

# Datorlingvistisk grammatik

## OH-serie 5

http://stp.lingfil.uu.se/~matsd/uv/uv09/dg/



Mats Dahllöf  
Institutionen för lingvistik och filologi  
Januari 2009

1

### SWB's grammatik översikt

Grammatiken "handlar om" särdragsstrukturer och (syntax)trääd. Bl.a. följande komponenter ingår:

- Typhierarkin, särdragsdeklarationer och begränsningar på särdrags värden (säger hur våra särdragsstrukturer får se ut).
- Lexikoningångar (licensierar [tillåter] trädens löv).
- Regler (licensierar andra noder i träden).
- Principer (mer generella begränsningar).

3

### Komplementval (COMPS)

- Kapitel 3: *itr*, *str* och *dtr*, atomära etiketter bara.
- Långt fler möjliga COMPS-värden måste finnas. (Exempel?)
- Kapitel 3-lösningen skulle ge motsvarande multiplikation av regler. Redundans. Brist på systematik.
- I semantiken kommer vi att behöva knyta ihop huvudordets betydelse med komplementens. Hur göra det?
- Lösning: Samla ihop särdragsstrukturer för *saknade* satsdelar i en lista.

5

Datorlingvistisk grammatik I — HT 2007 (Mats Dahllöf)

### Huvud-komplement-regeln (kap 4)

Nu får vi EN generell huvud-komplementregel av följande slag. (Fler detaljer kommer.)

$$\left[ \begin{array}{c} \textit{phrase} \\ \text{VAL} \left[ \begin{array}{c} \text{COMPS } \langle \rangle \\ \text{SPR } \boxed{A} \end{array} \right] \end{array} \right] \rightarrow \mathbf{H} \left[ \begin{array}{c} \textit{word} \\ \text{VAL} \left[ \begin{array}{c} \text{COMPS } \langle \boxed{1}, \dots, \boxed{n} \rangle \\ \text{SPR } \boxed{A} \end{array} \right] \end{array} \right] \boxed{1} \dots \boxed{n}$$

Notera *word/phrase!* Komplementen blir systrar till *huvudordet*. **H** pekar ut huvudet, så att HFP appliceras.  $\boxed{A}$ : SPR-listan, som "bara" perkolerar. (Behöver ej skrivas i regeln. Se Valensprincipen.)

7

## CIRKA FÖRELÄSNING 6-7

Obs! Olika formaliseringar i SWB kapitel 3 och 4! Gradvisa revisioner.

- Logiska uppbyggnaden hos SWB's grammatik.
- Särdragsstrukturer och syntax(träd).
- Listbaserad hantering av valens: lexikoningångar och syntaxregler.
- Fler principer: enklare regler.
- Valenshanteringen i SWB's grammatik.
- Frasstrukturreglerna (efter valensprincipen).
- Kasusstyrning, modifierare

2

### Specificerare och komplement (i SWB)

- Komplementen är systrar till ett ord (*word*-entitet).
- En specificerare är syster till en fras (*phrase*-entitet).
- Särdragen SPR och COMPS kopplas ihop som särdrag i särdragsstrukturer av typen *val-cat*, men innehållet ändras från kapitel 3 till kapitel 4.
- Dessa *val-cat*-entiteter är värden på särdraget VAL(ence), som hör ihop med *expression*-entiteter (alltså både ord och fraser). (VAL är ett systemsärdrag till HEAD.)

4

### COMPS-listor (från kap 3 till 4)

- "COMPS *itr*" blir "COMPS  $\langle \rangle$ ".
- "COMPS *str*" blir "COMPS  $\langle \text{NP} \rangle$ ".
- "COMPS *dtr*" blir "COMPS  $\langle \text{NP}, \text{NP} \rangle$ ".
- NP förkortning för särdragsmatrisen för nominalfraser.
- Nu kan vi plocka ihop COMPS-listor av valfri längd och med valfria komponenter.

6

Datorlingvistisk grammatik I — HT 2007 (Mats Dahllöf)

### Specificerar-huvud-regeln

$$\left[ \begin{array}{c} \textit{phrase} \\ \text{VAL} \left[ \begin{array}{c} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \rangle \end{array} \right] \end{array} \right] \rightarrow \boxed{1} \mathbf{H} \left[ \begin{array}{c} \text{VAL} \left[ \begin{array}{c} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \boxed{1} \rangle \end{array} \right] \end{array} \right]$$

$\boxed{1}$ : specificeraren.

Huvudet är COMPS  $\langle \rangle$ , d.v.s. eventuellt krävda komplement ingår.

Huvudet får vara *word* (utan komplement) eller *phrase* (med komplement).

8

## Valensprincipen

Följande default-princip antas:

### Valensprincipen

Såvida inte en regel säger annat, så är moderns värden på VAL-särdragen (SPR och COMPS) identiska med dem på huvuddottern.

Därmed kan vi förenkla våra reglers utseende.

Valensprincipen och HEAD-särdragsprincipen reglerar informationens perkolation från huvud till moder.

9

Datorlingvistisk grammatik I — HT 2007 (Mats Dahlöf)

### Regler, nya formuleringar

Valensprincipen tar hand om "oanvända" valenssärdrag.

#### Huvud-komplement-regeln

$$\left[ \begin{array}{c} \textit{phrase} \\ \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \end{array} \right] \rightarrow \mathbf{H} \left[ \begin{array}{c} \textit{word} \\ \text{VAL} \left[ \text{COMPS} \langle \boxed{1}, \dots, \boxed{n} \rangle \right] \end{array} \right] \boxed{1} \dots \boxed{n}$$

#### Specificerar-huvud-regeln

$$\left[ \begin{array}{c} \textit{phrase} \\ \text{VAL} \left[ \text{SPR} \langle \rangle \right] \end{array} \right] \rightarrow \boxed{1} \mathbf{H} \left[ \begin{array}{c} \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \\ \text{SPR} \langle \boxed{1} \rangle \end{array} \right]$$

Huvudet "väljer" specificerare och komplement.

11

## Valens, SWB, kapitel 4

- Begreppen *itr*, *str* och *dtr* i kapitel 3 har fått en analys.
- En regel täcker in alla kopplingar huvudord-komplement. Vi säger direkt och på en gång att ett huvudord måste kompletteras med alla sina komplement.
- De ytterligare valensramar (utöver *itr*, *str* och *dtr*) som behövs kan lätt fångas med denna "metod".
- Specificerare hanteras enligt samma princip som komplement. Huvuden kan "välja" komplement.

13

## Fraser och ord, ett exempel

Så, ett egennamn som *Pelle*...

- ... kan beskrivas av matrisen  $\left[ \begin{array}{c} \textit{word} \\ \text{HEAD} \textit{ noun} \\ \text{VAL} \left[ \begin{array}{c} \text{COMPS} \langle \rangle \\ \text{SPR} \langle \rangle \end{array} \right] \end{array} \right]$  OCH fungera som en NP.

Vi behöver revidera våra förkortningar av NP, NOM, S och VP i enlighet med detta.

15

## Valensprincipen

Följande default-princip antas:

### Valensprincipen

Såvida inte en regel säger annat, så är moderns värden på VAL-särdragen (SPR och COMPS) identiska med dem på huvuddottern.

Därmed kan vi förenkla våra reglers utseende.

Valensprincipen och HEAD-särdragsprincipen reglerar informationens perkolation från huvud till moder.

10

## Huvud-modifierar-regeln

För efterställda modifierare, som väljer sitt huvud via särdraget MOD.

$$\left[ \textit{phrase} \right] \rightarrow \mathbf{H} \left[ \boxed{1} \left[ \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \right] \left[ \begin{array}{c} \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \\ \text{MOD} \langle \boxed{1} \rangle \end{array} \right] \right]$$

Obs! Huvudet blir "valt".

12

## Fraser och ord

- Distinktionen *word* vs *phrase* får en minskad betydelse i SWB kapitel 4.
- Det viktiga är hur "mättade" (saturated) uttryck är. Mättad = alla "nödvändiga" specificerare och komplement är på plats, d.v.s. SPR ⟨ ⟩ och COMPS ⟨ ⟩.
- Vi behöver ingen regel som tar oss från ord med COMPS ⟨ ⟩ till motsvarande fras (likande den i kapitel 3).

14

## Nya förkortningar: NP

$$\text{NP} = \left[ \begin{array}{c} \text{HEAD} \textit{ noun} \\ \text{VAL} \left[ \begin{array}{c} \text{COMPS} \langle \rangle \\ \text{SPR} \langle \rangle \end{array} \right] \end{array} \right]$$

(vi specificerar inte *word* eller *phrase*)

Substantiv (ord) kan vara omättade både vad gäller SPR och COMPS.

16

## Nya förkortningar: NOM

$$\text{NOM} = \left[ \begin{array}{l} \text{HEAD } \textit{noun} \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle X \rangle \end{array} \right] \end{array} \right]$$

(X står för godtycklig särdragsstruktur.)

(vi specificerar inte *word* eller *phrase*)

Substantiv (ord) *kan* vara omättade både vad gäller SPR och COMPS.

17

## Nya förkortningar: S

$$\text{S} = \left[ \begin{array}{l} \text{HEAD } \textit{verb} \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \rangle \end{array} \right] \end{array} \right]$$

Verb (ord) *kan* vara omättade både vad gäller SPR och COMPS.

18

## Nya förkortningar: VP

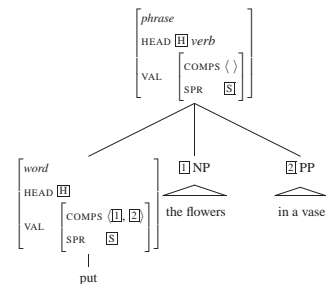
$$\text{VP} = \left[ \begin{array}{l} \text{HEAD } \textit{verb} \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle X \rangle \end{array} \right] \end{array} \right]$$

(X står för godtycklig särdragsstruktur.)

Verb (ord) *kan* vara omättade både vad gäller SPR och COMPS.

19

## Lokalt träd, huvud + två komplement (SWB, s. 97)



20

## Kasus och kasusstyrning

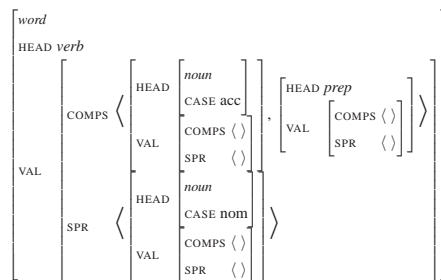
- Kasusdistinktionen visar sig endast på vissa personliga pronomena: Nominativ: *I, she, he, they*. Akkusativ: *me, her, him, them*. (Svenska: liknande situation.)
- Substantivshövadade NP underbestämde vad gäller kasus: *a letter* kan vara både subjekt och objekt.

TYPE	FEATURES/CONSTRAINTS	IST
<i>noun</i>	[CASE {nom, acc}]	<i>agr-pos</i>

21

## Verb som tar två komplement

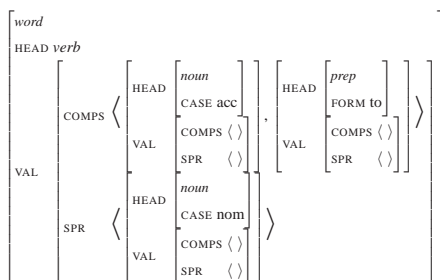
Som *put* i föregående träd, med kasusstyrning:



22

## PP-komplement, val av preposition

T.ex. *send* (FORM: SWB, 316. Fundera på *put!*):



23

## Huvud-modifierar-regeln

$$[\textit{phrase}] \rightarrow \mathbf{H} \left[ \textit{VAL} \left[ \textit{COMPS} \langle \rangle \right] \right] \left[ \textit{VAL} \left[ \textit{COMPS} \langle \rangle \right] \left[ \textit{MOD} \langle [1] \rangle \right] \right]$$

MOD anger till vilka huvuden uttrycket ifråga kan knytas som modifierare.

Obs! Bara efterställda modifierare!

Hur göra med framförställda och hur reglera ordföljd?

24

## Valenssärdragen, översikt

- SPR: huvudet väljer en eller ingen specificerare.
- COMPS: huvudet väljer en sekvens av komplement (0, 1, 2, 3, kanske fler, det är öppet).
- MOD: bestämningen (modifieraren närmare bestämt) väljer ett huvud. (Alltså: *omvänt* mot hur SPR och COMPS fungerar.)