Session 3: Collocation

http://tinyurl.com/669o4zt
The fundamentals of corpus linguistics?

- Corpus
- Concordance
- Collocation

(see John Sinclair, *Corpus, Concordance, Collocation*, 1991)
“I'm just going out to commit certain deeds”
Practical

Explore the collocates of borrow with BNCweb - see the handout
Some more examples for further study

What are the 'collocational profiles' of:

aftermath

budge

on the edge of
on the verge of
on the brink of

?
Ways of using collocation for research

- Corpus, concordance, collocation
- Colligation
- Semantic preference and semantic prosody
- Collocational behaviour of (near) synonyms
- Cross-linguistic comparative study

[http://applij.oxfordjournals.org/cgi/content/full/27/1/103]
A critical look at collocation

Some pitfalls with collocation studies:

Semantic preferences are sometimes based on crude semantic classifications: negative/neutral/positive

There is often a lack of a benchmark for 'negativity' in the corpus (e.g. “it's all bad news in the corpus”)

Disentangling collocations from various forms of multi-word expression is tricky

Some interesting theoretical issues:

Collocations are (often, largely) inaccessible to intuition and introspection – you have to look at the corpus

'Idiom principle' v. 'open choice principle' – what do collocations tell us about the selection of linguistic items when we produce language?

Further reading:

Next week...

4. Corpus Linguistics: Annotation

In this session we will be looking at key methods and problems of applying linguistic annotation to a corpus, and how to use the annotated corpus.

*Key topics:*

- linguistic annotation
- metadata
- tagging, tags and tagsets
- word classes
Tip of the Week

Oxford English Corpus from OUP – 2 billion words and counting
http://dws-sketch.uk.oup.com/bonito/home.html

Username and password on request from Martin Wynne