.
2. From **Settings**, choose **Choose Texts**. Locate the text or corpus from a drive. A small corpus created for this session can be found at Q:English/HistCorp, and details of its contents found on the last page of this handout. Within this folder, you will find two sub-folders called ‘Calvin’ and ‘Luther’. Select one folder, click ‘All’ and click ‘Store’, then select the other, click ‘All’ and click ‘Store’. The dialogue box should show that there are 8 files selected, in two directories. Click ‘OK’.
3. From the **File** menu, choose **Start**. Then click to “Make a WordList now”.
4. WordList will produce three windows:
  - new wordlist ordered by Frequency (F);
  - new wordlist ordered Alphabetically (A);
  - new wordlist Statistics (S). Among other information, the Statistics window will tell you how many tokens (total number of word-forms) and types (number of different word-forms) the corpus contains.
5. These lists may give you ideas of words to investigate in more detail. Making a WordList of an individual text will give you an impression of what the text is about.

---

### **B. Concord**

Concord is a concordancer, and presents all of the instances of a chosen word in their immediate linguistic context.


1. From the **Tools** menu of the WordSmith Tools Controller, select **Concord**.
2. From **Settings**, choose **Choose Texts**. WordSmith should have remembered the texts previously chosen. If not, reselect the corpus from the drive.

3. From **Settings**, choose **Search Word**. Type a word into the dialogue box which appears and then click “Go Now!”.


- You might find that the following are all interesting words in this corpus: predestination, providence / providence, Christ, presence, faith, scripture, papacy, marriage, iustification / justification. (Bear in mind that spellings may not be modern-day ones.)


\* If you want to find all inflected forms of a word or some variant spellings, try using the wildcard \* which replaces any number of characters, including zero. So \*stific\* will retrieve the variants ‘justification’, ‘iustification’ and ‘iustificacion’ (as well as one instance of ‘testificacion’).


4. Concord will retrieve all of the instances of the search word. The concordance can be re-

ordered using the **Re-sort** button:  Try reordering the concordance lines according to the word immediately to the left (1L), and then the right (1R), of the search word. This should highlight patterns of ‘collocation’ of the word, that is, its common lexical environment.

5. Complex re-sorting is possible: try ordering the concordance lines according to a MAIN SORT according to filename (which will group all the Luther and all the Calvin texts together) and a secondary sort according to the word immediately to the left or right of the search word. Do any differences of collocation stand out between the occurrences of your word in Luther and in Calvin?

6. Repeated collocates (lexical neighbours) of the search word can be computed by clicking on the **Collocate** button: 

7. (Close the Collocates window to return to the main concordance.) Repeated sequences of words from these concordance lines can be calculated using the **Cluster** button: 

8. (Close the Clusters window to return to the main concordance.) A visualisation of where in the text the occurrences of the word appear can be obtained with the **Plot** button: 

---

### **Some questions**

- What are the common left collocates of ‘presence’?

- Does ‘providence’ (NB spelling) occur proportionately more frequently in texts by Luther or by Calvin? (Remember that the Calvin subcorpus is roughly 2 ½ times larger.)

- Does either author refer to the other in his writing? Can you make any generalisations about the context in which reference is made?

- Are the occurrences of ‘Chirch’ (note spelling) in Calvin evenly spread through the four texts? (Hint: use the Plot function) How might this be explained?

- Can you see any applications of concordances for your own teaching or research?

### **C. KeyWords**

The KeyWords tool compares two previously created WordLists to determine which word-forms are unusually frequent in one compared to the other.

1. WordList files of Calvin and Luther respectively have been created and saved on the Q: drive.
  2. From the **Tools** menu of the WordSmith Tools Controller, select **KeyWords**.
  3. From the KeyWords **File** menu, select **Start**. In the “Choose Word Lists” dialogue box which appears, choose the Luther list in the left-hand box, and the Calvin list in the right-hand one. Click **OK**.
  4. From left to right, the KeyWord list returned shows:
    - the words in order of ‘keyness’;
    - the frequency of that word in the Luther corpus;
    - the percentage of occurrences of the word in the Luther subcorpus, compared with the corpus as a whole;
    - the frequency of the word in the Calvin corpus;
    - the percentage of occurrences of the word in the Calvin subcorpus, compared with the corpus as a whole;
    - an indication of ‘keyness’. Mike Scott (in the Help menu) defines a word which is positively key as one which “occurs more often than would be expected by chance in comparison with the reference corpus”.
  5. The three main tools in WordSmith are interconnected. Select a word in a KeyWord list, and then click on the large blue C from the buttons to be taken to a concordance of that word in the subcorpus.
-

### **Text Mining for Historians - WordSmith Tools session**

The following gives details of the small corpus put together for this session. All texts are from Early English Books Online - <http://eebo.chadwyck.com/home>.

#### **LUTHER – total 290,542 words**

##### **Luther 1 (116,434 words)**

**Author:** Luther, Martin, 1483-1546. **Date:** 1581 **Title:** A commentarie or exposition vpon the two Epistles generall of Sainct Peter, and that of Sainct Jude. First faithfullie gathered out of the lectures and preachings of that worthie instrumente in Goddes Church, Doctour Martine Luther. And now out of Latine, for the singuler benefite and comfort of the godlie, familiarlie translated into Englishe by Thomas Newton

##### **Luther 2 (138,373 words)**

**Author:** Luther, Martin, 1483-1546. **Date:** 1577 **Title:** A commentarie vpon the fiftene Psalmes, called Psalmi graduum, that is, Psalmes of degrees faithfully copied out of the lectures of D. Martin Luther ; very frutefull and comfortable for all Christian afflicted consciences to reade ; translated out of Latine into Englishe by Henry Bull.

##### **Luther 3 (20,330 words)**

**Author:** Luther, Martin, 1483-1546. **Date:** 1554 **Title:** A faithful admonition of a certeyne true pastor and prophete sent vnto the Germanes at such a time as certain great princes went about to bryng alienes into Germany, [and] to restore the papacy; the kingdom of Antichrist. [Et]c Now tra[n]slated into English for a like admonicion vnto all true Englyshe harts; wherby they may learn and know how to consider [and] receiue the procedinges of the English magistrates and bishops. With a preface of M. Philip Melancthon. ...

##### **Luther 4 (15,406 words)**

**Author:** Luther, Martin, 1483-1546. **Date:** 1570 **Title:** A very comfortable and necessary sermon in these our dayes made by the right reuerend father and faithfull seruaunt of Iesus Christ Martin Luther ; concerning the comming of our Sauior Christ to Iudgement and the signes that go before the Last Day, which sermon is an exposition of the Gospell appointed to be red in the church on the second Sunday in Aduent ; and is now newly translated out of Latin into English and something augmented and enlarged by the translator with certaine notes in the margent.

#### **CALVIN – total 754,568 words**

##### **Calvin 1 (28,846 words)**

**Author:** Calvin, Jean, 1509-1564. **Date:** 1548 **Title:** A faythfull and moost godlye treatyse concernyng the most sacred Sacrament of the blessed body and bloude of our sauioure Chryst, co[m]piled by Iohn Caluyn, a man of no lesse lernyng and lytterature then godlye studye, and example of liuyng. And translated into Latin by Lacijs a man of lyke excellencie. And nowe last of all, translated into Englyshe by a faythfull brother, no lesse desirous to profyt the weake brotheres then to exercise the talent of the Lorde to his honoure and glorye. In declaration whereof, he hath set before this lytle booke an epistle to the reader much more effectuous then in the fyrst edicion. Whereunto the order that the church and congregation of Christ in Denmarke doth vse at the receyuyng of baptisme, the Supper of ye Lorde, and wedlocke: is added. Myles Couerdale

##### **Calvin 2 (658,194 words)**

**Author:** Calvin, Jean, 1509-1564. **Date:** 1561 **Title:** The institution of Christian religion, vvrytten in Latine by maister Ihon Caluin, and translated into Englysh according to the authors last edition. Seen and allowed according to the order appointed in the Quenes maiesties iniunctions

##### **Calvin 3 (46,277 words)**

**Author:** Calvin, Jean, 1509-1564. **Date:** 1567 **Title:** A little booke of Iohn Caluines concernyng offences whereby at this daye diuers are feared, and many also quight withdrawen from the pure doctrine of the Gospell, a woorke very needefull and profitable. Translated out of Latine into Englishe by Arthur Goldinge.

##### **Calvin 4 (21,251 words)**

**Author:** Calvin, Jean, 1509-1564. **Date:** 1549 **Title:** Of the life or conuersation of a Christen man, a right godly treatise, wrytten [sic] in the latin tonge, by maister Iohn Caluyn, a man of ryghte excellent learnyng and of no lesse godly conuersation. Translated into English by Thomas Broke Esquier Paymaister of Douer.