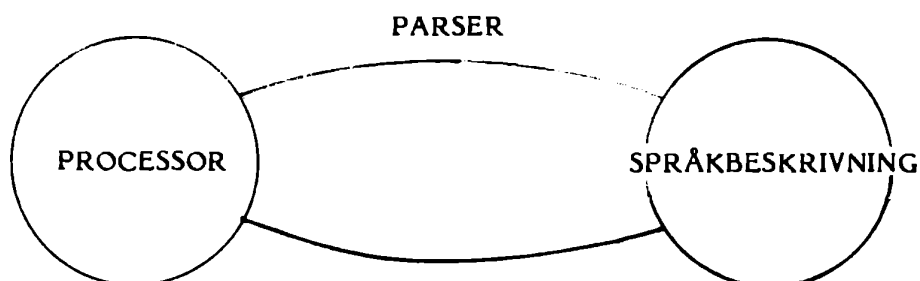


Anna Sågvall Hein  
Centrum för datorlingvistik  
Sturegatan 13 B, 5 tr  
752 23 UPPSALA

## REGELAKTIVERING I EN PARSER FÖR SVENSKA (SVE.UCP)

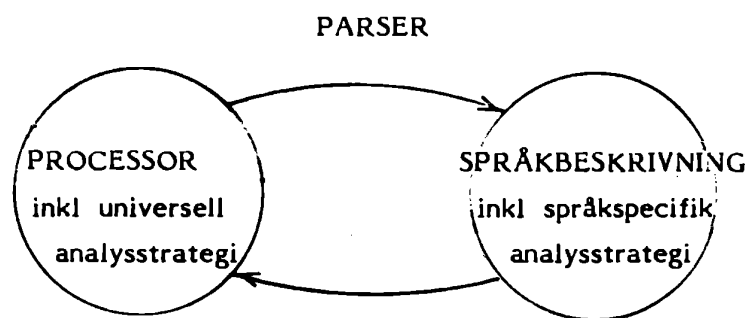
1. Inledning. Parsing är den process i vilken ett språkligt uttryck (ord, fras, sats, mening) tilldelas en lingvistisk beskrivning enligt en given språkbeskrivning. För det processerande maskineri, som sätter språkbeskrivningen i arbete i relation till det uttryck som skall analyseras, analysuttrycket, använder jag termen 'processor'. Termen introducerades i datorlingvistiska sammanhang av R Kaplan i hans artikel 'A General Syntactic Processor' (Kaplan 73). Utformningen av processorn är ur datorlingvistiskt perspektiv inte mindre intressant än den av språkbeskrivningen.



Figur 1

GSP inkluderar ett format för att representera analysuttrycket i dess olika bearbetningsstadier, dvs dess initiala form såväl som intermediära och finala resultat, 'the chart', och ett antal processprimitiver för uppbyggande av en nätverksgrammatik. Analysen förlöper som en följd av enkla bearbetningssteg, 'tasks'. Administrationen av de olika bearbetningsstegen handhas av en agenda. Vidare innefattar GSP en styrande algoritm, samt ytterligare ett antal kontrollstrukturer i algoritmens tjänst. I en av Kay föreslagen vidareutveckling av GSP (Kay 75), omvandlas charten från en passiv datastruktur för lagring av lingvistiska strukturer till en aktiv struktur, vilken övertar många

av de kontrollfunktioner som i den tidigare versionen vilade på särskilda hjälpstrukturer. Uppsala Chart Processor, som utvecklats och implementerats vid Centrum för datorlingvistik (se Sågvall Hein 80, 81, 82 och 84, under utgivning, samt Carlsson 81), bygger vidare på GSP-traditionen, med utnyttjande av iden om 'den aktiva charten'. Försedd med en språkbeskrivning för ett givet språk, fungerar UCP som en parser för språket i fråga. Sålunda förfogar vi redan över en parser för svenska, SVE.UCP, med begränsad kompetens (se Sågvall Hein 83). Pågående arbete koncentreras till en systematisk utbyggnad av den svenska kompetensen. Från sina föregångare skiljer sig UCP främst därigenom, att den utför både morfologisk och syntaktisk analys. Den morfologiska analysen förlöper i analogi med den syntaktiska. Såväl lexikonsökning som regeltillämpning (fonologisk, morfologisk och syntaktisk) realiseras som en följd av konceptuellt homogena bearbetningssteg. Att kunna integrera den morfologiska analysen med den syntaktiska är en förutsättning för korrekt ordigenkänning i de fall de lexikaliska enheterna är diskontinuerligt presenterade i texten. Ett annat av de utmärkande dragen för en UCP-parser har att göra med förhållandet mellan processorn och språkbeskrivningen. Sålunda inkluderar den sarspråkliga kompetensen inte enbart de lingvistiska enheter och regler, som beskriver språket i fråga, utan även regelaktiverande mekanismer. Ansatsen bygger på det antagandet, att man kan skilja mellan en universell och en språkspecifik analysstrategi. Den universella strategin ligger inom processorns domän, och den språkspecifika inom den sarspråkliga kompetensens område (fig 2).



Figur 2

När det gäller frågan hur analysstrategin skall inordnas i språkbeskrivningen har jag valt att integrera den i de enskilda lingvistiska reglerna som det alternativ som erbjuder störst flexibilitet och därmed också den bästa experimentmiljön.<sup>1</sup> Sålunda följer vi i UCP-formalismen upp den ursprungliga GSP-tanken på en procedural lingvistisk formalism. Språkbeskrivningen i en UCP-parser för-

delar sig på tre typer av konstrukter, nämligen grammatik, lexika och baser. Innehållet i de olika konstrukterna uttrycks i en och samma formalism, dvs den procedurala UCP-formalismen. UCP-formalismen är sålunda generell. Konstrukterna skiljer sig åt med avseende på organisationsform och accessmetod.

I sin nuvarande utformning inkluderar SVE.UCP: en grammatik innehållande morfologiska och syntaktiska regler (SVE.GRAM), ett huvudlexikon över rötter, stammar och oböjliga ord (SVE.DIC), ett antal affixlexikon och ett lexikon över de skiljetecken vilka betraktas som särskilda ord (SEP). Vidare omfattar den svenska språkbeskrivningen en lexembas och en karaktärbas. I lexembasen finns semantisk information avseende substantiv och verb. Karaktärbasen reflekterar en klassificering av karaktärerna i bokstäver, siffror och skiljetecken. Via särdragsmarkering anges för bokstäverna huruvida de svarar mot vokal eller konsonant. Dessutom finns ett antal fonetiska särdrag upptagna.

En av de stora fördelarna med en generell lingvistisk formalism, och en procedural sådan, är den att den underlättar samspelet mellan de olika lingvistiska kunskapskällorna under parsingprocessen. I SVE.UCP har vi sålunda utformat en analysgång för svenska meningar, som grovt ser ut som följer. Meningen presenteras för SVE.UCP som en följd av karaktärer exakt som den står i texten, dvs med skiljetecken och stora och små bokstäver. Analysen inleds med att UCP anropar begynnelseregeln i grammatiken. Dess effekt är att initiera två olika processer, nämligen sökning i huvudlexikonet, resp tillämpning av huvudsatsregeln. Lexikonsökningen, som avser karaktärer, igångsätts omedelbart, medan huvudsatsregeln måste invänta igenkänning av den typ av lingvistiska enheter, för vilken den är definierad, dvs ord och fraser. Grammatiska beskrivningar av dessa presenteras som resultat av det analysarbete som lexikonsökningen ger upphov till. Då t ex en substantivisk stam återfunnits i lexikonet anropas sålunda bl a den morfologiska regel som känner igen substantiv och bygger upp motsvarande morfologiska beskrivning. (Vidare initieras sökning i de substantiviska affixlexikonerna efter påföljande affix.) Substantivregeln i sin tur anropar en NP-regel i grammatiken, som känner igen och bygger upp beskrivningar av nominalfraser. Därigenom blir lämpligt arbetsmaterial tillgängligt för huvudsatsregeln, som kan sätta i gång sitt arbete.

Föreliggande presentation syftar inte till att beskriva enskildheterna i samspelet mellan grammatik, lexika och baser i vår svenska parser, utan till att visa på de faciliteter som UCP-formalismen erbjuder för att uttrycka ett sådant. Helt kort kan dock konstateras att den analysstrategi som vi byggt in i den

svenska kompetensen i vår parser, ligger i linje med den perceptuella strategi som Kimball förespråkar (Kimball 73).

2. UCP-formalismen. UCP-formalismen inkluderar, förutom ett format för att uttrycka lingvistiska regler, också ett format för att uttrycka grammatiska beskrivningar av lingvistiska enheter. De uttrycks som hierarkiska mängder av attribut-värde-par (Sågwall Hein 83 och Ahrenberg, denna volym). Grammatiken upptar ett antal regler, vilka var och en består av ett unikt namn, en regelkropp, en domänspecifikation och, optionellt, angivande av ett kontextuellt villkor för dess tillämpning (fig 3).

```
Regelnamn
  Regelkropp
  Domän
  (Villkor för dess tillämpning)
```

Figur 3

Regelkroppen är uppbyggd som en följd av enkla operationer, vilka exekveras i sekvens, till dess någon operation misslyckas (returnerar NIL), varvid exekveringen avbryts. Operationerna utförs i relation till analysuttrycket i dess chartrepresentation. I domänen anges på vilken typ av lingvistiska enheter regeln är tillämplig. Som första exempel på en grammatisk regel anförs den regel med vilken analysen av en svensk mening inleds (fig 4).

```
START.RULE
  PROCESS(SVE.DIC),
  PROCESS(MAIN);
  Domain: CHAR;
```

Figur 4

Regelns namn är START.RULE. Regelkroppen upptar två operationer, nämligen PROCESS(SVE.DIC) och PROCESS(MAIN). I det ögonblick startregeln anropas föreligger analysuttrycket som en följd av karaktärer, varför regelns domän specificeras till karaktärer (CHAR). SVE.DIC är namnet på det svenska huvudlexikonet. Det upptar rötter, stammar och ordformer. Exekvering av operationen PROCESS(SVE.DIC) har som effekt att sökning i huvudlexikonet initieras. MAIN är namnet

på huvudsatsregeln i grammatiken. Via operationen (MAIN) anropas denna regel. Operationen uttrycker sålunda en prediktion om en huvudsats. Startregeln anropas av UCP. Övriga regelanrop sker via operationer i grammatik och lexika. Fortsättningsvis skall jag koncentrera mig på vilka möjligheter som finns för detta. En förenklad version av vår svenska huvudsatsregel (fig 5) får fungera som illustration.

**2.1 Regelaktivering i den svenska huvudsatsregeln.** Som framgår av domänspecifikationen i fig 5, opererar huvudsatsregeln på fraser (PHRase.CAT) och ord (WORD.CAT). Huvudsatsens regelkropp består till stor del av s k subregelanrop. V.FIN, ADVERBIALS1, SUBJECT, ADVERBIALS2, FUNDCHECK och END.OF.SENT är mnemoniska namn på grammatiska lägen, subregler, i vilka faktiska operationer specificeras. Så snart processorn under exekveringen av en regel träffar på ett subregelnamn, så utförs de operationer som upptas under detta regelnamn. Först i huvudsatsregeln ligger dock ett antal operationer, som inleder uppbyggandet av en grammatisk beskrivning av satsen. Bland dessa finns en, som ansätter den första konstituenten i satsen som fundament.<sup>2</sup>

```

MAIN
  "ansätt aktuell konstituent som fundament,
    ... ",
  ADVANCE,
  V.FIN,
  ADVANCE,
  ADVERBIALS1,
  SUBJECT,
  ADVERBIALS1,
  DO(<& PRED VFIN VERB.FEAT LEX :VERBACTION>),
  ADVERBIALS2,
  FUNDCHECK,
  (END.OF.SENT, MINORSTORE, MAJORPROCESS(SENTENCE)/
  "test på konjunktion och el skiljetecken",
  MAJORPROCESS(SENTENCE.COORD), MINORSTORE, ADVANCE,
  PROCESS(MAIN));
  Domain: PHR.CAT / WORD.CAT;

```

Figur 5

Efter utförandet av de initiala ansättningsoperationerna flyttas processorns

uppmärksamhet ett steg framåt i satsen, dvs till slutet av den första konstituenten (fundamentet). Detta sker med hjälp av operatorn ADVANCE. Subregeln V.FIN predicerar ett finit verb i detta läge och tilldelar ev återfunnet sådant till predikatet, varpå ett avancemang sker till subregeln ADVERBIALS1.<sup>3</sup> ADVERBIALS1 beaktar optionella satsadverbial och gör motsvarande tilldelningar till den grammatiska beskrivningen. SUBJECT (fig 6) illustrerar möjligheten till parallell analys.

#### SUBJECT

```

("aktuell konstituent är en NP",
 "kasustest",
 "ansätt aktuell konstituent som subjekt",
 ADVANCE //
 "fundamentet är en NP",
 "kasustest",
 "ansätt fundamentet som subjekt");

```

Figur 6

När subjeksregeln tillämpas, har processorn sin uppmärksamhet riktad på den första konstituenten efter det finita verbet, såvida denna inte tolkats som ett satsadverbial, i vilket fall processorn fokuserar på den första konstituenten efter detta satsadverbial. Subjeksregeln upptar två av varandra oberoende parallella alternativ, åtskilda av '//'. I det första alternativet inspekteras aktuell konstituent, och i det andra fundamentet.<sup>4</sup> Om aktuell konstituent är en NP i grundformskasus, så ansätts (beskrivningen av) den som subjekt, och ett avancemang sker till nästa konstituent. I annat fall avbryts exekveringen av denna analysgren. Oberoende av dess utfall, exekveras den alternativa analysgrenen enl följande. Om fundamentet är en NP i grundformskasus, så ansätts den som subjekt. I annat fall avbryts exekveringen. Om vi sålunda har en sats av typen 'NP V NP', så förgrenar sig analysen via subjeksregeln i två av varandra oberoende alternativ, med olika värden på subjeksattributet. Vi talar här om parallella processer och om operationen 'independent disjunktion' ('//'). Båda processerna löper samman i sökandet efter ytterligare optionella satsadverbial (ADVERBIALS1). Därpå följer DO-operationen, som utifrån syntaktiska och semantiska egenskaper hos det finita verbet svarar för igenkänning av övriga grammatiska funktioner på satsnivå.

Med hjälp av DO-operatorn kan vi utföra s k variabelt subregelansrop. Då en

DO-operation exekveras, evalueras först dess argument till ett subregelnamn, varpå denna subregel i sin tur exekveras. Som Do-operationens argument är utformat i huvudsatsregeln, evalueras det till namn på vad jag kallar en 'verbaktionsregel', i vilken det finita verbets argumentstruktur definieras. DO-argumentet har den i UCP-formalismen gängse formen av ett s k path-uttryck, '<...>', (se Kay 77). Om vi bortser från dess sista med kolon (':') prefigerade element, så uttrycker det en referens till det finita verbets lexembeteckning i den grammatiska beskrivning av satsen som är under uppbyggnad. Om det finita verbet sålunda är 'bor', så returnerar den första delen av path-uttrycket lexembeteckningen '!bo'. Det med kolon prefigerade elementet (':verbaction') tolkas av processorn som en anmodan att söka efter värdet på egenskapen 'verbaction' under lexembeteckningen '!bo' i lexembasen. För exemplet i fråga returnerar processorn 'type.bo', som är namn på aktuell verbaktionsregel. Med hjälp av kolon-faciliteten kan vi sålunda signalera till processorn, att den skall hämta information ur någon av baserna. (Lexembasen inkluderar f n förutom verb med associerade verbaktionsregler även substantiv med associerade semantiska särdrag.) Fig 7 visar verbaktionsregeln för verb av typen 'bo'.

```

TYPE.BO
  PREDSUBJ('ANIM),
  ADVLOC;

```

Figur 7

Regel upptar två subregelanrop, varav det första illustrerar hur en subregel kan anropas med ett argument, s k subregelanrop med variabel. PREDSUBJ ställer krav på subjektet; i detta fall krävs att subjektet är animat. ADVLOC kräver ett lokationsadverbial och gör motsvarande ansättning till den grammatiska beskrivningen av satsen.

Låt oss gå återvända till huvudsatsregeln (fig 5). ADVERBIALS2 känner igen efterställda satsadverbial och FUNDCHECK kontrollerar att fundamentet fått en grammatisk tolkning. Därpå uttrycker huvudsatsregeln två alternativ, åtskilda av '/', ett uttryck för operationen 'dependent disjunktion' (jfr OR). Enligt det första alternativet är satsen (och meningen) slut, i vilket fall huvudsatskonstituenten skall avslutas. Detta sker via operationen MINORSTORE. Den fungerar på så sätt, att den avslutar, stänger, konstituenten före det element som processorn just inspekterar, i detta fall det skiljetecken som markerar slut på meningen. Den låter oss sålunda se ett element framåt, in-

nan vi bestämmer oss för att avsluta en konstituent. Då vi känt igen en huvudsats och påföljande stora skiljetecken, vet vi också att vi har att göra med en mening. I regeln SENTENCE byggs den grammatiska beskrivningen av meningen upp, i vilken beskrivningen av huvudsatsen ingår som en del.

Anropet till SENTENCE sker via operatorn MAJORPROCESS. Den har som effekt, att den regel vars namn står som dess argument (SENTENCE), initieras från samma punkt i analysuttrycket som den regel från vilken anropet görs (MAIN). Sålunda kommer sentence-regeln att tillämpas från meningens början. Därigenom skiljer den sig från den andra regelaktiverande operatorn PROCESS, som initierar en process från den punkt på vilken processorn fokuserar i det ögonblick anropet görs. Med andra ord, så svarar MAJORPROCESS för regelaktivering bottom-up och PROCESS för regelaktivering top-down.

Enligt huvudsatsregelns återstående alternativ är huvudsatsen men inte meningen slut. COORD.TEST söker på små skiljetecken och konjunktioner. Om testet ger ett negativt utfall, så avbryts exekveringen av regeln. I annat fall sker följande: Beskrivningen av huvudsatsen avslutas. Regeln för koordinerade meningar (SENTENCE.COORD) anropas bottom-up och huvudsatsregeln anropas top-down. Som framgår av det sistnämnda anropet tillåter formalismen, att en regel anropas rekursivt.

Huvudsatsregeln arbetar på fraser och ord. Igenkänning av dessa och uppbyggande av motsvarande grammatiska beskrivningar sker som en följd av den sökning i huvudlexikonet som startregeln initierar (fig 4). Låt oss titta närmare på hur detta realiseras.

**2.2 Lexikonsökning och igenkänning av ord och fraser.** Uppslagsorden i lexikonerna i en UCP-parser ligger lagrade som följder av teckenstråk, och sökningen går stegvis, tecken för tecken. Varje sådant söksteg är analogt med tillämpning av en grammatisk regel. En sökregel inkluderar en obligatorisk operation, nämligen jämförelse mellan aktuellt tecken i analysuttrycket och i lexikonstråket. Operationen i fråga betraktas som universellt giltig, varför den är inbyggd i processorn. Däremot får vi för varje enskild applikation bestämma vilken typ av tecken sökningen avser, dvs definiera sökregelns domän. Vidare kan vi specificera ytterligare operationer för de fall jämförelsen utfaller positivt. Denna information ges i ett antal meta-uppslagsord för varje lexikon. Fig 8 visar metauppslagsorden i SVE.DIC.



```

DOMAIN
  CHAR;

GATHERING-RULE
  "tilldelning av aktuell karaktär till attributet CHARS";

FORWARD-ACTION
  ADVANCE;

...

```

Figur 8

Domänen specificeras till karaktärer (bokstäver, skiljetecken, specialtecken, mellanslag och siffror). (Som alternativ kan man tänka sig att sökningen skall utföras på en fonematisk representation av analysuttrycket.) Under GATHERING-RULE anger vi, att uppgift om vilken karaktär som jämförts skall tilldelas till den grammatiska beskrivningen. Det är med andra ord ett sätt att spara uppgift om den initiala strängen. Under FORWARD-ACTION flyttar vi processorns uppmärksamhet ett steg (i detta fall karaktär) framåt. Då denna framflyttning sker oberoende av om jämförelseoperationen gäller den sista karaktären i ett uppslagsord eller inte, blir sökningen icke-deterministisk och alla i lexikonet förutsedda segmenteringar beaktas.

Då sökprocessen lett fram till ett uppslagsord, dvs vid positivt utfall av jämförelsen mellan en karaktär i analysuttrycket och uppslagsordets sista karaktär, ansluts de operationer som bildar uppslagsordets regelkropp till operationen under GATHERING-RULE. Fig 9 visar exempel på en lexikonartikel.

```

EVA
  "tilldelning av värdet NOUNSTEM till attributet MORPH.CAT",
  "tilldelning av värdet !EVA till attributet LEX",
  PATTERN.FIRST.NAME;

```

Figur 9

Regelkroppen består av tre operationer, dvs ansättning av morfologisk kategori (NOUNSTEM), ansättning av lexembeteckning (!EVA) och ett subregelanrop (PATTERN.FIRST.NAME).

```

PATTERN.FIRST.NAME
  SPECIFY.SEX,
  "tilldelning av värdet + till attributet PROPR",
  "tilldelning av värdet + till attributet FIRST (förnam)",
  "tilldelning av värdet SING till attributet NUMB",
  "tilldelning av värdet UTR till attributet GENDER",
  STORE.PROPR;

```

Figur 10

Anrop till PATTERN.FIRST.NAME (fig 10) görs från alla förnamn i lexikonet. Utgående från sista bokstaven i stammen och med utnyttjande av de uppgifter om karaktärerna som finns lagrade i karaktärbasen, fastställs via SPECIFY.SEX naturligt genus.<sup>5</sup> Sist återfinns ett anrop till subregeln STORE.PROPRium (fig 11).

```

STORE.PROPR
  MAJORPROCESS(NOUN),
  ADVANCE,
  (SPACE.OR.PUNCT / MINORSTORE, PROCESS(CASE));

```

Figur 11

STORE.PROPR inleds med ett anrop till substantivregeln, som känner igen substantivet och bygger upp dess morfologiska beskrivning. Från substantivregeln anropas vidare den NP-regel, som känner igen nominalfrasen och bygger upp dess syntaktiska beskrivning. STORE.PROPR avslutas med en disjunktiv (alternativ) operation. Det första alternativet, SPACE.OR.PUNCT, avser det fall då stammen omedelbart följs av en separator (ett mellanslag eller ett skiljetecken). Härvid avslutas beskrivningen av stamkonstituenten på erforderligt sätt, dvs inklusive mellanslaget och exklusive skiljetecknet. Därpå initieras sökning i huvudlexikonet (efter nästa ord), respektive i separatorlexikonet (efter aktuellt skiljetecken). I annat fall avslutas beskrivningen av stamkonstituenten via MINORSTORE, och sökning i kasuslexikonet igångsätts.

**2.3 Sammanfattning av de processororienterade faciliteterna i UCP-formalismen.** Sammanfattningsvis erbjuder UCP-formalismen följande möjligheter vad det gäller att uttrycka en språkspecifik analysstrategi:

- möjlighet att initiera sökning i visst lexikon från en regel i grammatiken eller från en lexikonartikel (PROCESS dic.name),
- möjlighet att anropa en regel i grammatiken top-down från en annan grammatisk regel eller från en lexikonartikel (PROCESS rule.name),
- möjlighet att anropa en regel i grammatiken bottom-up från en annan grammatisk regel eller från en lexikonartikel (MAJORPROCESS rule.name),
- möjlighet till enkel subregelaktivering från en grammatisk regel eller från en lexikonartikel ( t ex SUBJECT, PATTERN.FIRST.NAME),
- möjlighet till variabel subregelaktivering från en grammatisk regel eller från en lexikonartikel ( DO(<...>)), samt
- möjlighet till subregelaktivering med variabel från en grammatisk regel eller från en lexikonartikel ( t ex (PREDSUBJ 'ANIM)).

3. Slutsats. Den informella presentationen av UCP-formalismen ovan, utifrån dess tillämpning i vår svenska parser, har inriktats mot att visa formalismens processororienterade faciliteter.<sup>6</sup> Genom förekomsten av denna typ av faciliteter skiljer sig en procedural formalism från en deklarativ. Det är genom tillgång till dessa, som språkvetaren i utarbetandet av en språkbeskrivning för ett givet språk har möjlighet att uttrycka en språkspecifik analysstrategi. UCP-formalismen erbjuder stor frihet i den konkreta utformningen av språkbeskrivningen, vad gäller dess strukturering i grammatik, lexika och baser samt vid utformningen av de enskilda reglerna i grammatik och lexika. I kraft av sin processororienterade utformning öppnar den sig mot interaktion med andra komponenter i en språkförståelsemodell.

#### NOTER.

- (1) Ett annat alternativ vore att utforma analysstrategin som en egen komponent i språkbeskrivningen. Ett sådant företag bedöms dock som prematurt, då det bör vila på fördjupade insikter om olika analysstrategier. Vi räknar med att vi med vår ansats kommer att kunna experimentera oss fram till resultat

av värde i det sammanhanget.

- (2) För att inte belasta framställningen i onödan med de tekniska enskildheterna i UCP-formalismen, har jag valt att uttrycka ansättnings- och testoperationer informellt. Det informella uttryckssättet markeras via anföringstecken.
- (3) Om inget finit verb återfinns, så avbryts exekveringen av detta alternativ; en förklaring till misslyckandet kan vara att endast en delkonstituent av det egentliga fundamentet uppfattats som sådant; UCP svarar för att alla tänkbara alternativ, som grammatiken upptar beaktas.
- (4) Trots att processorn i detta analysmoment fokuserar på konstituenten efter det finita verbet (ev föregången av optionellt satsadverbial), så är fundamentet tillgängligt för inspektion i och med att det ansatts i den grammatiska beskrivning av satsen som är under uppbyggnad. - Aktuell grammatisk beskrivning under uppbyggnad är ständigt tillgänglig under analysen i form av en aktiv båge i charten.
- (5) För förnamn, som bryter mot de allmänna reglerna, anges genus individuellt i lexikonartikeln.
- (6) De av formalismens operatorer som svarar mot tilldelning, tester etc har avsiktligt förbigåtts. För en informell presentation av UCP-formalismen i sin helhet hänvisas till Sågvall Hein 84, under utgivning, och för en teknisk presentation till Carlsson 81.

#### REFERENSER

- Ahrenberg, L (1984), De grammatiska beskrivningarna i SVE.UCP (denna volym)
- Carlsson, M (1981), Uppsala Chart Parser 2. System Documentation.  
UCDL-R-81-1. Center for Computational Linguistics. Uppsala University.
- Kaplan, R M (1973), A General Syntactic Processor in Rustin, ed, Natural Language Processing. Englewood Cliffs, N J: Prentice Hall.
- Kay, M (1975), Syntactic Processing and the Functional Sentence Perspective.  
I: Schank, R & Nash-Webber, B L (utg), Theoretical Issues in Natural Language Processing, Cambridge, Mass

--- (1977), Reversible Grammar. Summary of the Formalism. Xerox Research Center. Palo Alto.

Kimball, J (1973), Seven Principles of Surface Structure Parsing in Natural Language. Cognition, vol 2, s 15-47

Sågvall Hein, A (1980), An Overview of the Uppsala Chart Parser Version 1 (UCP-1). UCDL-R-80-1. Center for Computational Linguistics. Uppsala University.

--- (1981), Uppsala Chart Parser, Version 2 (UCP-2) - En översikt. I: Lien, E (red), De Nordiske Datalingvistikdagene 1981. Foredrag fra en konferanse på Universitetscenteret på Dragvoll 22-23 oktober 1981.

--- (1982), An Experimental Parser. I: Proceedings of COLING 1982. Prag s 121-126

--- (1983), A Parser for Swedish. Status Report for SVE.UCP, February 1983. UCDL-R-83-1. Centrum för datorlingvistik. Uppsala Universitet.

--- (1984), Parsing by means of Uppsala Chart Processor. I: Bolc, L (ed), Natural Language Parsing Systems. Springer-Verlag. (Planerad utgivning 1984).