

Closing Conference

Words – Concepts – Meanings. Which Historical Semantics for the Middle Ages?

December 2014 will mark the end of the joint research project *Political Language in the Middle Ages. Semantic Approaches* at the Johann Wolfgang Goethe-University Frankfurt am Main, supported by the DFG with funding coming from the Gottfried Wilhelm Leibniz Award. As part of a conference, we will offer a *discussion of the project's progress* and intend to develop, in collaboration with our guests, *perspectives for the future* of digitally supported historical semantics of the Pre-Modern era. We will also introduce a *new tool* which promises to both transform and simplify the work of historians and philologists.

Forty Years after the “Historische Grundbegriffe”: Historical Semantics Today

Between 1972 and 1992, the famous seven volumes of “*Geschichtliche Grundbegriffe. Historisches Lexikon zur Politisch-Sozialen Sprache in Deutschland*” were published. In his introduction, Reinhart Koselleck had intended to direct the authors’ attentions to specific, pre-defined corpora of analysis, yet he later lamented that too many writers continued with approaches steeped in the past of the history of ideas. At the time, the authors lacked the technological means to analyse larger volumes of text, in particular when looking at something other than the classics of political theory. It is only today that we are able to meet the theoretical challenge set by Koselleck in the 1960s. Now, computer-based historical semantics examining broad corpora is possible. Methods and technologies employed, however, remain complex and require long lead times. The Leibniz Project on the semantics of political language in the Middle Ages was, among others, intended to comparatively simplify and accelerate computer-based corpus-linguistic processes. This was enabled by a close cooperation with the “[Study Group Text Technology](#)“, set up in 2008 at the Goethe University (Alexander Mehler).

New opportunities: Interpretation and Quantification

By looking at word use in the Latin Middle Ages, the Leibniz-Project “Political Language in the Middle Ages. Semantic Approaches” aimed at delivering the ground work for a systematic examination of political language in the Middle Ages. Headed by Bernhard Jussen, Jan Rüdiger (until 2011), and Gregor Rohmann (from 2011), the project had two missions: First, to provide corpus-based historical semantics with digital tools; second, to reappraise recent standards of historical semantics in several individual projects on medieval studies.

Two New Tools for Research – CHS and HSCM

Leibniz Project and Study Group Text Technology have developed two powerful, web-based tools for research which make it possible to semantically examine any given Latin text qualitatively and quantitatively in ways which had previously been impossible: on the one hand, a password-protected expert tool (*HSCM*), on the other, an open, simple and easy-to-use

tool (*CHS*) for historians unable to familiarise themselves with the former. The world's largest free Latin word-form based lexicon was built (*Frankfurt Latin Lexicon*) which can also be accessed online.

CHS – Computational Historical Semantics

CHS, a fast, password-free tool which is simple to use, currently allows quantitative examination of a constantly growing body of Latin texts (currently some 4,500), increasingly from high-quality editions (esp. MGH).

HSCM – Historical Semantics Corpus Management

HSCM is an expert tool which allows for a more wide-ranging, active work with Latin texts – integrating and annotating new texts, correcting existing annotations, expanding the word-form lexicon, collationing, etc. It requires a password and some time for familiarisation.

Approach

Both platforms share four key steps:

- Corpus-oriented searches make it possible to discover key usage aspects for specific words or word constellations and automatically compare these to usage in other texts. The option of automatically comparing makes rapid indexing much easier and has a considerable impact on hermeneutic inquiries.
- Automatic analysis of co-occurrence enables users to compare any individual usage with normal usage and in this way provides pointers for specific word use.
- Even though the software is currently nowhere near recognising Latin syntax, it already lets users easily capture grammatical structures of word usage.
- Finally, for each text or corpus, individual observations during interpretative reading, i.e. actual hermeneutics, may be quantitatively rated, offering a methodological control for the step from interesting to exemplary passages.

Word-form Lexicon

Software and user interfaces CHS and HSCM are constantly updated. Basis and at the same result of the process is the world's largest Latin word-form lexicon, the *Frankfurt Latin Lexicon*, which currently covers some 11m word forms. The lexicon can be accessed separately on the CHS pages and may be edited using HSCM.

Texts, Genres, Partners, Task Sharing

A large number (ca. 4,500) of commonly used texts have already been processed for the research community. The aim is to (a) prepare current, controllable editions for analysis and (b) coordinate the range of research activities in different locations as regards digitalisation. Initially, work began with the *Patrologia Latina*. We now have better material owing to collaboration with:

- MGH (first texts in CHS,
- Leipzig Corpus of Open Greek and Latin,

- the Academy Project *School of Salamanca* in Mainz (first texts in CHS),
- *Corpus Corporum*, a project developed at the University of Zurich (first online texts)
- the Bielefeld-based SFB 585 (Special Research Field, hosting a project considering Late Medieval texts on political theory)

Each text is classified into one or several categories based on the definitions of the *Typologie des Sources du Moyen Âge Occidental* to make it easier for users to create individual corpora.

Usability

In our projects, we have constantly reviewed the usability of developed tools and determined the possibilities and limits of computer-based semasiological and onomasiological queries. Some projects hence went beyond mere word usage histories from the start. They were able to leverage HSCM/CHS for first heuristic queries. Others relied more on corpus-linguistic methodologies, broadening the theoretical base for medieval studies. In this way, using exemplary individual cases, we have created new access trails into the historical semantics of politics in Medieval Latin Europe.

Challenges and Perspectives for Future Activities

Several follow-up projects will continue the work of the Leibniz Project. On the occasion of its conclusion, we want to initiate a debate on its current results, starting with the challenges of a semasiological-corpuslinguistic approach. We want to invite everyone who closely followed our work to debate the perspectives of developing a historical semantics of the Middle Ages.

Tools such as CHS and HSCM are not aimed at replacing reading literacy with quantitative analysis, but rather at accelerating heuristics, making it more precise, and providing a means of reviewing hermeneutic results. Searches for exemplary or unique passages have become more systematic (and considerably faster), “exciting” and “exemplary” are easier to distinguish; results are achieved which would have been almost impossible manually; findings are fully reproducible. Corpus linguistic processes offer hermeneutic access even to non-specialists, results are easier to understand. We do not advocate *distant reading*, but rather *close reading* with quantitative control. Often, results are similar to those which traditional textual work might also have achieved. Yet with the support of technology, we have also gained insights which challenge the doctrines of classical hermeneutics. In addition to this role as a corrective, there are new opportunities for the further development of the semasiological approach:

Theoretical Perspectives

- *Language vs. Image, Performance, and Practice*: To what extent does “historical semantics” imply a focus of linguistic forms of expression? Would semiotics – in a broader sense, which would include non-linguistic generation of meaning – be covered by this umbrella term? In semi-oral societies, which are also characterised by a linguistic barrier between scholars and non-scholars, politics are possibly not entirely transacted by means of language and definitely not to a larger part by written communication. In addition, linguistic communication is always accompanied by non-verbal signals. What, then, should be the place of computer-based textual analysis in our concept of political semantics in the Middle Ages? Does it reflect how political actors of the past perceived politics and how they acted politically?

- *Semasiology vs. Onomasiology*: When we look for words, we assume medieval authors to have used the expressions expected by our research design to describe the respective phenomenon (concept, term). How do we make sure that we are not looking at the wrong places? Do we have to pay closer attention to the differences between “political theory” and “political language”?
- *Single-word Analysis vs. Examination of Phrases and Syntactic Combinations*: How can we serially examine figures of speech and collocations? What can we gain from leveraging the opportunities offered by textual linguistics in analysing the role played by grammar as a semantic carrier?
- *Causal Logistic Linearity vs. Typology and Intertextuality*: To what extent does the historical-semantic approach assume the argumentative linearity considered the normal case in modern written culture? We know that text production and reception of medieval writers and readers were based on processes of typology (key word: four senses of recognition), intertextuality with canonical works (key word: exegesis), and meditative textual reception. Corpus-linguistic processes have to be (or could) be designed in such a way that they unearth, rather than hide, these specifications.

Methodological Perspectives

- *Current State of Research vs. Hermeneutics without Presuppositions*: How is the historical-semantic approach to find its corpora? Must it necessarily be underpinned by classifications of genre which the authors would have been unaware of? What is it that makes a text relevant when looking for “political language”? What does the number of hits when searching for a term say about the text’s heuristic interest?
- *Layout and Graphics as Semantic Carriers*: Some dimensions which carry meaning will be overlooked if an analysis fails to examine the manuscript. Differences between editions and manuscripts have to be scrutinised, in particular with regard to, e.g., serialities, forms of diagrammatics, or idiosyncrasies (e.g. acronyms) normalised in editions.
- *Latin vs. Popular Speech and Discourse of the Elites vs. Discourse of the Masses*: Examined texts are usually closely connected to the milieu of a scholarly, Latin-speaking clerical culture. The Latin we examine was acquired as a second language which, in particular in its literate versions, was more resistant to change than vernacular languages. Should we look for a second, possibly more sweeping, political discourse in addition to or underlying this scholarly discourse, a discourse not just of the “people”, but also of non-scholarly / non-clerical political elites? Do the tools of historical semantics offer opportunities for capturing this parallel discourse?
- *Translatability of Semantics*: For canonical texts, medieval scholarly life was based on translations. For a wide range of purposes, Latin reasoning was in turn translated into vernacular languages. What was the impact of these multiple processes of transformation on semantics?
- *Specific Semantics of Individual Text Genres*: Corpus-based process should not persist in using the lawnmower principle when examining large bodies of text, but also have to focus on the specific semantic structures of each genre. The linguistic features of encyclopaedias differ from those of homilies, those of chronicles from urbaria.
- *Latin Europe vs. Comparatistics*: In the long run, a historical semantics of political thought in the European Middle Ages would have to look at the Byzantine Empire or the Muslim-Arabian territories, to track down similarities, (inter-)dependencies, and differences in the development of political thought in Euro-Mediterranean cultural regions.

Technical Perspectives

- *Edition vs. Handwritten Records*: To what extent can text-linguistic analysis be made independent of the editions' varying quality and/or the editors' presuppositions? How is a study of the lexical field *rex* or the lexical field *regnum* to rely on editors which may have resolved the acronym *reg.* either way?
- *Usability of Textual Linguistics for Medieval Corpora*: Working with texts in Medieval Latin requires considerable preparation, as all graphic, orthographic, and grammatical variations have to be determined. IT users remain largely unaware of the individual steps taken in this background process, yet the relationship between costs and benefits has to be determined separately for each specific case.
- *Restriction to Medieval Latin*: With the above comments on the restriction to Latin Europe in mind, opening the HSCM for Western European vernacular languages, in particular Greek and Arabic, would be necessary. However, problems would not be limited to lexical indexing, but – at least as regards the specific grammar of Semitic languages – also include programming. Not until the project with its focus on Latin written language, however, was it possible to establish a significant basis for an extension of this kind.

Planned Panels

We would like to discuss these and other challenges to a historical semantics of the Middle Ages with everyone interested. We are looking for contributions to three panels, and we have consciously refrained from allocating the issues listed above to any specific panel:

Practical Application of CHS and HSCM: For a Corpus-based Historical Semantics of the Middle Ages

As the Leibniz Project nears completion, we do not intend to discuss only our own findings. We offer anyone interested in conceptual and perceptual history or semiotics of the Middle Ages to test CHS or HSCM for their work. The debates will result in new specifications for the further development of these new tools.

Beyond Vocabulary: Potentials of Historical Semantics for Medieval Studies

It is little helpful to suggest that all methods have their blind spots, and a history of word usage would remain a history of word usage. Potentials and limits of historical semantics for Medieval Studies have to be more precisely determined. When we consider the specific features medieval culture outlined above, what specific benefit does the analysis of linguistic records provide?

Between Manuscript and Editions: Text Technology and the Middle Ages

What are the experiences of our colleagues during the development of IT-supported processes for historical semantic medieval studies? How should and must computer-linguistic approaches handle the specifics of medieval texts and corpora? In what ways are empiricism on the one hand and design of corpora and text editing mutually dependent?

Information about the Conference

The conference of the Leibniz Project is, at the same time, intended at initiating a continued development of the selected approach. On the long road to a historical semantics of the Middle Ages simply debating the issues raised here at the conference will already be a key step.

Hosted by: Leibniz Project “Political Language in the Middle Ages. Semantic approaches at the Johann Wolfgang Goethe-University Frankfurt am Main” (Prof. Dr. Bernhard Jussen; PD Dr. Gregor Rohmann)

Date: 19th & 20th February, 2015

Location: Johann Wolfgang Goethe-University Frankfurt am Main, Campus Westend

Contact: Gregor Rohmann (g.rohmann@em.uni-frankfurt.de)

Conference languages will be English and German.

Please mail abstracts for talks (no more than 300 wds. in length) no later than October 31st, 2014 to: PD Dr. Gregor Rohmann (g.rohmann@em.uni-frankfurt.de).