
Topics and genre changes in Czech sociological articles

Radim Hládik

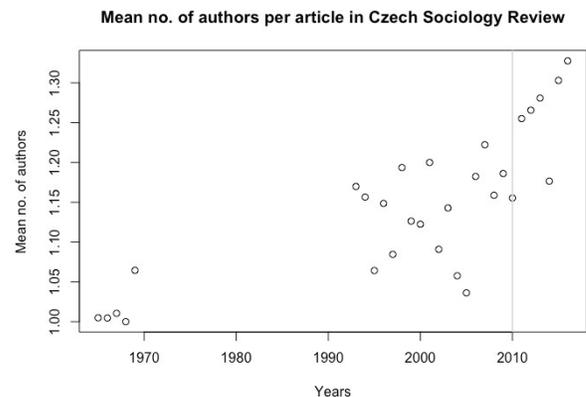
radim.hladik@fulbrightmail.org

Czech Academy of Sciences, Czech Republic

Introduction

Since Kuhn's distinction between "textbook" science and "article" science, studies of scientific texts have established the social nature of academic writing. Analysis of scientific texts can yield epistemological, disciplinary, and historical insights. These texts are the arena in which knowledge claims are raised (Myers 1985), trials of strength held (Latour 1987), intra-disciplinary boundaries drawn (Wolfe 1990), disciplinary histories traced (Bazerman 1988). Genres of scientific texts have thus been shown as socially enacted structures (Berkenkotter, Huckin, Ackerman 1994) rather than as transparent styles. Despite their many merits, the hitherto available empirical studies on scientific writing have been constrained by either a focus on early history (such as the original Philosophical Transactions of the Royal Society) or reductive sampling for traditional content analysis that allow researchers to grasp the otherwise immense textual data. Approaches inspired by digital humanities approaches offer new possibilities of studying disciplinary formations, as was demonstrated by Goldstone and Underwood (2014) in their distant reading of literary studies texts. This paper follows suit in reporting the results of applying digital humanities methods of text analysis to the corpus of research articles in sociology, specifically, in Czech Sociological Review. Writing in wide-scope disciplines, such as sociology, is of particular interest because it embodies the conflict between literature and the notion of social science (Lepenies 1988). Sociology has been revealed as a discipline of two writing cultures, monographic and journal (Wolfe 1990). The writing in sociology is also oscillating between the aspiration to the positivist ideal of science (Leenhardt 1992) and the acceptance of diverse styles (Agger 2002). Abbot and Barman (1997) have concluded, on the basis of sequence comparison, that research articles in sociology lack "rhetorical ri-

gidity". The discipline thus offers a particularly opportune resource for the analysis of genre and topical variations. Czech Sociological Review was chosen as an example of a "core" journal in the country. As Oromaner (2008) demonstrated, "core journals" in sociology have tendency to become central to the discipline's "intellectual integration". Thus the results of the analysis can be taken as indicative of the mainstream tendencies in Czech sociology. The focus on the Czech sociology has the additional advantage of representing a interesting example of a discipline undergoing substantial transformation in the wake of academic and wider societal changes that came about with the fall of the Communist Party regimes in 1989. Also, the journal offers open access to its content. The data for the analysis were scrapped from the journal's website in September 2016 and the resulting data set contains 3483 articles. A preliminary exploration of metadata revealed noteworthy patterns around the year in which a new policy for science evaluation had been introduced (cf. the figure).



The data collection and analysis is carried out using R language. Besides crude measures of the corpus, the paper will also report the analysis of the textual data, using text mining techniques to comment on the issues that have been raised in the available literature. Topic modeling through LDA model will be used to assess the topical changes across time. Annual frequencies of particular words will be used as indices of changes of the transforming disciplines (this includes, especially, the words relating contemporary sociology to its "communist" variety, such as references to Marx or "communism"). Multidimensional scaling will then be employed to reveal term clustering around further keywords that arguably important in sociology. Quantitative bias, or a lack of thereof, will be measured by the presence of numbers. The overall purpose of the

analysis is to address the questions raised in pre-existing literature using the specific example of a Czech social science discipline and to demonstrate the usefulness of text mining techniques in the analysis of scientific writing.

Bibliography

- Abbott, A., and Barman, E.** (1997). "Sequence Comparison Via Alignment and Gibbs Sampling: A Formal Analysis of the Emergence of the Modern Sociological Article." *Sociological Methodology* 27 (1): 47-87.
- Agger, B.** (2002). "Sociological Writing in the Wake of Post-modernism." *Cultural Studies / Critical Methodologies* 2 (4): 427-59.
- Bazerman, C.** (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.
- Berkenkotter, C., Huckin, T. N., and Ackerman, J.** (1994). "Social Context and Socially Constructed Texts." In *Landmark Essays on Writing across the Curriculum*, edited by Charles Bazerman and David R. Russell. London, New York: Routledge.
- Goldstone, A., and Underwood, T.** (2014). "The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us." *New Literary History* 45 (3): 359-84.
- Leenhardt, J.** (1992). "Writing and 'Scientific Discourse' in Sociology." *History of the Human Sciences* 5 (1): 63-71.
- Latour, B.** (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, Mass: Harvard University Press.
- Lepenies, W.** (1988). *Between Literature and Science: The Rise of Sociology. Ideas in Context*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.
- Wolfe, A.** (1990). "Books vs. Articles: Two Ways of Publishing Sociology." *Sociological Forum* 5 (3): 477-89.
- Oromaner, M.** (2008). "Intellectual Integration and Articles in *Core Sociology Journals*, 1960-2000." *American Sociologist* 39 (4): 279-89.