The British National Corpus and BNCweb

The British National Corpus (BNC) ([http://www.natcorp.ox.ac.uk/](http://www.natcorp.ox.ac.uk/)) is an electronic collection of a 100 million words of written (90%) and spoken (10%) British English from the early 1990s.

**Written data**
90 million words, including extracts from newspapers, academic books, popular fiction, letters and university essays

**Spoken data**
10 million words, composed of demographic data and context governed data

- *The demographic part*
  transcripts of about 900 everyday unscripted spoken conversations

- *The context-governed part*
  spoken language collected in public contexts – e.g. radio phone-ins, government meetings, classroom interactions

You can explore the BNC in a variety of ways. We are going to use BNCweb, which is based at Lancaster University.

---

1. **How to get a BNCweb username**

Start up an internet browser, and go to [http://bncweb.lancs.ac.uk/bncwebSignup](http://bncweb.lancs.ac.uk/bncwebSignup).

Click on the link that says “register for an account”.

Enter your details into the form.

Choose a username, and **use a password that is different from other passwords you use** for, say, your email. This is because BNCweb passwords are not secure. (Hint: it’s a good idea to write it down.)

Once you have submitted the form, open your email. You will get a mail containing a link to confirm your registration. Click on this – it will take you to BNCweb.

**From now on**, you can just start BNCweb by going to [http://bncweb.lancs.ac.uk](http://bncweb.lancs.ac.uk) and entering your username and password.

---

2. **Using BNCweb to investigate language use.**

We are going to use the BNC to investigate how certain words are used. This section provides instructions for using BNCweb.
2.1 What does literally mean?

The Concise Oxford Dictionary defines literally as an adverb deriving from the noun literal, which means ‘taking words in their usual or primary sense, without metaphor or allegory’.

Is this really how people use the word literally? Investigate this by carrying out the following search.

- From the BNCweb start page, type in the word/phrase you wish to investigate and then click on the ‘Start Query’ button.

BNCweb will display the results of your query in the form of a concordance, similar to that shown below:

- Click on the KWIC view button. What happens when you do this?
- If you want to go back to the previous screen, click the browser back button.
**TASK:** Examine the concordance lines carefully and work out what people mean when they use the word *literally*.

- When you have finished, click on the ‘Go!’ button, next to New Query.

<table>
<thead>
<tr>
<th>No</th>
<th>Filename</th>
<th>Hits 1 to 50</th>
<th>Page 1 / 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A01 181</td>
<td>The food is ‘miserable’ cold coffee in the morining, with half a piece of bread, lunch a thin soup of rotten vegetables, the evening meal boiled pasta and the same stinking soup.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>A01 229</td>
<td>Purple and blue, the hand shadows of the hollow breakers are cast upon the mist of night, which gathers cold and low, advancing like the shadow of death upon the guilty</td>
<td></td>
</tr>
</tbody>
</table>

2.2. **Are thin and slim synonyms?**

Let’s compare the synonyms thin and slim.

- From the BNCweb start page, type in *thin* and then click on the ‘Start Query’ button.
- Click on the drop-down menu that contains *New Query*, and select *Collocations*.
- Don’t forget to click on *Go!*

On the next page, just click *submit*.

The next screen shows all the collocates of *thin* that occur in a +5/-5 window. You will notice that there is a mixture of grammatical and lexical words.

We’re going to change the search parameters so that we only see which nouns occur immediately to the right of *thin*. 
Change the **collocation window span** to **1 Right** and **1 Right**.  
Change **and/or tag** to **any noun**  
Change **Statistics** to **Rank by frequency**  
Don’t forget to click on **Go!**

You should now see the output from your search. In the left-hand column you will see all those singular nouns that occur one place to the right of *thin*.

You can save these results by clicking on the drop down menu at the top right of your screen, and select **Download results**. (Don’t forget to click on **Go!**). The results will be downloaded to a plain text file, which you can open in **Notepad** (other text editors are available).

You can scroll through the results and **click on the word frequencies** if you want to see the context in which they occur.

**Now, repeat the whole process for** *slim*

Start by selecting **New query** and clicking **Go!**

**Questions:** What patterns can you see? When do we tend to use the word *thin* and when do we tend to use the word *slim*? Are the two words really synonyms? Do the results suggest any further hypotheses about the usage of these words, and if so, could you investigate these hypotheses using a corpus

Once you have answered these questions, repeat the process but this time, compare the words *slim* and *slender.*