

2018-01184	Nina Tahmasebi	Beredningsgrupp: HS-J
------------	----------------	-----------------------

Utlysningsnamn: Forskningsbidrag Stora utlysningen 2018 (Humaniora och samhällsvetenskap)	Bidragsform: Projektbidrag
Projekttitel (svenska): Mot automatiska metoder för att upptäcka språkförändring	Sökt inriktning: Fri

Novelty and originality

6

1 - Poor, 2 - Weak, 3 - Good, 4 - Very Good, 5 - Very good to excellent, 6 - Excellent, 7 - Outstanding

The project is in the forefront of computational linguistics, especially quantitative semantics, and has the potential to aid in reviving the field of diachronic semantics within linguistics. It aims at using vector-space models to establish word sense automatically for Swedish. Based on this, change of senses will be detected and thereafter word replacements. Lexical replacement has received very little attention in the computational literature, and including this adds to the project's novelty and originality.

Apart from the contribution this would give to the development of computational linguistics, it would be of great value for e.g. historical linguistics and historical studies generally and to digital humanities at large.

Scientific quality of the proposed research

5

1 - Poor, 2 - Weak, 3 - Good, 4 - Very Good, 5 - Very good to excellent, 6 - Excellent, 7 - Outstanding

The proposal demonstrates awareness of difficulties in current techniques for inducing word senses and for comparing lexical semantics over time, e.g. the data requirements, the (normal) existence of multiple word senses, the existence of different types of change, and the possibility that senses change variably. Promising lines of research are sketched with the potential of overcoming these. Thought has been given as to how the different lines of work may be evaluated.

It has to be admitted though, that for a non-expert it remains a bit hazy exactly how aims that have eluded previous investigations will be reached, and the project plan would have benefited from elaborating on this a little bit. In addition, it is somewhat difficult to exactly judge the novelty of this Swedish project as compared to what has previously been done on English. Furthermore, as the project target semantic changes for Swedish, one could have expected some more references to earlier research on this, e.g. by Anna Vogel.

Merits of the applicant(s)

6

1 - Poor, 2 - Weak, 3 - Good, 4 - Very Good, 5 - Very good to excellent, 6 - Excellent, 7 - Outstanding

The PI has a solid list of publications, and is generally a well-qualified researcher within the field. The other participants also have strong and relevant merits. Taken together, the project team clearly has the relevant background for the different parts of the project. National and international network should promote the anticipated theoretical and methodological development.

Feasibility

3

1 - Not feasible, 2 - Partly feasible, 3 - Feasible

The project is ambitious but seems to be feasible. There are preliminary results that indicate that the planned methods work. The infrastructure at Språkbanken provides both lexical data and language tools and the involved personnel are obviously qualified for the task. Some concern is raised by the fact that the project is divided into four work packages, which build on one another, and if WP1 (to develop reliable automatic word sense induction methods for Swedish) fails, the remaining WP: s will not be possible to carry out. However, in total, the project must clearly be evaluated as feasible.

Overall assessment of the application's scientific 5

quality*

1 - Poor, 2 - Weak, 3 - Good, 4 - Very good, 5 - Very good to excellent, 6 - Excellent, 7 - Outstanding

This project can contribute to a new foundation for diachronic semantics. It promises novelty in a number of dimensions and sketches the computational linguistic aspects in remarkable detail. Furthermore, the project has the potential to contribute not only to computational linguistics but also to be of applied value for digital humanities and historical studies. It might also provide insights into the mechanisms of linguistic changes.