

Advanced Content Analysis

Dr. Aude Bicquelet (course convenor)

a.j.bicquelet@lse.ac.uk

&

Andrea Gobbo (course assistant)

a.gobbo@lse.ac.uk

Course Description and Objectives

Building upon the theoretical and methodological foundations laid during the 2012 ECPR Summer School in Methods and Techniques, this course will introduce students with advanced approaches to Exploratory, Qualitative and Quantitative Content Analysis techniques. Advanced methods for analysing textual data such as Semantic Network Analysis, for example, will also be taught. Measurement techniques, Corpus Construction and sampling strategies will be discussed with two specific aims:

(a) How to address validation and reliability issues in Content Analysis?

(b) How to reduce the risks of making unwarranted inferences from textual data?

Students will be provided with textual data to apply models and techniques surveyed during the lectures but are strongly encouraged to bring their own documents (any type of text in digitised format) to practice on during lab sessions.

Learning Outcomes

Students taking this course will learn:

- 1) The epistemological roots and assumptions underlying the use of different CA methods
- 2) How to select, unitize and code texts for analysis
- 3) How to apply exploratory, inductive, hybrid, deductive approaches to textual data
- 4) How to evaluate validity, reliability and the robustness of CA results.

Students will be introduced to the use of software packages for the analysis of textual data (i.e.; Alceste, Iramuteq, QDA Miner, Wordstat,) but no prior knowledge (either technical or mathematical) is essential. **Students taking this course are required to bring a laptop** and will be asked to download the 30 day trial version of QDA miner/ Wordstat/ Simstat. For more information please see: Provalis package <http://www.provalisresearch.com/>.

Students who may already have previously downloaded the trial version (but not purchased the software subsequently) will be provided with new 30 day trial version.

Teaching Arrangements

The course comprises five lectures (90 mn) and five lab sessions (90 mn).

Lecture outlines and Readings

Day 1: Introduction – Approaches to Content Analysis & Corpus Construction.

Lecture: This lecture will introduce the course and its structure. It will explore key definitional and conceptual issues related to the use of content analysis in social science research. The emphasis will be on providing students with what can be thought of as the basics: epistemological roots and assumptions; research design and the methodology of content analysis. In its second part, the lecture will compare and contrast two avenues for obtaining a text corpus: statistical sampling *v.* corpus construction.

Seminar:

- (1) Presentation of Participants - discussion about participants' research projects, data and potential strategies to address research questions.
- (2) Hands on Exercise: Gathering and compiling data - corpus construction *via* the use of on-line datasets (i.e. the British National Corpus).
- (3) Take-Home Exercise: Building a Corpus (Analysing newspaper articles, or else).

Compulsory Readings

**Bauer M. W and Aarts, B. (2000) Corpus construction: a principle for qualitative data collection, in: Bauer MW & G Gaskell (eds) *Qualitative researching with text, image and sound. A practical handbook for social researchers*, London, Sage, (pp. 19-37).

*Neuendorf, K. A. (2002). *The Content Analysis Guidebook*. Sage, Thousand Oaks, CA. Sage, Thousand Oaks, CA (Chap. 1 & 2)

Optional

Schreier, M. (2012), *Qualitative Content Analysis in Practice* Sage, Thousand Oaks, CA. Sage, Thousand Oaks, CA (Chap. 1 & 2).

Howe, K.R (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher* 17 (10-16).

Rossi, P.H (1994). The war between the quals and the quants: Is a lasting peace possible? In: *The Qualitative-Quantitative Debate: New Perspectives*, C.S Reichardt and S.F Rallis, eds. San Francisco: Jossey-Bass (pp. 23-36)

DAY 2: Unitizing & Exploring – How to mine your data?

Lecture: The emphasis of this lecture will be on identifying the various units that need to be considered when collecting and analysing textual data, i.e. units of data collection; units of sampling; units of analysis and units of coding. The lecture will then provide an overview of exploratory, text-mining approaches for the analysis of large data sets introducing the use of the Alceste software and profiling examples of its various applications in social science research.

Seminar:

- (1) Preparing data, unitizing, coding and running an analysis with Alceste.
- (2) Hands on Exercise: Using Iramuteq (*R*) as a substitute for Alceste.
- (3) Take-Home Exercise: unitizing and coding newspaper articles using Alceste or Iramuteq and interpreting outputs.

Compulsory Readings

**Bicquelet, A. & Weale, A. (2011) ‘Coping with the cornucopia: Can Text Mining Help Handle the Data Deluge in Public Policy Analysis?’ *Policy & Internet* (Vol.3, N.4)

**Schonhardt-Bailey, C. (2008). The congressional debate on partial-birth abortion: Constitutional gravitas and moral passion. *British Journal of Political Science*, 383–410.

* Schonhardt-Bailey, C. and Bailey, A. (2008). Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution of 1979", *Political Analysis* 16: 4.

Optional

Bicquelet, A., Weale, A & Bara, J. ‘In a different Parliamentary Voice? An analysis of gender differences in UK parliamentary debates about abortion’ *Politics & Gender*.

Guérin-Pace, F (1998) ‘Textual Statistics. An Exploratory Tool for the Social Sciences’, *New Methodological Approaches in the Social Sciences*, 10:1, pp. 73-95.

Lahlou, L. (1996) ‘A Method to Extract Social Representations from Linguistic Corpora’, *Japanese Journal of Experimental Social Psychology*, 35, pp. 278-291.

Brugidou, M. (2003) ‘Argumentation and Values: An Analysis of Ordinary Political Competence via an Open-Ended Question’, *International Journal of Public Opinion Research*, 15:4, pp. 413- 430.

Bara, J., Weale, A. and Bicquelet, A. (2007) ‘Analysing Parliamentary Debate with Computer Assistance’, *Swiss Journal of Political Science*, 13:4, pp. 577-605.

DAY 3: Back to basics – Qualitative and Hybrid approaches.

Lecture: Analysing text from a qualitative perspective involves five crucial steps:

- (a) Discovering themes and subthemes;
- (b) Describing core and peripheral elements of themes;
- (c) Building hierarchies of themes and codebooks;
- (d) Applying themes,
- (e) Linking themes into theoretical models.

This lecture will look at these five steps in depth. The focus will be on practical techniques and models used in qualitative content analysis. The second part of the lecture will look at hybrid techniques explaining how they bridge the gap between interpretative and quantitative approaches to textual analysis.

Seminar:

- (1) Building a strong (reliable and transparent) codebook
- (2) Hands on Exercise: Building a coding frame (using QDA Miner or NVivo)
- (3) Take-Home Exercise: Building a codebook and a coding frame (for the analysis of newspaper articles, or else).

Compulsory Readings

**Hsieh, H.-F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.

**Feredey, J. and Cochrane, E.M. (2006) 'Demonstrating Rigor Using Thematic Analysis: A hybrid approach of inductive and deductive coding and theme development,' *International Journal of Qualitative Methods*, 5(1): 80-92.

*Attride-Stirling, J. (2001) 'Thematic Networks: an Analytic Tool for Qualitative Research', *Qualitative Research* 1(3): 385-405.

Optional

Boyatzis, R.E (1998) *Transforming Qualitative Information*. Sage, Thousand Oaks, CA. (Chapter 4 and 5).

Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1(2). Retrieved July 28, 2008, from <http://217.160.35.246/fqs-texte/2-00/2-00mayring-e.pdf>.

Braun, V, Clarke, V (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology* 3: 77- 101.

Russel, B. H. and Ryan, G.W. (2010). *Analysing Qualitative Data: Systematic Approaches*. Sage, Thousand Oaks, CA. (Chapter 3).

DAY 4: Classical Content Analysis – Constructing Codes & Assessing reliability.

Lecture: Content Analysis is a set of methods for systematically coding, analysing and drawing valid inferences from texts. These methods are used across the social sciences to explore explicit and covert meanings (also called manifest and latent content) and to test hypotheses about texts. This lecture will introduce the main analytical steps involved in Classical Content Analysis. The second part of the lecture will focus on Validity and Reliability issues and will consider various strategies to strengthen the robustness of the results obtained through classical content analysis.

Seminar:

- (1) Transforming qualitative information into quantitative data
- (2) Hands on Exercise: building a matrix & checking for inter-coder reliability
- (3) Take-Home Exercise: building a matrix & checking for inter-coder reliability

Compulsory Readings

** Cunningham, G. B., Sagas, M., Satore, M. L., Amsden, M. L., & Schellhase, A. (2004). Gender representation in the NCAA News: Is the glass half full or half empty? *Sex Roles: A Journal of Research*, 50(11/12), 861-870.

* Hirschman, E.C. (1987). People as products: Analysis of a complex marketing exchange. *Journal of Marketing*, 51. (98-108).

Optional

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Sage, Thousand Oaks, CA, 2nd edition (Chapter 1 and 2)

Woolley J T (2000) 'Using media-based data in studies of politics', *American Journal of Political Science*, 44: 156-173.

Banerjee, M., Capozzoli, M., McSweeney, L., and Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 27(1):3–23.

Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. 2008. "Gender differences in language use: An analysis of 14,000 text samples." *Discourse Processes* 45:211–236.

DAY 5: Content Analysis & beyond

Lecture: This lecture will recap and summarize the various approaches surveyed over the last four sessions. The aim here will not be to draw definitive comparisons but rather to make useful analytical distinctions between different techniques according to (a) their ambitions (b) acknowledged limitations. In the second part, the lecture will focus on innovative techniques for the analysis of textual data (i.e. Semantic network analysis).

Seminar:

- (1) Individual debrief on the course and individual advice on Research Project
- (2) Hands on Exercise: Using VOSviewer for Semantic Network analysis
- (3) Take-Home Exercise: Integrating Semantic Network analysis in Text Mining.

Compulsory Readings

** Hanneman, R. A. and M. Riddle (2005). *Introduction to social network methods*. Riverside, CA, University of California, Riverside (published in digital form at <http://faculty.ucr.edu/~hanneman/>) (Chap. 1 & 2)

* Muñoz-Leiva, F et al. (2012) An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality & Quantity*, 46:1077—1095.

Optional

Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, 7(3).

Degenne, A. and M. Forse (1999). *Introducing Social Networks*. London: Sage (Chap. 1 & 2)

Wasserman, S. and K. Faust. (1994) *Social network analysis: methods and applications*. Cambridge Cambridge University Press (Chapters 5, 6 & 7)

Snijders, T. A. B., G. G. de van Bunt, et al. (2010) 'Introduction to stochastic actor-based models for network dynamics' *Social Networks* 32 (1).

References:

- Attride-Stirling, J. (2001) 'Thematic Networks: an Analytic Tool for Qualitative Research', *Qualitative Research* 1(3): 385-405.
- Bara, J., Weale, A. and Bicquelet, A. (2007) 'Analysing Parliamentary Debate with Computer Assistance', *Swiss Journal of Political Science*, 13:4, pp. 577-605.
- Bauer M.W (2000) 'Classical content analysis: a review', in: Bauer, M. and Gaskell, G. (eds) *Qualitative Researching with Text, Image and Sound*, London: Sage, pp.131-151.
- Bicquelet, A., Weale, A & Bara, J. (2012) 'In a different Parliamentary Voice? An analysis of gender differences in UK parliamentary debates about abortion' *Politics & Gender*. 8 (1). 83-121,
- Bicquelet, A. & Weale, A. (2012) 'Coping with the cornucopia: Can Text Mining Help Handle the Data Deluge in Public Policy Analysis?' *Policy & Internet* 3 (4).
- Boyatzis, R.E (1998) *Transforming Qualitative Information*. Sage, Thousand Oaks, CA. (Chapter 4 and 5).
- Braun, V, Clarke, V (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology* 3: 77- 101
- Fereday, J. and Cochrane, E.M. (2006) 'Demonstrating Rigor Using Thematic Analysis: A hybrid approach of inductive and deductive coding and theme development,' *International Journal of Qualitative Methods*, 5(1): 80-92
- Franzosi R P (2004) *From words to numbers: narrative, data and social science*, New York, CUP.
- Guba, E.G. and Y.S. Lincoln (1994). Competing paradigms in qualitative research. In: *Handbook of Qualitative Research*, N.K Denzin and Y.S. Lincoln, eds. Thousand Oaks, CA: Sage publications (pp. 105-117).
- Hsieh, H.-F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Holsti O (1969) *Content analysis for the social sciences and humanities*, Reading, MA, Addison-Wiley [chapter on types of CA following the logic of signs: syntax, semantics, pragmatics]
- Howe, K.R (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher* 17: 10-16.
- Kelle U (2000) Computer-assisted analysis: coding and indexing, in Bauer MW & G Gaskell (ed) *Qualitative researching with text, image and sound. A practical handbook for social researchers*, London, Sage, pp282-298.

- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Sage, Thousand Oaks, CA, 2nd edition.
- Laver, M. and Garry, J. (2000). Estimating policy positions from political texts. *American Journal of Political Science*, 44(3):619–634.
- Laver, M., Benoit, K., and Garry, J. (2003). Estimating the policy positions of political actors using words as data. *American Political Science Review*, 97(2):311–331
- Miles, M., & Huberman, A.M. (1999). *Qualitative Data Analysis*. Thousand Oaks, CA: Sage Publications. 2nd Edition (Introduction).
- Neuendorf, K. A. (2002). *The Content Analysis Guidebook*. Sage, Thousand Oaks, CA. Sage, Thousand Oaks, CA.
- Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. (2008). “Gender differences in language use: An analysis of 14,000 text samples.” *Discourse Processes* 45:211–236.
- Roberts, C. W. (2000). A conceptual framework for quantitative text analysis. *Quality and Quantity*, 34(3):259–274.
- Rossi, P.H (1994). The war between the quals and the quants: Is a lasting peace possible? In: *The Qualitative-Quantitative Debate: New Perspectives*, C.S Reichardt and S.F Rallis, eds. San Francisco: Jossey-Bass (pp 23-36)
- Russel, B. H. and Ryan, G.W. (2010). *Analysing Qualitative Data: Systematic Approaches*. Sage, Thousand Oaks, CA. Sage, Thousand Oaks, CA.
- Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers*. Sage, Thousand Oaks, CA. Sage, Thousand Oaks, CA.
- Schonhardt-Bailey, C. (2008). The congressional debate on partial-birth abortion: Constitutional gravitas and moral passion. *British Journal of Political Science*, 38:383–410.
- Schonhardt-Bailey, C. and Bayley, A. (2008). Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution of 1979", *Political Analysis* 16: 4.
- Slapin, J. and Proksch, S.-O. (2008). A scaling model for estimating time series policy positions from texts. *American Journal of Political Science*, 52(8).
- Tausczik, Y & Pennebaker, J. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 29(1) 24–54.
- Tesch, R. (1990). *Qualitative Research: Analysis Types and Software Tools*, New York: Falmer Press (Types of Qualitative Research, pp 55-77).
- Woolley J T (2000) ‘Using media-based data in studies of politics’, *American Journal of Political Science*, 44: 156-173.