Examining Delayed Complements in Norwegian within an Incremental Left-Branching Grammar Framework

Petter Haugereid

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

Palacký University, Olomouc, July 8–10, 2024

Delayed complements

(1) So many people enrolled for the course that we had to move to a larger room.
(Huddleston and Pullum, 2002, 967)

Syntactic dependency

(2) More people walked into the room than they had
     expected

Norwegian data - minimal examples

(3)   a. Han er **så sen at   jeg smiler**.
             he   is so late that I   smile
             He is so late that I smile.

      b. **Så sen** er han **at   jeg smiler**.
             so late is he   that I   smile
             He is so late that I smile.

Data
○○○●○○○○

HPSG analysis
○○

Incremental analysis
○○○○○○○○○○○○○○○

References

John 3,16 in the Bible

(4) For **så høyt** har Gud elsket verden   **at   han ga**
     for so highly has God loved world-DEF that he   gave
     sin   Sønn, den enbårne   [...]
     REFL son,   the one and only [...]
     For God so loved the world that he gave his one and
     only Son [...]

Other licensing adverbs

Adverbs that require complement clauses in Norwegian:

- *så* and *såpass* 'so' function as degree adverbs modifying adjectives, adverbs, and prepositions
- *slik*, *sånn*, and *sådan* 'such' modify determiners or function alone, as adverbs.

Examples of *slik* and *sånn* 'such'

(5) Med **slik en kraft** traff   den **at jeg falt**.
     with such a   force struck it   that I   fell
     It struck with such a force that I fell.

(6) **Sånn en god lyd   i ørene hennes** var det **at hun**
     such a   good sound in ears   her   was it   that she
     ble varm.
     got warm
     It was such a good sound in her ears that she got warm

Example of *såpass* 'so'

(7) Såpass mange motsetninger finnes det i den   bibelske
    so   many contradictions exist in the biblical book
    boksamlingen at   akkurat det er et umulig   prosjekt.
    collection-DEF that exactly that is an impossible project
    So many contradictions exist in the biblical book
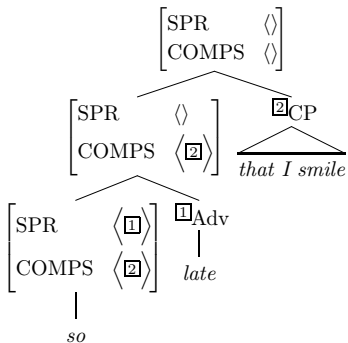   *collection that exactly that is an impossible project.*

Corpus searches – Leksikografisk bokmåskorpus (Fjeld *et al.*, 2020)

	Matches	Manual inspection	Estimate
så . . . at	59,671	29/50	34,609
såpass . . . at	1,346	46/50	1,238
slik . . . at	9,723	19/50	3,694
sånn . . . at	1,260	22/50	554
sådan . . . at	65	10/65	10
Total			40,105

Table: Estimated number of complement clauses licensed by adverbs in Leksikografisk bokmåskorpus

- About 1 in 40 is a delayed complement $\approx$ 1000

Data
○○○○○○○○

HPSG analysis
●○

Incremental analysis
○○○○○○○○○○○○○○○○○

References

ERG analysis (Flickinger, 2000)

MRS

MRS of the sentence *He was so late that I laughed.*

Incremental parsing

- Two constructions:
  - the delayed complement construction
  - the more regular construction, with the complement clause adjoined to the so-phrase
  - $\rightarrow$ uniform analysis
- Implementation:
  - Norsyg – an HPSG-inspired incremental typed feature structure grammar for Norwegian (Haugereid, 2009)
  - using the LKB system (Copestake, 2002) as a part of the Delph-In effort: https://github.com/delph-in/docs/wiki
- Design:
  - division between a parse tree and a constituent tree (Haugereid and Morey, 2012),
  - utterances are parsed in a bottom-up fashion, incrementally, from left to right
  - $\rightarrow$ completely left-branching tree structure

The analysis of delayed complement clauses

- The central assumption of the analysis drawn from Huddleston and Pullum (2002, 967):
  - the complement clause consistently appears at the end of the clause
  - it is a complement of clause structure, rather than the licensing adverbs
- Adverb licenses the complement clause via a feature LC (Licensed Complement)
- The feature ascends the tree until it triggers a rule, initiating the parsing of a complement clause

Data
○○○○○○○○

HPSG analysis
○○

Incremental analysis
○○●○○○○○○○○○○○○○

References

Anaysis of sentence with delayed complement
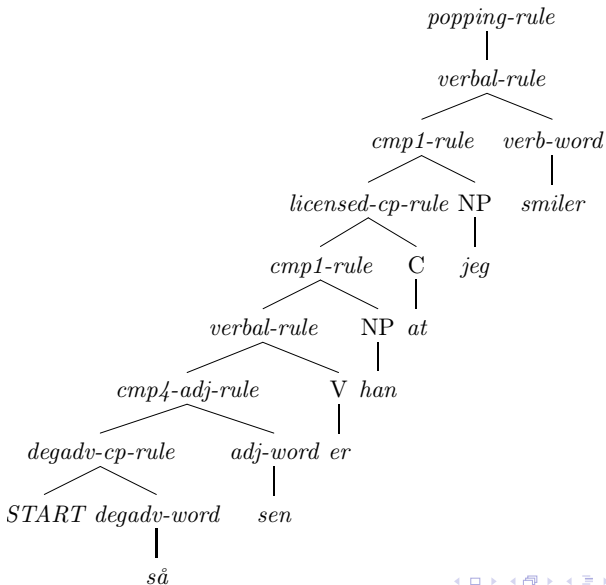
Analysis consists of five components

Five components:

1. lexical entries for the licensing degree adverbs
2. type hierarchy of predicate types for licensing adverbs
3. a rule for the licensing adverbs
4. a feature LC (Licensed Complement)
5. a rule for the licensed complement

1. Lexical entry for 'so'

$$(8) \quad \begin{bmatrix} \textit{degadv-word} \\ \text{STEM} \quad \left\langle \text{``så''} \right\rangle \\ \text{HEAD} \quad \begin{bmatrix} \textit{degadv} \\ \text{MOD} \left\langle \left[\text{HEAD } \textit{adj-adv-prep} \right] \right\rangle \end{bmatrix} \\ \text{KEYREL} \quad \left[\text{PRED } \textit{så\_deg} \right] \end{bmatrix}$$

2. Type hierarchy of predicate types for the degree adverb *så 'so'*

predsort

degadv+

comp–   comp+

så_deg

så_deg_rel   så_deg-cp_rel
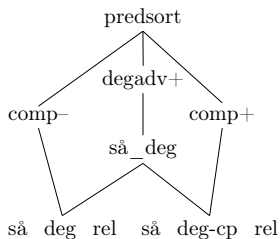
3. Rule for attaching degree adverb that requires a complement clause

$$
\begin{bmatrix}
\textit{degadv-cp-rule} \\
\text{CAT} \quad \boxed{5} \\[4pt]
\text{PREMOD} \quad \boxed{2}\left[\text{CONT}|\text{HOOK}\begin{bmatrix}\text{LTOP} & \boxed{3} \\ \text{INDEX} & \boxed{4}\end{bmatrix}\right] \\[12pt]
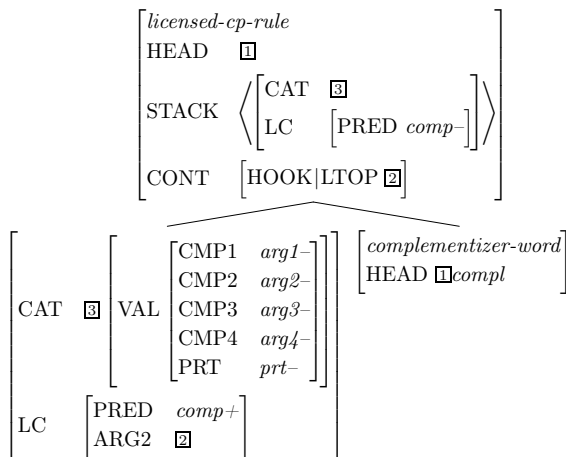\text{LC} \quad \boxed{1}\Big[\text{PRED} \quad \textit{comp+}\Big] \\[8pt]
\text{C-CONT} \quad \left\langle \, ! \, \boxed{1}\begin{bmatrix}\text{LBL} & \boxed{3} \\ \text{ARG0} & \textit{event} \\ \text{ARG1} & \boxed{4}\end{bmatrix} ! \right\rangle
\end{bmatrix}
$$

$$
\begin{bmatrix}\text{CAT} & \boxed{5} \\ \text{LC} & \Big[\text{PRED } \textit{degadv+}\Big]\end{bmatrix}
\quad
\begin{bmatrix}
\textit{degadv-word} \\
\text{HEAD}|\text{MOD}\big\langle\boxed{2}\big\rangle \\
\text{KEYREL } \boxed{1}\Big[\text{PRED } \textit{degadv+}\Big]
\end{bmatrix}
$$
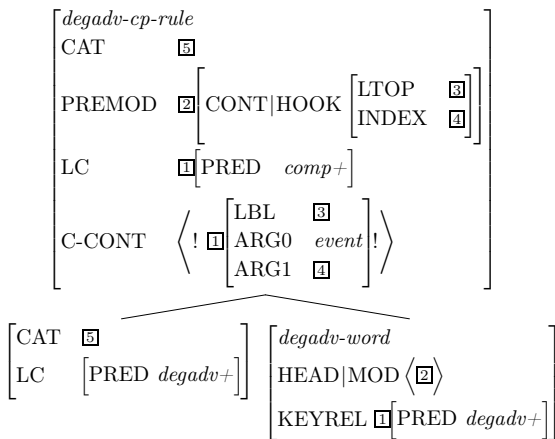
5. Rule for attaching complementizer initiating CP licensed by degree modifier
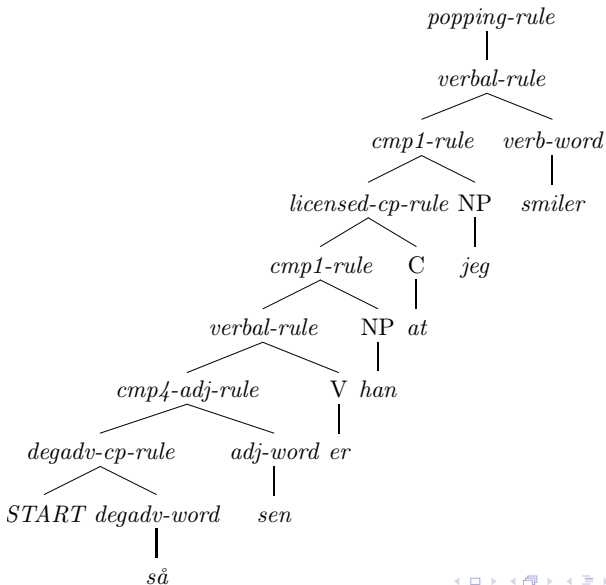
$$\begin{bmatrix} \textit{licensed-cp-rule} \\ \text{HEAD} \quad \boxed{1} \\ \text{STACK} \quad \left\langle \begin{bmatrix} \text{CAT} & \boxed{3} \\ \text{LC} & \left[\text{PRED } \textit{comp-} \right] \end{bmatrix} \right\rangle \\ \text{CONT} \quad \left[\text{HOOK}|\text{LTOP } \boxed{2} \right] \end{bmatrix}$$

$$\begin{bmatrix} \text{CAT} & \boxed{3} & \begin{bmatrix} \text{VAL} & \begin{bmatrix} \text{CMP1} & \textit{arg1-} \\ \text{CMP2} & \textit{arg2-} \\ \text{CMP3} & \textit{arg3-} \\ \text{CMP4} & \textit{arg4-} \\ \text{PRT} & \textit{prt-} \end{bmatrix} \end{bmatrix} \\ \text{LC} & \begin{bmatrix} \text{PRED} & \textit{comp+} \\ \text{ARG2} & \boxed{2} \end{bmatrix} \end{bmatrix} \quad \begin{bmatrix} \textit{complementizer-word} \\ \text{HEAD } \boxed{1}\textit{compl} \end{bmatrix}$$
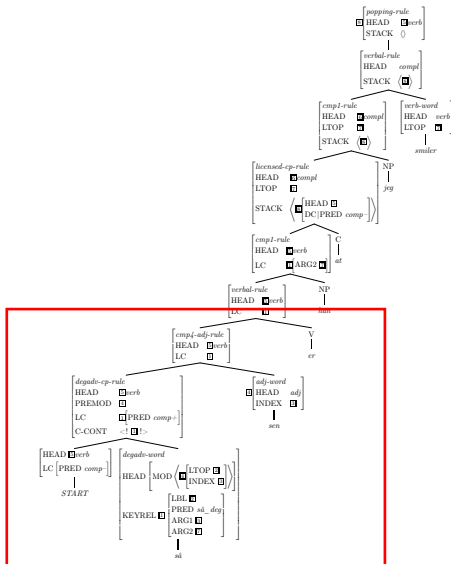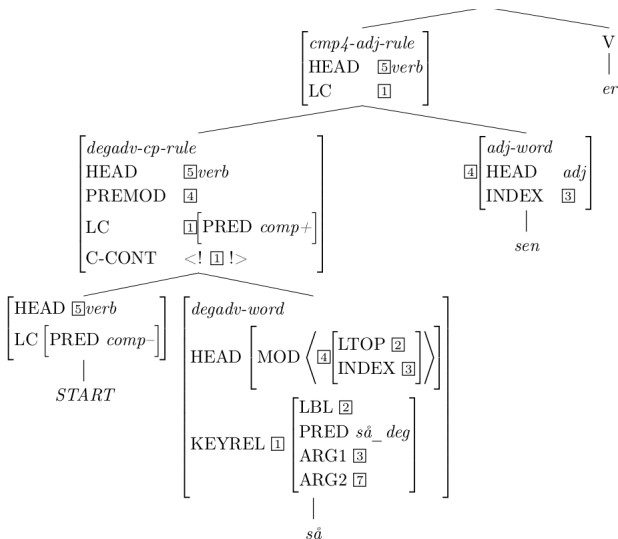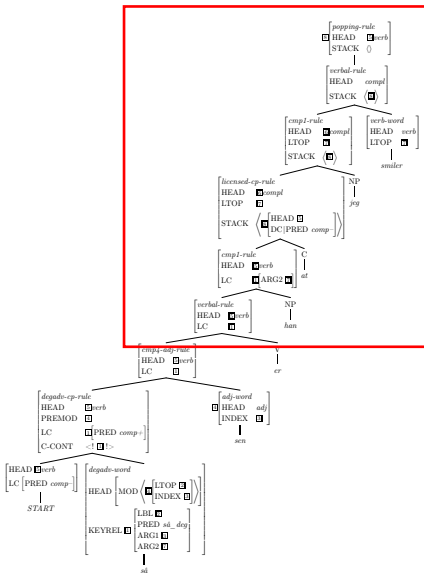
Anaysis of sentence with delayed complement

Anaysis of sentence with delayed complement

Anaysis of sentence with delayed complement

Anaysis of sentence with delayed complement

Data
○○○○○○○○

HPSG analysis
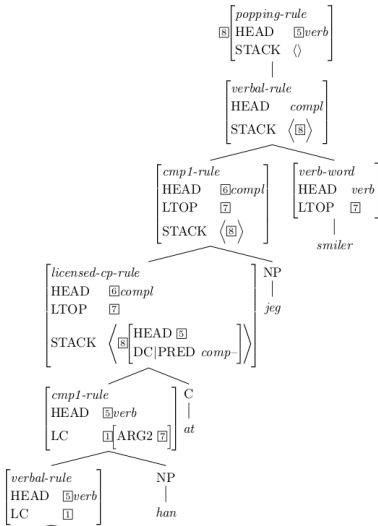○○

Incremental analysis
○○○○○○○○○○○○○●○○○○○

References
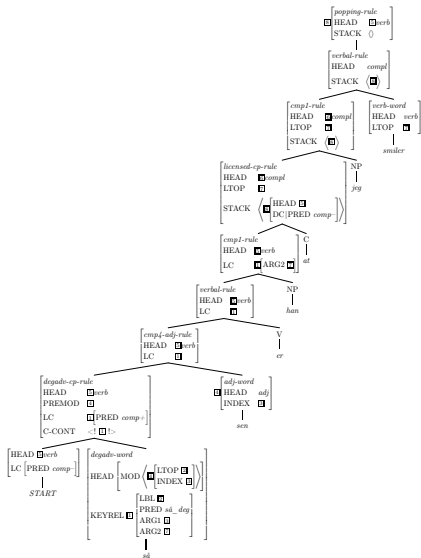
Anaysis of sentence with delayed complement

# Anaysis of sentence with delayed complement

Can this analysis be adapted by other HPSG grammars?

- When a grammar is not incremental, one cannot expect the LC feature to always go from first daughter to mother
- One would have to tag the LC feature with either the first or the second daughter
- But we cannot predict whether the LC feature in an NP comes from the first or the second daughter
- *a good sound*
  - From the right: *a so good sound*
  - From the left: *such a good sound*

Data
○○○○○○○○

HPSG analysis
○○

Incremental analysis
○○○○○○○○○○○○○○○●○

References

Getting the LC feature from the right daughter of an NP
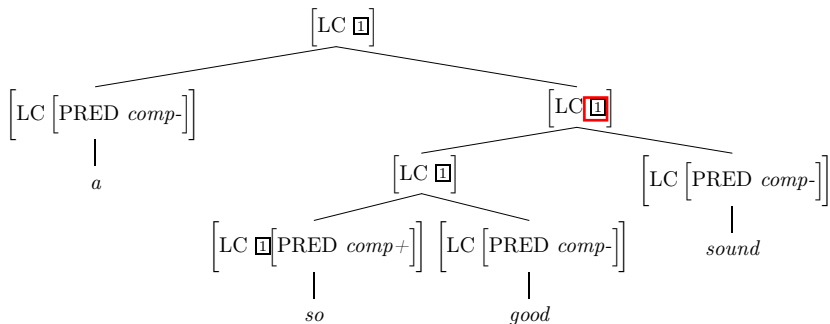
Getting the LC feature from the left daughter of an NP
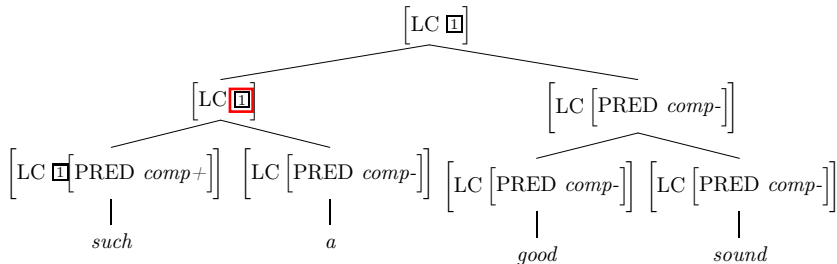
Conclusion

An analysis of delayed complements is presented

- CPs licensed by adverbs like *så* 'so' assumed to be complements of the clause rather than complements of the adverb

- a consistent analysis can be applied regardless of the proximity between the so-phrase and the CP

- accommodating scenarios where the phrase with the licensing adverb is either adjacent to or distant from the complement clause, while maintaining a uniform analysis throughout

- Not obvious how the analysis can be adopted in a regular HPSG grammar

Data
○○○○○○○○

HPSG analysis
○○

Incremental analysis
○○○○○○○○○○○○○○○○

References

References

Copestake, A. (2002). *Implementing Typed Feature Structure Grammars*. CSLI publications.

Fjeld, R. V., Nøklestad, A., and Hagen, K. (2020). Leksikografisk bokmålskorpus (LBK) – bakgrunn og bruk. *Oslo Studies in Language*, **11**(1), 47–59.

Flickinger, D. P. (2000). On building a more efficient grammar by exploiting types. *Natural Language Engineering*, **6**(1), 15–28.

Haugereid, P. (2009). *Phrasal subconstructions: A constructionalist grammar design, exemplified with Norwegian and English*. Ph.D. thesis, Norwegian University of Science and Technology.

Haugereid, P. and Morey, M. (2012). A left-branching grammar design for incremental parsing. In S. Müller, editor, *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, pages 181–194.

Huddleston, R. and Pullum, G. K. (2002). *The Cambridge Grammar of the English Language.* Cambridge University Press.