Studies of the Swedish term "handikapp", and related concepts, over time

Henrik Danielsson, Lars Ahrenberg, Lotta Holme, Staffan Bengtsson, Arne Jönsson
Linköping University, Sweden, Jönköping University, Sweden, The Swedish Institute for Disability Research
henrik.danielsson@liu.se

2018-11-07

LINKÖPING UNIVERSITY

Background

- In Swedish, the term "handikapp" was introduced as an umbrella term for the many different terms that denoted special types of disability.

- "Handikappad" was something a person was, but with the introduction of an environment related view, other words like "funktionsnedsättning" and later "delaktig" were more used.

- This is both a word replacement and a word sense replacement.
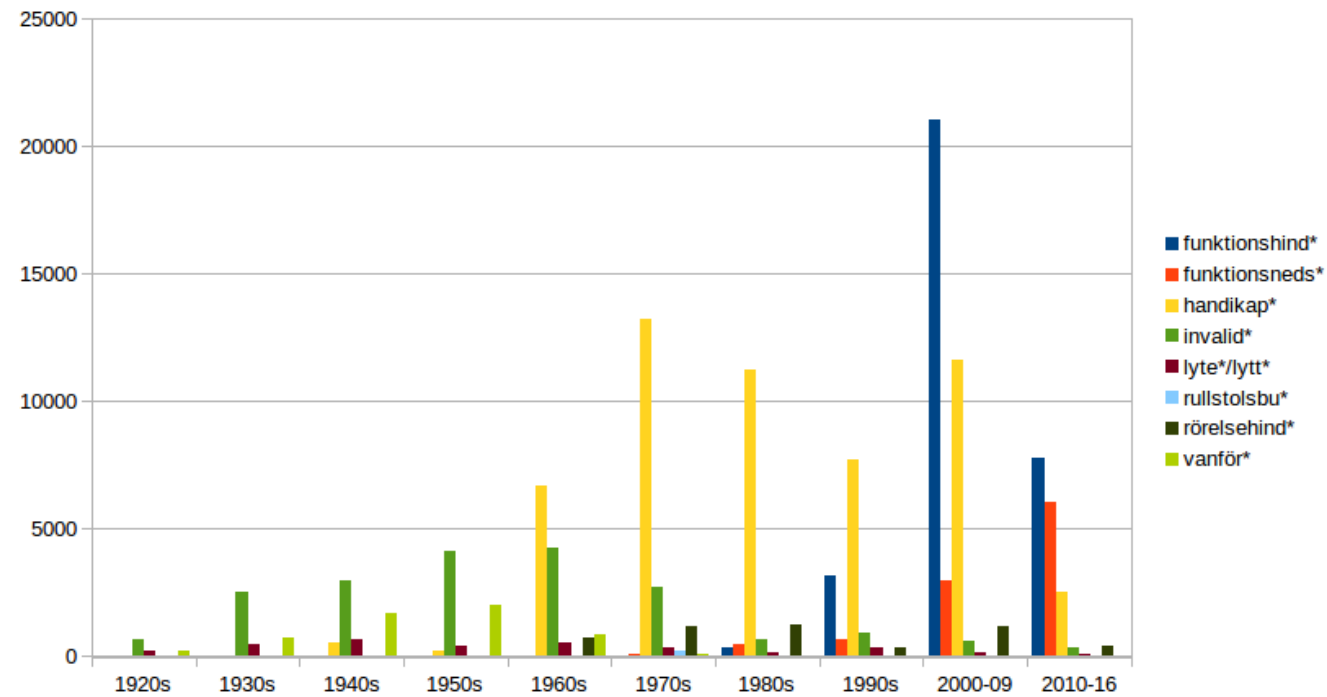
The project - ambition

- In this recently started project, language technologists, historians, and disability researchers are collaborating on exploring how different "handikapp"-related terms have been used.
  - Over time
  - Different corpora
  - Word frequencies
  - Word sense
  - Word polarity/sentiment
- Aim to first compare to previous qualitative studies and then learn new things
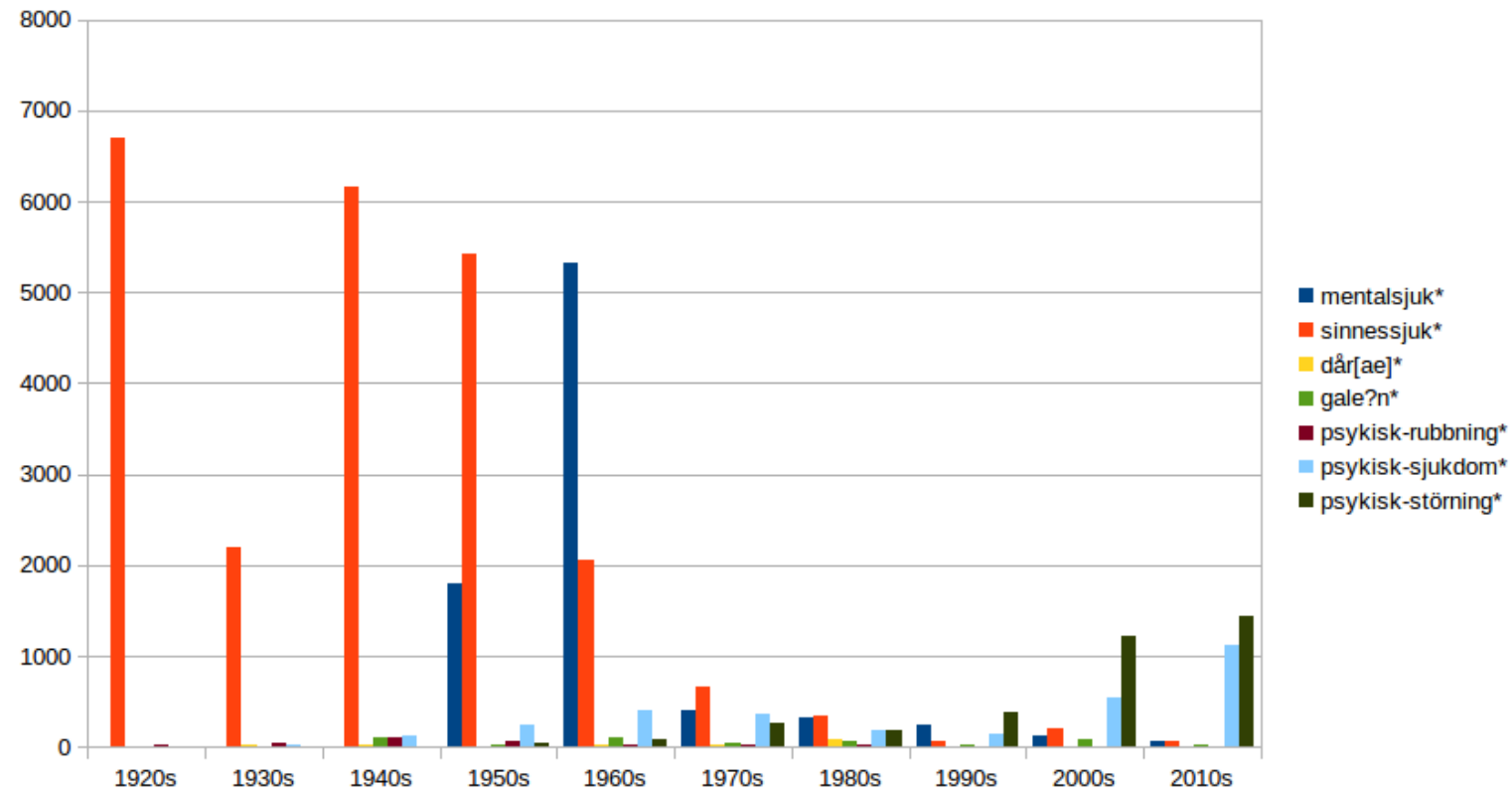
LINKÖPING
UNIVERSITY

The project – What we have done

- So far, word frequencies has been explored in one corpus, "Statens offentliga utredningar" (SOU) from 1922 to 2016 as found at Språkbanken
- The corpus was split into sub-corpora for each decade
- Division in three types of terms:
  - Physical disabilities
  - Mental disabilities
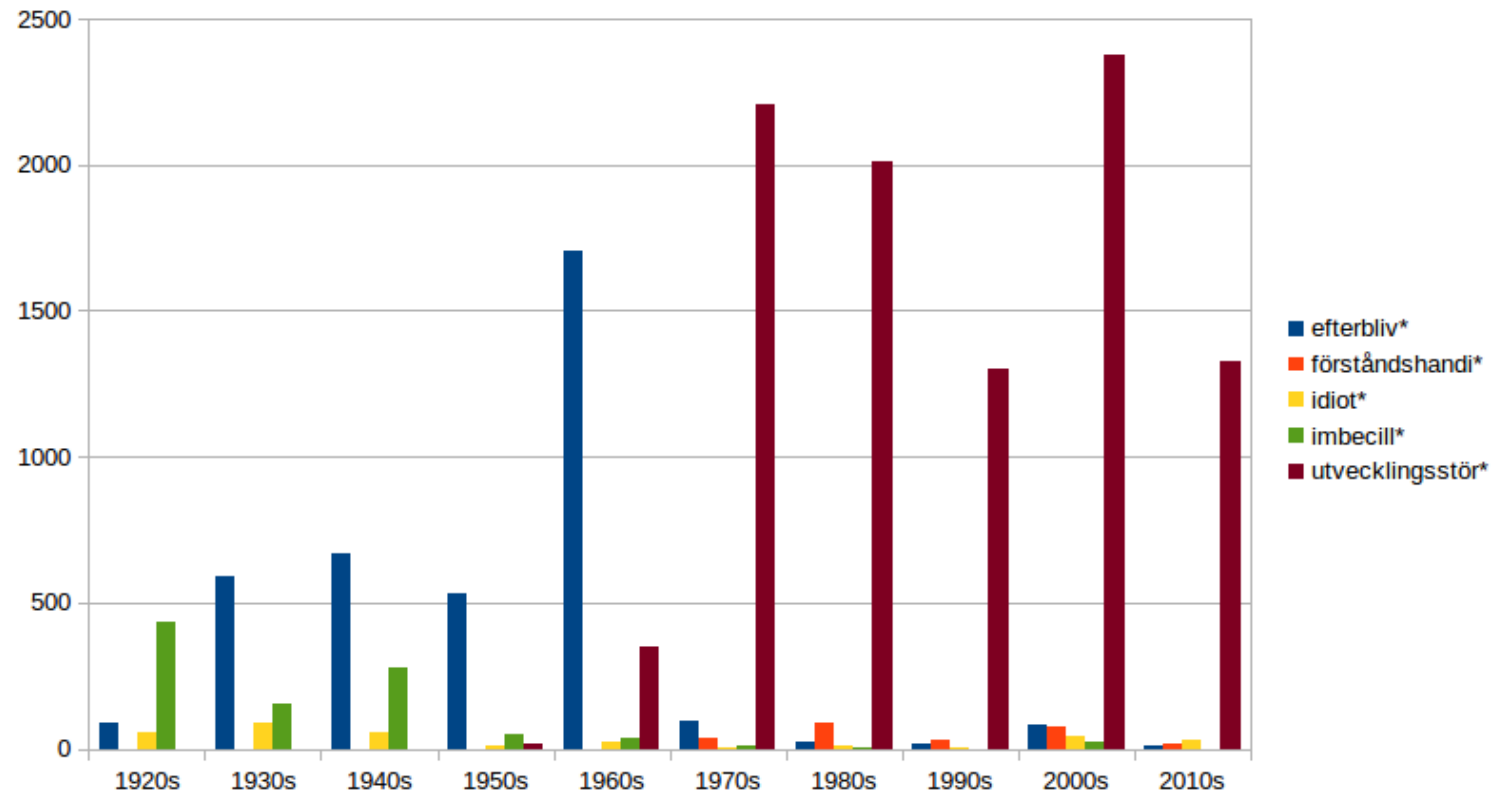 - Intellectual disabilities

LINKÖPING UNIVERSITY

Word frequencies – Physical disabilities/general

Word frequencies – Mental disabilities

Word frequencies – intellectual disabilities

Word similarities

- We have performed some experiments on word similarities using Word2vec.

- We used the default parameters for the Gensim package and the function wv.most similar to generate the most similar vectors in the models.

- Words were tokenized, stop words removed etc

LINKÖPING
UNIVERSITY

Word similarities 1970-1979

Funktionshinder	Funktionsnedsättning	Handikapp
Funktionsrubbningar	Psykiska	Handikappade
Funktionsrubbning	Handikapp	Rörelsehinder
Cirkulationsorganen	Psykisk	Handikappet

Word similarities 1980-1989 – Convergence

Funktionshinder	Funktionsnedsättning	Handikapp
Handikapp	Handikapp	Handikappade
Handikappade	Funktionsnedsättningar	Rörelsehinder
Funktionsnedsättningar	Funktionsnedsättningen	Funktionshinder

LINKÖPING
UNIVERSITY

Word similarities – "Delaktighet" Not convergence

1970-1979	Gemenskap	Trygghet	Samhörighet
1980-1989	Gemenskap	Samhällsarbetet	Ansvarstagande
1990-1999	Jämlikhet	Demokrati	Delaktiga
2000-2009	Jämlikhet	Självbestämmande	Respekt
2010-2019	Jämlikhet	Egenmakt	Autonomi

LINKÖPING
UNIVERSITY

What we want to do

- More sources (which?) –how long time before descisions on terminology change have implications in lay texts

- Word sense change – How to do that?

  - Is "funktionshinder" in relation to "miljö" and "funktionsnedsättning" in relation to the person?

  - Has it changed over time?

- Terminology has changed often in Swedish/Scandinavian but not in English. Is that because terms get negative sentiment over time and then changed?

LINKÖPING UNIVERSITY

How can it be done?

- How can we show that one term is taking over the role of another?
    - Change point detection (Kulkarni et al., 2015)
    - Temporal word analogies (Szymanski, 2017)
    - Joint learning of time-aware embeddings and alignments (Yao et al., 2018)
- How can we show that a term is changing its polarity?
    - Without document level polarities?
    - Using a time-static lexicon such as senSaldo?

LINKÖPING
UNIVERSITY

Problems

- **Mistakes in scanning of documents**
- **Hyphenations**

LINKÖPING
UNIVERSITY

Henrik Danielsson

henrik.danielsson@liu.se

www.liu.se

LiU LINKÖPING
UNIVERSITY