Computational Semantics and Pragmatics 2011

Raquel Fernández – ILLC, University of Amsterdam

Homework #2   Due: 17 October 2011

Exercise 1: Distributional meaning

Have a look at this online tool: http://www.scottishcorpus.ac.uk/corpus/bnc/compair.php.

1.1 What do you think this tool does? Explain it in your own words.

1.2 The English adverbs *'utterly'* and *'absolutely'* are typically considered synonymous by most thesauri (http://thesaurus.com/browse/utterly). Compare them using this tool. What do you observe? What can you conclude about their meaning and usage?

1.3 Compare at least one more pair of words and comment on the results.

Exercise 2: Word usages across genres

The web interface of the Corpus of Contemporary American English (COCA) http://corpus.byu.edu/coca/ allows you to compare the *collocates* of a word (those words which appear frequently nearby) across different genres. This snap shot shows the results of a search for adjectival collocates of the word *'chair'* in the FICTION and ACADEMIC sections of the corpus:
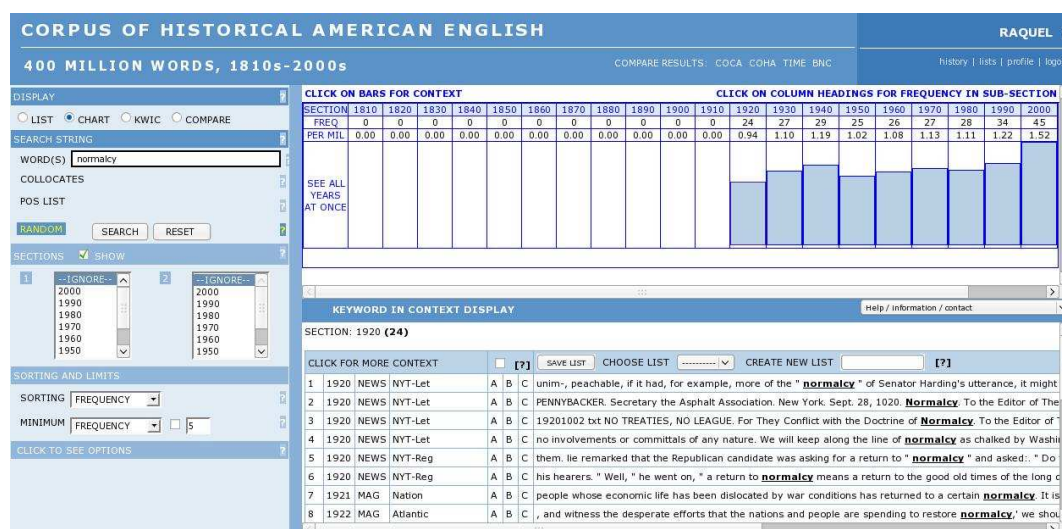

2.1 Reproduce this search on the COCA website and explore the results: you can click on the collocate words to see the contexts where they appear, and change the value in the POS LIST field to look for collocates with Part-of-Speech other than adjective. Given these distributional results, what can you conclude about the meaning of the word *'chair'*? Back up your claims with examples.

2.2 Compare the collocates of two other words across genres and explain your findings.

1

Exercise 3: Word usages from a diachronic perspective

Extracting the collocates of a word in different time periods can give us an indication of how the meaning of a word has changed over time. The web interface of the Corpus of Historical American English (COHA) http://corpus.byu.edu/coha/ allows you do this and much more. For instance, the following snap shot shows the nominal collocates of the word *'web'* in the 1950s and the 2000s.

3.1 Using the COHA web interface, compare the collocates of another word or construction across different time periods and explain your findings.

3.2 Read this post on the Language Log http://languagelog.ldc.upenn.edu/nll/?p=2585. The snap shot below shows how to use the COHA interface to obtain the graph with the frequency of use of the word *'normalcy'* that appears in Mark Liberman's reply to one of the comments to the post.

Using the Chart display of the COHA web interface, find:

(a) a word or construction whose frequency has increased to the present time;

(b) a word or construction whose frequency has decreased since the 1800s.

Provide evidence for your claims and explain your findings.