

Lars Ahrenberg  
Institutionen för lingvistik  
Uppsala universitet  
Box 513  
751 20 UPPSALA

## DE GRAMMATISKA BESKRIVNINGARNA I SVE.UCP

1. Inledning. SVE.UCP är en parser för svenska som är under utveckling vid Uppsala Centrum för Datorlingvistik (UCDL). Grammatik och lexikon i SVE.UCP är proceduralt formulerade i en särskild formalism anpassad till parserns arbetande del, processorn, kallad UCP:n (the Uppsala Chart Processor)<sup>(1)</sup>. Varje sträng som SVE.UCP accepterar ansätts en eller flera grammatiska beskrivningar. Ämnet för den här korta artikeln är just de grammatiska beskrivningarna, i synnerhet deras formella egenskaper och de möjligheter och begränsningar som ligger i dessa. Här och där kommer jag också att göra jämförelser med andra grammatiska beskrivningar av samma format, i första hand med de som används i Lexical Functional Grammar (LFG; Bresnan, 1982).

En grammatisk beskrivning kan utföras med olika syften. Den kan t ex utgöra ett steg på vägen i ett program som simulerar textförståelse eller vara ett datum i en textundersökning. Syftet bestämmer givetvis beskrivningens utseende till stor del. SVE.UCP utvecklas för närvarande inte med inriktning på någon speciell tillämpning, men är tänkt att vara användbar för bl a de nämnda syftena. Dess grammatiska beskrivningar är tämligen innehållsrika och omfattar både morfologi och syntax. Vad gäller syntaxen har utgångspunkten för beskrivningarna varit Telesman (1974) även om vi ibland har funnit det motiverat att avvika från hans analyser. I alla händelser är beskrivningarna innehållsligt sett mycket traditionella. i (1) visas den grammatiska beskrivning som SVE.UCP ger meningens Eva läser det<sup>(2)</sup>.

2. Format och innehåll. SVE.UCP:s grammatiska beskrivningar kan enklast betraktas som en strukturerad mängd av påståenden. Ett påstående har formen av ett par vars första element vi kallar attribut och vars andra element kallas värde. Ett attribut är alltid en teckensträng medan ett värde kan vara en teckensträng, en grammatisk beskrivning eller en referens till ett annat värde. i (2)-(4) illustreras påståenden av alla tre typerna hämtade från (1).

(2) NUMB = SING

(3) OBJ = (PHR.CAT = NP  
NP.FEAT = (GENDER = NEUTR  
NUMB = SING  
PERS = 3  
CASE = NIL<sup>(3)</sup>  
LEX = !DEN)  
HEAD = (CHARS = (1 = D  
2 = E  
3 = T)  
WORD.CAT = PRO  
CASE = NIL))

(4) SUBJ = (\* SENT FUND)

(1) EVA LASER DET.:

(\* = (SENT.CAT = SENT  
SENT.FEAT = (CPLX = SIMPLE)  
SENT = (PHR.CAT = CL  
CL.FEAT = (TYPE = MAIN  
MODE = DECL  
TENSE = PRES)  
FUND = (PHR.CAT = NP  
NP.FEAT = (PROPR = +  
NUMB = SING  
GENDER = UTR  
CASE = NIL  
LEX = !EVA)  
NAME1 = (WORD.CAT = NOUN  
NOUN.FEAT = (SEX = FEM  
FIRST = +  
CASE = NIL)  
STEM = (CHARS = (1 = E  
2 = V  
3 = A)  
MORPH.CAT = NOUNSTEM)))  
PRED = (VP.FEAT = (FIN = +  
LEX = !LXSA)  
VFIN = (WORD.CAT = VERB  
VERB.FEAT = (FIN = +)  
STEM = (CHARS = (1 = L  
2 = X  
3 = S)  
MORPH.CAT = VERBSTEM  
CONJ = -ER)  
FLEX = (CHARS = (1 = E  
2 = R)  
MORPH.CAT = VERB.FLEX  
CONJ = -ER))  
OBJ = (PHR.CAT = NP

```

NP.FEAT = (GENDER = NEUTR
           NUMB = SING
           PERS = 3
           CASE = NIL <S>
           LEX = !DEN)
HEAD = (CHARS = (1 = D
                 2 = E
                 3 = T)
        WORD.CAT = PRO
        CASE = NIL)))
SUBJ = (* SENT FUND)))

```

(2) kan utläsas "Numerus är singularis.", men det är då inte klart vilket elements numerus som avses. Detta framgår emellertid av den plats som påståendet befinner sig på i beskrivningen, närmare bestämt av den väg, eller följd av attribut som leder fram till det. Det finns två vägar genom (1) som leder till ett attribut/värde-par av typen (2), nämligen **⟨\* SENT FUND.NP.FEAT⟩** och **⟨\* SENT PRED OBJ NP.FEAT⟩**. I det första fallet är det således fundamentets numerus som anges, i det andra objektets.

Uttryck av typen **⟨\* SENT PRED OBJ NP.FEAT⟩** eller **⟨\* SENT FUND⟩** som anger vägar genom en grammatisk beskrivning har själva värden. Deras värde är identiskt med det sista attributets värde. (4) anger alltså att värdet för attributet SUBJ är identiskt med värdet för attributet FUND, dvs att subjektet är identiskt med fundamentet i den givna meningen. Påståendet (3) ska tolkas så att objektet är ett element med den grammatiska beskrivning som anges av uttrycket till höger om likhetstecknet. Vilket objektet vi talar om framgår som förut av den väg som leder fram till beskrivningen, i detta fall **⟨\* SENT PRED⟩**. Märk att varje väg börjar i en och samma punkt, nämligen det särskilda attributet '\*'. En grammatisk beskrivning för en given mening kommer alltid ut som värde till detta attribut.

För varje attribut i den grammatiska beskrivningen finns exakt en väg som förbinder det med \*. Den grammatiska beskrivningen har således trädstruktur, med attributen som icke-terminala noder, \* som rotnod och "enkla" värden, dvs teckensträngar och referenser som terminala noder. Till skillnad från gängse frasstrukturträd är dock döttrarna till en given nod inte ordnade i förhållande till varandra, vilket gör att den linjära ordningen mellan konstituenterna i en mening inte representeras strukturellt. Hur man kan göra istället ska jag återkomma till nedan.

De grammatiska beskrivningarna i SVE.UCP lyder under ytterligare ett villkor. Det är att inget attribut får ha mer än ett värde eller, om vi tänker i trädstrukturer, att om ett attribut (en nod) förgrenar sig så är alla döttrar skilda attribut. Detta villkor ska vi kalla unicitetsprincipen<sup>(4)</sup>.

Trädstrukturen hos de grammatiska beskrivningarna gör det möjligt och naturligt att representera hierarkiska relationer i meningen strukturellt. Satsled representeras av delträd. (Omvändningen gäller dock inte, ett delträd behöver inte alltid svara mot ett satsled - se nedan). Avsaknaden av linjär ordning mellan beskrivningens delar medför att satsleden primärt beskrivs utifrån sina grammatiska funktioner (som subjekt, adverbial, etc) men dessutom ges uppgift om följande fyra egenskaper:

- 1/ Position i satsen,
- 2/ Morfosyntaktisk kategori,
- 3/ Grammatiska egenskaper (förutom kategori),
- 4/ Delkonstituenten.

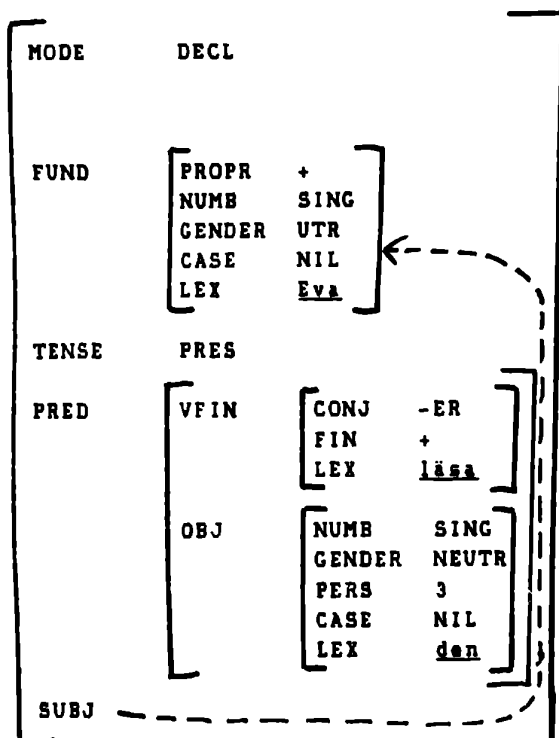
Position är inte direkt kodifierat i (1). Istället finns ett attribut CHARS, som anges för de minsta leden (morferna) och som anger den teckensträng som svarar mot dessa. För den nänskliga användaren är dock detta i de flesta fall tillräckligt för att man ska kunna avgöra vad som blivit analyserat som vad i beskrivningen. I (1) finns fyra förekomster av CHARS svarande mot de fyra morfer som strängen analyserats i.

Attribut som anger morfosyntaktisk kategori har alla sufficet CAT. Vi skiljer på fraskategorier (NP, PP, ADJP, ...), ordkategorier (NOUN, PREP, ADJ, ...) och morfkatgorier (NOUNSTEM, ADJSTEM, VERBFLEX, ...). Attribut som anger grammatiska särdrag är samlade under ett attribut med suffixet FEAT och ett förled som utgörs av ett kategorinamn. Om fraskategorin för ett led är NP samlas således dess övriga egenskaper upp i en beskrivning under attributet NP.FEAT, så som i (4). Attribut som anger delkonstituenten har inga särskilda kännetecken man ges så långt möjligt namn som antyder konstituenternas grammatiska funktion. På satsnivå har vi således attribut som SUBJ, PRED, OBJ och ADV.LOC, på frasnivå DET, MOD, HEAD, etc och på ordnivå STEM och FLEX. Attribut som gäller ett och samma led skrivs under varandra i en kolumn och ett värde skrivs alltid till höger om sitt attribut.

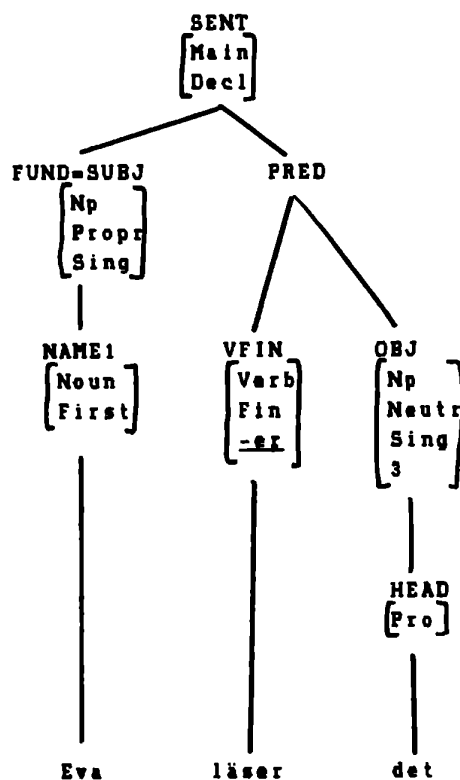
Formatet på SVE.UCP:s grammatiska beskrivningar representerar en utveckling av Kays (1977) "structures", som också är oordnade träd begränsade av unicitetsprincipen. Användningen av oordnade träd som representationsformat för grammatisk struktur har varit sparsam inom modern teoretisk linvistik, som i stället föredragit ordnade träd<sup>(5)</sup>. Det ökade intresset för grammatiska funktioner har emellertid lett till att oordnade träd börjat användas alltmer. De sk f-strukturer som används inom LFG har precis samma format men något annorlunda restriktioner på vad som kan vara ett möjligt värde. Kays (1982)

Functional Descriptions är också väsentligen oordnade träd. Det kan sägas att SVE.UCP:s utskriftsformat inte är det mest transparenta för att återge trädstruktur. En lådstruktur av den typ som används för LFG:s f-strukturer, illustrerad i (5), eller en graf av typen (6) som för skillnad mellan olika typer av attribut är åskådligare, men helt ekvivalenta representationsformer.

(5)



(6)



3. Unicitetsprincipen. Unicitetsprincipen har en komputationell motivering men också en lingvistisk. Den komputationella motiveringen är att unifieringsoperationen som bygger upp grammatiska beskrivningar inte vore väldefinierad annars. Den lingvistiska är att det helt enkelt normalt verkar vara så att paradigmatiska val bara görs en gång inom en given överordnad struktur. Många villkor som formulerats för lingvistisk struktur på olika beskrivningsnivåer kan ses som specialfall av denna princip. En av de mera kända är väl kasusgrammatikens One-case-per-clause-principle (Fillmore, 1968:21). I s k Relationell Grammatik finns motsvarande villkor formulerat för grammatiska funktioner (Johnson, 1977:155, Perlmutter, 1980:211). Inom X-barteori anses varje fras ha ett entydigt bestämt huvud (Jackendoff, 1977:53) och ibland begränsas dess "specificerare" explicit till högst en av varje slag (ibid. 104). Men det finns många fall där man inte gärna kan, eller vill, upprätthålla den. Det gäller i första hand samordnade konstruktioner, som i (7) och (8),

för dels är det omotiverat att ge de olika leden olika funktion, dels finns det ingen gräns för hur många de kan vara. Det finns också andra fall där det är rimligt att säga att vi har att göra med flera element av samma funktion, t ex vid sekvenser av satsadverbial som i (9), eller vid sekvenser av namn, som i (10). Perifera adverbial, som i (11), kan också i princip vara obegränsat många och man undviker många problem om man låter bli att skilja på dem.

(7) Man säljer ägg, potatis, grönsaker, blommor och mycket annat där.

(8) Han har en fin, billig, liten lägenhet.

(9) De har nog tyvärr antagligen redan glömt det.

(10) Tager du, Karl Teodor Erik Jakob Olsson, denna Eva Margareta ...

(11) Igår åt jag middag med Svante på Wermlandskällaren.

Det finns olika sätt att komma runt unicitetsprincipen. I LFG tillåter man mängder av f-strukturer som värden. F-strukturen för (11) kommer således att innehålla ett attribut ADJUNCTS vars värde är en mängd av tre f-strukturer, var och en svarande mot ett av adverbialen. I UCP:n används en speciell variabel över attributnamn, kallad :NEW, vid uppbyggnaden av beskrivningar som kan innehålla ett uppräkneligt antal led av samma slag. :NEW instansieras som heltalet '1' första gången det ges ett värde i en viss regel och därefter som ett en enhet större heltal då det på nytt tilldelas ett värde av samma regel. Attributen under CHARS i (1) illustrerar en användning av :NEW-variabeln. En annan illustreras i (12), som förenklat visar hur samordningen i (7) beskrivs av SVE.UCP.

(12) FUND = (PHR.CAT = NP  
NP.FEAT = (COORD = +)  
1 = ägg  
2 = potatis  
3 = grönsaker  
4 = blommor  
CONJ = och  
5 = mycket annat

Med detta sätt att beskriva samordningar och liknande sekvenser blir den linjära ordningen mellan de ingående leden också inkodad. Denna information är väsentlig om man använder beskrivningen som utgångspunkt för semantisk

tolkning (något som i LFG anses vara f-strukturens viktigaste uppgift eller om den används för generering. Anaforiska pronomen tolkas gärna olika beroende på var i en samordning de uppträder, vilket (13) och (14) illustrerar, och ordningsföljden mellan adverbial kan vara avgörande för deras relativa "scope". (15) har två tolkningar, men (16) har bara en.

(13) Olle och hans fru var här igår.

(14) Hans fru och Olle var här igår.

(15) På grund av kylan gick han inte hem.

(16) Inte gick han hem på grund av kylan (i alla fall).

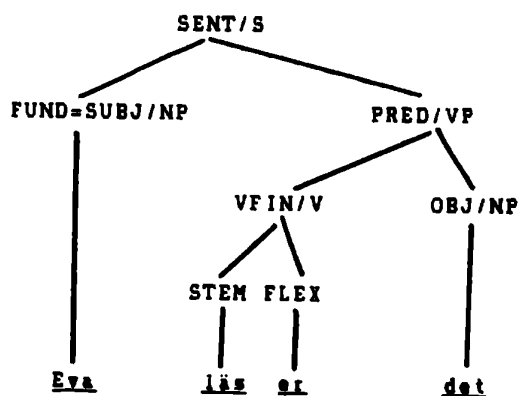
4. Frasstruktur. Eftersom SVE.UCP:s grammatiska beskrivningar inte återger linjär ordning strukturellt kan de naturligtvis inte heller återge frasstruktur strukturellt. Däremot kan frasstruktur beskrivas på annat sätt. Vi kan införa ett attribut ORDER vars värde ger information om ordningsföljden mellan delkonstituenten. Ett problem är då att vi inte vet vilka attribut som representerar delkonstituenten och vilka som representerar annan information om konstituenten. Men detta kan antingen specificeras globalt eller genom att vi känner igen delkonstituenten på att de har ett värde som anger morfosyntaktisk kategori.

Med införandet av ORDER måste vi också införa en ny typ av värden. I det enklaste fallet är värdet en ordnad lista av de attributnamn (vägar) som representerar delkonstituenten, svarande mot en viss permutation av dem. (1) kan t ex kompletteras med följande enkla beskrivningar för att vi ska ha en fullständig specifikation av frasstrukturen:

under SENT: ORDER = ⟨FUND, PRED⟩  
under PRED: ORDER = ⟨VFIN, OBJ⟩  
under VFIN: ORDER = ⟨STEM, FLEX⟩

Vi antar då att konstituenten med multipla funktioner ordnas via det attribut som har deras grammatiska beskrivning till värde, så att t ex SUBJ i (1) inte behöver ordnas explicit eftersom dess värde är en referens till FUND. Den information som då finns under ORDER och CHARS svarar precis mot den som ges av frasstrukturträdet (17), som vi för tydlighets skull också kan etikettera med morfosyntaktiska kategorier.

(17)



I det allmänna fallet finns emellertid ingen perfekt korrespondens mellan funktionella enheter och fraser, i alla fall om de senare uppfattas som sammanhängande. Många funktionella enheter är diskontinuerliga. Vi har inte bara fall av extraherade led utan också många fall av mera lokala diskontinuiteter. (18)-(22) visar några olika typer. De extraherade leden ger egentligen minst problem eftersom de ordnas efter sina tematiska funktioner, dvs som fundament, fokus etc. För beskrivningen av de andra fallen kan vi införa variabler av olika slag (se Kay, 1981). Ta (19) som exempel. På satsnivå kan vi anta att vi analyserat (19) i ett tidsadverbial (då), ett predikat (gav upp), ett subjekt (han) och

- (18) Glasögonen brukar hon glömma var hon har lagt. (EXTRAHERADE LED)  
(19) Då gav han äntligen upp. (VERB + PARTIKEL)  
(20) En bättre åkare än Stenmark finns inte. (KOMPARATIVT ATTRIBUT)  
(21) Jag såg en duva igår som var alldeles vit. (KORRELAT + RELATIV)  
(22) Nästa år, sägs det, kommer kungen att abdikera. (PARENTETISKT INSKOTT)

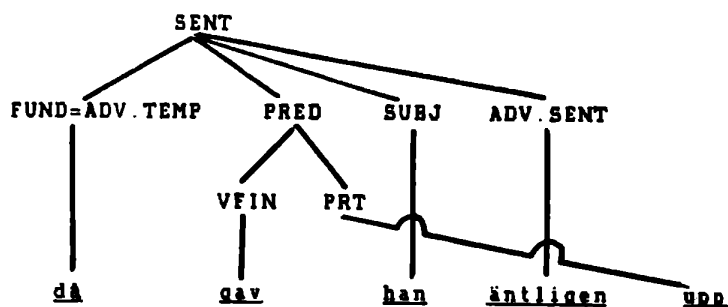
ett satsadverbial (äntligen) och att predikatet analyserats i ett finit verb och partikel. De ordningsbeskrivningar vi behöver kan ges följande form, där C är en variabel över precis en konstituent, X en variabel över en godtycklig

under SENT: ORDER = <# FUND C SUBJ ADV.SENT X>  
under PRED: ORDER = <# C VFIN X PRT>

sekvens av konstituenterna, och # en satsgräns. Dessa två ordningsbeskrivningar tillsammans specificerar den givna ordningen mellan konstituenterna och ingen annan, men ger inte ett ordnat frasstrukturträd av gängse typ utan ett träd av utseendet (23). Det talar om för oss precis vilka funktionella enheter som identifieras och hur de fördelar sig på inputsträngen. Denna typ av träd är inte så vanlig som representationsformat, kanske för att de gör ett osnyggt intryck, men de kan väl motiveras. Deras formella egenskaper är heller inte svårare att specificera än ordnade träd<sup>(6)</sup>.



(23)



Givet den stora användning som frasstrukturträd fått i modern lingvistik är det kanske frestande att anta att meningars inneboende struktur helt enkelt är "frasstrukturell" och att varje mening måste tilldelas (minst) en frasstruktur. Detta förefaller att vara den generativa grammatikens position, utan att man därför besvärat sig så mycket att argumentera för den<sup>(7)</sup>. Typiskt är då att man försöker återinföra beskrivningen av grammatiska relationer på frasstrukturkonfigurationer. Också LFG, som annars kraftigt tar avstånd från frasstrukturträd som generellt beskrivningsformat, behåller en nivå av kontextfritt genererbar frasstruktur, så att varje mening grammatisk beskrivs med två strukturer, en frasstruktur (eller c-struktur som man föredrar att kalla det) och en funktionell struktur, f-strukturen. Varför är det då enligt LFG nödvändigt med en c-struktur, som explicit sägs vara helt irrelevant för semantisk tolkning? Man ger två huvudskäl; det ena är att en sådan struktur behövs som input till den fonologiska komponenten, det andra är att den jämte lexikoninformation behövs som input vid härledningen av f-strukturer. Det första skälet ska jag inte beröra, men det andra som i princip är Aspectteorins ide att frasstrukturella konfigurationer är definierade för grammatiska relationer (Chomsky, 1965:68 f). Så associerar t ex Kaplan & Bresnan (1982:184) med frasstrukturelgen (24) en tilldelning av subjektsfunktionen till NP.

(24) S → NP VP

Dock skiljer sig LFG:s användning av dessa konfigurationella definitioner från Aspectsteorins på några väsentliga punkter:

- 1/ De är språkspecifika. Subjekt, objekt osv är universella primitiva begrepp som inte kan reduceras till frasstrukturkonfigurationer, men som i vissa språk uttrycks konfigurationellt, i andra kanske morfologiskt;
- 2/ De hänför sig direkt till det språkliga uttrycket (ytstrukturen);

3/ Varje instans av en grammatisk funktion behöver inte uttryckas på detta sätt. Den kan t ex vara given lexikalt.

I och med att den tredje möjligheten finns uppstår frågan om när vi har att göra med t ex ett konfigurationellt givet subjekt och när vi har ett subjekt som är givet på annat sätt. LFG:s svar tycks vara att vi har det förra fallet dels när subjektet står där det ska stå enligt definitionen (naturligt nog) eller då det extraherats ut ur sin sats som i en fråga eller en relativsats. I s k kontrollstrukturer får däremot komplementets verb sitt subjekt från det överordnade verbets lexikoninformation. Lova har således en specifikation som talar om att subjektet i komplementsbisatsen är identiskt med dess eget i en mening som (25).

(25) Pelle lovade komma.

(26) Det läser Eva.

Skälet till att extraherade led ges en grammatisk funktion via ett tomt element i en definierande position är att det finns strukturella villkor på när sådana extraktioner är möjliga (Ross, 1969). Men det finns också många andra villkor på dem, t ex av pragmatisk art (Allwood, 1982; Engdahl, 1982) utan att det anses nödvändigt att upprätta en särskild pragmatisk representation som i härledningen föregår f-strukturen. Överhuvudtaget gäller för bestämningen av grammatiska funktioner att annan information än den rent morfosyntaktiska måste komma med. Meningen (26) ges av SVE.UCP bara en beskrivning och de strukturellt möjliga alternativen slås ut på ett tidigt stadium när verbets lexikala egenskaper jämförs med de båda NP:nas lexikala egenskaper. En sekvens NP - V - NP kan åsättas flera olika c-strukturer, t ex de i (27), svarande mot olika f-strukturer, men det finns såvitt jag kan se ingen anledning att generera alla dessa i sin helhet för att först i efterhand slå ut flertalet av dem på grundval av lexikal information, vilket i princip är den procedur som LFG använder sig av. (28) är ett annat exempel som lätt leder till en explosion av c-strukturer och f-strukturer men som tillåter en (eller två) tolkningar i ett textsammanhang.

- |                        |                             |                            |
|------------------------|-----------------------------|----------------------------|
| (27)a. NP - VP(V - NP) | SUBJ - PRED (VERB - OBJ)    | /Erik badar Lisa/          |
|                        | alt SUBJ - PRED (VERB - PF) | /Erik badar en glad gosse/ |
| b. NP - S(V - NP - e)  | FUND=OBJ - VFIN - SUBJ      | /Henne badar Erik/         |
| c. NP - VP(V) - NP     | SUBJ - PRED (VERB) - ADV    | /Erik badar varje dag/     |
| d. NP - S(V - NP)      | FUND=ADV - PRED - SUBJ      | /Varje dag badar Erik/     |

(28) Eva och Erik Lind bor en trappa upp i ett gult hus vid Stortorget i Nora.

En annan fråga är om strukturella villkor måste uttryckas i termer av frasstrukturkonfigurationer. Den enda funktion av dem som vi hittills diskuterat som kan ges en entydig frasstrukturdefinition är fundamentet medan däremot både subjekt och objekt kräver flera olika. Om vi däremot uttrycket strukturella positioner i termer av fält (Didrichsen, 1957) och fältgränser kan vi få mera användbara identifikationskriterier, förslagsvis

- 1/ SUBJEKT är ett NP, eller ett S i nexusfältet eller i fundamentställning, som står i nominativ eller i ommarkerat kasus och som uppfyller de lexikala villkor som verbet ställer:
- 2/ OBJEKT är ett NP, eller ett S i verbalfältet, eller i fundamentställning, som har ackusativt eller ommarkerat kasus som uppfyller de lexikala villkor som verbet ställer.

Dessa definitioner är inte så enkla som de frasstrukturella, men har fördelen att vara direkt tillämpbara på språkliga data. De är också möjliga att formalisera; i SVE.UCP finns två regler som bestämmer subjekt i en sats. Den ena, som utnyttjar kategoriell och morfologisk information, anropas i huvudsatsregeln efter det att det finita verbet och ett eventuellt satsadverb identifierats (i nexusfältet, om vi så vill; (Jf Sågvall Hein, 1983b), och plockar upp ett subjekt i den positionen eller från fundamentet. Den andra, som kontrollerar att det sålunda identifierade subjektet har vissa särdrag, anropas från den regel som söker efter verbets argument och som hämtas ur lexikonet. Denna verbregel svarar också för objektssökning och kan på samma sätt lägga semantiska villkor på detta.

#### NOTER

- ⟨1⟩ För information om UCP:n och UCP-formalismen se Sågvall Hein (1980, 1983a) samt Carlsson (1982).
- ⟨2⟩ (1) är inte den fullständiga beskrivningen utan en något förkortad version. Skillnaden är dock betydelselös för den här framställningens syften.
- ⟨3⟩ CASE = NIL innebär att ledets kasus är ommarkerat. LEX = !DEN är en hänvisning till ett lexem, vars egenskaper är rudimentärt specificerade i en särskild lexembas.
- ⟨4⟩ Jf Kapland & Bresnan, 1982:181.
- ⟨5⟩ Syntaktiska analysträd av den typ som används i Montaguegrammatik är dock undantag.
- ⟨6⟩ Se t ex Karlgren, 1972 och McCawley, 1982, för argument och formaliseringar av sådana alternativa representationsformer.

⟨7⟩ Chomsky & Miller (1963:288-89) skriver: "We assume that such a tree graph must be a part of the structural description of any sentence; we refer to it as a phrase-marker (P-marker). A grammar must, for adequacy, provide a P-marker for each sentence." Se också Chomsky (1965:123f). För några fördelar med frasstrukturrepresentation, se Gazdar (1982).

#### REFERENSER

- Allwood, J. (1982) The Complex NP Constraint in Swedish. I Engdahl & Ejerhed (1982) s 15-32.
- Bresnan, J. (1982) (utg.) The Mental Representation of Grammatical Relations, Cambridge, Mass. The MIT Press.
- Carlson, M. (1982) Uppsala Chart Parser 2, System Documentaion. Uppsala, UCDL.
- Chomsky, N. (1965) Aspects of the Theory of Syntax. Cambridge, Mass. The MIT Press.
- Chomsky, N. & Miller, G. (1963) Introduction to the Formal Analysis of Natural Language. I Luce, Bush & Galanter (utg.) Handbook of Mathematical Psychology s. 269-322.
- Diderichsen, P. (1957) Elementaer dansk grammatik, 2 upp,. Köpenhamn, Gyldendal.
- Engdahl, E. (1982) Restrictions on Unbounded Dependencies in Swedish. I Engdahl & Ejerhed (1982) s. 151-174.
- Engdahl, E. & Ejerhed, E. (1982) (utg.) Readings on Unbounded Dependencies in Scandinavian Languages, Umeå, Almqvist & Wiksell international.
- Gazdar, G. (1982) Phrase Structure Grammar, I P. Jacobson & G. K. Pullum, (utg.), The Nature of Syntactic Representaion, Reidel, Dordrecht.
- Jackendoff, R. (1977) X-bar Syntax, Cambridge, Mass. The MIT Press.
- Johnson, D. (1977) On Relational Constraints on Grammars. I Syntax & Semantics, Vol 8, s. 151-177.
- Karlgren, H. (1972) Why Trees in Syntax? Interim Report No. 41, KVAL, Stockholm.
- Kaplan, R. & Bresnan, J. (1982) Grammatical Representation. I Bresnan, 1982.
- Kay, M. (1977) Reversible Grammar; Summary of the Formalism. Xerox Research Center, Palo Alto.
- Kay, M. (1981) An Algorithm for Compiling Parsing Tables from a Grammar. Xerox Research Center, Palo Alto.
- McCawley, J. D. (1982) Parentheticals and Discontinuous Constituent Structure. Linguistic Inquiry, Vol 13, s. 91-106.
- Perlmutter, D. M. (1980) Relational Grammar. I Syntax & Semantics, Vol 13, s. 195-230, New York, Academic Press.
- Ross, J. R. (1967) Constraints on Variables in Syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Sågwall-Hein, A-L. (1980) An Overview of the Uppsala Chart Parser Version 1. Uppsala, UCDL.

Sågwall-Hein, A-L. (1983a) A Parser for Swedish. Status Report for SVE.UCP, February, 1983, Uppsala, UCDL.

Sågwall-Hein, A-L. (1983b) Regelaktivering i en parser för svenska (SVE.UCP). I denna volym.

Teleman, U. (1974) Manual för grammatisk beskrivning av talad och skriven svenska. Lund, Studentlitteratur.