

### Lecture 3. Semantics and discourse: Collocations, keywords and reliability of manual coding

Aim: The lecture is based on Brezina (2018), Chapter 3. It explores meanings of words in context, which is an area important to both linguistic and social analyses. Topics discussed are collocations, keywords and manual coding of concordance lines; these play a key role both in the study of semantics ('dictionary' meanings of words) and in discourse analysis.

► Key terms: collocation collocation network association measure contingency table keywords lockwords inter-rater agreement judgement variable

Time:

1-hour lecture.

2-hour computer lab session with exercises and Lancaster Stats Tools online (optional).

1-hour individual study (readings).

Statistical tools: [Collocation calculator](#), [#LancsBox](#) and [Agreement calculator](#).

Practical exercises: [Chapter 3 Exercises and answers](#).

Data: [Chapter 3 data](#).

Readings: Brezina 2018 Chapter 3 + optionally selected Advanced readings recommended in the book.

Outline:

1. Collocations and association measure
2. Collocation graphs and networks: Exploring cross-associations
3. Keywords and lockwords
4. Inter-rater agreement measures
5. Things to remember

Main points – 'Things to remember':

- There are many association measures each highlighting different aspects of the collocational relationship (e.g. frequency or exclusivity). There is no one best association measure.
- Collocations can be presented in a tabular (table) or visual form (graph).
- Collocation networks show complex cross-associations in texts and discourses.
- The keyword procedure in its essence is a comparison which depends on a number of parameters. There is no such thing as one set of keywords.
- For judgement variables inter-rater agreement statistic should be reported. Gwet's AC<sub>1</sub> and AC<sub>2</sub>, Cohen's  $\kappa$  and Fleiss'  $\kappa$  as well as Interclass correlation can be used depending on the situation.