Bergen update

Petter Haugereid

Department of Language, Literature, Mathematics and Interpreting
Western Norway University of Applied Sciences

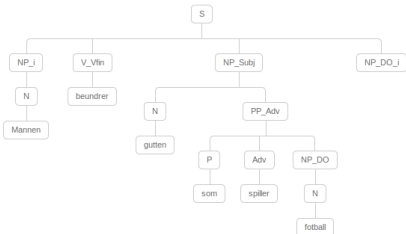Cambridge, July 14, 2019

A web application for automatic syntactic analysis

- 4.5 years of 3-year project over
- Aim of the application
  - take a sentence as input
  - parse it
  - give as output the kind of analysis that is requested by the student
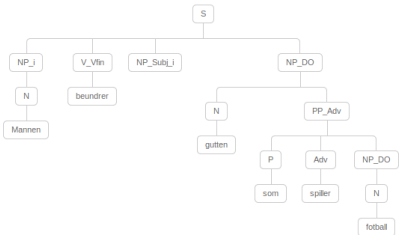
A web application for automatic syntactic analysis

- A simple Web site has been set up
- Produces constituent tree and a field analysis
- http://eple.hib.no/~phau/ (currently the server is down ...)

A web application for automatic syntactic analysis

Comparing dependency structures

Den første perioden var veldig vanskelig. SCORE: 0.875

- MATCH:
    - være SUBJ periode
    - periode ADJUNCT første
    - være PREDLINK vanskelig
    - vanskelig ADJUNCT veldig
- MISSING:
- XLE:
    - være SUBJ periode
    - periode ADJUNCT første
    - være PREDLINK vanskelig
    - vanskelig ADJUNCT veldig
- ACE:
    - være SUBJ periode
    - periode ADJUNCT første
    - være PREDLINK vanskelig
    - vanskelig ADJUNCT veldig
    - *periode ADJUNCT vanskelig*

Comparing dependency structures – scores

- 1.0: 99
- 0.9–1.0: 7
- 0.8–0.9: 24
- 0.7–0.8: 28
- 0.6–0.7: 39
- 0.5–0.6: 89
- 0.4–0.5: 39
- 0.3–0.4: 76
- 0.2–0.3: 87
- 0.1–0.2: 72
- 0.0–0.1: 48
- Below 0: 411
- Total: 1023

Plan

- Lexicon
  - + Converting the Norwegian LFG lexicon (NorGram) into an HPSG lexicon
- Statistical model
  - + Parsing treebanked sentences from the INESS Norwegian LFG treebank project with Norsyg
  - +/– Comparing dependency structures and create a HPSG treebank
- +/– Application
  - Using bash
  - Hosted by the Western Norway University of Applied Sciences

Demophin

http://clarino.uib.no/demophin/norsyg/