Two Swe-Clarin pilot projects in Språkbanken

A big-data approach to the study of rhetorical history

Rhetorical history is traditionally studied through rhetorical treatises or selected rhetorical practices, for example the speeches of major orators. Although valuable sources, these do not give us the answers to all our questions. Indeed, focus on a few canonical works or the major historical figures might even lead us to reproduce cultural self-identifications and false generalizations. However, thanks to increasing availability of relevant digitized texts, we are now at a point where it is possible to see how new research questions can be formulated – and how old research questions can be addressed from a new angle or established results verified – on the basis of exhaustive collections of data, rather than small samples, but where a methodology has not yet established itself.

In this pilot project – a collaboration between Jon Viklund, rhetorical historian at Uppsala University and the Språkbanken group of Swe-Clarin – we study the historical development of attitudes towards rhetoric as expressed in public discourse over the last 200 years. The aim of the data for the preliminary investigation conducted in the pilot project is a massive collection of texts: one billion words of digitized Swedish newspapers from the 19th and early 20th century.

The research questions were addressed using Språkbanken’s existing corpus infrastructure Korp. This allowed us both to start to explore the material quickly and to note where added functionality would be needed to that provided by Korp, which was designed primarily with linguistic research in mind.

Nevertheless, Korp functions such as the word picture and trend graph proved to be useful tools for this kind of investigation.

Thus, the word picture (shown to the left in the picture below) was used to investigate significant modifications of the noun vältalighet ‘eloquence’, illustrating general attitudes to rhetoric as well as salient conceptual metaphors employed in characterizing it, and the trend graph (to the right in the picture below) revealed (1) a cyclically varying preoccupation with ‘speech’ and ‘speakers’ following the three-year intervals between the sessions of the old-style parliament (abolished in 1866), and (2) a markedly increasing interest in political matters towards the end of the 19th century.

The Swe-Clarin steering committee is finally in place!

The Swe-Clarin steering committee, consisting of 8 distinguished members of the Swedish and Norwegian research community, has been chosen to represent a wide range of fields in digital humanities as well as language technology. The steering committee’s first meeting was held in Gothenburg on February 19th and there are plans for another meeting already this spring. In the coming newsletters we will present each member of the steering committee as well as the members of the national coordination team.

Our next physical consortium meeting will be held in Uppsala on April 4-5. We encourage any researchers interested in Swe-Clarin to submit proposals for discussion topics, future work or collaborations. The proposals can be submitted via info@sweclarin.se latest on March 24th.

Calendar

14–15 Mars: NCN workshop in Oslo on Big language data
4–5 April: Consortium meeting in Uppsala.
Spring: Swe-Clarin-on-tour stops in Stockholm. The theme is the Nordic Museum questionnaires.
November: Another Swe-Clarin-on-tour stop in conjunction with SLTC, the Swedish Language Technology Conference, in Umeå.

Future News

If you have information for the newsletter, please send it to the Coordination Team (info@sweclarin.se). If you do not get this newsletter automatically, you can sign up for the news list here:
http://lists.sweclarin.se/mailman/listinfo/news_lists.sweclarin.se