Exploring the BNC and COCA online using BYU interfaces

Corpora at Brigham Young University (BYU) [http://corpora.byu.edu/](http://corpora.byu.edu/)

BYU- BNC:
Simple but powerful interface to the BNC. You can search using POS and lemma, but you can't search or reveal distributions relating to all of the contextual and socio-linguistic features.

COCA – Corpus of Contemporary American English.
Very big, balanced corpus. Searchable by genre and time period. Same interface as BYU-BNC so easy to repeat searches on the two corpora.

Features of the BYU interfaces:
- Search for word, words, or string
- Search on POS, genre, context
- Restrict search to genre and/or (for some corpora) time period
- Integrated thesaurus
- Display results as Chart, List (default) or Compare words (search for two words/phrases)
- Many more advanced features – consult documentation in bottom right frame of interface.
- No information about speaker or non-linguistic features

Read more about the corpora and how to use the interface in the bottom right frame of the BYU interface. [http://corpora.byu.edu/](http://corpora.byu.edu/)

---

Ylva Berglund Prytz
<ylva.berglund@oucs.ox.ac.uk>
**Task 1. What are you like?**

**What does ‘and I’m like’ mean? When is it used? By whom?**

Go to [http://www.americancorpus.org/](http://www.americancorpus.org/)

In the left-hand panel:

1. Tick **CHART**

2. Under **SEARCH STRING**, type in *and I’m like* (note - space before ‘m).

3. Click on the **SEARCH** button (or just press Enter)

4. Explore the result. What is it showing? What conclusions can you draw? *(Does use vary with genre? How is the frequency distribution over time?)*

5. Click on the left-most column in the chart (spoken). Explore the concordances *(Are they all the same (kind)? Can you tell anything from the kind of texts where they are used?)*

6. Now make type in `[c*] [p*] [be] like` in the search box and run the search *(what do you think you will get?)*
   Can you see what the search string means, what it searches?
   In what way are the results different? *(what kind of phrases? Frequency?)*

**Compare to the BNC**

Now go to [http://corpus.byu.edu/bnc/](http://corpus.byu.edu/bnc/)

Make the same search(es)
What results do you get?
What does this mean (do you think)?
Warming up with BYU-BNC - "fast" concordancing

How many distinct meanings does the word fast have in English? You may want to look in a dictionary to see whether you have thought of all of the possibilities. There should be at least two, and maybe many more, depending on how you (or the dictionary) choose to classify them. In this activity, you will look in the corpus for examples of the word fast in order to see whether the evidence in the corpus supports what you (and the dictionary) think about this word.

1. Type the word fast in the box next to "Word(s)" in the 'Search' section on the left, then click on the 'search' button at the bottom of this section.
2. In the upper box on the right-hand side, there is a list telling you that the word 'fast' occurs 7357 times. Click on the word fast, and some concordance lines will appear in the lower box on this side.
3. 7357 entries are too many to look through. Click on 'SAMPLE: 100 ENTRIES' to reduce the number of concordance lines to a more manageable 100 lines.
4. Read the concordances. You can make this window wider by clicking on one of the small arrows on the left, and scroll down to see more lines.
5. Click on one of the titles for more information about the source and more context. (Note that for licensing reasons, the amount of co-text is limited here.) Right click and select the command for going back in this frame to get back to the list of concordance lines.
6. To get a bigger window, right-click on 'SAMPLE: 100 ENTRIES' and select the option to open in a new tab or browser window.
7. In the left-hand frame, under 'Display' at the top, click on 'Chart' and then press the 'search' button again. You will now see some information about the distribution of the word 'fast' in various types of text in the corpus.
8. 'Quick' could be considered a synonym or near-synonym for fast. Does quick have the same range of meanings? Is it only a synonym for some of the meanings of fast? Which ones? What words do you think might co-occur more frequently with quick than with fast? Do you expect to find any idioms involving quick? Discuss and/or write down your answers to this before doing a search. It's hard not to say with hindsight that the results are what you were expecting! To compare the usage and distribution of fast and quick, click on 'Compare words' in the 'Display' section, and enter fast and quick on separate lines under 'Words'.
9. BYU-BNC implements a synonym search facility, using the Wordnet ontology for English. Try searching for ['=fast'] and examine the results.
10. You can see all inflected forms of a word by making use of the lemma tags in the BNC. Searching for [fast] should give you all the forms of the word.
11. You can also restrict searches by part-of-speech. For example, you can restrict the results to fast as a verb by searching for fast.[v*]. You can also combine lemma (or synonym) searches with part-of-speech searches. What does a search for [fast],[v*] produce? Look at 'Query syntax' in the help files under 'More information...' in the bottom right-hand frame for more information and a useful reference guide on query options.

Ylva Berglund Prytz
<ylva.berglund@oucs.ox.ac.uk>