BNCweb – a custom-made interface to the BNC
Ylva Berglund Prytz & Martin Wynne
Version: January 2015

BNCweb is an advanced and flexible online interface to the British National Corpus (BNC). The BNC is a snapshot of the English language in Britain in the late 20th century, with examples of speech and writing of a wide variety of types. More information about the BNC can be found at http://www.natcorp.ox.ac.uk/. This document contains some exercises to help you get started to use BNCweb to explore the usage of words in English.

Task 1. Who uses the word borrow and in what kind of constructions and contexts?

1.1 Getting started

What does your intuition say about the use of borrow? We're going to use the corpus evidence to explore how the word is used. But before logging in, think about the following questions, and write down some of your expectations. We can then compare your intuition to what we find in the corpus.

What is borrowed and by whom?

What words do you expect to find together with borrow?
Can these words be grouped in some way, for example based on their word class, function, or meaning?
Where would you expect these words (before or after borrow? Immediately adjacent or not?)
Who do you think uses the work borrow? In what context or type of language would you find borrow?
Are there any words that are NOT used with borrow?

Once you have thought about these questions, and maybe discussed them with others, start to follow the tasks below.

1. In a web browser, go to http://ota.oerc.ox.ac.uk/ and follow the link to the BNCweb service
2. Log in using your University of Oxford single-sign on credentials
3. In the Standard Query box, type in the word borrow.
4. Click the ‘Start Query’ button to run the search (see the image below)
5. Explore the concordance lines which are now displayed on your screen. Here are some tips on things that you can try.

- Take a look at the summary information at the top of the browser page, and try to answer the following questions:
  
  How many occurrences of borrow are there in the corpus?
  What is the relative frequency (hits per million words)?
  What other information do you get at first glance?

- Can you get a feel for how borrow is used? What words do you find before or after borrow? (Click on the high-lighted borrow to see more context)

- Is the word used in particular types of texts or contexts? (Hover over the 'Filename' to see information about the text).

- Click on the 'Show KWIC view' button to get an alternative visualization of the fragments of text. This might help to spot patterns of words around borrow.

- Try the 'Show in random order' button as well. Does this help?

- Place your mouse cursor over the search work on one concordance line. What information do you get? (This works on JavaScript enabled browsers only.)

So far it is difficult to get an overview of the results, since you have only looked at one page at a time. Let’s take a look at some ways of re-ordering the results to it easier to spot patterns of words, and some ways of extracting some information about frequencies and distributions.

1.2 Sorting

The BNCweb interface allows sorting on one position, but has some other sorting options which you do not see in all tools.

Above the concordance lines is a box which says [New Query]. Click in it and select [Sort] from the list that pops up. Click [Go!]

- Explore the sort options.
- What does the ‘Position’ option do?
- Why would you use ‘Starting with letter’?
- How many of your examples are immediately preceded by a verb? (Hint: Select ‘1 Left’ and Tag restrictions ‘any verb’)
• Look at ‘Spoken or Written’. What are the most important and significant figures here which would allow you to make a statement about the distribution of *borrow* in speech and writing?
• Look at ‘Derived text type’ which should give a better breakdown of types of speech and writing events – in which category is the work most frequent?
• At the top of the page, select ‘Speaker: Age’ in the first box. Click ‘Show distribution’.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Speaker: Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categories (for crosstabs only):</td>
<td>no crosstabs</td>
</tr>
</tbody>
</table>

Note that this category only applies to the spoken texts, as you might expect. Would you say that the usage of *borrow* is strongly age-related?

• What about differences between men and women (choose ‘Speaker: Sex’ in the top box)?
• You can also cross-tabulate two categories, for example age and sex.

Explore the results. Does any group use the work more than others? What happens if you look at social class?

<table>
<thead>
<tr>
<th>Categories</th>
<th>Speaker: Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categories (for crosstabs only):</td>
<td>Speaker: Sex</td>
</tr>
</tbody>
</table>

**Further hints and tips on distributions**

The ‘File-frequency extremes’ function will identify the file which has the highest proportion of a search term. Where is *borrow* used most?

(‘Speaker-frequency extremes’ identifies the speaker who uses a word/phrase most)

**When you want to make another search, click ‘Go’ next to ‘New Query’.

You can view previous queries under ‘Query history’ on the start page (left-hand margin)

Use the ‘User settings’ to change the default display options.