

BENGT SIGURD

Inst för Fonetik och Lingvistik, LUNDS Universitet  
Helgonabacken 12, S-22362 LUND

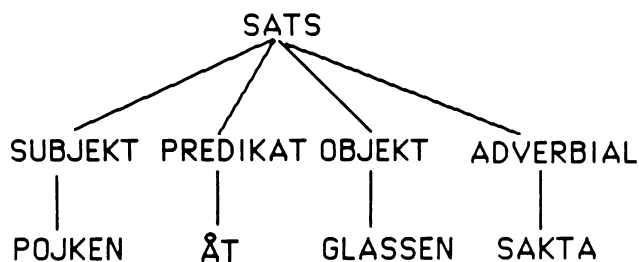
## REFERENTGRAMMATIK - EN KORT PRESENTATION

### Satsdelsanalys (funktionell analys)

I analys av meningar och satser enligt traditionell skolgrammatik skiljde man mellan identifiering av ordklasser, t.ex. substantiv, verb, adverb och identifieringen av satsdelar, t.ex. subjekt, predikat, objekt, adverbial. I satsen "Pojken åt glassen sakta" skulle man enligt klassisk skolgrammatik säga att "pojken" och "glassen" är substantiv, "åt" är verb och "sakta" är adverb. Om man ombads ta ut satsdelarna skulle man säga att "pojken" är subjekt, "åt" är predikat, "glassen" är objekt och "sakta" är adverbial. Om man skulle rita ett diagram var det fråga om ett satsdelsdiagram och det kunde då bli som nedan där en motsvarande parentesnotation också givits.

Funktionell analys (satsdelsanalys)

*Trädrepresentation*



*Parentesrepresentation*

s(subj(pojken),pred(åt),obj(glassen),advl(sakta))

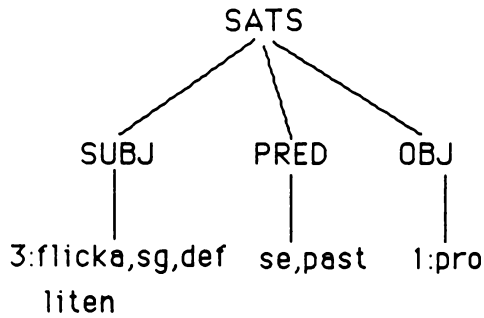
Referentgrammatik gör samma uppdelning i en ordklassanalys och en satsdelsanalys, men kallar den senare funktionell analys i anslutning till modernt internationellt språkbruk. Varken traditionell eller modernare grammatikforskning har givit terminologi och metodik som gör att man är överens i alla detaljer om hur olika ord och ordgrupper skall benämnas. När man utvecklar en grammatik som skall kunna tillämpas av en dator måste man

emellertid bestämma sig för terminologi och representationssätt och vara konsekvent. Vi skall beröra några av dessa problem.

I parentesnotationen ovan har satsdelsbeteckningarna förkortats och bl.a. dessa förkortningar återkommer i bilagda utskrifter av datorkörningar: s=sats, subj=subjekt, obj=objekt, pred=predikat, advl=adverbial, sadvl=satsadverbial.

De funktionella representationerna är avsedda att vara ett interface mot logik och semantik. Att döma av den framgång man haft när man beskrivit språk efter denna traditionella mall har de en avsevärd universalitet. Det är det som gör det möjligt att använda dem som en mellanrepresentation vid översättning mellan språk. Det är en fördel om ett standardformat då kan användas och i den datorimplementering av referentgrammatik som gjorts av SWETRA (Swedish Automatic Translation Group, Lund) har man preliminärt bestämt sig för ett format omfattande maximalt 10 konstituenten (också nämnda i denna ordning i funktionella representationer): subj, pradv, pred, dobj, obj, sadvl, sadvl, advl, advl, advl. I följande sats förekommer t.ex. ett predikat, ett partikeladverbial (pradv), två satsadverbial samt två vanliga adverbial: "Pojken kanske inte sprang in till staden igår. Motsvarande funktionella representation skulle bli: s(subj(pojken),pradv(in), pred(sprang), sadvl(kanske),sadvl(inte), advl(till, staden),advl(igår)).

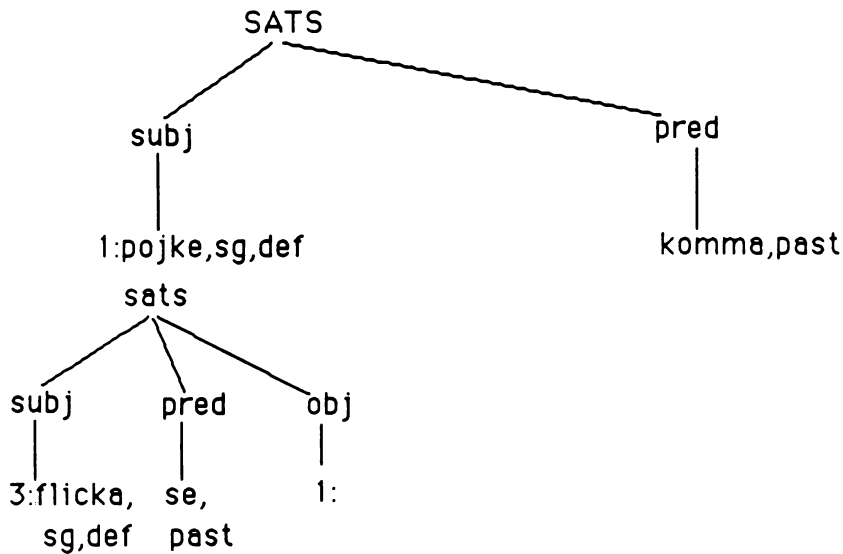
Bestämningar till substantiv inom en nominalfras kan föras upp (efter "nom" =namn i datorprogrammet) som ytterligare upplysningar efter huvudordet, det ord som avgör kongruensen inom nominalfrasen. Man kan också analysera orden vidare morfologiskt och t.ex. sätta "pojke,sg,best" för "pojken" eller, med användande av angliserande semantisk representation "boy,sg,def". På samma sätt kan man representera "åt" med "eat, past". I datorimplementeringar av referentgrammatik avsedda för automatisk översättning användes normalt sådana angliserande semantiska representationer (machineese). En funktionell analys av "Den lilla flickan såg honom" är följande.



s(subj(3:flicka,sg,def,liten),pred(se,past),obj(1:pro))

I denna funktionella representation har också satts in de referentnummer som givit referentgrammatik dess namn. Subjektet räknas som referent nummer 3 i den text som de två exempelmeningarna bildar: "Pojken(1) åt glassen(2) sakta. Den lilla flickan(3) såg honom(1)". Objektet i den sista satsen har fått nummer 1, eftersom det är fråga om samma referent som tidigare, vilken hänvisas till med ett pronomen (pro) denna gång.

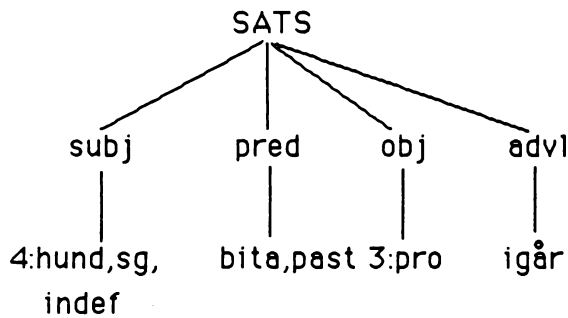
En relativsats är också en bestämning inom en nominalfras och den kan ge en ytterligare upplysning om referenten ifråga. En funktionell analys av "Pojken som flickan såg kom" är följande:



s(subj(1:pojke,s(subj(3:flicka),pred(se,past),obj(1))),pred(kom))  
sg,def sg,def

Dessa representationer visar att bland bestämningarna till referent nummer 1 finns en sats där den tidigare nämnda flickan (referent nummer 3) är subjekt och referent nummer 1 objekt. Detta stämmer väl med hur man brukar uttrycka sig i fråga om relativsatser. Man skulle här säga att objektet i relativsatsen "som flickan såg" är "som", vilket i sin tur syftar på samma sak som ordet "pojken" (korrelatet). Referentgrammatik konkretiserar genom sina nummer den referent som traditionell och modern grammatik talar om.

Satsen "Igår bet en hund henne" har nedanstående funktionella representation, där ordningen i den funktionella representationen blir: subj, pred, obj, advl, även om satsdelarna inte kommer i den ordningen i den föreliggande satsen. Ordet "henne" antas syfta på den tidigare nämnda flickan och hunden blir då den fjärde (nominella) referenten i texten.



s(subj(4:hund,sg,indef),pred(bita,past),obj(3:pro),advl(igår))

Referentgrammatisk analys ger också uppgift om satsens modus (deklarativ, fråga, imperativ) och fokuserad satsdel - i satsen ovan "igår". Satser kan ha samma funktionella representationer, men olika satsdel i fokus (som fundament för att följa Diderichsens terminologi). Funktionella representationer för några satser med s.k. formellt subjekt ("det") är följande, där "\_" markerar tom plats:

*Funktionell representation*

s(subj(\_),pred(regna,past))

s(subj(sparv),pred(sitter),advl(där))

s(subj(s(subj(du),pred(kom)),pred(bra))

*Sats*

Det regnar

Det sitter en sparv där

Det var bra att du kom

Man kan själv välja hur detaljerad eller specifik man vill ha sin representation: om man vill sätta ut referentsiffror, om man vill ge detaljerad ordanalys etc. I satsen "Det var bra att du kom" har inte tempus (past) representerats (vilket man kan göra genom att tillägga "past" vid "bra"). Den motsvarande satsen "Att du kom var bra" har samma funktionella representation i referentgrammatik, men i den är fokuseringen en annan. Satsen "En sparv sitter där" skulle få samma funktionella representation som "Det sitter en sparv där", men har inte samma konstituent i fokus. Många småord t.ex. "som", "att", "det" syns inte i den funktionella representationen.

Man kan utforma den funktionella representationen med tanke på de syften man har med den. Det är naturligt att låta frågor få en funktionell representation som motsvarar svaret. Det gör att den funktionella representationen motsvarande "Vem bet hunden igår?" bör vara:  $s(\text{subj}(\text{vem}), \text{pred}(\text{bet}), \text{obj}(\text{hunden}), \text{advl}(\text{igår}))$  och den funktionella representationen motsvarande "När bet hunden henne?" bör vara:  $s(\text{subj}(\text{hunden}), \text{pred}(\text{bet}), \text{obj}(\text{henne}), \text{advl}(\text{när}))$ . Frågan "Bet hunden henne igår?" kan ha samma funktionella representation som "Hunden bet henne igår"; skillnaden noteras i modusvärdet som är "d"(=deklarativ) för ett påstående, "q"(=question) för en fråga. I frågor med frågeord anses de stå i fokus, i ja/nej-frågor anses det finita verbet stå i fokus. Referentgrammatik är förberedd för textlingvistisk analys genom möjligheten att hålla reda på referenterna (diskursreferenterna) i texten och registrera vad som fokuseras i serier av meningar i en text.

En funktionell representation för frågan "Vem sade pojken att hunden bet?" bör identifiera "vem" som objekt till "bet", men dessutom bör den ta hänsyn till att en frågan "Vad sade pojken?" bör ha "vad" som objekt och att att-satsen bör vara objekt i svaret "Pojken sade att en hund bet flickan". En lämplig funktionell representation för "Vem sade pojken att hunden bet?" är då följande:  $s(\text{subj}(\text{pojken}), \text{pred}(\text{sade}), \text{obj}(s(\text{subj}(\text{hunden}), \text{pred}(\text{bet}), \text{obj}(\text{vem}))))$

Liksom ord t.ex. "som" kan sakna motsvarighet i den funktionella representationen så kan den funktionella representationen innehålla element som saknas i den föreliggande satsen. Exempel på detta ger s.k. kontrollverb, t.ex. "lova". Satsen "Per lovade (att) komma" brukar analyseras så att man säger att subjektet i den överordnade satsen (Per) kontrollerar och är subjekt till infinitiven "komma". I

referentgrammatik kan man visa detta i den funktionella representationen genom att sätta in ett subjekt vid infinitiven:  
s(subj(Per),pred(lovade),obj(s(subj(Per),pred(komma))))

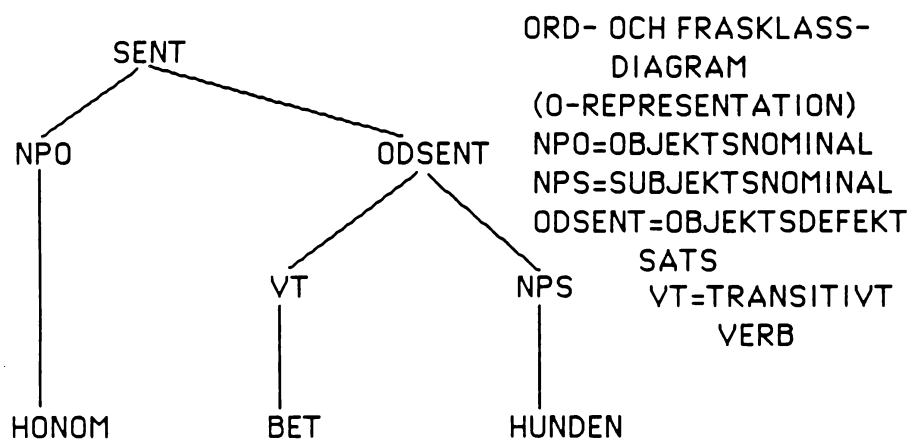
### Ord- och frasklassanalys (kategorianalys)

Som nämnts skiljer referentgrammatik liksom traditionell grammatik mellan funktionell analys i subjekt, predikat, objekt etc. och analys i ordklasser. Det finns inte en så väletablerad modern term som funktionell analys för ordklassanalys och vi väljer att tala om ord- och frasklassanalys, eftersom analysen utvidgas till fraser. Vi räknar med både fulla och defekta fraser i referentgrammatik.

Uppslaget att räkna med defekta fraser kommer från Gazdar. En kategori (fras eller sats) säges vara defekt därför att den saknar en "väntad" konstituent. En prepositionsfras som "på båten" blir defekt då den bara innehåller prepositionen "på" som i satsen: "Vad tittade Per på?" (Uppenbarligen är "Vad" den saknade konstituenten). I satsen "Flickan trodde pojken att hunden bet?" är bisatsen "att hunden bet" defekt därför att det transitiva verbet "bet" saknar ett objekt - det som motsvaras av "Flickan". I den funktionella representationen är dessa saknade element placerade på sina platser såsom vi har sett. Gazdar intresserar sig i sin grammatik, som kallas Generaliserad frasstrukturgrammatik, dock bara för vad vi kallar ord- och frasklassanalys inte för en särskild funktionell representation. Referentgrammatik har övertagit en rad benämningar för ord- och fraskategorier och försökt systematisera dem så att de kan användas i datorprogram för analys och syntes av meningar. Kategorierna väljs och definieras så att de kan användas i de program av grammatiska regler som härleder en funktionell representation från en föreliggande sats.

Meningen "Honom bet hunden" har klart en funktionell representation: s(subj(hunden),pred(bet),obj(honom)) och en analys enligt referentgrammatikens principer ger också upplysningen att det är en deklarativ sats och att "honom" står i fokus. En referentgrammatik för svenska måste skilja mellan subjektsnominalfraser (nps) och objektsnominalfraser (npo) även om bara pronomen visar denna skillnad genom skilda former. I den föreliggande meningen identifierar vi "honom" som ett objektspronomen (proo i internationell terminologi), "bet" som ett

transitivt finit verb (vt) och "hunden" som en form som kan stå som subjektsnominal (nps). Tillsammans utgör de sista två orden "bet hunden" en objektsdefekt sats och vi benämner en sådan sats "odsent" där o=objekt, d=defekt,sent=sentence. Vi konstaterar att en objektsdefekt sats innehåller ett transitivt verb följt av ett nps. En motsvarande ord- och frasklassrepresentation (o-representation) i parentesform blir då: sent(np(honom),odsent(vt(bet),nps(hunden))). Vi använder "sent" som namn på rotnoden i o-representationen. Detta motsvarar nedanstående träd-diagram under vilket vi också skrivit den funktionella representationen.



*Funktionell representation*

s(subj(hunden),pred(bet),objekt(honom))

Ord- och frasklassanalysen kan ses nerifrån (motsvarande en bottom-up-process) som identifikation av "honom" som ett NPO, av "bet" som ett VT och "hunden" som ett NPS. Ett VT följt av ett NPS kan sedan enligt en regel identifieras som en ODSENT och ett NPO följt av en ODSENT kan identifieras som en SENT. Sett uppifrån (som en top-down-process) kan diagrammet sägas visa att en sats (SENT) kan bestå av ett objektsnominal (NPO) om det följer en objektsdefekt sats (ODSENT) efter. En objektsdefekt sats kan bestå av ett transitivt verb (VT) följt av ett subjektsnominal (NPS).

I satsen "Vem bet hunden?" kan vi på motsvarande sätt se "Vem" som ett objektsnominal (ett frågeobjektsnominal: npqo) följt av en sdsent, om vi tänker på en sats med den funktionella representationen: s(subj(hunden),pred(bet),obj(vem)). Samma sats har uppenbarligen en annan möjlig -om än oväntad- tolkning

motsvarande:  $s(\text{subj}(\text{vem}), \text{pred}(\text{bet}), \text{hunden})$  . Den senare tolkningen motsvarar o-analysen:  $\text{sent}(\text{nps}(\text{vem}), \text{sdsent}(\text{vt}(\text{bet}), \text{npo}(\text{hunden})))$ . Här betecknar "sdsent" subjektsdefekt sats - vi skall se närmare på de olika kategorierna nedan.

### Referentgrammatiska regler

De grammatiska reglerna i referentgrammatik beskriver hur olika kategorier kan kombineras till högre kategorier, vilken ordning kategorier skall komma i och vilka ord som tillhör dem. Dessutom talar de referentgrammatiska reglerna om hur den motsvarande funktionella representationen ser ut. Reglerna härleder den funktionella representationen samtidigt som de tillämpas på en föreliggande sats. Omvänt kan reglerna härleda de sekvenser av fraskategorier och slutligen ord som motsvarar en funktionell representation. Regler för referentgrammatik skrivs i ett för lingvister bekvämt format kallat DCG (Definite Clause Grammar) som vanligen finns tillgängligt i nyare implementeringar av programmeringsspråket Prolog. Sådana regler förstår datorn direkt och reglerna kan alltså användas både för analys (parsning) och syntes (generering) av meningar. De regler som behövs för att göra de ovan beskrivna analyserna av satsen "Honom bet hunden" är följande (något förenklade, bl.a. utan referentnummer; "\_" kan stå på en plats för en variabel som inte är relevant (anonym) för tillfället):

```
sent(d,_,X,F) --> npo(X),odsent(,_,X,F).
odsent(,_,X,s(subj(Y),pred(Z),obj(X)) --> vt(Z),nps(Y).
npo(honom) --> [honom].
nps(hunden) --> [hunden].
vt(bet) --> [bet].
```

Om man skriver:  $\text{sent}(\text{M}, \_, \text{T}, \text{F}, [\text{honom}, \text{bet}, \text{hunden}], [])$  till datorn (vederbörligen laddad) så returnerar den den önskade funktionella representationen såsom ett värde hos F och "d" som ett värde för modus (M). Reglerna ger inte någon o-representation. Om man skriver  $\text{sent}(\text{d}, \_, \text{honom}, \text{s}(\text{subj}(\text{hunden}), \text{pred}(\text{bet}), \text{obj}(\text{honom})), \text{X}, [])$  så ger programmet satsen "honom,bet,hunden" som ett värde för X. ([] - den tomma listan - kan sägas uttrycka att det inte blir några ord kvar).



I reglerna, liksom allmänt i Prolog, användes stora bokstäver för att beteckna variabler. Man kan beskriva hur reglerna tillämpas i analys (parsning) på följande sätt. Se efter om det första ordet (X) är en npo. Om så är fallet se efter om det kommer en odsent efter. Ta i så fall dess funktionella representation (F) och sätt på motsvarande plats i sent. Odsent saknar en konstituent och kommer att sätta in det funna X som objekt i sin funktionella representation. Den funktionella representation som hamnar på plats i sent kommer därför att vara komplett. I den aktuella meningen hittar programmet "bet" som är ett transitivt verb (vt) och blir värdet för Z. Sist hittar programmet "hunden" som är ett nps och då blir Y. Värdena av X,Y,Z sätts då in i den funktionella representationen till vänster om pilen. Prolog har en inbyggd parser som försöker utföra den uppgift man beskrivit i reglerna. Prolog kallas ett deklarativt språk (i motsats till proceduralt språk). Man kan emellertid följa sökprocessens gång genom att använda inbyggda kommandon som "trace" och "spy".

Lägg märke till att referentgrammatik inte behöver några transformationsregler. Formalismen tillåter en att specificera vilken representation man vill till vänster om pilen som en motsvarighet till ord- och fraskategorierna till höger om pilen. Man kan placera satsdelarna i vilken ordning man vill och kalla dem vad man vill. Likaså behöver inte referentgrammatik några spår (trace, e) eller noll-konstituenten som EST-, GB- och GPSG-grammatik. Flyttade konstituenten kan sägas åka med på vissa specialplatser i fraskategorierna såsom framgått. Dessa specialplatser fungerar som tillfälliga landningsplatser och motsvarar i vissa fall de COMP-noder som andra modeller utnyttjar för liknande ändamål (se Sell,1986).

Vi skall emellertid inte gå in längre på reglerna här, men tala om att de såvitt vi kan se kan hantera alla de problem som behöver hanteras: ordföljd, optionella konstituenten, kongruens, långa flyttningar (unbounded dependencies), ordböjning. Det finns ganska omfattande referentgrammatiska regler (program) för svenska och engelska som styrker detta påstående. Dessa moduler är utarbetade inom SWETRA där de används för översättning mellan språken via den funktionella representationen.

### Mening, huvudsats och bisats

Vi har hittills talat om satser och meningar som man brukar utan att göra klar åtskillnad. I själva verket måste man betrakta mening (meng) som en högre enhet som kan bestå av en eller flera huvudsatser (sent) vilka i sin tur kan innehålla en eller flera bisatser (sunt) som satsdelar - i regel objekt eller adverbial. Vi räknar med att det egentligen är meningen som innehåller interpunktionstecknet (punkt, frågetecken, utropstecken). Meningen "Men pojken sprang och flickan hoppade." anses ha följande o-representation:

meng(konj(men),s(subj(pojken),pred(sprang),konj(och),  
s(subj(flickan),pred(hoppade)))).

Vi kallar "men" och "och" för konjunktioner i anslutning till traditionell grammatik här. Lägg märke till att den inledande konjunktionen "men" inte leder till omvänd ordföljd i den efterföljande satsen vilket ett inledande adverb i en sats (sent) skulle ha gjort. Samordnade satser där subjektet i den senare satsen saknas (strukits, eliderats) kan i referentgrammatik naturligen beskrivas genom att den senare satsen betraktas som subjektsdefekt. Så kan man t.ex. analysera: "Pojken sprang och sjöng" där "sjöng" då representerar den subjektsdefekta satsen, men man kan i många fall tveka om man skall analysera en sådan sats såsom innehållande två samordnade verb i stället.

Det är nödvändigt att skilja mellan huvudsats (sent) och bisats (sunt) i svenskan (däremot inte i t.ex. engelskan och polskan), eftersom som bekant satsadverben placerar sig efter det finita verbet i huvudsats men före i bisats. Vi illustrerar de olika ordningarna av konstituenten som karakteriserar olika satstyper i appendix. Där visar vi också hur man generellt kan se svenska satser såsom bestående av ett initialt element följt av en konstituent som saknar detta element. Man kan betrakta "Honom bet hunden" som ett npo(honom) följt av en odsent(vt(bet),sps(hunden)). Men vi betraktar inte bara objektsinledda och adverbialsinledda satser på detta sätt utan också subjektsinledda. Sålunda analyserar vi "Hunden bet flickan" som: sent(nps(hunden),sdsent(vt(bet),npo(flickan))).

Det betraktelsesätt vi anlagt ovan är högst naturligt för nordiska lingvister, som känner till Paul Diderichsens satsschema där man just avskiljer ett första element, det s.k. fundamentet från

resten. Diderichsen observerar som bekant sedan att de följande konstituenterna alltid kommer i samma ordning och att det finns en tom plats från vilken man kan säga att konstituenten i fundamentet flyttats. Diderichsen är emellertid inte så uppmärksam på de serier av konstituenten som verkligen kan följa på ett visst fundament. Det är dessa serier eller mönster som identifieras av referentgrammatik och kallas subjeksdefekt sats (sdsent), objektsdefekt sats (odsent), adverbdefekt sats (adsent), etc.

### Nominalfras och relativsats

Typiska nominalfraser är: "den lille pojken", "flickan vid fönstret", "barnet som lekte". Ur ordklassynpunkt kan vi urskilja den framförställda artikeln "den", adjektivet "lille", det relativa pronomet "som" samt gängse verb och andra ordklasser i relativsatsen. En nominalfras kan enligt referentgrammatisk analys bestå av ett huvud (nph) vilket består av ett substantiv föregånget av en eller flera bestämningar inklusive artikel. Substantivet i huvudet styr kongruensen inom nominalfrasen genom att referenten berikas med de grammatiska drag (definithet, genus, numerus) som behövs. Efter substantivet kan förekomma en eller flera prepositionsfraser som kallas "ppa" och dessutom kan det förekomma en eller flera relativsatser "relcl". Förenklat kan vi visa detta genom följande regel som t.ex. kan beskriva: "Barnet på gården som sprang" (de olika upplysningarna A,B,C samlas vid "nom"=namn.

$np(R, \text{nom}(A, B, C)) \rightarrow \text{nph}(R, A), \text{ppa}(R, B), \text{relcl}(R, C).$

Regeln visar att samma referentnummer (berikad med grammatiska drag) återfinnes inom nph, ppa och relcl. Vi skall nu se närmare på relativa satser. Som nämnts tidigare är det naturligt att betrakta relativa satser som defekta och relativmarkören som en förmedlare av den saknade konstituenten. Den saknade konstituenten kan vara olika satsdelar och det är då lämpligt att urskilja olika relativa markörer: rels (subjektsrelativmarkör), relo (objektsrelativmarkör), relg (genitivrelativmarkör) och rela (adverbrelativmarkör). Dessa kombineras med motsvarande defekta kategorier: (sdsunt, odsunt, adsunt, odpp, gdnp) såsom exemplifieras av nedanstående analyser.

1. Flickan som sprang

*o-representation*

np(nph(flickan),relcl(rels(som),sdsunt(vi(sprang))))

*funktionell representation*

np(1,(flickan,s(subj(1),pred(sprang))))

2. Flickan som hunden bet

*o-representation*

np(nph(flickan),relcl(relo(som),odsunt(nps(hunden),vt(bet))))

*funktionell representation*

np(1,(flickan,nom(s(subj(2,hunden),pred(bet),obj(1))))

3. Flickan som hunden sprang till

*o-representation*

np(nph(flickan),relo(som),odsunt(nps(hunden),vi(sprang),odpp(till))))

*funktionell representation*

np(1,(flickan,s(subj(2,hunden),pred(sprang),advl(till,1))))

4. Flickan till vilken hunden sprang

*o-representation*

np(nph(flickan),relcl(rela(p(till),relo(vilken)),adsunt(nps(hunden),  
vi(sprang))))

*funktionell representation*

np(1,(flickan,s(subj(2,hunden),pred(sprang),advl(till,1))))

5. Flickan vars hund sprang

*o-representation*

np(nph(flickan),relcl(rels(relg(vars),n(hund)),sdsunt(vi(sprang))))

*funktionell representation*

np(1,(flickan,s(subj(2,(hund,poss(1)),pred(sprang))))

6. Flickan till vars mor hunden sprang

*o-representation*

np(nph(flickan),relcl(rela(p(till),relo(relg(vars),n(mor)),  
adsunt(nps(hunden),vi(sprang))))

*funktionell representation*

np(1,(flickan,s(subj(2,hunden),pred(sprang),  
advl(till,np(3,(mor,poss(1))))))

I ovanstående uppställning är förkortningen "odpp"= objektsdefekt prepositionsfras, vilken då bara representeras av en preposition. Vi har representerat genitivens motsvarighet i den funktionella representationen med "poss()".

Vi observerar att "som" aldrig kan strykas (vara []) före sdsunt, men väl före odsunt, ett känt faktum. Vi observerar också att "som" inte kan vara objektsrelativ (relo) efter preposition. Man kan inte säga "flickan till som hunden sprang". Det finns ett antal relativa adverbial och konjunktioner som kan analyseras på motsvarande sätt, t.ex. "varest", "varigenom", "då", "när".

### Satser med hjälverb eller particip

I referentgrammatik betraktas infinitiver och particip som "minor sentences", satser som saknar en eller flera konstituenten, vilka kan fyllas i från omgivningen. En sats som "Per vill simma" får denna funktionella representation:

$s(\text{subj}(\text{Per}), \text{pred}(\text{vill}), \text{obj}(s(\text{subj}(X), \text{pred}(\text{simma}))))$

I denna representation kan man direkt sätta Per som X eftersom det ju är Per som skall simma. Subjektet i sådana satser med kedjor av verb är aktuellt som subjekt även i senare infinitiver som visas av: "Per vill försöka sluta röka". Analysen med infinitiven som objekt stöds också av frågor som: "Vad vill Per?" Ord- och frasanalysen av "Per vill simma" är:  $\text{sent}(\text{nps}(\text{Per}), \text{sdsent}(\text{mod}(\text{vill}), \text{isent}(\text{simma})))$  där "mod" betecknar modalt verb och "isent" betecknar infinitivsats.

En infinitivsats kan sakna en satsdel (förutom subjektet) såsom framgår av en mening som: "Vem vill hunden bita?" där "Vem" måste vara objekt till "bita" i den funktionella representationen:  $s(\text{subj}(\text{hunden}), \text{pred}(\text{vill}), \text{obj}(s(\text{subj}(\text{hunden}), \text{pred}(\text{bita}), \text{obj}(\text{vem}))))$ . För att hantera detta räknar referentgrammatik med defekta infinitivsatsen också, i detta fall en "odisent" (objektdefekt infinitivsats).

På motsvarande sätt behandlas particip. Satsen: "Per har hämtat bilen" har den funktionella representationen:  $s(\text{subj}(\text{Per}), \text{pred}(\text{har}), \text{obj}(\text{pfsent}(\text{subj}(\text{Per}), \text{pred}(\text{hämtat}), \text{obj}(\text{bilen}))))$ . I satsen "Bilen är hämtad" måste däremot "bilen" bli objekt till participet "hämta".

### Avslutning

Som framgått kan man använda referentgrammatik i menings- och textanalys "för hand", men dess styrka är att den är så systematiskt och konsekvent formulerad att den också kan användas för datoranalys och datorsyntes av meningar. Vi har antytt ovan hur det går till. De fragment av svenska engelska och några andra språk som

programmerats och kan köras på VAX eller PC kan också användas för automatisk översättning, men det krävs ofta transferregler där den funktionella representationen för meningen i ett språk överförs till den motsvarande funktionella representation i det andra språket. Även om de funktionella representationerna är abstrakta och avlägsnar mycket av de enskilda språkens "lokala" idiosynkrasier så är de ofta inte tillräckligt genereralla. Vi skall emellertid inte gå in på dessa problem här.

### *LITTERATUR*

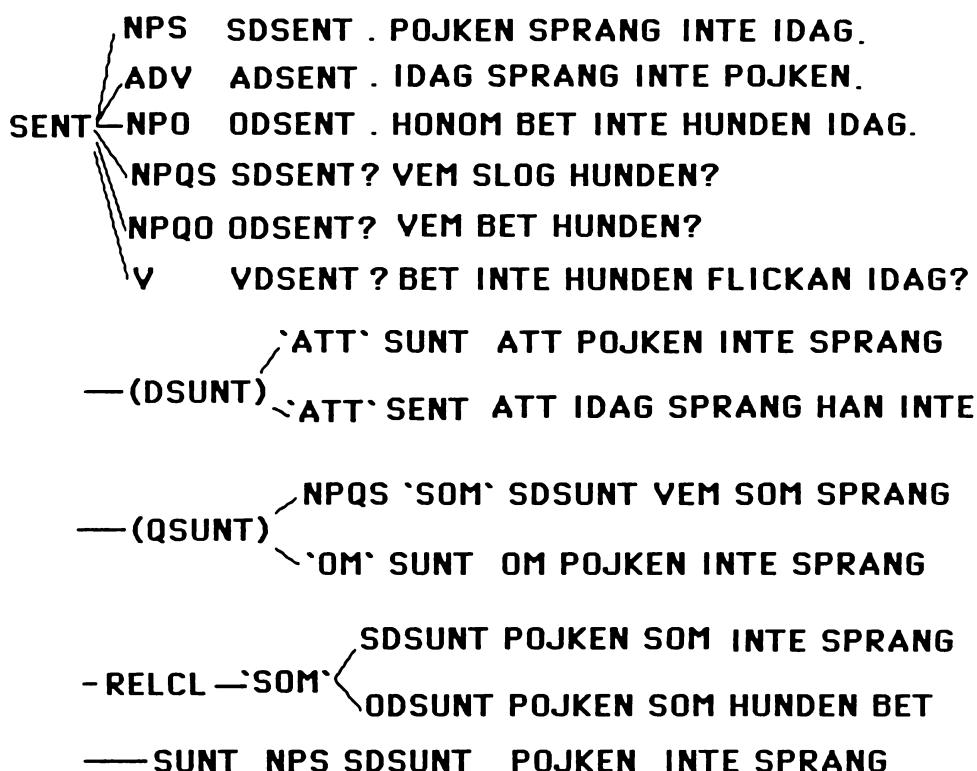
- Chomsky, N. Lectures on Government and Binding. Dordrecht  
(1981:Foris)
- Diderichsen, P. Elementaer dansk grammatik. København (1946)
- Gazdar, G., E. Klein, G. Pullum & I. Sag, Generalized Phrase Structure  
Grammar. Oxford (1985: Basil Blackwell)
- Sells, P. Lectures on contemporary syntactic theories. CSLI,  
Stanford (1985)
- Sigurd, B. Referent grammar (RG). A generalized phrase structure  
Grammar with built-in referents. *Studia Linguistica*  
1987:2
- Sigurd, B. A referent grammatical analysis of relative clauses.  
Working Paper Dept of Linguistics and Phonetics,  
Lund 1988.
- Gawronska-Werngren, B. A referent grammatical analysis of relative  
clauses in Polish (1988: manuscript)

Föredrag vid Nordiska Datalingvistdagarna i Köpenhamn 3-4  
nov, 1987

I SWETRA (Swedish Automatic Translation Group, Lund) arbetar  
också Mats Eeg-Olofsson och Lars Gustafsson (med stöd av HSFR).

APPENDIX (1)

B.SIGURD.REFERENT GRAMMAR (1987)



SOME BASIC TYPES OF MAIN SENTENCES (SENT) AND  
SUBORDINATE SENTENCES (SUNT).

SDSENT= SUBJECT DEFECTIVE SENTENCE

ODSENT= OBJECT DEFECTIVE SENTENCE

ADSENT= ADVERBIAL DEFECTIVE SENTENCE

SDSUNT= SUBJECT DEFECTIVE SUBORDINATE SENTENCE

NPS= SUBJECT NP, NPO= OBJECT NP

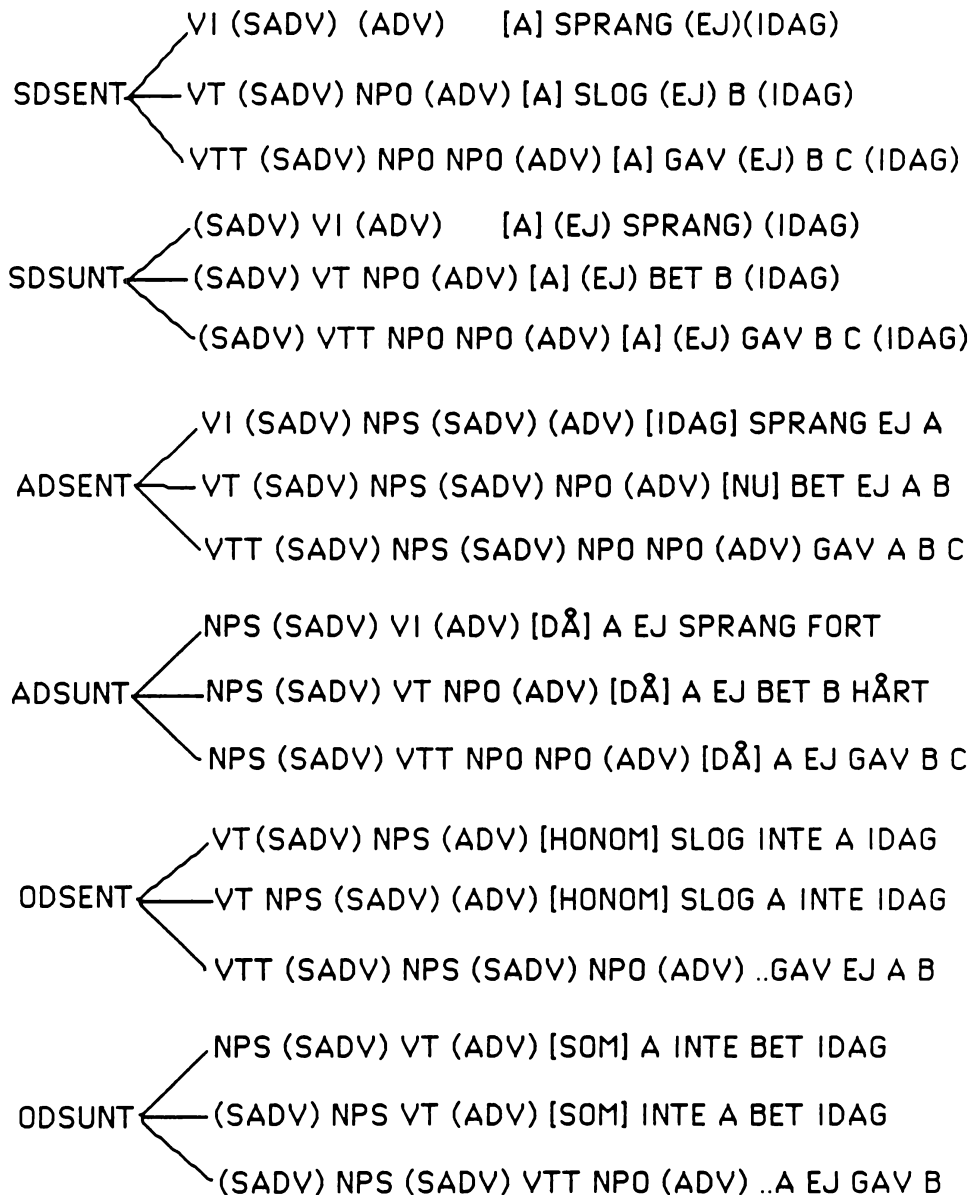
NPQS= SUBJECT QUESTION WORD (VEM,VAD)

NPQO= OBJECT QUESTION WORD (VEM,VAD)

ADV= ADVERB

NOTE THE DIFFERENT PLACEMENT OF SENTENCE  
ADVERBS IN MAIN AND SUBORDINATE SENTENCES

APPENDIX (2)      **B.Sigurd,Referent Grammar (1987)**



WORD (CONSTITUENT) ORDER WITHIN DEFECTIVE MAIN (SENT) AND SUBORDINATE (SUNT) DEFECTIVE SENTENCES. SDSENT= LACKING SUBJECT, NPS=SUBJECT NP, NPO=OBJECT NP, VI=INTRANSITIVE VERB, VT=TRANSITIVE VERB, VTT= DOUBLY TRANSITIVE VERB, SADV=SENTENCE ADVERB



UTSKRIFTER AV REFERENTGRAMMATISKA DATORANALYSER.

CHECK BEGÅR ANALYS. FÖRSTA RADEN GER MODUS (d,q).

ANDRA RADEN GER FOKUSERAD KONSTITUENT.

TREDJE RADEN GER FUNKTIONELL ANALYS.

?- check([pojken,lovade,flickan,att,komma,.]).

```
d
np(_1, nom(m(boy, sg), m(def)))
s(subj(np(_1, nom(m(boy, sg), m(def))))), pred(m(promise, past)), dobj(np(_2,
  nom(m(girl, sg), m(def))), aobj(s(subj(np(_2, nom(m(girl, sg),
  m(def))), pred(m(come))))))
```

```
d
np(_1, nom(m(boy, sg), m(def)))
s(subj(np(_1, nom(m(boy, sg), m(def))))), pred(m(promise, past)), dobj(np(_3,
  nom(m(girl, sg), m(def))), aobj(s(subj(_2), pred(m(come))))))
```

?- check([pojken,vill,komma,.]).

```
d
np(_1, nom(m(boy, sg), m(def)))
s(subj(np(_1, nom(m(boy, sg), m(def))))), pred(m(will, pres)), obj(s(subj(np(_1
  , nom(m(boy, sg), m(def))), pred(m(come))))))
no
```

?- check([pojken,vill,kunna,komma,.]).

```
d
np(_1, nom(m(boy, sg), m(def)))
s(subj(np(_1, nom(m(boy, sg), m(def))))), pred(m(will, pres)), obj(s(subj(np(_1
  , nom(m(boy, sg), m(def))), pred(m(can, inf)), obj(s(subj(np(_1,
  nom(m(boy, sg), m(def))), pred(m(come)))))))))
no
```

?- check([pojken,har,sprungit,.]).

```
d
np(_1, nom(m(boy, sg), m(def)))
s(subj(np(_1, nom(m(boy, sg), m(def))))), pred(perf), obj(s(subj(np(_1,
  nom(m(boy, sg), m(def))), pred(m(run, perf))))))
```

?- check([barnet,som,pojken,som,hunden,bet,slog,sprang,.]).

```
d
np(_1, nom(m(child, sg), m(def), s(subj(np(_2, nom(m(boy, sg), m(def),
  s(subj(np(_3, nom(m(dog, sg), m(def))), pred(v(m(bite, past))),
  obj(_2))))), pred(v(m(hit, past))), obj(_1))))))
s(subj(np(_1, nom(m(child, sg), m(def), s(subj(np(_2, nom(m(boy, sg), m(def),
  s(subj(np(_3, nom(m(dog, sg), m(def))), pred(v(m(bite, past))),
  obj(_2))))), pred(v(m(hit, past))), obj(_1))))), pred(v(m(run,
  past))))))
no
```

UTSKRIFTER AV REFERENTGRAMMATISKA DATORANALYSER.

CHECK BEGÄR ANALYS. FÖRSTA RADEN GER MODUS (d,q).

ANDRA RADEN GER FOKUSERAD KONSTITUENT.

TREDJE RADEN GER FUNKTIONELL ANALYS.

?- check([vem,sade,pojken,sprang,?]).

```

q
np(_1, nom(who))
s(subj(np(_2, nom(m(boy, sg), m(def))))), pred(m(say, past)), obj(s(subj(np(_1,
  nom(who))), pred(v(m(run, past))))))

```

?- check([vem,trodde,pojken,att,flickan,hoppades,att,hunden,bet,?]).

```

q
np(_1, nom(whom))
s(subj(np(_2, nom(m(boy, sg), m(def))))), pred(m(believe, past)), obj(s(subj(np
  (_3, nom(m(girl, sg), m(def))))), pred(m(hope, past)), obj(s(subj(np(_4
    , nom(m(dog, sg), m(def))))), pred(v(m(bite, past))), obj(np(_1,
      nom(whom))))))))

```

```

?-
?- check([honom,sade,pojken,att,flickan,hoppades,att,pojken,sade,att,hunden,bet
.]).

```

```

_1
np(_2, nom(m(him)))
s(subj(np(_3, nom(m(boy, sg), m(def))))), pred(m(say, past)), obj(s(subj(np(_4,
  nom(m(girl, sg), m(def))))), pred(m(hope, past)), obj(s(subj(np(_5,
    nom(m(boy, sg), m(def))))), pred(m(say, past)), obj(s(subj(np(_6,
      nom(m(dog, sg), m(def))))), pred(v(m(bite, past))), obj(np(_2,
        nom(m(him))))))))))

```

?- check([flickan,som,pojken,sade,att,han,slog,sprang,.]).

```

d
np(_1, nom(m(girl, sg), m(def), s(subj(np(_2, nom(m(boy, sg), m(def))))),
  pred(m(say, past)), obj(s(subj(np(_3, nom(m(he))))), pred(v(m(hit,
    past))), obj(_1))))))
s(subj(np(_1, nom(m(girl, sg), m(def), s(subj(np(_2, nom(m(boy, sg), m(def))))),
  , pred(m(say, past)), obj(s(subj(np(_3, nom(m(he))))), pred(v(m(hit,
    past))), obj(_1)))))), pred(v(m(run, past))))

```

?- check([vem,trodde,pojken,sade,att,hunden,bet,honom,?]).

```

q
np(_1, nom(who))
s(subj(np(_2, nom(m(boy, sg), m(def))))), pred(m(believe, past)), obj(s(subj(np
  (_1, nom(who))), pred(m(say, past)), obj(s(subj(np(_3, nom(m(dog,
    sg), m(def))))), pred(v(m(bite, past))), obj(np(_4, nom(m(him))))))))))

```