

LSE Department of Methodology,
MY428/528 - LT 2014

Qualitative Text Analysis

Course Convenor: Aude Bicquelet
(a.j.bicquelet@lse.ac.uk)

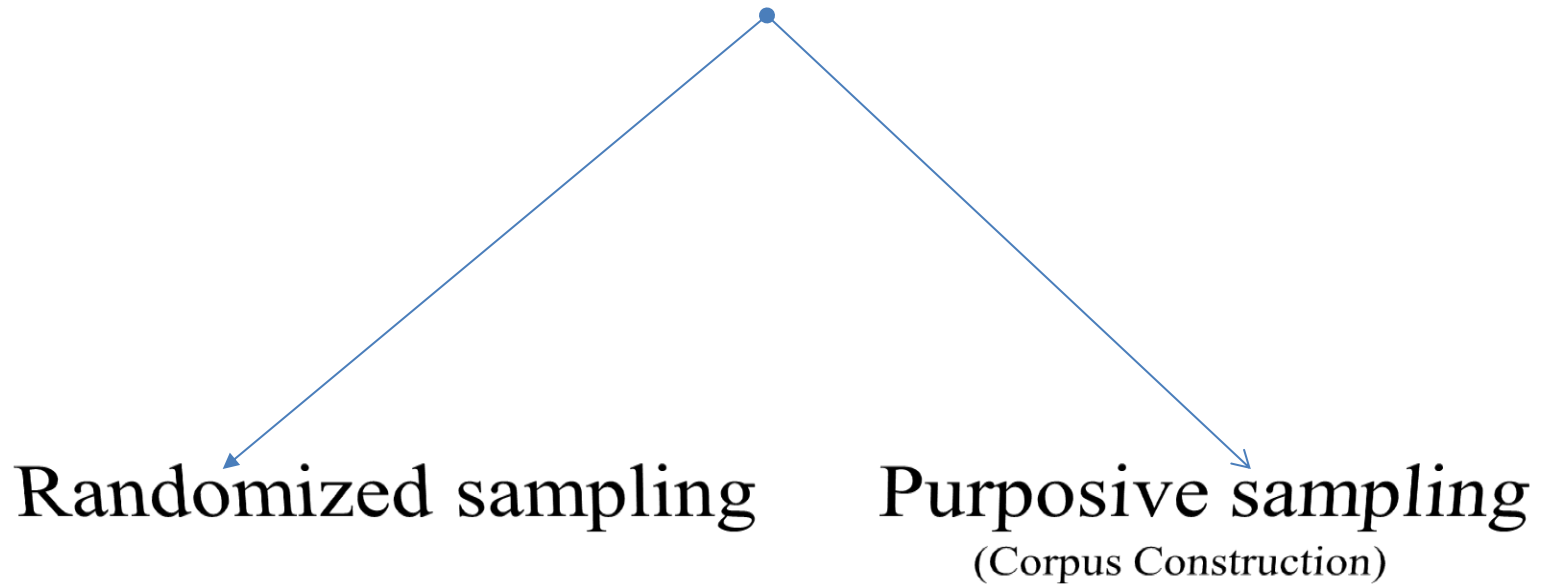
Office Hours:
Thursday 11:30-13:30

Seminar Activity Sampling & Unitizing

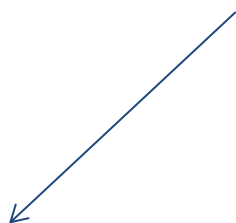
Week 2

SAMPLING

Text Selection



Text Selection



Randomized sampling

Purposive sampling
(Corpus Construction)

Sampling and Message Units

Defining the population

The population is the sets of units being studied (sets of units to which the researcher wishes to generalize)

- This is often a set of messages (or a population of people, whose messages are collected and then analysed).
- Once the population is defined it must serve as the basis for any sampling.
- In the case of a small population, there may be no need to draw a smaller, representative sample of the population (all units in the population may be included in the study, which would then be called a *census*)..

Sampling and Message Units

Example 1

Breen (1997) searched the NEXIS database for all newspaper articles in major papers between 1991 and 1994 that included the key search terms ‘catholic’, ‘priest’ or ‘clergy’. The search resulted in a set of articles that served as the population from which he then drew a sample.

Example 2

Keenan (1996) searched US TV news archives for the terms ‘public relations’ and ‘PR’ between 1980 and 1995. A total of 79 stories were identified. This entire population was studied in a *census* content analysis.

Sampling and Message Units

Sampling Strategies (1)

Cluster Sampling: Any random sampling in which a group or set of messages are sampled together.

Example: Lin (1997) collected a full week of broadcast network prime time TV commercials, with the month and the week randomly selected.

Sampling and Message Units

Sampling Strategies (2)

Multistage Sampling: Any random sampling technique in which two or more sampling steps are used.

Example: Hill and Hughes (1997) employed a two stage technique to analyse newsgroups discussions on the Internet. They first randomly sampled newsgroups from a sampling frame – a list derived from USENET. They then randomly sampled threads of discussion from those newsgroup selected.

Sampling and Message Units

Sampling Strategies (3)

Stratified Sampling: The sampling frame is segmented according to categories on some variable (s) of prime interest to the researcher.

Example: Smith (1999): Interested in comparing women's role portrayals in movies during the Golden Age of Hollywood (the 1930s and 1940s) with contemporary images (the 1990s) Smith stratified according to decade. She first constructed three different sampling frames of films featuring women, (one on each decade), and then conducted a systematic random sampling on each.

UNITIZING

Unitization

Whether you are using an inductive or a hypothetico-deductive approach to textual analysis, you need to be clear about:

(1) What your Units of Sampling (or Units of data collection) are.

- Units of sampling serve as the basis for identifying the population and drawing on a sample (or constructing a corpus).

(2) What your units of analysis are.

- Units of analysis are identifiable message or message component (words, sentences, paragraphs...) on which variables are being measured (or on which text is being analysed)

Unitization

Example 1:

Weyls (2001) collected news stories dealing with adult entertainment from NEXIS and analysed each story using the computer programme Diction. His ultimate goal was to track changes in coverage by *year*.

Unit (s) of sampling = News story

Unit (s) of analysis = Year

Unitization

Example 2:

Bicquelet et al. (2011) collected parliamentary debates on abortion to check whether male and female parliamentarians used the same type of vocabulary in the House of Commons.

Unit (s) of data collection = Parliamentary debates

Unit (s) of analysis = 'speech acts' (paragraphs – multiple coding allowed).

Exercise

Exercise

- **Read the forum discussion about the possibility of the Queen retiring.**

Code the data using:

- (1) Words as Units of Analysis
- (2) Sentences as Units of Analysis
- (3) Paragraphs as Units of Analysis

- **Compare the different types of results you obtain.**