



Mededelingen van de Afdeling Letterkunde, Nieuwe Reeks, Deel 54 no. 2

Deze Mededeling werd in verkorte vorm uitgesproken in de vergadering van de Afdeling Letterkunde, gehouden op 12 november 1990

PIETER A.M. SEUREN

# Grammatika als algoritme: rekenen met taal

KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN  
NOORD-HOLLANDSCHE, AMSTERDAM/NEW YORK/OXFORD/TOKYO, 1991

Copyright van deze uitgave © 1991 Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke wijze dan ook, zonder voorafgaande schriftelijke toestemming van de rechthebbende, behoudens de uitzonderingen bij de wet gesteld.

Zetwerk en druk: Casparie Heerhugowaard bv

## I. CAPTATIO BENEVOLENTIAE

In de vergadering van 10 juni 1963 van de Afdeling Letterkunde van deze Akademie hield de het jaar daarop overleden Evert Willem Beth zijn intussen vermaarde voordracht 'Konstanten van het wiskundige denken'. Hierin zette hij uiteen wat het begrip 'algorithm' tegenwoordig inhoudt. Aanvankelijk stond het voor de vier klassieke rekenkundige procedures van optellen, aftrekken, vermenigvuldigen en delen, zoals die in de 9e eeuw waren ontworpen door Mohammed Ibn Musa, hoogleraar te Beiruth. Van diens bijnaam *Al Huarizmi*, ('de Khivarees') is ook het woord *algorithm* afgeleid, dat oorspronkelijk *algorisme* was, maar later werd verbasterd vanwege een verwarring met het Griekse woord *arithmós*, 'getal'. Sinds echter in de dertiger jaren van deze eeuw het begrip rekenen werd uitgebreid tot elke formeel precies bepaalde operatie op symboolreeksen heeft ook het begrip algorithm een analoge veralgemening ondergaan. Thans omvat het elke formeel bepaalde operatie op een input van één of meer symboolreeksen met als output wederom symboolreeksen.

Beth liet vooral zien dat algorithmes ook gebruikt kunnen worden om verzamelingen van symboolreeksen of, zoals vaak gezegd wordt, van *talen*, precies te definiëren. Met behulp van analytische en structuurverlenende technische hulpsymbolen kan een output worden gekarakteriseerd die beantwoordt aan wat men kan noemen de in grammatikaal opzicht welgevormde zinnen van een taal. Beth verried ons dat de in 1963 inmiddels alom bekende zg. *generatieve grammatika* in feite niets anders is dan een *algorithm* om de welgevormde zinnen van natuurlijke talen te genereren en tegelijk van een structuur te voorzien. De man die hoofdzakelijk verantwoordelijk was voor de verbreiding van het begrip generatieve grammatika, de Amerikaan Noam Chomsky, had dit aspekt nimmer belicht, en het was daarom dat Beths voordracht in die tijd zo uitermate verhelderend werkte. Het is mijn bedoeling om, in het voetspoor van Beth, meer in detail te laten zien hoe sterk de klassieke algorithmes zoals die lang geleden door Al Huarizmi voor de rekenkunde zijn ontworpen in één systeem passen met de transformationele rekenwijzen zoals die sinds de vroege jaren 60 zijn ontwikkeld en thans met name in de semantische syntaxis gebruikt worden.

Zulke overeenkomsten tussen binnen de wiskunde en/of de logika ontwikkelde formele systemen enerzijds en kennelijk psychologische reële, dus in de cognitie werkelijk toegepaste, rekenmethodes anderzijds zijn niet alleen in wetenschappelijke maar ook in filosofische zin uitermate frappant. Hoe komt het dat zuiver wiskundige formele systemen, die toch alleen maar ten doel hebben de metafysisch noodzakelijke structuur van elk mogelijk universum te specificeren, in zeer verwante vormen worden teruggevonden in organismen die door biologische evolutie tot stand zijn gekomen? En direct hierop aansluitend komt de vraag hoe het toch te verklaren en te begrijpen valt dat de, laat ons zeggen, 'biologische' systemen steeds toch op detailpunten van de wiskundige systemen afwijken en er vaak zelfs een, overigens wiskundig ge-

zien meestal oninteressante, verfijning van blijken te zijn? Kortom, de vraag is: hoe komt een biologisch organisme ertoe rekenprocedures te ontwikkelen die zijn eigen functioneren mogelijk maken? Deze filosofische vraagstelling, en tegelijk dit inzicht, van de verrassende doch nooit totale overeenkomsten tussen wiskundige en 'biologische' rekensystemen past weliswaar in het geheel van de problematiek die de laatste tijd in 'philosophy of mind' actueel is geworden, maar heeft daar toch niet de aandacht gekregen die ze verdient.

Bij een dergelijke vraagstelling dient uiteraard voorop te staan dat de in het wetenschappelijk onderzoek aangetroffen 'biologische' systemen inderdaad zo zijn als we denken dat ze zijn. Hoewel, zoals bekend, die zekerheid praktisch gesproken niet voor de volle honderd procent te verkrijgen valt, zijn er toch maatstaven ter toetsing van de plausibiliteit van wetenschappelijke theorieën. Welnu, bij de juist geformuleerde filosofische vraagstelling dienen de gepostuleerde biologische rekensystemen aan de strengst mogelijke toetsen onderworpen te worden. Immers, zo dat niet gebeurt is er te weinig garantie dat de filosofische vraagstelling in feite geen ijdele is.

In deze voordracht zal ik trachten aannemelijk te maken dat in het geval van de grammatika de vraagstelling inderdaad niet ijdel is, en wel omdat het grammatikale rekensysteem dat ik in de taal meen aan te treffen een ijzersterke empirische basis heeft in de taalfeiten en bovendien significante en vaak verrassende generaliseringën bevat over het empirische materiaal. Uiteraard kan men, zeker in kort tijdsbestek, van een algemene theorie niet 'bewijzen' dat ze correct is. Wel kan men frappante voorbeelden laten zien, en wel bij voorkeur van constructies die in andere bestaande theorieën geen bevredigende oplossing gevonden hebben. Het is deze overweging die geleid heeft tot de keuze van de onder behandelde constructies, waarbij ook aandacht geschonken zal worden aan ogenschijnlijk onbelangrijke details. De functie van zulke gedetailleerde observaties is precies te laten zien dát er een systeem werkzaam is, en welk. De filosofische vraagstelling wordt hiermee onderbouwd en acuut gemaakt.

Maar laat ons terugkeren naar Beths voordracht van 1963. Een van de dingen die hij daar liet zien was dat de toen gangbare vormen van generatieve grammatika stilzwijgend genoeg namen met generatieve algorithmes die geen zg. beslissingsprocedure impliceerden. Men kan immers een onderscheid maken tussen die algorithmisch gegenereerde talen waar men voor elke willekeurige symboolreeks  $S$  in het vocabulaire van een taal  $T$  in een eindig aantal stappen kan *beslissen* of  $S$  al dan niet tot  $T$  behoort, en die talen waarvoor dat niet geldt. Welnu, de oudere vormen van generatieve grammatika laten een dergelijke beslissingsprocedure geheel buiten beschouwing. Dit is onbevredigend, daar sprekers van een taal in het algemeen (twijfelgevallen daargelaten) wel degelijk in staat zijn welgevormde van onwelgevormde zinnen te onderscheiden, zodat men mag aannemen dat zij over een soort beslissingsprocedure beschikken.

Inmiddels is er heel wat gebeurd in de grammatikawetenschap. Zo is men nu in brede kring overtuigd van de wenselijkheid om een generatieve gram-

matika tevens van een beslissingsprocedure te voorzien, die men gewoonlijk opvat als een automatische ontleedprocedure voor zinnen, het zg. *parsing*. Maar ook in andere opzichten zijn de ontwikkelingen stormachtig geweest. Vlak na Beths voordracht brak het inzicht door (Katz & Fodor 1963; Katz & Postal 1964), dat een formele (dus algorithmische) grammatika niet alleen maar tot taak had de verzameling welgevormde zinnen van een taal puur vormelijk te karakteriseren, maar dat zij tevens een verantwoording moest leveren van de wijze waarop de betrokken zinnen in semantische zin worden verstaan. Men was immers een algorithmische grammatika gaan zien als een wetenschappelijke hypothese over, of reconstructie van, de bij elke taalgebruiker aanwezige beheersing van zijn taal. En het is nu eenmaal een integraal onderdeel van wat men beheersing van een taal noemt dat geweten wordt wat de zinnen van de taal in kwestie betekenen. Een grammatika heeft daarom in deze opvatting, behalve de puur syntactische taak van de specificatie van alle welgevormde zinnen, bovendien de semantische taak van een specificatie van de mogelijke betekenissen van elke welgevormde zin.

Omstreeks 1970 was het in de taalkunde een algemeen geaccepteerd standpunt dat een grammatika haar semantische taak slechts kon volbrengen door zinnen te relateren aan zg. *semantische analyses*, d.w.z. door ze te 'vertalen', of te 'parafaseren' in termen van een bepaalde semantisch-analytische taal, die al gauw in hoofdzaak overeen bleek te stemmen met de formele taal die tegenwoordig in de predikatencalculus als onderdeel van de logica wordt gebruikt. De ontdekking van deze verwantschap bracht een schok teweeg. Immers nadat de invoering van de moderne predikatencalculus aan het begin van de eeuw had geleid tot een pijnlijke verwijdering van taalwetenschap en logica, daar de logische taal elke grammatische of syntactische relevantie leek te ontberen, bleek nu ineens dat, op een ander en dieper niveau, deze relevantie toch in hoge mate aanwezig was. Logica en taalkunde bleken toch onlosmakelijk verbonden.

Het relateren van zinsstructuren aan de aan hen beantwoordende semantische analyses houdt in dat de ene soort structuren in de andere wordt *ge-transformeerd* via een formeel precies vastgelegde procedure, een algorithmische dus, dat aan de betekenis van de betrokken zinnen niets mag veranderen. We moeten dus een *systeem van betekenisbehoudende transformaties* ontwerpen. Daar echter ook in de generatieve syntaxis, dus de puur vormelijke specificatie van welgevormde zinnen, om onafhankelijke redenen al transformaties waren ingevoerd, rees de vraag naar de verhouding tussen de puur syntactische en de semantische transformaties. En op dit punt ontstonden al gauw zeer uiteenlopende opvattingen.

Eenzijds ontwikkelde zich de opvatting dat de syntactische en de semantische specificaties van zinnen geheel los van elkaar gezien moesten worden. Dit is de opvatting der *autonome syntaxis*, hoofdzakelijk verdedigd door de reeds genoemde Chomsky en zijn volgelingen. Daartegenover ontstond de theorie der *semantische syntaxis* (aanvankelijk meestal *generatieve semantiek* ge-

noemd), die stelt dat het formuleren van de preciese transformationele afbeeldingsrelaties tussen zinnen ('oppervlaktestrukturen') en hun semantische analyses, de semantische taak der grammatika dus, tegelijk de gewenste syntaxis oplevert: een grammatika is niets anders dan een algorithmische transformationele afbeeldingsprocedure tussen oppervlaktestrukturen van zinnen en hun semantische analyses.

Intussen kwam geen der beide scholen toe aan de eigenlijke semantiek. Alles wat hun theorieën vermochten te bewerkstelligen was een 'vertalen' van zinnen in een logisch-analytische taal. Zulk 'vertalen' levert echter alleen weer een nieuwe talige structuur op, en geen specificatie van de betekenis in taalonafhankelijke termen. Hoe een eigenlijke semantiek er uit zou kunnen zien, hoe de semantiek zich zou kunnen bevrijden uit de vicieuze 'synonieme cirkel', is een vraag die tot op de huidige dag binnen de linguïstiek in strikte zin niet of nauwelijks is beantwoord. Daarvoor dient men zich bij andere vakken, zoals de logika of de psycholinguïstiek, te vervoegen. In § 2 zal iets meer hierover worden gezegd.

Momenteel wordt de taalkundige markt in hoge mate overheerst door vormen van autonome syntaxis. Het gedachtenkader van de semantische syntaxis ontmoet op dit moment weinig waardering. Ik ben echter geneigd dit eerder toe te schrijven aan sociologische processen op het 'marktplein' dan aan de *inhoudelijke* krachtverhoudingen van de betrokken theorieën, een neiging die ik baseer op het feit dat er tot op heden geen enkel steekhoudend argument naar voren is gebracht waaruit zou blijken dat de autonome syntaxis in wetenschappelijke zin superieur zou zijn aan haar semantische zuster. Hoe dit echter ook zij, ik zal niet proberen deze kwestie hier te beslechten. De zuiver inhoudelijke wetenschappelijke afweging zal ongetwijfeld een keer plaatsvinden, en dan zal blijken aan welke kant de grootste verdiensten liggen. Wat ik wel zal doen is trachten een idee te geven van de wijze waarop men, in grammatisch opzicht, kan rekenen met taal, en van de verhelderende en verklarende werking die daarvan uit kan gaan. Het is onvermijdelijk dat ik dat doe in het kader vna de semantische syntaxis, het kader waarin ik sinds 25 jaar werk en aan de opbouw waarvan ik heb bijgedragen.

## 2. WAAROM GEEN 'OPPERVLAKTE-SEMANTIEK'?

Alvorens we aan het rekenen kunnen slaan dient nog een enkel woord gezegd te worden over de noodzaak, althans de wenselijkheid, van het aannemen van een apart niveau van semantische analyse als onderscheiden van de oppervlaktestructuur van zinnen. Binnen de taalkunde is dit weliswaar een zeer algemeen geaccepteerd standpunt, maar dit geldt niet voor de logika, waar sinds de jaren 30 eigen opvattingen zijn ontwikkeld over procedures van semantische interpretatie. Daar heeft men steeds af willen zien van een transformationele omzetting van zinnen in hun semantische analyses of omgekeerd. De 'vertaling' in termen van predikatencalculus vindt daar plaats door aan taalkundige eenheden (woorden) een lexicaal vastgelegde interpre-



tatie mee te geven in termen van predikatencalculus. Men maakt daarbij gebruik van geavanceerde logische technieken, met name van de zg. lambda-calculus. Op de technische aspecten hiervan zullen we echter niet ingaan.

De logika heeft het grote voordeel dat ze inderdaad toekomt aan een eigenlijke vorm van semantiek, die uitstijgt boven het ‘vertalen’ van zinnen in een logisch-analytische taal. In de logika bestaat de semantiek in eigenlijke zin uit een zg. modeltheoretische interpretatie, die voor elke zin een waarheidswaarde oplevert gegeven een onafhankelijk gegeven stand van zaken. De betekenis van een zin wordt aldus (althans in de meest gangbare vormen van modeltheoretische semantiek) een functie van standen van zaken naar waarheidswaarden.

Ongeacht de grote en niet weg te cijferen verdiensten van de logische, of ‘formele’, semantiek, moet toch gesteld worden dat het afwijzen van een separaat niveau van semantische analyse voor zinnen, transformationeel verbonden met hun oppervlaktestructuur, niet raadzaam lijkt. Op grond van hun beroepservaring neigen professionele linguïsten ertoe dit methodologische uitgangspunt der formele semantici te riskant te vinden. Hun voornaamste reden berust op het al sinds Plato’s dialoog *Cratylus* erkende feit dat de oppervlaktestructuren van zinnen de neiging hebben hun betekenissen op allerlei manieren te maskeren. Zinnen zijn semantisch onregelmatig. De voorbeelden liggen voor het grijpen wanneer men er een oog voor heeft ontwikkeld. Ik zal er enkele geven. Vergelijk eens de twee Nederlandse zinnen.

- (1) a. Ik heb Karel in geen week gezien.  
b. Ik heb Karel in geen bibliotheek gezien.

De tweede zin betekent iets als ‘er is geen bibliotheek waarin ik Karel gezien heb’. Wanneer we echter een analoge betekenisweergave voor de eerste zin opstellen (‘er is geen week waarin ik Karel gezien heb’) krijgen we een verkeerd resultaat. De eerste zin betekent heel duidelijk ‘het is een week dat ik Karel niet gezien heb’. Omgekeerd krijgen we een foutief resultaat wanneer we de betekenisanalyse van de eerste zin op de tweede toepassen. We zouden dan als betekenis van zin (1b) krijgen: ‘Er is een bibliotheek waarin ik Karel niet gezien heb’, wat natuurlijk niet is wat deze zin betekent. Aangezien (1b) het normale geval vertegenwoordigt, is er kennelijk met (1a) iets speciaals aan de hand.

Voorbeelden uit andere talen zijn eveneens gemakkelijk te vinden. Het Franse *Ça ne doit pas être gai* betekent niet ‘dat hoeft niet leuk te zijn’, hoewel het wel voorwendt dat te betekenen, maar ‘dat is vast niet leuk’ of ‘dat kan onmogelijk leuk zijn’, m.a.w. de zin wordt verstaan alsof er in plaats van het woord *doit* zijn modale tegenpool *peut* gelezen moest worden. Of neem het Londense Cockney Engels, waarin de negatie, soms veelvuldig, wordt gecopieerd in het vervolg van de zin:

- (2) E ain’t never been no good to no woman, not never.

Deze zin bevat zes negaties, afgewisseld door kwantoren. De logika leert ons

dat een letterlijke verlating van de zin in de taal van de predikaten calculus iets oplevert als: 'hij is wel eens goed geweest voor een of andere vrouw'. Maar dat is uiteraard niet wat deze zin betekent. In een fel artikeltje in het tijdschrift *Mind* van 1957 wil Bertrand Russell ons doen geloven dat de natuurlijke taal daarom onlogisch, onbetrouwbaar en slordig is,<sup>1</sup> maar wij houden het er vooralsnog liever op dat elke natuurlijke taal zich aan bepaalde regels houdt die er soms ook verantwoordelijk voor kunnen zijn dat zinnen een andere betekenis tonen dan ze hebben. Aangezien een competente spreker van een taal deze regels, op een impliciete manier, 'kent', is hij echter in staat de werkelijke betekenis feilloos te reconstrueren.

Er is echter nog meer aan de hand. In hetzelfde Cockney dialect vinden we ook zinnen als:

- (3)            There was hardly no money, nor hardly no hope.<sup>2</sup>

Hier vinden we iets wat er uitziet als dezelfde negatie-copiëring als in zin (2). Alleen, er is geen negatie zichtbaar die als uitgangspunt van het copiëeringsproces gefungeerd kan hebben. In standaard Engels luidt deze zin:

- (4)            There was hardly any money, and hardly any hope.

Deze zin draagt zijn betekenis iets 'eerlijker' op zijn gezicht, maar hij bevat geen negatie, althans niet zichtbaar. Wanneer we nu een generalisatie willen treffen en het verschijnsel van negatie-copiëring ook op zin (3) van toepassing willen verklaren, dan ligt het voor de hand dat we voor zin (3), bij wijze van wetenschappelijke hypothese, een 'onzichtbare' negatie postuleren. Welnu, in de semantische syntaxis doen we dat door een logisch-semantische analyse voor zin (3) te postuleren die de benodigde negatie inderdaad bevat. De regels van de syntaxis van de betrokken taal (of taalvariant) zorgen er dan voor dat die onderliggende negatie netjes gecopieerd wordt op die plaatsen waar de regels dat voorschrijven. Alles spreekt voor een dergelijke analyse, met name het feit dat het woord *hardly* in de Engelse grammatika bekend staat als een 'semi-negatief' woord. Negatieve en semi-negatieve woorden hebben o.a. de eigenschap dat ze inversie veroorzaken van onderwerp en persoonsvorm wanneer ze vooraan in de zin staan:

- (5)    a. *Never had she* experienced such a feeling of real power.  
       b. The fog was heavy. *Hardly could we* distinguish the contours of the house.

<sup>1</sup> Russell is op dit punt zeer uitgesproken. Zo schrijft hij in dit korte artikeltje (1957): 'This brings me to a fundamental divergence between myself and many philosophers with whom Mr. Strawson appears to be in general agreement. They are persuaded that common speech is good enough not only for daily life, but also for philosophy. I, on the contrary, am persuaded that common speech is full of vagueness and inaccuracy, and that any attempt at being precise and accurate requires modification of common speech both as regards vocabulary and as regards syntax. ... I have no wish to claim the support of common usage ... I agree, however, with Mr. Strawson's statement that ordinary language has no exact logic.'

<sup>2</sup> Regel uit een Londens volksliedje.

Wat ligt er meer voor de hand dan te postuleren dat het woord *hardly* in z'n logisch-semantische analyse een negatie bevat, zodat het geanalyseerd wordt als bv. 'bijna niet'? De genoemde verschijnselen kunnen dan allemaal aan het vóórkomen van de negatie in de semantische analyse worden gekoppeld.<sup>3</sup>

Een buitengewoon sprekend geval trof ik onlangs aan in een artikel van 1974, van de hand van de Amerikaanse linguïst Paul Schachter, over seriële werkwoordsconstructies. Schachter geeft daar het volgende voorbeeld uit het Akuapem dialect van de Westafrikaanse Kwa-taal Akan:

- (6) Me-de aburou mi-gu msu-m.  
 ik-neem mais ik-stroom water-in  
 ('Ik strooi mais in het water')

Het werkwoord *gu* ('stromen') slaat in dit geval semantisch ondubbelzinnig op het *mais* dat het water in gaat, niet op de ik-figuur. Toch is *mi* ('ik') het grammaticale subject van *gu*, en niet *aburou* ('mais'), en dat terwijl het werkwoord *gu* in normale gevallen niet eens een subject toelaat dat naar een enkel individu verwijst (net als bv. het Nederlandse werkwoord *zich verspreiden*: dat alleen van een groep of een niet-vaste stof, niet van een individu, gezegd kan worden). Het is duidelijk dat de syntaxis hier fors ingegrepen heeft in de oorspronkelijke semantische analyse van de zin. Allereerst is het echte subject van *gu*, nl. *aburou*, weggelaten ('gedeleerd'), ongetwijfeld onder invloed van het feit dat het net als object van *de* ('nemen') is genoemd. En vervolgens is het subject van het hogere werkwoord *de* gecopieerd bij het lagere werkwoord *gu*. Dergelijke copiëringen zijn kennelijk puur mechanisch, en dus niet betekenisveranderend, zij het wel betekenisverdoezelend.

Professionele linguïsten zijn beter vertrouwd met zulke observaties dan hun collega's van logischen huize. Op grond van deze beroepservaring achten zij het vertrouwen der logici op hun formele middelen, zoals de lambda-calculus, hoe geavanceerd en hoe slim bedacht ook, toch misplaatst. Zij vrezden dat het 'vertalen' van zinnen in hun predikaatlogische equivalenten zonder transformaties en slechts met gebruikmaking van logische calculi onmogelijk zal blijken, met name omdat logische transformaties nooit, zoals de syntactische, semantisch bedriegen. In elk geval zullen zij een dermate grote formele rompslomp meebrengen dat inzicht in de onderhavige kwesties eerder gehinderd dan bevorderd wordt.

Deze vrees wordt versterkt door het besef dat modeltheoretische calculi in het algemeen niet uitmunten op het stuk van 'psychologische plausibiliteit'. Deze calculi gooien geen hoge ogen wanneer het gaat om een realistische re-

<sup>3</sup> Dergelijke gevallen komen meer voor. Neem bv. de Cockney zin:

(i) He is making more money than you'll never make.

In standaard Engels leest men *ever* in plaats van *never*. Ook hier is prake van negatie-copiëring, terug te voeren op een 'verborgen' negatie in de comparatiefconstructie. In z'n semantische analyse geven we deze zin (ruwweg) weer als 'he makes money to an extent [to which not] you will ever make money'. Het element *to which not*, dat we laten beantwoorden aan *than*, bevat de negatie die in *never* is gecopieerd.

constructie van de rekenprocessen die zich afspelen in de hersenen van sprekers en hoorders bij het uiten en verstaan van zinnen. Wat van deze processen bekend is wijst in de richting van nogal andersoortige calculi dan die welke in de logische modeltheorie gangbaar zijn.

Tenslotte is er de overweging dat het omzetten van zinnen in hun semantische analyses, en omgekeerd, met behulp van puur syntactische transformatieregels inmiddels zo succesvol verloopt, en steeds zo goed blijkt te passen op allerlei eigenaardige feiten die op het eerste gezicht geheel onsystematisch lijken, dat we het vertrouwen beginnen te krijgen dat we op de een of andere manier kennelijk greep hebben op het systeem dat feitelijk aan de geobserveerde verschijnselen ten grondslag ligt.

Nu we ons dan zo sterk gemaakt hebben is het wellicht tijd om iets te laten zien en over te gaan tot het eigenlijke werk, het syntactisch rekenen met taal.

### 3. ENKELE BASISBEGRIPPEN

Binnen het kader van de grammatika is een transformatieregel ('T-regel') in de meest algemene zin een functie van zinsstructuren naar zinsstructuren. Als we zinsstructuren het etiket S geven kunnen we dus zeggen dat een T-regel  $T_n$  de vorm heeft van  $S \Rightarrow_{T_n} S'$ , of, zo men verkiest, van  $T_n(S, S')$ .  $T_n$  voert een algorithmische operatie uit op S en levert S' af. Dit is echter een zo algemene en dus weinigzeggende karakterisering van het begrip T-regel, dat, met name sinds eind jaren 60, de behoefte gevoeld is om dit begrip nader in te perken. In de school van Chomsky wordt dit vooral gemotiveerd met het argument dat een algemene grammatikatheorie een reconstructie moet zijn van het algemeen-menselijke ingeboren vermogen om tijdens de eerste levensjaren specifiek een menselijke taal te leren. Dit is in beginsel ongetwijfeld juist, zij het wellicht wat eenzijdig in het veel bredere kader van onderzoek naar mentalen als zeer specifieke communicatiesystemen. Hoe dan ook, het heeft alle zin het juist gegeven algemene regelschema voor T-regels aan nadere restricties te onderwerpen. Dit geeft ons niet alleen een beter inzicht in de soortspecifieke eigenschappen van mentalen. Het verschaft ook een verder methodologisch middel om tussen concurrerende hypothesen te beslissen: die hypothese die zich beter houdt aan een geheel van zo specifiek mogelijke algemene restricties op T-regels krijgt daardoor winstpunten.<sup>4</sup>

Allereerst moet gesteld worden dat alle S-structuren, op elk niveau van analyse, als *boomdiagrammen* zijn weer te geven. De notie 'boomstructuur' is in de jaren 30 door Leonard Bloomfield in de taalkunde geïntroduceerd

<sup>4</sup> De restricties die in de semantische syntaxis nuttig zijn gebleken verschillen in essentiële opzichten van die welke in het door Chomsky sinds 1970 ontwikkelde kader van autonome syntaxis, de theorie van 'Government and Binding' (GB), zijn gaan gelden (Chomsky 1982). Veel theoretische grammatici zien het als een groot bezwaar van het GB-kader dat het bij toepassingen op nieuwe talen telkenmale nieuwe aanpassingen behoeft, met als gevolg dat aan het beoogde universele karakter van de theorie in steeds toenemende mate afbreuk wordt gedaan.

(Bloomfield 1933), en was waarschijnlijk door hem ontleend aan de psycholoog Wilhelm Wundt, die reeds in zijn *Logik* van 1880 zinnen van boomdiagram-analyses voorzag. Een T-regel is dus een omzetting van de ene *boomstructuur* in de andere. Een boomstructuur (als onderscheiden bv. van een matrix-structuur) is in principe een structuur van hiërarchische geordende constituenten onder één ‘hoogste’ overkoepelende constituent, die beantwoordt aan de structuur als geheel. In de syntaxis heeft de hoogste constituent steeds het etiket S (‘sentence’). De hoogste constituent bestaat uit een of meer lagere constituenten. Deze bestaan op hun beurt weer elk uit een of meer lagere constituenten, enzovoort totdat de laagste constituenten zijn bereikt, de elementaire bouwblokken van de gehele structuur.

Het is sinds de jaren 50 algemeen bekend dat boomstructuren direkt genereerd kunnen worden door algorithmes in de vorm van zg. *formatie-* of *expansieregels*. Dit was het soort algorithm dat Beth voornamelijk voor ogen stond in 1963, en dat ten grondslag lag aan het originele concept van generatieve grammatika.<sup>5</sup> De thans gangbare opvatting is dat de formatieregels de structuren definiëren die als input dienen voor het transformationele mechanisme, terwijl het eigenlijke syntactische rekenwerk door de T-regels wordt verricht.<sup>6</sup> Sedert de genoemde publicaties van Katz, Fodor en Postal uit het begin der jaren 60 wordt het bovendien in brede kring aanvaard dat de enige vorm van recursie die in de syntaxis is toegelaten bestaat uit het terugkeren van het hoofdsymbool S, zodat dan de met S verbonden formatieregels opnieuw kunnen gaan werken.<sup>7</sup> Deze algemene restrictie op syntactische recursiviteit heeft uitermate vruchtbaar gewerkt en is de bron geweest van vele nieuwe inzichten.

De gedachte dat taalstructuren *boomstructuren* zijn is sinds Bloomfield algemeen aanvaard.<sup>8</sup> Wat minder algemeen wordt ingezien is dat boomstructuren bij uitstek geschikt zijn voor rekenoperaties. Het zijn, om zo te zeggen,

<sup>5</sup> Dit concept van generatieve grammatika in de vorm van boomstructuur-opbouwende expansie- of formatieregels is voor het eerst te vinden in het in 1951 gepubliceerde maar reeds enkele jaren daarvoor voltooide werk *Methods in Structural Linguistics*, van Bloomfields leerling Harris, met name op de blzn. 365-373. Het is aan dit werk van Harris dat Chomsky de gedachte van een generatieve grammatika heeft ontleend.

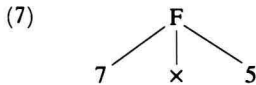
<sup>6</sup> Het is hierbij geenszins uitgesloten dat er ook bepaalde formatieschema's bestaan voor de oppervlaktestructuren van een bepaalde taal. Dit zijn dan geen regels in de zin dat zij de oppervlaktestructuren volledig zouden determineren, maar eerder restricties op bepaalde aspecten van deze structuren. Men spreekt dan gewoonlijk van *structurele filters*.

<sup>7</sup> Het symbool S kan met name terugkeren als argumentterm bij bepaalde predikaten (functoren) en bovendien, maar dan met minstens één variabele en een operator die deze bindt, binnen een argumentterm ter nadere specificatie van zijn descriptieve inhoud.

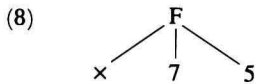
<sup>8</sup> Een enkele keer treft men linguïsten aan die volhouden dat in hun theorie geen boomstructuren voorkomen. Het blijkt dan echter steeds dat zij niet meer bedoelen dan dat ze geen boomdiagramnotatie gebruiken, en vergeten daarbij dat ook in Bloomfield (1933) geen enkel uitgetekend boomdiagram voorkomt, terwijl hij toch de boomstructuur-analyse heeft geïntroduceerd! Mij is geen serieuze modern linguïstisch werk bekend dat niet op het begrip van boomdiagram-analyse, zoals in de tekst gedefinieerd, is gebaseerd. Het doet er daarbij minder toe welke notatie men gebruikt om zulke analyses weer te geven.

ideale rekenramen. Dit blijkt al uit het feit dat alle wiskundige formules als boomstructuren te analyseren zijn, ook al worden ze zelden zo weergegeven. We zullen nu eerste even stilstaan bij een paar zeer eenvoudige rekenkundige formules en de bijbehorende rekenwijzen. Straks zal dan blijken hoezeer ze overeenkomen met taalkundige boomstructuren en de rekenwijzen die daarop worden toegepast.

Een formule als  $7 \times 5$  kan op de volgende wijze als boomdiagram worden weergegeven:

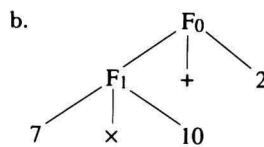
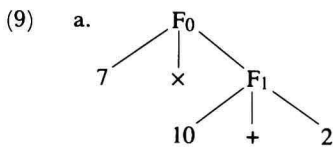


(waar 'F' staat voor 'formule'). Soms schrijft men ook  $\times(7,5)$  met de functor voorop en de argumenten er tussen haakjes achter. Dit heet de *Poolse notatie*. Het boomdiagram is dan:

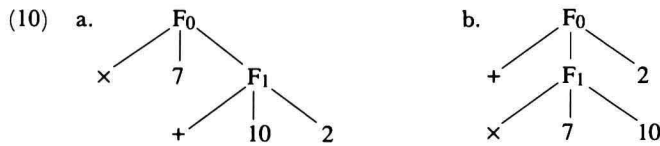


Het verschil tussen (7) en (8) is puur syntactisch: men kan z'n formules zus of zo definiëren, alnaargelang de eigen voorkeur of misschien, in sommige gevallen, op grond van bepaalde praktische voordelen.

Verschillen in hiërarchieën der constituenten worden in boomdiagrammen uitgedrukt door verschillende vertakkingen, maar haakjes kunnen hetzelfde werk doen. Zo drukken we het verschil tussen bv.  $7 \times (10 + 2)$  en  $(7 \times 10) + 2$  uit door de verschillende vertakkingen in respectievelijk (9a) en (9b):



In de Poolse notatie krijgen we respectievelijk (10a) en (10b):



Met deze bomen kan gerekend worden zodra is vastgelegd wat de functoren zijn en wat de argumenten, en welke functies door de functoren worden gedenoteerd. Wanneer we weten dat '×' staat voor de functor 'product' en dat '7' en '5' argumenttermen zijn die getallen denoteren, kunnen we het getallenpaar <7,5> invoeren in de functie 'product'. De waarde van de functie voor deze argumenten, de 'uitkomst' dus, is uiteraard 35, zodat we voor de F in (7) en (8) het getal 35 kunnen schrijven. Op analoge wijze kunnen we uitrekenen dat F<sub>1</sub> in (9a) en (10a) de waarde 12 krijgt, en F<sub>0</sub> de waarde 84, terwijl in (9b) en (10b) F<sub>1</sub> de waarde 70 krijgt en F<sub>0</sub> de waarde 72.

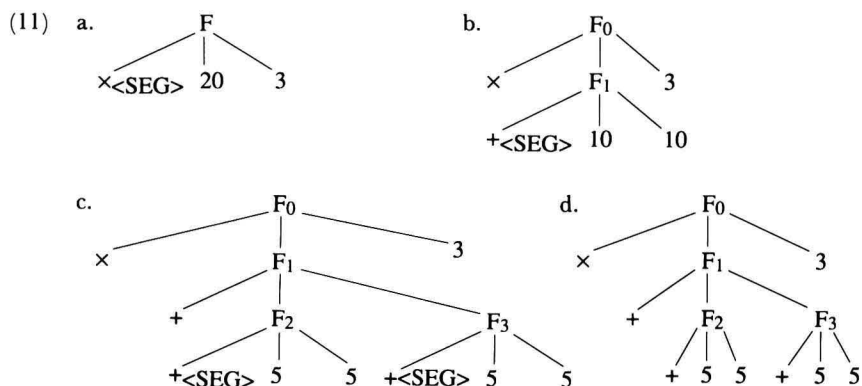
Dit alles lijkt kinderlijk eenvoudig, maar laat ons voorzichtig zijn. Allereerst kan worden opgemerkt dat, in termen van de zojuist gegeven boomdiagramanalyses, de enige vorm van recursie die in de klassieke rekenwijzen aangenomen behoort te worden recursie is van het symbool F, dat behalve als etiket van een gehele formule ook de plaats kan innemen van een argument bij een functor. De overeenkomst met de kort hierboven vermelde exclusieve S-recursie in de syntaxis is frappant, te meer daar de ontdekking van deze restrictie in de syntaxis op generlei wijze door de rekenkunde is geïnspireerd.

Verder laten de zojuist gegeven rekenvoorbeelden (7)-(10) zien hoe men waarden van functies kan berekenen, om zo te zeggen, door de boom heen naar boven toe. Rekenwijzen die zich van onder af door de boom heen naar boven werken totdat ze de uiteindelijke waarde voor de hoogste constituent afleveren worden gewoonlijk *compositional* genoemd. Het principe der compositionaliteit volgt direct uit de boomstructuur en de te volgen rekenwijze. Immers, (9a) heeft bij voorbeeld de vorm van een vermenigvuldiging: de functor × heeft het grootste bereik. Deze werkt echter alleen op getallenparen, en de boom geeft maar één getal; op de andere plaats staat het symbool F<sub>1</sub> en geen getal. Pas wanneer eerst het getal is uitgerekend dat aan F<sub>1</sub> beantwoordt, nl. 12, kan de vermenigvuldiging worden uitgevoerd.

In de boven gegeven voorbeelden is de waarde van elke berekening weer een getal, en kan dus weer als input dienen voor een volgende, 'hogere', berekening, totdat de uiteindelijke waarde wordt gevonden, die wederom een getal is. Deze wijze van rekenen is een directe parallel van wat in de formele semantiek gebeurt met linguïstische boomstructuren. Elke zich vertakkende constituent van een linguïstische boomstructuur heeft steeds één functor onder zich en verder één of meer argument-constituenten. De laatste functor levert een waarheidswaarde af voor elke onafhankelijk gedefinieerde stand

van zaken, dus, met andere woorden, een verzameling van mogelijke standen van zaken. In de formele semantiek worden dus op compositionele wijze verzamelingen van mogelijke standen van zaken berekend voor S-structuren. Rekenwijzen waarbij de uiteindelijke waarde zich bevindt in het veld der dingen waarop de in de structuren voorkomende constituenten betrekking hebben, kan men *semantisch* noemen.<sup>9</sup> De geschetste rekenwijzen voor de structuren (7)-(10) zijn dus niet alleen compositioneel maar ook semantisch.

Men kan zich echter ook een andere toepassing van hetzelfde rekenbegrip in-deken, een toepassing die met meer recht *syntactisch* genoemd mag worden. Een enkel voorbeeld van een *syntactische transformatie in de rekenkunde* moge volstaan. Men denke zich in dat het, om de een of andere reden die er nu niet toe doet, nodig is om bij rekenkundige notaties even getallen uit te sluiten: alleen oneven getallen mogen of kunnen geschreven worden. Even getallen moeten eerst gesplitst worden in hun twee gelijke helften, net zo lang tot elke helft oneven is. Voor zo'n notatie moet elke functor worden voorzien van een instructie die hem gebiedt z'n argumenttermen, indien ze even getallen aanduiden, door 2 te delen en de beide helften op te tellen. Laten we deze instructie 'SEG' noemen ('Splits Even Getallen'), en haar tussen punthaakjes als subscript aan de functoren toevoegen zolang zij niet is uitgevoerd. Een vermenigvuldiging van 20 met 3 komt er dan, in de Poolse notatie, uit te zien als (11a). Uitvoering van de instructie SEG levert (11b) op, dat echter zelf weer een functor met de instructie SEG bevat, waarvan uitvoering (11c) oplevert. Uitvoering van de laatste twee SEG-instructies levert niets nieuws op daar alle argumenttermen nu oneven getallen aanduiden. (11d) is dus het uiteindelijke resultaat.



<sup>9</sup> Gedurende de laatste tien jaar is het in toenemende mate geaccepteerd rekenwijzen eveneens semantisch te noemen wanneer zij waarden opleveren die een tussenfunctie vervullen tussen de syntactische structuren van de zinnen enerzijds en de dingen waarop de syntactische structuren en hun constituenten betrekking hebben, hun extensies. Men is zich steeds meer gaan realiseren dat het voor een adequate semantiek noodzakelijk is een dergelijk tussenstadium aan te nemen, dat in cognitieve zin geïnterpreteerd wordt: de relatie van zinnen en hun elementen tot datgene waarover ze gaan wordt immers steeds bemiddeld door de cognitie. Vormen van semantiek die een 'dergelijk cognitief' tussenstadium in de calculus opnemen worden gewoonlijk aangeduid met de term 'discourse semantiek'. Zie bv. Kamp (1981); Seuren (1985.)



Men merke hierbij op dat de transformationele processen in kwestie in beide richtingen automatisch uitvoerbaar zijn, niet alleen met (11a) als input en (11d) als output maar ook omgekeerd.

Iets dergelijks is er nu in de taal ook aan de hand, alleen gaat de taal nog iets verder. Terwijl in het voorbeeld (11a-d) alle vier de structuren steeds dezelfde waarden opleveren wanneer de betrokken functies semantisch worden uitgerekend, nl. de waarde 60, is het in de taal zo dat er veelal syntactische transformaties optreden die geen semantische rekenarij meer toelaten omdat dan hetzij de uitkomst verkeerd wordt hetzij de rekenwijze niet meer toepasselijk is vanwege het optreden van nieuwe symbolen waarvoor geen rekenmatige eigenschappen zijn vasgelegd. Anders gezegd; terwijl wiskundige formules hun betekenis steeds eerlijk op het gezicht dragen zien zinnen van natuurlijke talen er vaak wat hun betekenis betreft misleidend uit, ook al weten competente sprekers uiteraard precies wat de zinnen in kwestie betekenen. In het dagelijks taalgebruik komt het vaak voor dat men zich verspreekt: men zegt iets anders dan wat men bedoelt, en meestal wordt dit door de begrijpende hoorder onderkend. Waar het hier om gaat is dat in natuurlijke talen zulke 'versprekingen' geïnstitutionaliseerd voorkomen en deel uitmaken van de legitieme grammatika. In de taal is het daarom essentieel dat de oppervlaktestructuur eerst wordt gereduceerd tot de semantische analyse alvorens er met semantisch rekenen wordt begonnen.

Vanwege de aan het begin vermelde behoefte aan grammatika's die niet alleen zinnen genereren maar tevens een ontleedprocedure voor zinnen meebrengen is het een eerste eis bij syntactische T-regels dat ze in beide richtingen uitvoerbaar zijn, van S naar S' en omgekeerd. Gewapend met de kennis dat bepaalde functoren bepaalde T-regels induceren moet het effect van elke T-regel steeds 'herstelbaar' zijn, d.w.z. de input moet aan de hand van de output gereconstrueerd kunnen worden. We spreken hier van de *eis der herstelbaarheid* van T-regels.

Dan de Poolse notatie, die boven enkele malen is gedemonstreerd. Deze is in de syntaxis der natuurlijke taal meer dan een willekeurige notationale variant. Sinds het befaamde in 1970 in het tijdschrift *Language* verschenen artikel van de Amerikaanse linguïst James McCawley, 'English as a VSO-language', is het duidelijk dat, althans binnen de semantische syntaxis, de T-regels aanzienlijk vereenvoudigd worden en zich veel beter aan algemene restricties laten onderwerpen wanneer wordt aangenomen dat voor vele talen de grondstructuur van elke S in de semantische analyses van zinnen bestaat uit het predikaat (de functor), gevolgd door de erbij behorende argumenttermen. Voor andere talen is het meer geëigend aan te nemen dat in de semantische analyses de functor volgt op de argumenttermen. Aangezien het predikaat tot de categorie V (Verb) behoort, en de standaard argumenttermen subject (S) en object (O) zijn, spreekt men van respectievelijk een VSO-volgorde of een SOV-volgorde der zinsconstituenten. In veel talen vindt men de semantische VSO- dan wel SOV-volgorde ook in de oppervlaktestructuren terug. Maar heel vaak, zoals bv. in het Engels en de

Romaanse talen, wordt ook de volgorde SVO als standaardvorm aangetroffen in de oppervlaktestructuren. Het blijkt nu in termen van de semantische syntaxis een uiterst eenvoudige operatie te zijn om onderliggende VSO-structuren om te vormen tot SVO-structuren. (Voor onderliggende SOV-structuren ligt dat iets, maar niet veel, gecompliceerder). Straks zal gedemonstreerd worden hoe die overgang in z'n werk gaat zonder dat ook maar enige extra assumptie behoeft te worden ingevoerd *pour le besoin de la cause*. Inmiddels is de gelijkenis met rekenkundige formules in de Poolse notatie frappant.

Een verdere overeenkomst wordt gevonden in het zeer beperkte arsenaal aan categorieën waaruit enerzijds de rekenkundige formules en anderzijds de semantische analyses worden opgebouwd. Bij de formules treffen we als categorieën aan F ('formule'), en verder *functor* ('predikaatsterm') voor de functies van optellen, aftrekken, vermenigvuldigen en delen, en *getallen naam* voor die argumenttermen die direct als getallen naam gegeven zijn. Afgezien van de interne structuur van de getallen namen zijn er dus slechts drie categorieën. Wonderlijk genoeg is de situatie in de semantische syntaxis geheel analoog, ook al is er van beïnvloeding vanuit de rekenkunde nooit enige sprake geweest. Sinds het begin der jaren 70 is het, als gevolg van baanbrekend werk van de al eerder genoemde James McCawley<sup>10</sup>, een in de semantische syntaxis aanvaarde stelling dat semantische analyses, afgezien van de interne structuur der argumenttermen, slechts drie categorieën kennen, nl. S (voor zinsstructuren), V (voor predikaten), en NP (voor normale argumenttermen). Alle andere categorieën die in oppervlaktestructuren onderscheiden worden, zoals adjectieven, bijwoorden, voorzetsels, komen tot stand als gevolg van transformationele operaties. Het is belangrijk zich te realiseren dat in de semantische syntaxis, nog steeds in het voetspoor van McCawley, die elementen in zinsstructuren die beantwoorden aan logische operatoren, zoals de negatie, de kwantoren, de modaliteiten, de werkwoordstijden, in semantische analyses alle worden geanalyseerd als predikaten. Zowel in rekenmatig als in logisch opzicht is dit geheel correct. Immers, ook in de logika zijn operatoren steeds functoren die voor hun argumenten een waarde, om zo te zeggen, naar boven afleveren.

Inmiddels begint een beeld op te doemen van syntactische boomstructuren die niet alleen in hoge mate beantwoorden aan die welke gangbaar zijn in de predikaten calculus maar die bovendien rekenwijzen toelaten die een directe verwantschap vertonen met die welke in de gewone rekenkunde al lang bekend zijn. In dit beeld past geheel het sinds midden zestiger jaren bekende en aanvaarde principe der *zg. cyclische toepassing der T-regels*. Op zuiver taalkundige gronden werden theoretische linguïsten in die tijd gewaar van het feit dat een groot deel der in de grammatika benodigde T-regels *cyclisch* geordend zijn. Dit houdt in dat zij eerst toegepast worden op de diepst ingebedde S-constituent, vervolgens op die daarboven, en zo door totdat de

<sup>10</sup> Zie met name zijn opstel 'A program for logic', gepubliceerd in McCawley (1973:285-319), en in Davidson & Harman (1972:498-544).

hoogste S-knoop is afgewerkt. Op dat moment is een structureel representatieniveau bereikt dat met de benaming van *ondiepe structuur* (shallow structure) wordt aangeduid. Tussen de ondiepe structuur en de oppervlaktestructuur liggen de *postcyclische T-regels*, die hoofdzakelijk neerkomen op verplichte of optionele volgordeverschuivingen en/of herstructureringen.

Het cyclische principe is in feite niets anders dan het principe der compositionaliteit. Dit kan op de volgende manier worden toegelicht. Elke S-structuur in de semantische analyse bestaat uit een functor (predikaat) met het categorie-etiket V, en de bijbehorende argumenttermen, die of NP zijn of wederom S. Analooq aan (11) zijn met functoren soms syntactische instructies (T-regels) verbonden, met name wanneer een S als één van de termen optreedt. Of er kan een taalspecifieke instructie zijn die in werking treedt wanneer zich een bepaalde S-configuratie voordoet (bv. de invoering van een dummy-subject als *het* of *er*, onder bepaalde condities, in het Nederlands). We spreken respectievelijk van lexicaal en structureel geïnduceerde regels. Deze regels hebben betrekking op de S in kwestie, dus de V en z'n termen, en resulteren in een structuurverandering in die S. Omdat die verandering pas kan worden aangebracht wanneer de betrokken V en z'n termen structureel welbepaald zijn, moeten alle regels die nog in eventueel ingebedde S-termen voorkomen eerst worden uitgevoerd. Zo worden automatisch eerst de regels die werken op de diepst ingebedde S-structuren toegepast, dan die van het eersthogere niveau van inbedding, en zo door tot het hoogste niveau is bereikt. Dit soort syntactische T-regels ordent zich daarom automatisch cyclisch.

#### 4. DE VOORNAAMSTE UNIVERSELE RESTRICTIES OP CYCLISCHE LEXICAAL GEÏNDUCEERDE T-REGELS

We zullen ons hier beperken tot cyclische lexicaal geïnduceerde T-regels. Van dergelijke T-regels bestaat er slechts een handvol in het universele syntactische arsenaal. Voor elk predikaat (niet per se een werkwoord in de oppervlaktestructuur) wordt in het lexicon gespecificeerd welke T-regels het, verplicht of optioneel, induceert. We spreken van het *inducerende predikaat*. Deze T-regels houden zich alle aan bepaalde algemene restricties. We zullen deze nu vermelden, althans voorzover we vermoeden dat ze inderdaad als universele restricties geldigheid hebben. Zij moeten onvermijdelijk eerst in betrekkelijk abstracte termen worden gespecificeerd. Concrete toepassingen op feitelijk taalmateriaal volgen in § 5.

Vooraf moeten twee eenvoudige universele zg. *herstructureringsprincipes* besproken worden. Het eerst houdt in, dat elke S die door een of andere cyclische regel beroofd wordt van zijn subject wordt gedegradeerd tot een onvolledige S, genoteerd als '/S'. Een /S beantwoordt precies aan wat in de traditionele structuralistische en generatieve grammatika 'VP' (verb phrase) genoemd wordt. Het tweede herstructureringsprincipe zegt dat elke S die door een cyclische regel wordt beroofd van zijn V zelf verdwijnt: de S-knoop in

kwestie wordt verwijderd, en alle materiaal dat door de verdwenen S werd gedomineerd wordt nu aangehecht aan de hogere S, in de gegeven volgorde.

Wat de regels zelf betreft valt allereerst op dat er in lexicaal geïnduceerde cyclische T-regels vier soorten operaties worden aangetroffen: (a) deletie, en verder (b) stijgende, (c) dalende, en (d) horizontale verplaatsing van constituenten. Stijging wordt gewoonlijk *Raising* genoemd, en voor daling is de term *Lowering* in zwang.

Er is maar één vorm van cyclische lexicaal geïnduceerde *deletie*. Deze is altijd beperkt tot het *subject* van een niet-finiete ingebedde S (over dit begrip meer in § 5.2). We spreken van de regel van *Subject-Deletie* (SD). Wanneer een subject wordt gedeleerd is het lexicaal leeg: het vervult slechts een nominale functie. In standaard gevallen wordt het *gecontroleerd* door een andere NP. De controlerrelatie houdt in een (constante of variabele) referentiële identiteit van de controlerende en de gedeleerde NP. De controlerende NP is meestal het subject bij het inducerende predikaat (verticale SD) of het subject van een subject-S ingebed onder het inducerende predikaat (horizontale SD). In sommige gevallen is de controlerende NP (in)direct object bij het inducerende predikaat.

Voorbeelden van subject-gecontroleerde verticale SD zijn er te over. Het zijn in het algemeen die gevallen waarbij een ingebed infinitiefcomplement een 'verzwegen' subject heeft dat referentieel overeenstemt met het hogere subject, zoals bij voorbeeld in *Karel hoopt volgend jaar zijn examen te halen*. Het verzwegen subject bij *halen* is referentieel gelijk aan het hogere subject *Karel*. Objectgecontroleerde verticale SD vindt men in gevallen als *Karel haalde Mien over om mee te gaan*, waar het verzwegen subject bij *meegaan* referentieel gelijk is aan het hogere object *Mien*. Horizontale SD wordt aangetroffen in zinnen als *Door te sporten behield Sofie haar conditie*. Hier is het verzwegen subject bij *sporten* referentieel gelijk aan het subject van de hele zin, *Sofie*. De analyse van zulke zinnen komt onder ter sprake.

Het is bovendien waarschijnlijk noodzakelijk aan te nemen dat er ook niet-standaard gevallen van SD zijn die niet door een andere NP gecontroleerd worden. Dit met het oog op die gevallen waar sprake is van een niet-specifiek gedeleerd subject, zoals bij voorbeeld in *Zo'n aanbod te weigeren is onmogelijk*. Hier is het subject van *weigeren* gedeleerd zonder dat er van een controlerende NP sprake is.

Afgezien van ongecontroleerde SD, ziet het regelschema van SD er dus uit als in figuur 1, waar 'NP<sub>x</sub>' de controlerende NP weergeeft en NP[x] de gecontroleerde NP. De te deleren NP is vetgedrukt.

Wat de bepalende factor is bij de vraag of de controlerende NP subject of object is bij het inducerende predikaat is nog in hoge mate duister. Ook is nog betrekkelijk weinig bekend over ongecontroleerde SD. Ware dat niet zo, dan zou het wellicht mogelijk zijn SD automatisch te laten voortvloeien uit de structurele en misschien andere eigenschappen van de input-bomen. SD zou dan niet langer een lexicaal geïnduceerde regel zijn, maar een, nog steeds cyclische, aan de boom afleesbare regel. Aangezien dit een verminde-

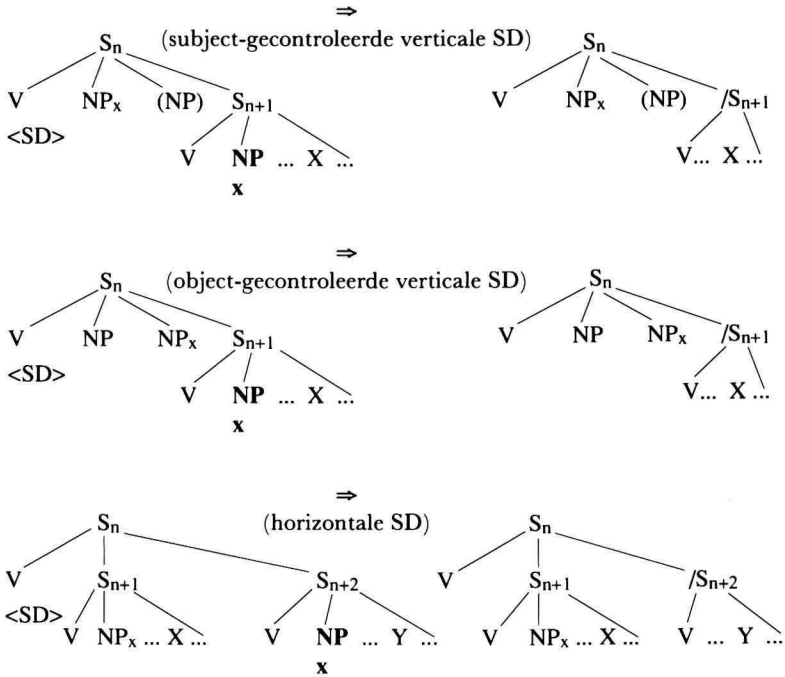


Fig. 1. Regelschema's voor Subject-Deletie (SD)

ring zou betekenen van het aantal idiosyncratische assumpties in de theorie, zou een dergelijke stap een theoretische vooruitgang betekenen. Momenteel zijn we echter nog niet zo ver.

Lexicaal geïnduceerde cyclische *stijging* (Raising) kan alleen betrekking hebben op de V of de subject-NP aan een onder het inducerende predikaat ingebedde S. Stijging van V wordt *Predicate Raising* (PR) genoemd.<sup>11</sup> Bij PR wordt de lagere V aan de hogere V gehecht, en wel via het proces van *adoptie*. Adoptie van een deelboom onder het etiket B door een deelboom onder het etiket A houdt in dat de adopterende knoop A een copie van zichzelf maakt en deze vervolgens onmiddellijk boven zich plaatst. De copie neemt dan de deelboom onder B aan als zusterconstituent van de oorspronkelijke deelboom A, en wel hetzij rechts hetzij links van A. Adoptie wordt gedemon-

<sup>11</sup> De regel Predicate Raising werd voor het eerst voorgesteld in McCawley (1968) – ook in McCawley (1973:155-166), maar alleen voor interne lexicale analyses, niet voor de ‘open’ syntaxis. In Seuren (1972) wordt getoond dat PR in veel talen, waaronder met name het Frans, het Nederlands en het Duits, veelvuldig optreedt in de ‘open’ syntaxis. Evers (1975) bevat de eerste, en tot op heden niet essentieel verbeterde, analyse en beschrijving van de betrokken verschijnselen in het Nederlands en het Duits in termen van PR (zijn benaming voor PR is ‘Verb Raising’).

streerd in de volgende figuur, waat 'LA' staat voor linkse adoptie en 'RA' voor rechtse adoptie. De te adopteren deelboom is vetgedrukt:

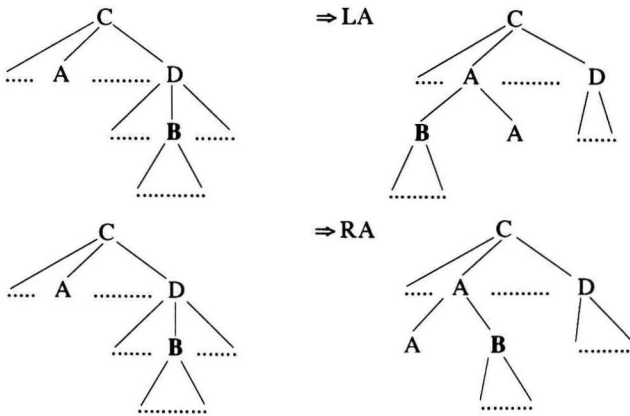


Fig. 2. Adoptie van deelboom B door A

Het regelschema van PR ziet er derhalve uit als in figuur 3, waar de omhoog te halen V vetgedrukt staat.

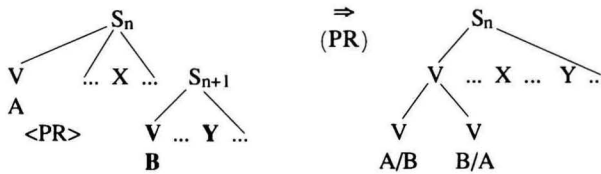


Fig. 3. Regelschema voor Predicate Raising (PR), met rechtse en linkse adoptie

Hierbij moet overigens worden aangetekend dat het waarschijnlijk nuttig is aan te nemen dat een V-deelboom die door PR omhoog wordt verplaatst ook in bepaalde gevallen van categorie (woordsoort) kan veranderen. Op deze wijze kunnen Engelse constructies als *He took down the ladder* met behulp van PR behandeld worden. Alleen verandert de lagere V in een partikel na PR. Vanuit de onderliggende semantische analyse  $s[v[\text{take}_{\langle \text{SD}, \text{PR} \rangle}]_{\text{NP}}[\text{he}]_{\text{NP}}[\text{the ladder}]_s[v[\text{down}]_{\text{NP}}[\text{x}]]]$  levert toepassing van SD, gevolgd door PR, op:  $s[v[v[\text{take}]_{\text{partikel}}[\text{down}]]_{\text{NP}}[\text{he}]_{\text{NP}}[\text{the ladder}]]_{\text{partikel}}[\text{down}]$ . Wanneer de optionele PR niet wordt toegepast krijgen we  $s[v[\text{take}]_{\text{NP}}[\text{he}]_{\text{NP}}[\text{the ladder}]]_{\text{partikel}}[\text{down}]$ , hetgeen beantwoordt aan de eveneens correcte zin *He took the ladder down*.

Raising van de subject-NP van de ingebedde S heet *Subject Raising* (SR). SR is een regel die in het Engels zeer veel voorkomt. Men vindt hem in zinnen als *I expect Jack to win* (SR uit object-S) of *Jack is likely to win* (SR uit subject-S).<sup>12</sup> Het procédé is wederom zeer eenvoudig. De subject-NP van de

ingebede (subject- of object-) S komt te staan op de plaats van de eigen S, die dus tot /S wordt gedegradeerd. Deze /S schuift één positie naar rechts. Het regelschema is gegeven in figuur 4, waar de omhooggehaalde NP vetgedrukt staat.

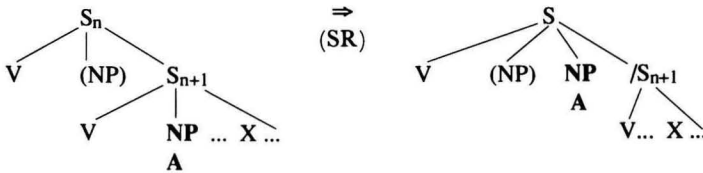


Fig. 4. Regelschema voor Subject Raising, uit subject- of object-S

*Lowering* (L) is alleen van toepassing op het inducerende predikaat zelf. Dit wordt, wanneer het L induceert, ergens in de ingebede  $S_{n+1}$  of  $/S_{n+1}$  geplaatst. Wanneer de hogere  $S_n$ , de S dus van het inducerende predikaat, geen ander materiaal domineert dan alleen de ingebede  $S_{n+1}$ , wordt hij 'gesnoeid', ingevolge het algemene voorschrift dat 'ijdele' knopen (d.w.z. knopen die alleen een identieke knoop domineren) automatisch uit elke boom worden verwijderd.

Lowering is typisch voor 'abstracte' predikaten, d.w.z. predikaten die in de oppervlaktestructuur beantwoorden aan een bereikdragende logische operator, zoals de negatie, de kwantoren, modale operatoren, werkwoords-tijden. De vertegenwoordigers van logische operatoren zijn in oppervlakte-structuren zelden werkwoorden. Ze behoren meestal tot andere soms idiosyncratische woordsoorten. Het woord *veel* bij voorbeeld, dat een grote hoeveelheid aanduidende existientiekwantor vertegenwoordigt, is hybridisch: het vertoont eigenschappen van substantieven naast die van adjectieven.<sup>13</sup> Voor elke V die L induceert moet in het lexicon gespecificeerd staan welke oppervlaktewoordsoort hij aanneemt.

De plaats waarop een aan L onderhevige V terechtkomt, de 'landings-plaats', loopt nogal uiteen per operator en per taal. Alleen kwantoren en andere variabelen bindende operatoren hebben steeds een welbepaalde landingsplaats: zij komen per se terecht op de plaats van de door hen gebonden variabele. Verder is er de zg. *Scope Ordering Constraint* (SOC), die inhoudt dat in beginsel (d.w.z. afgezien van enkele dispenserende clausules) de oppervlaktevolgorde het logisch bereik (de hiërarchische status) van bereikdragen-de operatoren uitdrukt in termen van volgorde: de hogere operator blijft steeds links van de lagere.

Afgezien van SOC loopt de landingsplaats van die operatoren die geen variabelen binden nogal uiteen. De negatie, bij voorbeeld, gedraagt zich in

<sup>12</sup> Het Engels lijkt in dit opzicht sterk op het Latijn, dat zowel de constructie *accusativus cum infinitivo* dus SR vanuit een object-S, kent als de constructie *nominativus cum infinitivo*, oftewel SR vanuit een subject-S.

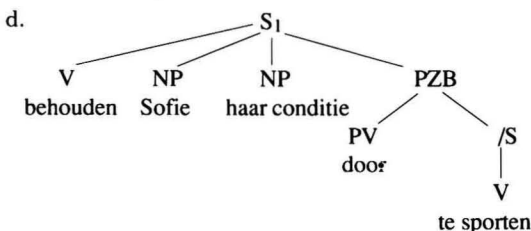
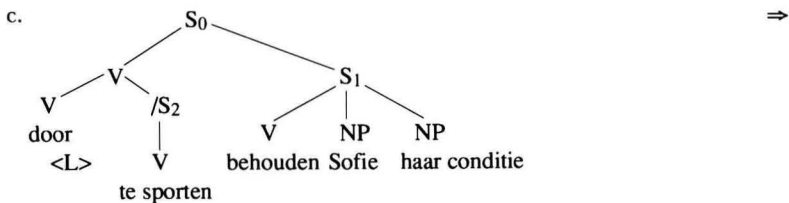
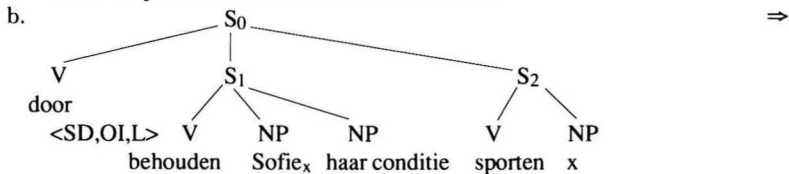
<sup>13</sup> Het heeft substantivische eigenschappen in bv. *Veel moet nog gedaan worden*. De adjectivische trekken blijken uit het feit dat *veel* trappen van vergelijking kent (*meer, meest*), dat het voorafgegaan kan worden door *erg* of *zeer*, en ook bv. door de slot-*e* in gevallen als *het vele eten*.

dit opzicht nogal verschillend in het Nederlands en het Engels. In het Nederlands zoekt de negatie steeds de meest rechtse toegestane positie in de lagere  $S_{n+1}$ . De regel hiervoor is dat *niet* helemaal rechts in  $S_{n+1}$  terecht komt, maar nooit over een daarin ingebedde S of /S heen, en ook steeds vóór een eventuele prepositiebepaling. SOC verbiedt voorts dat de negatie voorbij een eerder in  $S_{n+1}$  neergelaten element heen terecht komt: hij blijft dan vlak voor dat element steken. In het Engels daarentegen zoekt de negatie zo veel mogelijk de verbale cluster op in  $S_{n+1}$ . Daar nestelt *not* zich op een heel speciale wijze, waarover onder meer, binnen het verbale complex.

Tenslotte is er de *horizontale verplaatsing* als regelprocedé in de cyclus. Voorzover tot op heden is gebleken is er slechts één vorm van cyclische lexicaal geïnduceerde horizontale verplaatsing, de regel *Object Incorporatie* (OI). Deze bewerkstelligt dat de object-constituent van een S, de meest rechtse dus, via adoptie wordt ingelijfd bij het predikaat (V). Zo krijgt men complexe predikaten als *bier drinken*, *z'n geld verliezen*, *fietsen repareren* of in het Engels *take care of*, *keep tabs on*, etc. Dat hier inderdaad sprake is van een syntactische regel blijkt o.a. uit het bekende feit dat er passieve Engelse zinnen zijn van het type *She was taken good care of by her uncle*, of *I was never taken any notice of*. Gezien de regels voor het passief in het Engels kunnen zulke zinnen alleen onder de genoemde assumptie verklaard worden.

OI treedt op bij de beregeling van prepositionele zinsbepalingen (PZB), d.w.z. een prepositioneel voegwoord (PV), zoals *om* of *door*, gevolgd door een infinitief. We beregelen zin (12a) (afgezien van de werkwoordstijd en de postcyclische vooropplaatsing van de PZB) met behulp van SD, OI en L. Gegeven de semantische analyse (12b), wordt na SD en OI (12c) verkregen. Na L ontstaat de structuur (12d):

(12) a. Door te sporten behield Sofie haar conditie.





Bij het transformeren van semantische analyses naar zinnen en omgekeerd zijn voor de behandeling van de abstracte logische operatoren geen regels nodig die niet ook al hun bruikbaarheid bewezen hebben in de meer direct zichtbare syntaxis. Regels die al in de syntaxis waren opgesteld nog voordat er sprake was van een convergentie van logische en syntactische structuren, blijken keer op keer, mits in het juiste formaat gegoten, direct en ongewijzigd toepasbaar op de behandeling van logische operatoren. De ‘verlogisering’ van de syntaxis heeft dus geen meerkosten meegebracht in termen van extra regels. L en OI, bij voorbeeld, zijn nodig bij de behandeling van logische operatoren. Maar ze hebben hun nut al bewezen bij het beschrijven van positionele zinsbepalingen, zoals in (12).

Verder dient opgemerkt te worden dat het globale effect van de cyclus bestaat in (a) een veel vlakkere boomstructuur, met (b) veel meer verschillende categorieën. De functionaliteit hiervan is duidelijk wanneer men bedenkt dat semantische bomen als rekenramen moeten fungeren, terwijl zinnen geschikt moeten zijn voor snelle akoestische transmissie. Bij rekenramen zijn we gediend met veel en eenvormige inbeddingen. Bij akoestische transmissie is de lineaire volgorde uiteraard van primair belang.

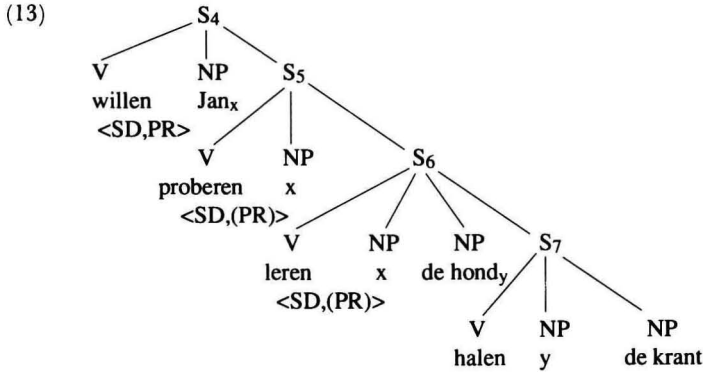
Een laatste punt in dit verband is het volgende. Het is steeds opnieuw verbazend hoe een klein verschil in lexicale regeltoekenning kan leiden tot grote syntactische verschillen tussen talen. Minieme wijzigingen in een regel of regelschema hebben vaak grote oppervlakteverschillen tot gevolg. We zullen daar nog enkele voorbeelden van tegenkomen. Dit is zeker een goed teken. Immers, de verschillen tussen talen onderling worden nu gereduceerd tot kleine variaties in een hoogst uniform systeem.

## 5. ENKELE VOORBEELDEN VAN REGELTOEPASSINGEN

### 5.1. De regels *SD*, *PR* en *SR*

Beschouwen we, bij wijze van voorbeeld, een semantische analyse van een Nederlandse bijzin met enkele ingebedde S-knopen, als in (13). We geven de hoogste S, betreffelijk willekeurig, het subscript 4. De T-regels zijn bij elk predikaat aangegeven tussen punthaken. Ronde haken duiden aan dat de betrokken regel optioneel is. We gaan hierbij uit van de VSO-volgorde, die dus beantwoordt aan de Poolse notatie.<sup>14</sup>

<sup>14</sup> In Koster (1975) wordt betoogd dat voor het Nederlands en het Duits een onderliggende SOV-volgorde aangenomen moet worden. Zijn argument is niet dwingend, mede omdat het slechts gericht is tegen de assumptie van een onderliggende SVO-volgorde, terwijl de VSO-hypothese onbesproken blijft. Koster's these is anderzijds ook niet onhoudbaar: geplaatst tegenover de VSO-hypothese eindigt Koster's SOV-hypothese met ongeveer gelijke punten. Er zijn voors en tegens voor beide standpunten, en een beslissend argument is tot op heden nog niet gevonden. Hoe dan ook, voor de hier geschetste beregeling is dit verschil irrelevant, daar immers in Koster's SOV-volgorde dezelfde regels steeds analoog toepasbaar zijn.

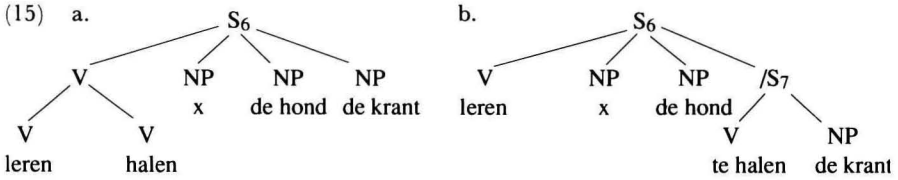


Deze structuur beantwoordt aan de volgende varianten, aannemende dat de bijzin in de verleden tijd staat:

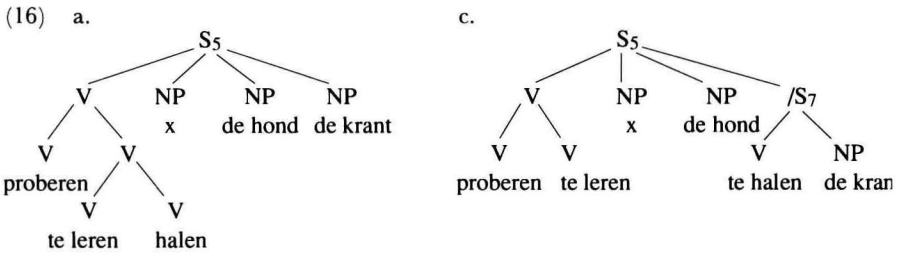
- (14) a. ...dat Jan de hond de krant wilde proberen te leren halen  
 b. ...dat Jan de hond wilde proberen te leren de krant te halen  
 c. ...dat Jan wilde proberen de hond de krant te leren halen  
 d. ...dat Jan wilde proberen de hond te leren de krant te halen

Alle drie der S-inbeddende predikaten in (13), *leren*, *proberen* en *willen* induceren SD. Bij *willen* en *proberen* is de controlerende NP het subject van deze werkwoorden. Bij *leren*, in de hier gebruikte betekenis van 'onderwijzen', is de controlerende NP het (indirect) object. De tweede T-regel die in (13) voorkomt is PR. Door PR wordt de V-knoop van de ingebedde S uit die S losgemaakt en via rechtse adoptie opnieuw aangehecht aan de V die de regel induceert. Er gelden nog de lexicale idiosyncrasieën dat het werkwoord *proberen* altijd het lagere werkwoord voorziet van het verder betekenisloze partikel *te*, terwijl bij *leren* dit *te* alleen optreedt wanneer PR niet is toegepast.

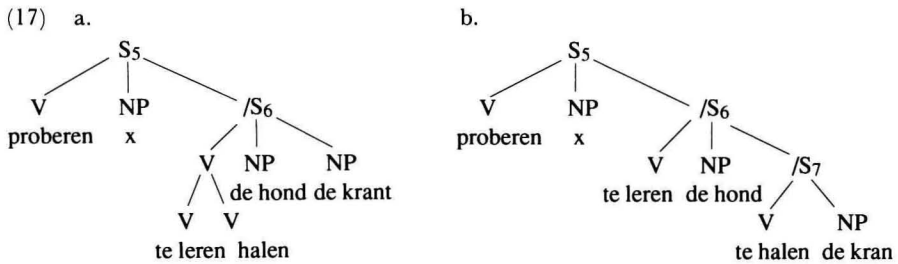
Wanneer we nu de instructies verbonden met het hoogste predikaat in (13), de V-constituent *willen*, dus de regels SD en PR, willen uitvoeren, stuiten we op de moeilijkheid dat de S<sub>5</sub>, waarvan het subject moet worden gedeleerd en de V moet worden geadopteerd, de vorm waarop deze instructies betrekking hebben nog niet heeft. Immers, in S<sub>5</sub> moeten nog de instructies SD en, optioneel, PR worden uitgevoerd op S<sub>6</sub>. En in S<sub>6</sub> geldt opnieuw dat SD en, optioneel, PR moeten worden uitgevoerd. Dit laatste is nu echter wel mogelijk, omdat S<sub>7</sub> inderdaad de vorm heeft waarin hij in een instructie kan worden ingevoerd. Het subject NP[y] wordt gedeleerd, met als gevolg dat S<sub>7</sub> verandert in /S<sub>7</sub>. Vervolgens wordt, als tenminste voor PR wordt gekozen, v[halen] rechts geadopteerd door v[leren], met als gevolg dat de knoop met het etiket /S<sub>7</sub> verdwijnt. Het overgebleven object NP[de krant] wordt aangehecht rechts van NP[de hond] onder S<sub>6</sub>. Het resultaat is als in (15a) bij toepassing van PR, en als in (15b) wanneer PR niet wordt toegepast:



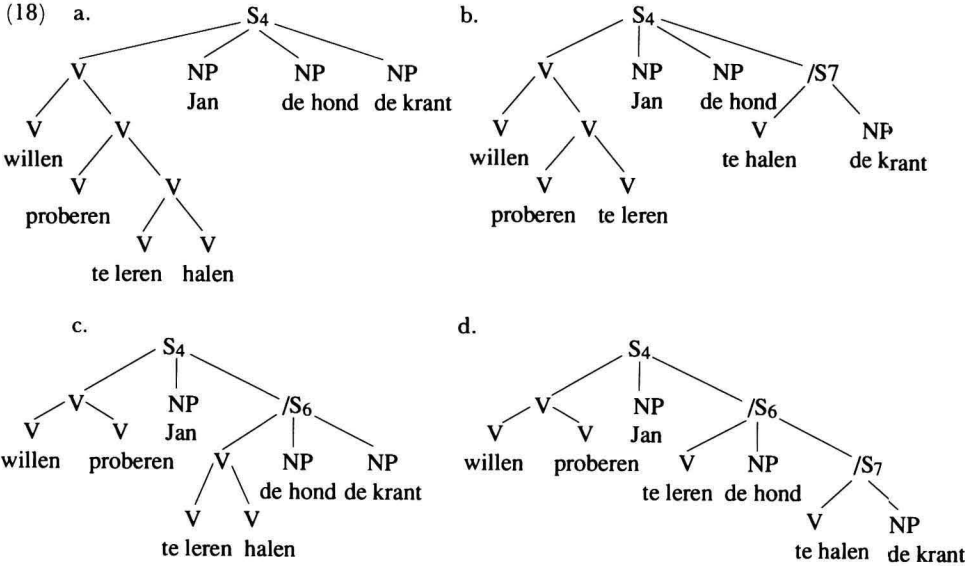
Nu heeft  $S_6$  twee mogelijke vormen, die beide kunnen worden ingevoerd in de T-regels die voor de V van  $S_5$  zijn gespecificeerd. Door SD verdwijnt de  $NP[x]$  van  $S_6$  in (14a) en in (14b), en  $S_6$  wordt  $/S_6$ . Bij toepassing van de hier opnieuw optionele regel PR op (15a) ontstaan respectievelijk (16a) en (16b):



Niet-toepassing van PR op de  $S_5$ -cyclus geeft voor (15a) en (15b) respectievelijk (17a) en (17b):



Op de  $S_4$ -cyclus moeten zowel SD als PR worden toegepast. Het resultaat bestaat derhalve uit vier verschillende structuren, (18a-d), die zullen gaan beantwoorden aan respectievelijk (14a-d):



Afgezien van de toekenning van de werkwoordstijd, waarover straks meer, is er nu nog de waarschijnlijk postcyclische regel V-Final, waardoor in bijzinnen de gehele V-cluster, en in hoofdzinnen alleen het niet-finiëte deel (zie § 5.3), naar rechts wordt verplaatst. Dat wil zeggen, de V-cluster in elke S of /S gaat naar uiterst rechts, maar nooit over een ingebedde S of /S heen. Het resultaat is nu gemakkelijk af te lezen: (14a-d) beantwoorden precies aan (18a-d). bovendien ziet men dat een ongrammatikale volgorde, als in:

(19) ... dat Jan de hond wilde proberen de krant te leren halen

niet gegenereerd kan worden.<sup>15</sup> Tot dusver is deze analyse dus betrekkelijk succesvol.

Een frappant aspect van deze analyse is dat ze praktisch ongewijzigd van toepassing is op het Duits. De groep Duitse werkwoorden die PR, verplicht of optioneel, induceren komt grotendeels overeen met de corresponderende

<sup>15</sup> Verschillende linguïsten, met name Den Besten en Rutten (1989), hebben erop gewezen dat (bij)zinnen als:

(i) ...dat hij die man heeft geweigerd te groeten

niet zonder meer als ongrammatikaal kunnen worden afgedaan. Deze observatie is zeker in principe terecht. Ik heb zelf de volgende twee gevallen geobserveerd:

(ii) Ik doe niet anders dan die jongen proberen z'n gevoelens te sparen. (Rinus Michels, TV)

(iii) Om die reden heeft men het probleem verzaakt goed aan te pakken. (TV-nieuws)

Zulke constructies passen in geen van de bestaande analyses. Den Besten & Rutten spreken van 'de derde constructie' en nemen een regel aan, die zij *Scrambling* noemen, en die in een laat stadium (in onze termen dus postcyclisch) een NP over een V heen naar links verplaatst. Dat het hier inderdaad gaat om een betrekkelijk marginaal verschijnsel dat 'laat' beregeld dient te worden (en dat bovendien sterk gebonden is aan het register van informele spreektaal), blijkt uit de juist gegeven zin (ii), waar de NP *die jongen* is losgemaakt uit de overkoepelende object-NP *die jongen z'n gevoelens* om vervolgens over *proberen* heen naar links verplaatst te worden.

groep Nederlandse werkwoorden (in het Nederlands is de groep iets ruimer uitgevallen). Het ene opvallende verschil is echter dat het Duits links adopteert waar het Nederlands het aan de rechterkant doet. Het gevolg is dat bij toepassing van PR de Duitse volgorde automatisch een spiegelbeeld is van die in Nederlandse zinnen, zoals blijkt uit (20a-d), die beantwoorden aan respectievelijk (14a-d):

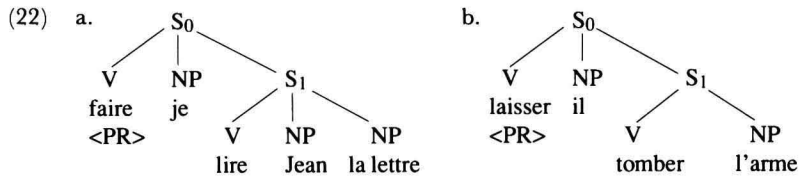
- (20) a. ...daß Jan den Hund die Zeitung holen zu lehren versuchen wollte
- b. ...daß Jan den Hund versuchen wollte zu lehren, die Zeitung zu holen
- c. ...daß Jan versuchen wollte, den Hund die Zeitung holen zu lehren
- d. ...daß Jan versuchen wollte, den Hund zu lehren, die Zeitung zu holen

Steeds waar PR is toegepast is de Duitse volgorde het spiegelbeeld van de Nederlandse. Waar PR niet is toegepast, dus op de cycli van *lernen* en *versuchen*, waar ook in het Duits PR optioneel is, komt een ingebedde /S (= VP) tot stand, en deze wordt in de Duitse orthografie steevast door een komma voorafgegaan. In (20b) en (20c) is PR steeds één keer wel en één keer niet toegepast. Daar vinden we danook maar één komma in de zin. In (20d), waar in beide gevallen PR niet is toegepast, treffen we twee komma's aan. Zowel de syntaxis als de orthografie van het Duits bevestigen aldus de gegeven analyse.

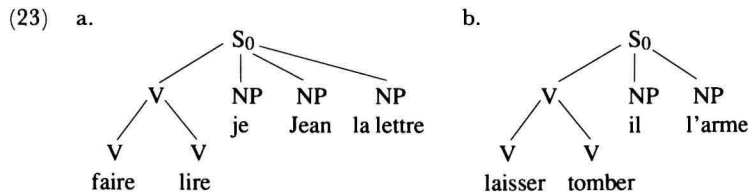
Het Frans kent PR bij een betrekkelijke kleine klasse van werkwoorden, waarvan *faire* en *laisser* de voornaamste vertegenwoordigers zijn.<sup>16</sup> De Franse zinnen:

- (21) a. Je ferai lire la lettre à Jean.
- b. Il laissa tomber l'arme.

zijn voorbeelden van PR-toepassing. Afgezien van de werkwoordstijden beantwoorden ze aan de onderliggende semantische analyses (22a) en (22b):



Toepassing van de verplichte PR-regel geeft respectievelijk:



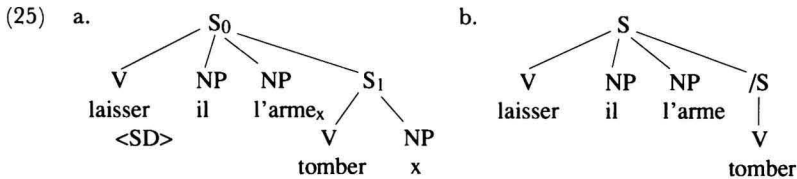
<sup>16</sup> Zie voor een gedetailleerde analyse van de Franse *faire*-constructie Seuren (1972).

Evenals vele andere talen heeft het Frans de regel dat wanneer het subject gevolgd wordt door nog twee NP-argumenten, de NP die onmiddellijk op het subject volgt een meewerkend voorwerp (dativus) wordt. De Franse dativus wordt, als hij niet pronominaal is, uitgedrukt door de prepositie *à* gevolgd door de NP in kwestie, en geplaatst achter de object-NP (Datief Extrapositie). (23a) krijgt dus nu de volgorde:  $s[v[v[\text{faire}]] v[\text{lire}]]_{NP}[\text{je}]_{NP}[\text{la lettre}]_{PP}[\text{à Jean}]$ .

Het Franse werkwoord *laisser* komt echter ook voor zonder dat PR is toegepast, en wel met een licht betekenisverschil:<sup>17</sup>

(24) Il laissa l'arme tomber.

Deze zin verschilt van (21b) doordat (21b) zegt dat de hij z'n wapen losliet (in het Engels: *He dropped his weapon*), terwijl (24) meedeelt dat de hij de val van het wapen niet onderbrak. Het verschil kan worden uitgedrukt door een andere argumentstructuur voor *tomber*, zoals getoond in (25a). Met deze argumentstructuur induceert *tomber* niet PR, maar de boven reeds besproken regel SD, met als resultaat (25b).



Het verschil in argumentstructuur (valentie) bestaat uit het feit dat in (25) *laisser* naast een nominaal subject ook een nominaal object heeft (*l'arme*), maar bovendien nog een toegevoegde object-S. Het nominale object controleert de deletie van het subject van de ingebedde object-S.

Als frappant detail kan hier worden opgemerkt (zie Seuren 1986) dat men dezelfde variatie in argumentstructuur, corresponderend met een analogo semantisch verschil, ook aantreft bij bv. het Nederlandse werkwoord *helpen*, zoals blijkt uit de volgende twee zinnen:

- (26) a. Ik heb Jan helpen failliet gaan.  
b. Ik heb Jan geholpen failliet te gaan.

Zin (26a) is duidelijk ambigu: hij betekent ofwel 'ik heb ertoe bijgedragen dat Jan failliet ging', ofwel 'ik heb Jan bijgestaan bij het bewerken van zijn eigen faillissement'. Zin (26b) kan alleen de laatste betekenis hebben. In (26b) hebben we de onderliggende argumentstructuur:

$s[v[\text{helpen}]_{NP}[\text{ik}]_{NP}[\text{Jan}_x]_{S}[v[\text{failliet gaan}]_{NP}[\text{x}]]]$ ,

terwijl het ambiguë (25a) hetzelfde argumentstructuur heeft, hetzij:

$s[v[\text{helpen}]_{NP}[\text{ik}]_{S}[v[\text{failliet gaan}]_{NP}[\text{Jan}]]]$ .

<sup>17</sup> Deze observatie staat op naam van de Belgische linguïst Nicolas Ruwet.

Wanneer we spreken van respectievelijk *helpen*<sup>3</sup> en *helpen*<sup>2</sup>, behoeven we nu nog maar te stellen dat *helpen*<sup>3</sup> PR optioneel, en *helpen*<sup>2</sup> PR verplicht induceert (en dat het lagere werkwoord *te* krijgt wanneer PR niet is toegepast). SD is uiteraard verplicht voor *helpen*<sup>3</sup>. Het verschil met het Franse *laisser* is dat, kennelijk, *laisser*<sup>3</sup> geen optionele PR meebrengt. *Laisser*<sup>2</sup> heeft, net als *helpen*<sup>2</sup>, verplichte toepassing van PR. Deze bijna volledige parallelie tussen *helpen* en *laisser*, gekoppeld aan de valentie (een argument meer of minder), wordt ook in andere talen gevonden bij andere werkwoorden.<sup>18</sup>

Er zijn nog veel meer frappante verschijnselen te noemen. Zo is er bij voorbeeld het feit dat bepaalde talen voor de behandeling van S-complementatie uitsluitend, of althans sterk, PR-gericht zijn terwijl andere talen voor analoge constructies uitsluitend, of zo goed als uitsluitend, de regel SR gebruiken. Het Nederlands en het Duits zijn (afgezien van de tempusroutine – zie § 5.2) nagenoeg exclusieve PR-talen.<sup>19</sup> Het Engels is een zo goed als exclusieve SR-taal, en het Frans heeft een sterke voorkeur voor PR, maar bezwijkt een enkele keer voor SR: daar treft men een zekere spanning met SR aan. Grevisse (1986, § 873) citeert een aantal gevallen, veelal met een accusatief pronomen, waar kennelijk SR is toegepast in plaats van de verplichte regel PR, zonder dat hij zich uiteraard van deze terminologie bedient. (De citaten (27a-d) zijn ontleend aan Grevisse; (27e) is door mijzelf geobserveerd):

- (27) a. ...j'ignore ... la nature des évidences...qui faisaient CE DÉCOURAGEMENT lui apparaître.  
 b. ...ces quelquea mots qu'il fait NISUS adresser à son Euryale.  
 c. Le courage de gorille...qui fait l'HOMME dominer sa peur...  
 d. Des nouvelles un peu moins bonnes LES firent précipiter leur départ.  
 e. Marié à une pharmacienne, qui L'a fait délaïsser sa passion pour la musique,...Martin rencontre un jour Éva.

Zulke gevallen zijn gemarkeerd: ze worden opgemerkt door professionele taalbespieders.<sup>20</sup> Een verklaring ligt meestal voor de hand, vaak in de vorm van een stilistische reden om de accusativus te gebruiken in plaats van de dativus. Dit blijkt wanneer men de zinnen omzet in hun syntactisch regelmatige vorm:

- (28) a. ...j'ignore...la nature des évidences...qui lui faisaient apparaître ce découragement.  
 b. ...ces quelques mots qu'il fait adresser à Nisus à son Euryale.  
 c. Le courage de gorille...qui fait dominer sa peur à l'homme...  
 d. Des nouvelles un peu moins bonnes leur firent précipiter leur départ.  
 e. Marié à une pharmacienne, qui lui a fait délaïsser sa passion pour la musique,...Martin rencontre un jour Éva.

<sup>18</sup> Zie bij voorbeeld Perlmutter (1970).

<sup>19</sup> Zie Seuren (1990a:99-109) voor een gedetailleerd argument.

<sup>20</sup> Grevisse (l.c.) merkt op dat de accusativus-constructie, dus SR, met *faire* in 17e-eeuws Frans veel frequenter was dan tegenwoordig.

In (28a) paradeert *lui* als subject van de ingebedde S, terwijl het ook daar een echte dativus is, en *ce découragement* ziet eruit als ingebed object, terwijl het ingebed subject is. (28b) en (28d) zijn onwelluidend vanwege respectievelijk de twee opeenvolgende *à*-bepalingen en de twee dicht bijeen staande voorkomens van *leur*. (28b) is bovendien ambigu naargelang *Nisus* dan wel *Euryale* als ingebed subject wordt opgevat. Waarom (28c) en (28e) onplezierig klinken is niet meteen duidelijk. Misschien heeft dit te maken met het feit dat de ingebedde S een mentaal proces weergeeft: z'n angst bedwingen, of z'n passie laten varen. Het is bekend dat predikaten die mentale processen weergeven prominente kandidaten zijn voor wat wel oneerbiedig genoemd wordt 'Dative Flip'. Dit is een taalveranderingsproces waarbij een oorspronkelijke dativus subject wordt. Men denke bij voorbeeld aan het Engelse *I like it*, dat in ouder Engels *It likes me* was. Of aan het Nederlandse *Mij mankeert niets* naast *Ik mankeer niets*. Dat dit op de een of andere manier in de taal ingebakken zit blijkt bv. uit een zin als *Ik maak niks uit*, van een Marokkaanse gastarbeider, die bedoelde te zeggen *Het maakt mij niets uit*. Het zou nader onderzocht moeten worden of en in hoeverre de dativus als gevolg van PR interfereert met het proces van Dative Flip.

Het Engels is daarentegen, zoals gezegd, een SR-taal. Toch vindt men ook daar een zekere spanning met PR. Het is bij voorbeeld goed Engels om te zeggen:

- (29) a. Then I let go the line.  
 b. The hostages must be let go immediately.

Zin (29a) is volstrekt acceptabel,<sup>21</sup> en verschilt semantisch van *Then I let the line go* op een wijze die geheel analoog is aan het verschil tussen de boven geciteerde Franse zinnen (21b) en (24). De woordvolgorde van (29a) en bovendien het passief in (29b) duiden op PR; in elk geval is SR in deze zinnen niet toegepast.

Dit is alles bij elkaar betrekkelijk interessant omdat hieruit niet alleen blijkt dat de geciteerde anomale gevallen toch blijken te passen in een universeel syntactisch kader, maar ook dat PR en SR kennelijk tot op zekere hoogte onderling inwisselbaar zijn. Dit laatste wordt weer dramatisch geïllustreerd door het Mauritiaans, de van het Frans afgeleide Creolentaal van het eiland Mauritius in de Indische Oceaan. Daar wordt zowel PR als SR aangetroffen bij dezelfde werkwoorden. Nader onderzoek (Seuren 1990b) heeft uitgewezen dat het criterium dat bepaalt of PR dan wel SR wordt gekozen ligt in de vraag of bij toepassing van PR de oorspronkelijke volgorde van subject en object in de ingebedde S bewaard blijft. Het Mauritiaans Creools kent, evenals het Frans, verplichte Datief Extrapositie, en het kan dus heel wel gebeuren dat, na toepassing van PR, een oorspronkelijk subject als da-

<sup>21</sup> In Amerikaans Engels is de combinatie *let go* zo zeer gelexicaliseerd dat het geheel als één werkwoord wordt opgevat, geconstrueerd met een prepositiebepaling ingeleid door *of*, zoals in *I let go of the line*.



tief-bepaling achter het oorspronkelijke object terecht komt, zoals bij voorbeeld in de boven geciteerde Franse zin (21a). In zulke gevallen mag PR niet worden toegepast, en wordt in plaats daarvan SR gebruikt, dat nooit zulke verspringingen tengevolge heeft.<sup>22</sup> Dit alles duidt er niet alleen op dat hier een systeem aan het werk is, maar ook dat we dit systeem in reële zin inderdaad op het spoor zijn.

### 5.2. SVO en de werkwoordstijden

We kunnen dit systeem nu moeiteloos uitbreiden om te verantwoorden hoe uit de onderliggende VSO- (of SOV-)structuur de oppervlaktestructuur SVO (precieser: NP-VP) ontstaat. In de semantische syntaxis gebeurt dit via de werkwoordstijden, of *tempora*.

Sinds Reichenbach (1947) wordt algemeen aangenomen dat finiete werkwoordsvormen ('persoonsvormen') het resultaat zijn van twee temporele operatoren. Dit kan in de semantische syntaxis als volgt beregeld worden. De hoogste,  $t_1$ , plaatst de zin in kwestie in het heden ('PRES') of een contextueel bepaald verleden ('PAST'), en in sommige talen ook de toekomst ('FUT'). Deze tempus is dus in zekere zin *deictisch*, doordat hij naar situationeel bepaalde tijdsaspecten verwijst. De lagere tempusoperator,  $t_2$ , plaatst de zin vervolgens in een relatie van gelijktijdigheid ('SIM') of temporele precedentie ('PREC') t.o.v. de door  $t_1$  genoemde tijd. De combinatie PRES-SIM levert de gewone tegenwoordige tijd op, en PAST-SIM de gewone verleden tijd. PRES-PREC geeft de voltooid tegenwoordige tijd (perfectum), en PAST-PREC geeft de voltooid verleden tijd (plusquamperfectum).

In de Westeuropese talen wordt SIM in het algemeen gerealiseerd als nulmorfeem ('Ø') dat al tijdens de cyclische behandeling wordt gesnoeid. PREC daarentegen wordt gerealiseerd als hulpwerkwoord van de voltooid tijd, doorgaans *have* in het Engels, en in het Nederlands *hebben* of *zijn*, alnaar gelang het hoofdwerkwoord. Deze hulpwerkwoorden dragen tegelijk de instructie 'VD' mee, die inhoudt dat het hoofdwerkwoord normaal gesproken de vorm aanneemt van voltooid deelwoord. Verder induceert  $t_2$  de regel L (met linkse adoptie door de lagere V), en  $t_1$  de regels SR en L (met opnieuw linkse adoptie door de lagere V), PREC blijft V bij Lowering, maar PRES en PAST krijgen de categorie 'Affix' (Af). Met een ingangsstructuur van de vorm VSO geeft deze beregeling automatisch de vorm SVO, of liever NP-VP, zoals blijkt uit figuur 5. Daar wordt een VSO-structuur in stappen omgebouwd tot een NP-VP structuur, ongeveer zoals een kikkervisje zich in stappen tot kikker ontwikkelt. We gebruiken de notatie  $S_n''$  voor een  $S_n$  met

<sup>22</sup> Dat een dergelijk criterium in een Creolentaal een rol speelt heeft vermoedelijk te maken met het feit dat Creolentalen, om historische redenen, 'semantische transparantie' nastreven. Daarmee wordt bedoeld dat de T-regels van de syntaxis de betekenis van de zinnen zo min mogelijk moeten verdoezelen, zodat het aanleren van de taal in kwestie zo gemakkelijk mogelijk verloopt en ook bij onvolkomen beheersing de verstaanbaarheid zo lang mogelijk gewaarborgd blijft. Zie Seuren & Wekker (1986).

twee tempora,  $S'_n$  voor een  $S_n$  met één tempus, die alleen  $t_2$  kan zijn, en  $S''_n$  voor een  $S_n$  zonder tempus. Na de tempusroutine worden de superscripten weggelaten. 'V<sub>M</sub>' staat voor 'hoofdwerkwoord'.

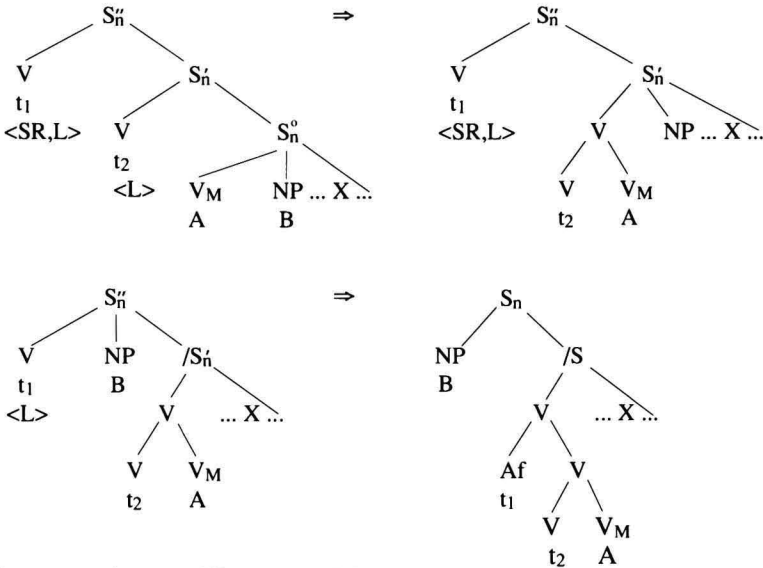


Fig. 5. Tempusroutine: van VSO naar NP-VP

Laten we een ogenblik stilstaan bij de verbale cluster in de output-structuur. Deze is, althans voor het Nederlands, van de vorm zoals die is gegeven in figuur 6.

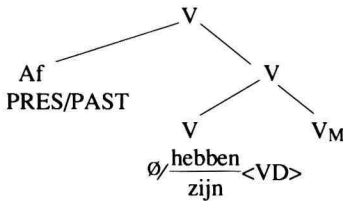


Fig. 6. Verbale output-cluster in het Nederlands

De regel VD bewerkt dat de onmiddellijk volgende V de vorm van een voltooid deelwoord krijgt, echter op één voorwaarde, die twee mogelijke formuleringen toelaat. We zeggen ofwel (a) dat  $V_M$  (de V van  $S''_n$ ) geen cluster mag zijn met meer dan één V, ofwel (b) dat wat de V van  $S''_n$  was in de semantische analyse uiterst rechts moet staan in een eventueel geclusterde  $V_M$ . Voor het Nederlands maakt het geen empirisch verschil of we de formulering (a) of (b) gebruiken. In beide gevallen wordt voorkomen dat een voltooid deelwoord optreedt wanneer de  $V_M$  zelf een complexe V-cluster is, ontstaan

bij voorbeeld door toepassing van PR. Dit verklaart waarom we in het Nederlands zeggen *Dat heeft mijn ogen doen tranen*, en niet *\*Dat heeft mijn ogen gedaan tranen*. Immers,  $V_M$  bestaat hier uit een V-cluster van de vorm  $v_M[v[\text{doen}]_V[v[\text{tranen}]]]$ , zodat noch aan conditie (a), noch aan conditie (b) is voldaan.

Voor het Duits echter maakt het wel degelijk verschil of de formulering (a) dan wel (b) wordt gekozen. In het Duits vinden we immers zowel *Ich habe dich tanzen gesehen* als *Ich habe dich tanzen sehen*. Bij PR heeft het Duits, zoals we gezien hebben, linkse adoptie. De  $V_M$  heeft hier dus de vorm  $v_M[v[\text{tanzen}]_V[v[\text{sehen}]]]$ , waar *sehen* de V is die in de semantische analyse direct onder PREC stond, dus de V van  $S_n^o$  in de semantische analyse. Nu is aan conditie (a) niet voldaan, maar aan conditie (b) wel. Het is verleidelijk te veronderstellen dat in sommige varianten van het Duits de formulering (a) van toepassing is, met als gevolg dat geen voltooid deelwoord optreedt, terwijl in andere varianten de formulering (b) geldt, zodat er wel een voltooid deelwoord optreedt.

Een opvallende factor bij de tempusroutine is dat slechts één kleine wijziging volstaat om oppervlaktezinnen te genereren van de vorm VSO. Daarvoor is slechts nodig dat  $t_1$  alleen de regel L induceert, en niet SR, zoals gemakkelijk valt na te gaan. Met een ingangsstructuur van de vorm SOV zal men, bij twee keer L, in de oppervlaktestructuur opnieuw SOV krijgen. Laat men daar echter  $t_1$  ook SR induceren, dan ontstaat weliswaar een NP-VP-structuur, maar de V bevindt zich dan uiterst rechts in de VP (/S), zodat in puur lineaire termen de volgorde dan ook weer SOV is. Het is nog de vraag of beide mogelijkheden gerealiseerd zijn in de talen van de wereld, en zo ja, waaraan het verschil empirisch afgelezen kan worden.

Een verder punt is dat we nu, gegeven het systeem van twee tempora, S-inbeddingen kunnen onderverdelen in inbeddingen met twee tempora, dus met een ingebedde  $S'_{n+1}$ , inbeddingen met één tempus, dus met een  $S'_{n+1}$ , en tempusloze inbeddingen met  $S_n^o$ . Het verschil is empirisch gemakkelijk vast te stellen. Inbeddingen met  $S'_{n+1}$  zijn altijd volledige bijzinnen met een finiet werkwoord (persoonsvorm). Inbeddingen met één tempus of zonder tempus zijn altijd infinitief, dus in de onbepaalde wijs. Wanneer de inbedding één tempus specificieert is het steeds mogelijk een infinitief te krijgen met een hulpwerkwoord van de voltooide tijd, zoals in *Ze moet nu gegeten hebben*. We stellen dan in het lexicon dat het epistemische *moeten* een subject- $S'_{n+1}$  inbedt. Anders ligt dit bij voorbeeld bij een werkwoord als het causatieve *doen*: daar kan men niet zeggen *Dat doet mijn ogen getraand hebben*. Alleen de tempusloze infinitief *tranen* is daar op z'n plaats. Het causatieve *doen* moet derhalve in het lexicon beschreven worden als een werkwoord dat een  $S^o$ , dus een tempusloze S, als objectterm neemt.

Talen zonder infinitief, zoals de meeste Balkantalen, laten alleen  $S'$ -inbeddingen toe. Immers, een ingebedde  $S'$  of  $S^o$  mist de  $t_1$  en de V van zo'n S krijgt dus nooit een finiete vorm: een ingebedde  $S'$  of  $S^o$  kan alleen maar infinitivaal zijn. In het Nieuw-Grieks, bij voorbeeld, heeft men danook geen

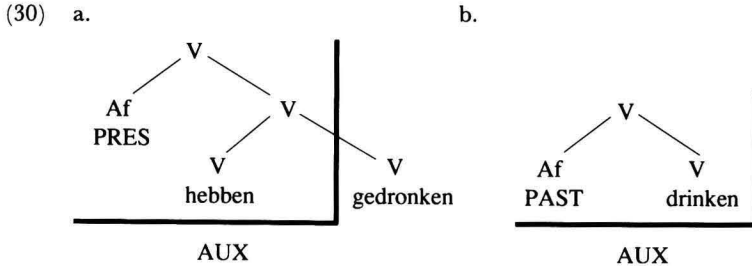
equivalent van bv. *Ik wil eten*, maar moet men zeggen 'ik wil dat ik eet' (*thélonafá*).

### 5.3. AUX

Dit systeem kan nu verder gedemonstreerd worden aan een aantal verschijnselen als V-Final in het Nederlands, inversie in het Nederlands en het Engels, of de plaatsing van de negatie in de Engelse werkwoordscluster, beide laatste regels in het Engels gecombineerd met de regel *Do-Support*. We zullen daarbij zien dat het Engels en het Nederlands in deze opzichten veel meer op elkaar lijken dan hun oppervlaktestructuren doen vermoeden.

Het gaat hier om verschijnselen die zich afspelen rondom de verbale output-cluster (VOC), en wel speciaal vlak voor of aan het eind van de cyclus. Het is bekend dat hier in veel talen sprake is van een geheel eigen soort verschijnselen, waarvoor lang niet altijd een inzichtelijke beregeling is gevonden. In de semantische syntaxis is, zeker voor een aantal bekende talen, een elegante oplossing voorhanden.

Het geheim schuilt in het opknippen vna de VOC in een finiet en een niet-finiet deel. Het finiete deel noemen we AUX, in overeenstemming met een gangbare traditie in de literatuur.<sup>23</sup> In het Nederlands wordt bij de vorming van AUX de VOC doorgeknipt vlak onder de tweede V van boven. In VOCs als (30a) of (30b) valt de caesuur dus als aangegeven:

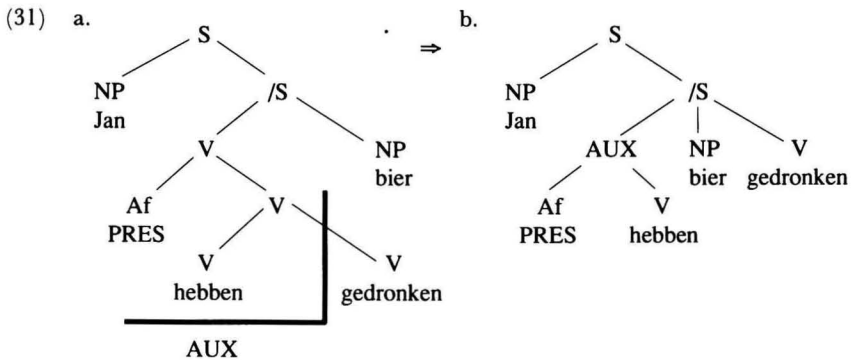


In (30a) bestaat AUX uit de VOC minus  $v$ [gedronken]. In (30b) beslaat AUX de gehele VOC, daar er onder de tweede V van boven niets meer volgt. Wanneer de AUX van (30a) door toepassing van een of andere regel wordt gescheiden van het niet-finiete deel  $v$ [gedronken], wordt hij vereen-

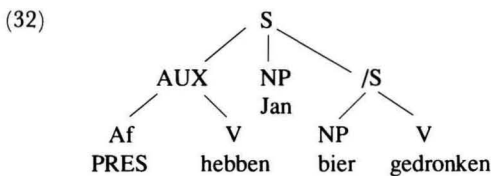
<sup>23</sup> Al in vroege stadia van de generatieve grammatika treft men een constituent AUX aan, juist omdat een onderscheid tussen finiet en niet-finiet deel van de werkwoordscluster steeds onmisbaar blijkt. De literatuur over AUX is overvloedig (zie bv. Klima 1964; Heny & Richards 1983), en elementen van de hier gepresenteerde behandeling vindt men her en der terug. Wat nieuw is aan deze behandeling is het losmaken van de AUX-constituent uit de verbale cluster zoals die eruit ziet na de tempusroutine. Hiermee wordt op eenvoudige wijze het verband tussen de syntaxis en de semantiek van de AUX-verschijnselen blootgelegd.

voudigd en omgedoopt tot  $AUX[Af[PRES]_V[hebben]]$ . De zuster-V van  $Af[PRES]$  is overbodig gezien de bovengenoemde algemene conventie waardoor automatisch knopen gesnoeid worden die niets anders domineren dan een knoop van dezelfde categorie.<sup>24</sup> Een latere morfologische regel zorgt ervoor dat  $Af$ , dus  $PRES$  of  $PAST$ , van de eerstvolgende  $V$  een persoonsvorm in de tegenwoordige of verleden tijd maakt.  $PRES$  gevolgd door *hebben* wordt dan  $_V[heeft]$  (mits het subject derde persoon enkelvoud is), en  $PAST$  gevolgd door *drinken* wordt  $_V[dronk]$ .

De boven al genoemde regel V-Final zorgt er in hoofdzinnen voor dat het niet-finiete deel van de  $VOC$  naar achteren wordt geplaatst, d.w.z. zo ver mogelijk naar rechts, maar niet over een ingebedde  $S$  of  $/S$  heen. Hierdoor wordt bij voorbeeld een structuur als (31a), mits hoofdzin, omgezet in (31b):

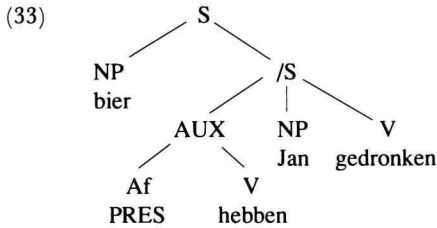


$AUX$  is kennelijk onmisbaar bij het formuleren van V-Final in hoofdzinnen. (In bijzinnen gaat de gehele V-cluster naar rechts: daar is  $AUX$  dus, voor deze regel althans, niet nodig.) Ook bij andere regels is  $AUX$  onmisbaar. In vragen treedt vooropplaatsing op van  $AUX$ . Gegeven (31a) is de vraagvorm:

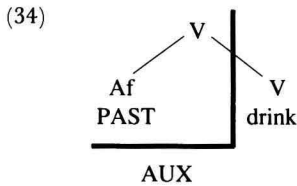


<sup>24</sup> De conditie voor het snoeien van ‘ijdele’ knopen moet in feite iets verruimd worden (al zijn de algemene termen daarvoor nog niet geheel duidelijk). Dit blijkt bij voorbeeld uit gevallen waar een aan L onderhevige  $V$  wordt gehercategoriseerd, zoals het geval is bij de Lowering van de negatie, die het etiket ‘adverbium’ krijgt. Onder wordt gedemonstreerd hoe in het Engels de negatie zich in de  $VOC$  nestelt en daarbij ook tot adverbium wordt omgedoopt. Ook in zulke gevallen dient de ‘ijdele’ zuster- $V$  van  $Af$  te worden weggesnoeid. Vermoedelijk luidt de juiste formulering dat lexicale categorieknoten, dus knopen die het etiket van een woordsoort dragen (zoals bv. ‘ $V$ ’), gesnoeid worden wanneer ze niet ook een woord van die soort domineren.

Ook bij vooropplaatsing van andere zinsdelen dan VOC is AUX een essentieel hulpmiddel. Het Nederlands kent immers het algemene structuurprincipe ('filter') dat de persoonsvorm, dus AUX, op de tweede plaats komt in de zin. Dit filter noemt men gewoonlijk 'V2'. Het bewerkstelligt de inversie van AUX en subject. Een dergelijke vooropplaatsing kan alleen in hoofdzinnen en moet beregeld worden na V-Final. Uitgaande van een structuur als (31b) krijgen we, bij vooropplaatsing van NP[bier], de structuur:



Het Engels is zeker zo interessant in dit opzicht. Daar komt AUX wederom tot stand door VOC af te knippen vlak onder de tweede V van boven. Er is echter een extra conditie: AUX mag nimmer het hoofdwerkwoord  $V_M$  bevatten. Een AUX als in (30b) is dus in het Engels uitgesloten. De VOC als gegeven in (34) laat  $V_M$  buiten AUX:

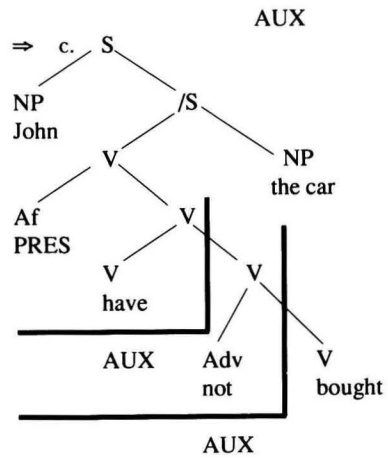
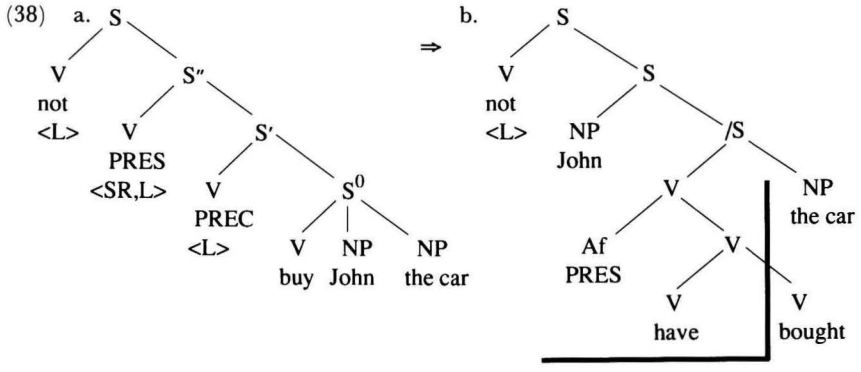
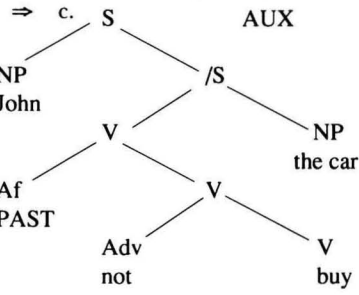
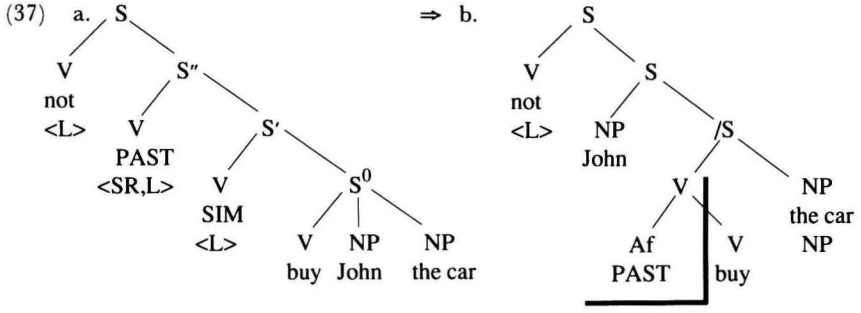


AUX is een constituent die in verschillende postcyclische regels optreedt. Zo wordt, evenals in het Nederlands, bij vragen AUX vooropgeplaatst. Bij de zg. *polar tag-questions* wordt de AUX na de zin herhaald en van negatie voorzien wanneer de VOC in de zin zelf geen negatie bevat, en omgekeerd, als in:

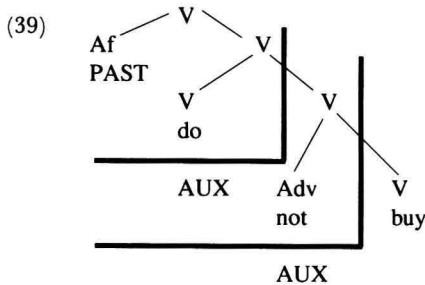
- (35) a. She is at home, isn't she?  
b. She isn't at home, is she?

Verder is AUX essentieel bij de beregeling van de negatie. De negatie-operator induceert de cyclische regel L, en de landingplaats is precies tussen AUX en  $V_M$ , waarbij *not* wordt gehercategoriseerd tot Adv (adverbium). Hieruit blijkt dat, tenminste in het Engels, AUX al moet bepaald worden tijdens de cyclus, en wel direct na Lowering van  $t_1$ . Het proces van negatie-Lowering wordt in de afleidingen (37) en (38) gedemonstreerd voor respectievelijk de zinnen (36a,b).

- (36) a. John didn't buy the car.  
b. John hasn't bought the car.



Nu moeten er twee verdere voorzieningen getroffen worden. Ten eerste moet de regel Do-Support worden ingevoerd.<sup>25</sup> Do-Support houdt in dat wanneer (a) AUX geen V bevat boven lexicaal materiaal en (b) Af (dus PRES of PAST) van  $V_M$  gescheiden is (geen zusterconstituent van  $V_M$  is), dan wordt  $v[do]$  ingevoerd vlak onder  $Af[PRES]$  of  $Af[PAST]$ . Dit nieuw ingevoegde dummy werkwoord  $v[do]$  wordt dan meteen onderdeel van AUX. De regel wordt uitgevoerd zodra de boomstructuur er aanleiding toe geeft. Men ziet nu dat Do-Support moet worden uitgevoerd op (37c). De VOC van (37c) wordt nu:



De tweede voorziening betreft de status van *not* nadat het op z'n landingsplaats is terechtgekomen. Hier laat het Engels twee mogelijkheden open: ofwel  $Adv[not]$  wordt onderdeel van AUX (inclusie), ofwel het blijft buiten AUX (exclusie). Het verschil blijkt uit de verschillende behandeling van *not* in bij voorbeeld:

- (40) a. Didn't John buy the car?  
 b. Did John not buy the car?
- (41) a. Hasn't John bought the car?  
 b. Has John not bought the car?

In de (a)-zinnen is *not* meegenomen met AUX, maar in de (b)-zinnen is het er buiten gelaten. Hoewel beide mogelijkheden door de grammatika worden toegelaten, is inclusie in AUX toch in modern Engels gangbaarder dan exclusie. Exclusie is sociolinguïstisch gemarkeerd als 'wat ouderwets' of 'eigen aan hogere socio-economische klassen'. In (38c) en (39) zijn beide vormen van AUX-afbakening aangegeven.

De bespreking van AUX kan hier slechts summier blijven. Allerlei verschijnselen, zoals met name de modale hulpwerkwoorden in het Engels, de hulpwerkwoorden van het passief of van de 'progressive form', moeten hier onbesproken blijven.<sup>26</sup> Het zal echter duidelijk zijn dat de boven geschetste aanzet, hoe onvolledig ook, tot optimisme stemt.

<sup>25</sup> De eerste die zowel de term als de regel invoerde was Klima (1964). Zijn beregeling wijkt echter nogal af van de hier gepresenteerde.

<sup>26</sup> Een uitvoerige publicatie, waarin o.a. AUX vollediger zal worden besproken, is in voorbereiding.



## 6. CONCLUSIE

Het is uiteraard ondoenlijk een alomvattende grammatikatheorie in het bestek van luttele bladzijden uit de doeken te doen. Wat wel doenlijk is, en dat is hier geprobeerd, is een betrekkelijk afgerond beeld geven van een stelsel van regels en regelschema's, waarvan op grond van ervaring en inductie wordt aangenomen dat ze behoren tot het universele arsenaal aan syntactische rekenwijzen, en wel als harde kern van de syntaxis van elke taal, en vervolgens aan een aantal sprekende voorbeelden te laten zien hoe feitelijke constructies in bepaalde talen met behulp van de gegeven apparatuur beregeld kunnen worden. Het aardige daarbij is dat sommige van de besproken constructies, zoals met name de Nederlandse PR-constructie (behandeld in § 5.1), die voor elke concurrerende theorie hetzij onoplosbaar hetzij behoorlijk weerbarstig zijn gebleken, in het kader van de semantische syntaxis moeiteloos geanalyseerd en beregeld worden. Het wachten is nu op gevallen die het omgekeerde beeld vertonen, constructies dus waarvoor andere theorieën gladde en moeiteloze beregelingen verschaffen, maar die niet of alleen ten koste van extra ad hoc theoretische 'investeringen' in de semantische syntaxis te behandelen zijn. Dan kan wellicht, eindelijk, worden overgegaan tot een inhoudelijke afweging van de verdiensten en de nadelen van de verschillende theoretische taalkundige kaders die elkaar momenteel de markt betwisten.

Inmiddels blijven de diepere algemene vragen bestaan. Welk complex van factoren heeft ertoe geleid dat betrekkelijk verticale semantische structuren van veel S-inbeddingen en weinig categorieën door de syntaxis worden omgezet in overwegend horizontale, 'vlakke' boomstructuren, met veel categorieën en een belangrijke rol voor volgordeverschijnselen. Men kan hier een zekere functionaliteit vermoeden, daar de 'steile' semantische bomen zich beter lenen voor semantisch rekenwerk en de 'vlakke' beter geschikt zijn voor snelle akoestische transmissie. Maar dan blijft toch nog altijd de vraag naar de eventuele functionaliteit van die oppervlakteverschijnselen die een semantisch verdoezelende werking hebben. Of zijn juist die verschijnselen niet functioneel, maar, bij voorbeeld, een bijwerking van de betrokken software of hardware?

De meest algemene vraag in dit verband, die naar de biologische oorsprong van rekensystemen in hersenen en hun verband met en relatie tot wiskundige rekensystemen, dringt zich nu met kracht op. Voor wat de syntaxis betreft begint zich nu een duidelijk beeld af te tekenen van zo'n biologisch rekensysteem. Wij zijn in de ban van een nieuwe verwondering over de behendigheid van de natuur.

## LIJST VAN GECITEERDE WERKEN

- den Besten, H. en J. Rutten, – *On Verb Raising, Extraposition and free word order in Dutch*. In: Jaspers et al. blz. 41-56. 1989
- Beth, E.W., – *Konstanten van het wiskundige denken*. Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen, afd. Letterkunde, Nieuwe Reeks, deel 26, No. 7. 1963
- Bloomfield, L., – *Language*, Holt, New York, 1933
- Chomsky, N. – *Lectures on Government and Binding. The Pisa Lectures*, Foris, Dordrecht, 1982
- Davidson, D. en G. Harman (red.), – *Semantics of Natural Language*. Dordrecht, Reidel, 1972
- Evers, A. – *The Transformational Cycle in Dutch and German*. Diss. RU-Utrecht. 1975
- Fodor, J.A. en J.J. Katz (red.), – *The structure of Language, Readings in the Philosophy of Language*. Prentice Hall, Englewood Cliffs, NJ. 1964
- Grevisse, M., – *Le bon usage*. Duculot, Parijs-Gembloux. 1986<sup>12</sup>
- Groenendijk, J.A.G., T.M.V. Janssen en M.B.J. Stokhof (red.), – *Formal Methods in the Study of Language*. Mathematisch Centrum, Amsterdam. (2 delen) 1981
- Harris, Z.S., – *Methods in Structural Linguistics*. Chicago University Press; Chicago. 1951
- Heny, F. en B. Richards (red.), – *Linguistic Categories: Auxiliaries and Related Puzzles*, Reidel, Dordrecht. (2 delen) 1983
- Jacobs, R.A. en P.S. Rosenbaum (red.), – *Readings in English Transformational Grammar*, Ginn, Waltham, Mass. 1970
- Jaspers, D., W. Klooster, Y. Putseys en P. Seuren (red.), – *Sentential Complementation and the Lexicon*. Foris, Dordrecht. 1989
- Kamp, H., – A theory of truth and semantic representation. In: Groenendijk et al. blz. 277-322. 1981
- Katz, J.J. en J.A. Fodor, – The structure of a semantic theory. *Language*, 39.2, 170-210. 1963
- Katz, J.J. en P.M. Postal, – *An Integrated Theory of Linguistic Descriptions*. MIT Press, Cambridge, Mass. 1964
- Klima, E.S., – Negation in English. In: Fodor en Katz blz. 246-323. 1964
- Koster, J., 1975 – Dutch as an SOV-language. *Linguistic Analysis* 1, 111-136. 1975
- McCawley, J.D., – Lexical insertion in a transformational grammar without deep structure. *Chicago Linguistic Society*, 4, 71-80. 1968
- McCawley, J.D., – English as a VSO-language. *Language* 46.2, 286-299. 1970
- McCawley, J.D., – A program for logic. In: Davidson en Harman blz. 498-544. 1972
- McCawley, J.D., – *Grammar and Meaning. Papers on Syntactic and Semantic Topics*, Taishukan, Tokyo. 1973
- Muysken, P.C. en N. Smith (red.), – *Substrata versus Universals in Creole Genesis. Papers from the Amsterdam Creole Workshop, April 1985*. Benjamins, Amsterdam. 1986
- Perlmutter, D.M., – The two verb *begin*. In: Jacobs en Rosenbaum blz. 107-119. 1970
- Reichenbach, H., – *Elements of Symbolic Logic*. Macmillan, Londen. 1947
- Russell, B., – Mr. Strawson on referring. *Mind*, 66, 385-389. 1957
- Schachter, P., – A non-transformational account of serial verbs. *Studies in African Linguistics, Supplement* 5, 253-370. 1974
- Seuren, P.A.M., – Predicate Raising and dative in French and sundry languages. Ongepubliceerd. Magdalen College, Oxford/L.A.U.T., Trier. 1972
- Seuren, P.A.M., – *Discourse Semantics*. Blackwell, Oxford. 1985

- Seuren, P.A.M., – Helpen en helpen is twee. *Glott* 9.1/2, 110-117. 1986
- Seuren, P.A.M., – *Filosofie van de taalwetenschappen*, Nijhoff, Leiden. 1990a
- Seuren, P.A.M., – Verb Syncopation and Predicate Raising in Mauritian Creole. *Linguistics*, 28.4, 809-844. 1990b
- Seuren, P.A.M. en H.Chr. Wekker, – Semantic transparency as a factor in Creole genesis. In: Muysken en Smith blz. 57-70. 1986
- Wundt, W., – *Logik. Eine Untersuchung der Prinzipien der Erkenntnis und der Methoden Wissenschaftlicher Forschung*. Enke, Stuttgart. 1880

