

Karin Friberg

Ph.D. Candidate
Department of Swedish Language
Göteborg University
Sweden

karin.friberg@svenska.gu.se
<http://spraakdata.gu.se/svekf>



Research Summary

Dealing with compounds in the Swedish language is something very different from dealing with them in the English language. Not only is Swedish a much more inflecting language than English, but also the morphological process of compounding is extremely productive. New compounds are coined all the time. Around 10% of words in running text are compounds, and most of these compounds will not be found in dictionaries. Furthermore, compounds are written as one word, making interpretation more complex than in languages that write the compound components as separate words. Most research in Information Retrieval is done in English, but language specific features, such as the ones mentioned above, makes this research not suitable to be applied to languages such as Swedish.

This project is about Swedish medical compounds in information retrieval. It deals with strategies for handling domain specific compounds in index and query expansion. The plan is to make use of domain information contained in the compounds, letting this information guide the choice of what to include in the index or in an expanded query.

There is currently no Swedish test collection for the medical domain. Therefore the Department of Swedish Language has decided to construct one. Supervision of the building of this test collection is part of the research project. A corpus of medical documents is being collected and shortly the work with construction of topics and with assessments of relevance will be initiated.

Academic Biography

Karin received a Master of Arts in Computational Linguistics from Göteborg University. Her Master's Thesis was on the subject of decomposing Swedish compounds. It included adapting to Swedish a decomposition tool from Lernout & Hauspie. After graduation Karin had a position at Nordisk Språkteknologi AS working on the Swedish version of Voice Xpress, a commercial dictation system from Lernout & Hauspie Speech Products.

Prior to her studies in Computational Linguistics and NLP Karin received a university degree in Journalistics and worked for some time at the daily newspapers Helsingborgs Dagblad and Tranås Tidning. Karin has also worked as a translator from English to Swedish and as a teacher in Swedish for Foreigners.

Karin graduated from the high school Natural Science Program with the top grade in every subject. She has achieved the highest possible score in the Swedish Scholastic Assessment Test (SweSAT), something which is accomplished by less than 1% of the participants.