The British National Corpus and BNCweb

The British National Corpus (BNC) (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

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 and entering your username and password.

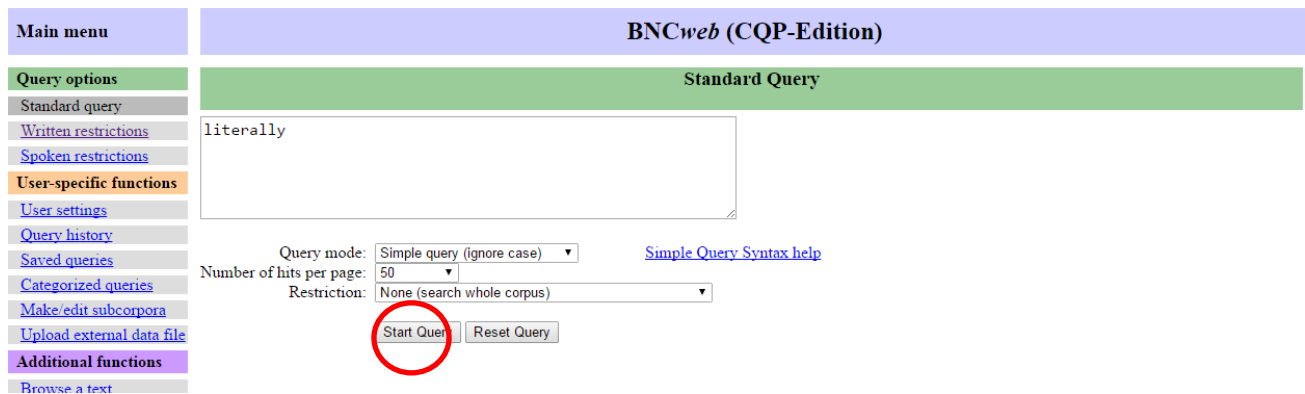## 2.   Using BNCweb to investigate language use.

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.

Main menu	**BNC*web* (CQP-Edition)**
Query options	**Standard Query**
Standard query	literally
Written restrictions	
Spoken restrictions	
User-specific functions	
User settings	
Query history	Query mode: Simple query (ignore case) ▼ Simple Query Syntax help
Saved queries	Number of hits per page: 50 ▼
Categorized queries	Restriction: None (search whole corpus) ▼
Make/edit subcorpora	
Upload external data file	Start Query Reset Query
Additional functions	
Browse a text	

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

Your query "literally" returned 1931 hits in 1080 different texts (98,313,429 words [4,048 texts]; frequency: 19.64 instances per million words) [0.361 seconds]

|◁ ◁◁ ▷▷ ▷| Show Page: 1   Show KWIC View   Show in random order   New Query ▼ Go!

No	Filename	Hits 1 to 50 Page 1 / 39
1	A06 2323	Absolute total recall — it was Portia's famous 'The quality of mercy is not strained' from The Merchant of Venice — and mercy was what I was asking for, **literally**, at that moment.
2	A08 2494	**Literally**, of course, not metaphorically.
3	A0A 24	Yeast is made up of millions of tiny fungus cells which **literally** go berserk when confronted by a liquid rich with sugars.
4	A0F 3371	Compassion's an interesting word — when you break it down it **literally** means 'to suffer with'.
5	A0G 1774	And those flowers are **literally** buzzing with honey-making, pollinating bees and greenfly-eating hoverflies.
6	A0G 2565	Greenfly can **literally** suck a plant dry
7	A0H 1348	But it is the glider pilot who **literally** has the life of the tow pilot in his hands.
8	A0H 1513	It was obvious from this that in going into storms we were quite **literally** playing with fire.
9	A0L 15	She'd taken him **literally**, although now she questioned whether that had been wise.
10	A0L 920	**Literally**,' said Lucy, sparkling into the conversation.
11	A0M 1004	However, the kick enjoys a measure of success because it can **literally** curl around an otherwise effective guard and thump into the back of the opponent's head.
12	A0P 479	But it should not be taken too **literally**.
13	A0P 1487	It is placing an artificial, not to say architectonic, view on the whole work so to argue; it is to take his 'mythology' too seriously, too **literally**; to fail to see that the 'excess' was part of the message.
14	A0T 632	I have tested **literally** hundreds of children on conservation tasks over a period of twenty years and have confirmed (to my own satisfaction at least) that fundamentally — in terms of the central systems — one does not succeed in drawing their attention to non-salient features.
15	A0T 1047	What I now want to suggest is that there is a lot of very successful science currently being conducted which **literally** depends on functionalism being true.
16	A0T 1147	It seems to me to be perverse in the extreme to attempt to resist the claim that the brain is **literally** carrying out these functions.
17	A0T 1478	It is therefore especially ironical that the neurophysiological CTP, which locates the basis of perception **literally** in the head, at a particular place away from the perceived object, actually sharpens the mystery of intentionality.
18	A0U 17	I had **literally** run away — from Englishness!
19	A0X 1516	I can only volunteer that this is done by designers who have an imperfect understanding of the nature and properties of wood, or who are **literally** prepared to risk it.
20	A11 1518	Their 1st runs at the end of 1981 were witnessed by scenes not seen since the end of steam when **literally** thousands either embarked on farewell trips or crowded platform ends for a last glimpse of the 3,300hp machines that were as popular as the LNER Gresley steam Pacifics they replaced.

➢ 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.

Your query "thin" returned 4959 hits in 1441 different texts (98,313,429 words [4,048 texts]; frequency: 50.44 instances per million words) [0.083 seconds - retrieved from cache]

| I< | << | >> | >I | Show Page: | 1 | | Show KWIC View | | Show in random order | | New Query | ▼ | Go! |

No	Filename	Hits 1 to 50 Page 1 / 100
1	A03 885	The food is 'maigre'; cold coffee in the morning, with half a piece of bread, lunch a **thin** soup of rotten vegetables, the evening meal boiled pasta and the same stinking soup.
2	A04 229	Purple and blue, the lurid 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

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**.

our query "thin" returned 4959 hits in 1441 different texts (98,313,429 words [4,048 texts]; frequency: 50.44 instances per million words) [0.083 seconds - retrieved from cache]

| < | << | >> | >I | Show Page: | 1 | | Show KWIC View | | Show in random order | | New Query | | Go! |

No	Filename	Hits 1 to 50 Page 1 / 100
		New Query
		Thin...
1	A03 885	The food is 'maigre'; cold coffee in the morning, with half a piece of bread, lunch a **thin** soup of r~ Frequency breakdown ng meal boiled pasta and the same stinking soup. Distribution
2	A04 229	Purple and blue, the lurid shadows of the hollow breakers are cast upon the mist of night, which ga Sort ing like the shadow of death upon the guilty ship as it labours amid the lightening of the sea, its **thin** mast written upon the sky, in lines of bloo Collocations Download...
3	A04 428	Gombrich's scholarly work includes many papers on meaning and interpretation in the visual arts, Categorise hits... cope of his story of art is instructive; by writing a survey he inevitably limited himself to a narrow range of comment. Save current set of hits...
4	A04 741	There are returns to the impasto, but **thin** painting certainly predominates towards the end of the century.
5	A04 1037	Writing about a small bust on a column, a **thin** sculpture in a series of four busts of which the others were fuller in form, she comments that 'the extremely thin proportions of the head contrast with the solidity of the base.
6	A04 1037	Writing about a small bust on a column, a thin sculpture in a series of four busts of which the others were fuller in form, she comments that 'the extremely **thin** proportions of the head contrast with the solidity of the base.
7	A06 785	And on old Hiems' **thin** and icy crown,

➢ Don't forget to click on **Go!**

On the next page, just click **submit**.

BNC Collocation Settings	
Calculate over sentence boundaries:	no ▼
Include lemma information:	no ▼
Maximum window span:	+/- 5 ▼
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**.

Collocation parameters:				
Information:	collocations ▼	Statistics:	Rank by frequency ▼	
Collocation window span:	1 Right ▼ - 1 Right ▼	Basis:	whole BNC ▼	
Freq(node, collocate) at least:	5 ▼	Freq(collocate) at least:	5 ▼	
Filter results by:	Specific collocate:	and/or tag: any noun ▼	Submit changed parameters ▼	Go!

➢ 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).

Collocation parameters:				
Information:	collocations ▼	Statistics:	T-score ▼	
Collocation window span:	1 Right ▼ - 1 Right ▼	Basis:	whole BNC ▼	
Freq(node, collocate) at least:	5 ▼	Freq(collocate) at least:	5 ▼	
Filter results by:	Specific collocate:	and/or tag: any noun ▼	Download all results ▼	Go!

You can scroll through the results and **<u>click on the word frequencies</u>** 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!**

Collocation parameters:				
Information:	collocations ▼	Statistics:	Rank by frequency ▼	
Collocation window span:	1 Right ▼ - 1 Right ▼	Basis:	whole BNC ▼	
Freq(node, collocate) at least:	5 ▼	Freq(collocate) at least:	5 ▼	
Filter results by:	Specific collocate:	and/or tag: any noun ▼	New Query ▼	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*.