Corpus Linguistics & the BNC

Course: English for Academic Purposes
Tutors: Frenzel & May
Presenter: Ye Wang
Date: 20\textsuperscript{th} Nov, 2006

Outline

1. What is corpus linguistics?
2. Corpus linguistics & EAP
3. The British National Corpus (BNC)

1 What Is Corpus Linguistics

1.1 What is a corpus?

- a collection of spoken or written texts to be used for linguistic analysis
- computerised electronic form
- various types of corpora:
  - synchronous vs. diachronic
  - written vs. spoken
  - general vs. specific
- sources for corpora:
  - ICAME
  - LDC
  - ELRA

1.2 Corpus vs. archive

- Corpus: systematic, planned, structured compilation of text; designed for a particular “representative function”
- Archive: opportunistically collected text; no particular purposes of research

1.3 Corpus Linguistics

- language description & explanation
- corpus as a source of evidence
- possibilities and probabilities in language use
- What do corpus linguists do?
  - corpus compiling
  - analysis tools developing
  - languages describing
  - language teaching, translation & so on
2 Corpus Linguistics & EAP

- computer analysis of corpora: evidence-based approach to language teaching
- more reliable, more accurate, supplement to human intuition
- genre analysis & description
- more awareness of options in particular disciplines
- contrastive approach to the learning of collocations
- methods: frequency & concordance

2.1 Frequency

- general academic word lists
- basic items of certain genre
- grammatical features (tagging)
- stylistic features

2.2 Concordance

- word in its linguistic environments
- typical patterns of collocation
- ‘semantic prosody’
- specific meanings of words
- change of meaning

3 The British National Corpus (BNC)

General facts of the BNC:

- year of publishing: 1994
- size: 100 million word collection
- monolingual: modern British English
- synchronic: late 20th century
- general: spoken (10%) & written (90%) samples of many different styles
- Design of the Corpus

4 References:


http://ell.phil.tu-chemnitz.de/compPhil/corpus.html

http://www.natcorp.ox.ac.uk/