

# CGN Syntactische Annotatie

Heleen HOEKSTRA, Michael MOORTGAT, Bram RENMANS, Machteld SCHOUPPE,  
Ineke SCHURMAN, Ton VAN DER WOUDE

December 2003

## Voorwoord

In dit document bespreken we de CGN syntactische annotatie in termen van dependentielabels (functionele informatie) en knooplabeis (categorie-informatie).

In §2 kijken we naar lokale dependencies, zoals die optreden in het domein van lexicale inhoudswoorden. Vervolgens behandelen we de annotatie van niet-lokale dependencies in §3. We benutten daartoe de mogelijkheid om aan een constituent meerdere dependentierollen toe te kennen. Meervoudige dependentierelaties spelen eveneens een rol in §4, bij de behandeling van de samentrekkingsverschijnselen die bij nevenschikking optreden. In §5 behandelen we enkele dependentierelaties die betrekking hebben op het discourse-niveau.

In §1 geven we om te beginnen een overzicht van enkele strategische uitgangspunten van de gekozen annotatiemethode.

De paragrafen worden afgewisseld met **AANTEKENINGEN** (in goud), waarin niet behandelde fenomenen, keuzeproblemen e.d. worden gedocumenteerd waar de annotatoren in de loop van het project op zijn gestuit, alsmede de beslissingen die er in die gevallen genomen zijn.

Wij danken Gosse Bouma, Gertjan van Noord, Richard Piepenbrock en Michaela Poß voor hun commentaar op eerdere versies van dit document.

## Inhoudsopgave

<b>1</b>	<b>Uitgangspunten</b>	<b>4</b>
1.1	Dependentiestructuren	4
1.2	Spreektaalfenomenen	6
1.3	Dekking	7
<b>2</b>	<b>Lokale dependentiestructuren</b>	<b>8</b>
2.1	Het verbale domein	8
2.1.1	Dependentielabels	8
2.1.2	Categorielabels	23
2.1.3	Extensies van verbale domeinen: complementeerdere	25
2.2	Het adjectiefdomein	30
2.2.1	Complementatiepatronen	31
2.2.2	Modificatie	32
2.3	Het prepositionele domein	34
2.3.1	Prepositionele hoofden	34
2.3.2	(Voorlopig) object van een prepositie	35
2.3.3	Absolute <i>met</i> -constructies	36
2.3.4	Modificerende elementen	37
2.3.5	Vorzetsluitdrukkingen	37
2.4	Het nominale domein	39
2.4.1	De nominale kern	40
2.4.2	De determinator	40
2.4.3	Complementen binnen de nominale constituent	45
2.4.4	Modificerende elementen	46
2.4.5	Bijstellingen enz.	50
2.4.6	Deverbale nominalisaties	52
2.4.7	Vocatieven enz.	53
2.5	Restgevallen	57
<b>3</b>	<b>Niet-lokale dependencies</b>	<b>58</b>
3.1	Relatieve zinnen	59
3.2	'Hoofdloze' relatieve zin	60
3.3	Constituentvraag: hoofd- en bijzin	61
<b>4</b>	<b>Nevenschikking</b>	<b>62</b>
4.1	Asyndetische nevenschikking	64
<b>5</b>	<b>Discourse-structuren</b>	<b>65</b>
5.1	Discourse-markeerdere	65
5.2	Uitbreidingen van de zin	66
5.3	Asyndetische constructies	68

5.4	Anakoloeten en dergelijke	69
5.4.1	Spiegelconstructies	69
5.4.2	Ingebedde dislocaties	70

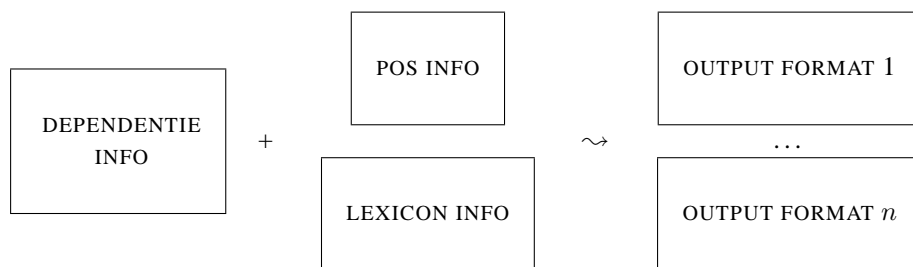
# 1 Uitgangspunten

De basisfilosofie van de annotatie is erop gericht een optimale input-outputverhouding te bereiken, in de volgende zin:

**Input** Aan de inputkant willen we de gehanteerde annotatieschema's zo eenvoudig mogelijk houden, om daarmee de werklust van de annotatie- en correctieprocedures te minimaliseren.

**Output** Aan de outputkant willen we de CGN-gebruikers een zo rijk mogelijke annotatie-informatie aanbieden, in een formaat dat voor verschillende gebruikersgroepen verschillende vormen kan aannemen.

Om dit doel te realiseren, gaan we uit van een zo theorie-neutraal mogelijk primair annotatieniveau in termen van dependentiestructuren. Die primaire annotatie kan worden verrijkt met informatie uit POS-tagging en, via de lexicologische koppeling, uit het CGN-lexicon. De combinatie van deze drie informatiebronnen leidt dan tot een aantal uitvoerformaten toegesneden op de wensen van verschillende gebruikersgroepen.



Wat betreft afgeleide outputformaten kan men denken aan:

- verrijking van categorielabels (voortaan: *c*-labels) met morfosyntactische features;
- verrijking van dependentielabels (voortaan: *d*-labels) met 'diepe' afhankelijkheden (zoals semantische controle);
- oppervlakte-constituentenbomen in een gebruikersvriendelijke notatie (met of zonder 'lege elementen' etc.);
- presentatiekwesties: keuzemogelijkheden voor de 'taal' van de label sets (Nederlands, Engels, ...)

Het is niet mogelijk om de outputkant van de annotatie los te zien van andere annotatieniveaus (zoals: lexicologische koppeling, prosodische annotatie) en van het projectonderdeel 'Exploitatiesoftware'. In dit stuk beperken we ons daarom tot de primaire dependentie-annotatie. Hieronder lichten we enkele kernbegrippen toe.

## 1.1 Dependentiestructuren

Een CGN-dependentiestructuur  $D = \langle V, E \rangle$  is een gelabelde gerichte, acyclische graaf (DAG). We beschikken over disjuncte verzamelingen CAT en DEP voor de labeling van, respectievelijk, knopen  $V$  en bogen ('edges')  $E$ .

- Knopen: CAT = POSCAT  $\cup$  PHCAT: categorielabels (*c*-labels), de vereniging van lexicale (part-of-speech) en frasale labels.
- Bogen: DEP: dependentielabels (*d*-labels).

We onderscheiden *gelede* en *ongelede* dependentiestructuren. Een ongelede dependentiestructuur is simpelweg een knoop met een *c*-label uit POSCAT. De elementaire bouwstenen van gelede dependentiestructuren noemen we lokale dependentiedomeinen. De moederknoop van een dependentiedomein is gelabeld met een frasaal label uit PHCAT. De dochters hebben *c*-labels uit CAT. De *d*-labels voor de moeder-dochtertakken worden gevormd door een *hoofd*, samen met de *complementen* en de *modificeerders* van dat hoofd.

**Hoofd** Het hoofd van een dependentiedomein projecteert het *c*-label van de moederknoop.

**Complementen** Het complementatiepatroon bepaalt de interpretatie van het hoofd in termen van thematische structuur. Een complementlabel komt per domein hoogstens één keer voor.

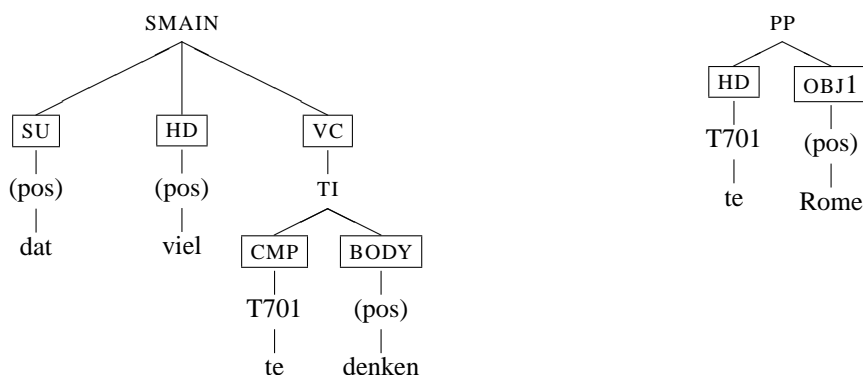
**Modificeerders** Modifierende elementen laten het *c*-label van de moederknoop ongemoeid; ze zijn (in principe) weglaatbaar zonder effect op de thematische structuur. Een modificeerderlabel kan binnen een domein meerdere voorkomens hebben.

**Commentaar** Het samennemen van complementatie en modificatie binnen één dependentiedomein leidt tot ‘ondiepe’ annotatiestructuren. Enkele gevolgen:

- een nieuw domein (hiërarchisch niveau) wordt pas geopend als een nieuw hoofd daar aanleiding toe geeft;
- complementatie en modificatie zijn *relaties* die frases onderhouden met betrekking tot een hoofd: als er geen complementen of modificeerders zijn, is er ook geen aanleiding tot niet-vertakkende projecties;
- dependentiedomeinen zijn, in het standaardgeval, *lexicaal verankerd*: het *c*-label van het hoofd valt samen met de POS-tag;
- de eis dat het hoofd het *c*-label van de moederknoop moet projecteren, houdt in dat we, waar de POS-informatie ondergedetermineerd is voor dat doel, het *d*-label van het hoofd kunnen laten desambigueren.

VOORBEELD 1. Ondiepe verbale projecties. We onderscheiden op *c*-label niveau *finiete* en *niet-finiete* verbale projecties. De persoonsvorm (=POS-tag) is hoofd van de finiete zinstypen; de infinitief of het deelwoord (=POS-tag) van de niet-finiete. Er is dus in de finiete zin geen behoefte aan een tussenliggend VP-niveau.

VOORBEELD 2. Op POS-niveau wordt geen onderscheid gemaakt tussen *te* als hoofd van een voorzetselgroep (PP), en als hoofd van een niet-finiete verbale projectie, de *te*-infinitief (TI): beide worden T701–VZ(init) gelabeld. We disambigueren door middel van het *d*-label voor *te*: voor het hoofd van de TI krijgt het het label CMP (zie §2.1.3).<sup>1</sup>

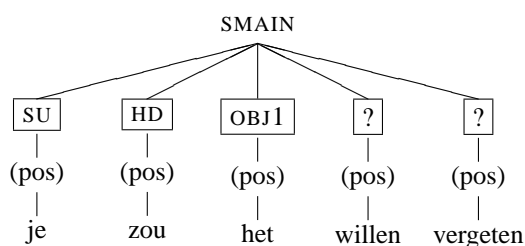


**Dependentie- versus constituentenstructuur** Uit het bovenstaande zal duidelijk zijn dat de dependentie-annotatie niet samenvalt met een (oppervlakte-)constituentenstructuur. In de CGN-annotatiefilosofie is die laatste een *afgeleide* van de dependentie-annotatie. De dependentie-annotatie, gecombineerd met de lexicologische koppeling, dient wel voldoende informatie te leveren om automatisch een (oppervlakte-)constituentenstructuur als exportformaat af te leiden. Welke vorm die constituentenstructuur dan aanneemt, kan afhangen van de beoogde gebruikersgroep, en van de rol die aan een (oppervlakte-)constituentenstructuur wordt toebedeeld in het geheel van de CGN-annotatieniveaus.

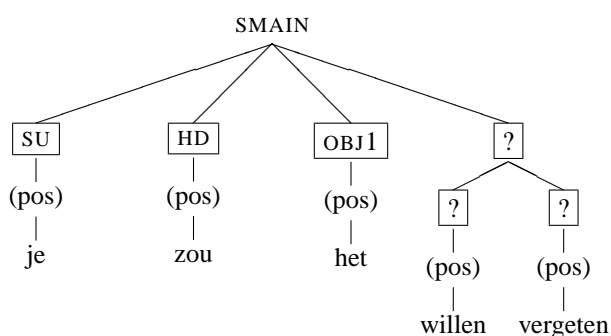
<sup>1</sup>De annotatievoorbeelden in dit document hebben de *c*-labels in klein kapitaal; de *d*-labels zijn ingelijst. De *d*-labels versieren de *bogen* van de annotatiegraaf: het zijn dus geen *knopen*. De POS-informatie wordt ongewijzigd van de POS-annotatie overgenomen, en hier meestal niet verder uitgespeld.

VOORBEELD. In de CGN-annotatie hebben werkwoordsgroepen op het dependentieniveau een geneste structuur, gemotiveerd door de onderscheiden subcategorisatie-eisen van de samenstellende hoofden. Op het niveau van (oppervlakte)-constituentenstructuur kan die geneste structuur als een ‘platte’ reeks werkwoorden worden uitgevoerd (bijv. voor een HPSG-gebruiker) of juist als een verbale clusterconstituent (bijv. als men aan de constituentenstructuur informatie over prosodische frasering wil ontfutselen).

Mogelijke uitvoer: plat



Mogelijke uitvoer: V-cluster



## 1.2 Spreektaalfenomenen

Gesproken taal kent een aantal constructies die in verzorgde geschreven taal in het algemeen als onwelgevoemd worden beschouwd en daarom zelden in de grammatische literatuur besproken worden (maar vergelijk [de Vries \(1911\)](#); [Jansen \(1981\)](#); [de Vries \(2001\)](#)). Zinsdelen kunnen bijvoorbeeld om discourse- of performance-redenen weggelaten of verdubbeld worden:

- (1) doen we ('topic-drop')
- (2) zijn vader beschuldigt hem dat hij zijn moeder vermoord heeft (voorlopig vz.vw. 'ontbreekt')
- (3) je moet heel snel moet je weer wisselen ('spiegelconstructie')
- (4) ik vind inmaak vind ik veel lekkerder ('spiegelconstructie')

Deze zinstypen worden ‘gewoon’ ontleed en krijgen zo nodig ‘gewoon’ twee onderwerpen, verbale hoofden enz. – zie paragraaf 5.4 voor details.

Dit soort constructies dient overigens scherp onderscheiden te worden van een ander spreektaalfenomeen, ‘performance errors’ en de daarop volgende ‘reparaties’:

- (5) bij een huwelijk was het vroeger gemakkelijk gezegd: tot de dood hon\*u ons scheidt hè<sup>2</sup>

<sup>2</sup>Als dat nodig lijkt voor een goed begrip van de voorbeeldzinnen, worden leestekens ingevoegd die in het CGN zelf niet gebruikt worden. In het CGN worden alleen de leestekens punt (.), vraagteken (?) en beletselteken (...) gebruikt.

(6) en het is waarschijnlijk het uh misschien het eten van sushi of ik weet niet wat

In (5) zegt de spreker *hon* waar hij of zij kennelijk *on(s)* bedoelt; de fout wordt in elk geval onmiddellijk hersteld. In (6) begint hij of zij met *het is waarschijnlijk*, bedenkt vervolgens dat dat tot een te sterke uitspraak gaat voeren, en vervangt het bijwoord door een zwakker woord, te weten *misschien*.

In dit soort gevallen nemen we aan dat de zinsfragmenten die gecorrigeerd worden daarmee automatisch niet tot de zin behoren; daarom worden ze in de syntactische annotatie ook buiten beschouwing gelaten.

Wordt een verspreking niet gecorrigeerd, dan wordt het ‘foute’ woord wel in de annotatie betrokken: in zin (7) functioneert *zonderlijke* als modificeerder van het zelfstandig naamwoord, *uitroeking* is hoofd van de lijdend-voorwerps-NP in (8).

(7) er lag geen spoor van humor in die zonderlijke vragen

(8) [...] waarin hij de uitroeking van de Pliniaanse volkeren aanbeval [...]

## AANTEKENINGEN

Als een woord herhaald wordt, wordt in het algemeen alleen het meest rechtse voorkomen in de boom betrokken. Bij verbetering mag hetgeen eerst gezegd was evenmin in de boom betrokken worden. Maar anderzijds willen we ook zo veel mogelijk tekst van een syntactische analyse voorzien. Dat betekent dat er soms wel een boom wordt gebouwd voor hetgeen later wordt verbeterd, maar dat dat ‘foute’ boompje niet wordt verbonden met de hoofdboom. Bijvoorbeeld:

(9) ik zeg ik wil uh ik wil dat wel gaan doen

We maken een klein boompje voor het eerste voorkomen van *ik wil*, dat niet vastzit aan de rest. We zijn daarbij wel pragmatisch: als het een erg rare boom oplevert, dan doen we het niet.

Hetzelfde geldt voor inhoudelijke verbetering (spreker beseft dat het vier jaren zijn ipv drie jaren):

(10) Maar ik heb lang in de drie in de vier jaren lesgegeven

In een zin als

(11) dat gebeurt eens in de drie vier jaar

is geen sprake van zelfcorrectie: *drie vier* betekent ‘(ongeveer) drie of vier’. We maken er een LIST van (zie 4.1) die DET is (zie 2.4.2).

Je hebt ook herhaling als stijlmiddel:

(12) hij liep langzaam, langzaam op het beest af

(13) we drinken drinken drinken totdat we zinken zinken zinken

Bij gebrek aan redelijke alternatieven maken we daar ook lijsten (LIST) van.

## 1.3 Dekking

Hoewel in dit document een fors aantal constructies van het Nederlands wordt beschreven, is het vanzelfsprekend onmogelijk om de complete Nederlandse grammatica hier uitputtend te behandelen. Niet behandelde fenomenen dienen zo veel mogelijk in de geest van dit document te worden geanalyseerd. De annotatoren documenteren dit soort gevallen, alsmede de genomen beslissingen.

## 2 Lokale dependentiestructuren

We bespreken achtereenvolgens dependentiedomeinen met werkwoordelijke, adjectivale, prepositionele en naamwoordelijke hoofden. Voor deze verschillende domeinen gebruiken we labels als *d*-labels HD (hoofd) en MOD (bepaling). Deze *d*-labels krijgen dus een ‘context-afhankelijke’ interpretatie: de POS-informatie van het hoofd bepaalt de moederknoop van het domein in kwestie (NP, AP, PP, of een van de werkwoordelijke projecties van Figuur 2); het *c*-label van de moederknoop, op zijn beurt, maakt duidelijk of MOD een bijvoeglijke dan wel bijwoordelijke bepaling is. Ook voor de *d*-labels van complementen maken we, zo mogelijk, van deze context-gevoelige interpretatie gebruik. Zo is OBJ1 het meest prominente argument in werkwoordelijke zowel als naamwoordelijke domeinen (‘lijdend voorwerp’ bij een verbaal hoofd, ‘oorzakelijk voorwerp’ bij een adjectief-hoofd, etc.).

### 2.1 Het verbale domein

#### 2.1.1 Dependentielabels

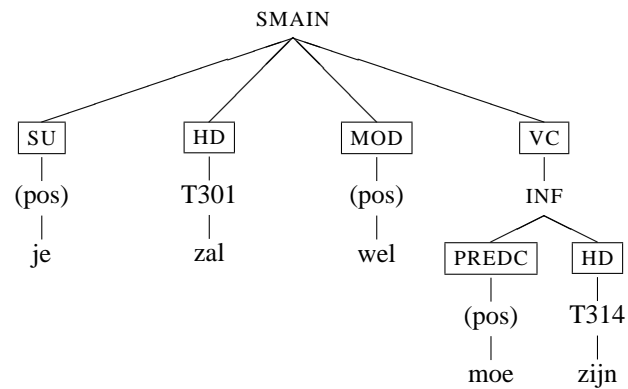
	<i>d</i> -LABEL	OMSCHRIJVING
<i>hoofd</i>	HD	werkwoordelijk hoofd (finit of niet-finit)
<i>complementatie</i>	SU	subject, onderwerp
	SUP	voorlopig subject
	OBJ1	direct object, lijdend voorwerp
	POBJ1	voorlopig direct object
	SE	verplicht reflexief object
	OBJ2	secundair object (meewerkend, belanghebbend, ondervindend)
	PREDC	predicatief complement
	VC	verbaal complement, werkwoordelijk deel van gezegde
	PC	voorzetselvoorwerp
	ME	maat (duur, gewicht, ...) complement
	LD	locatief of directioneel complement
SVP	scheidbaar deel van werkwoord	
<i>modificatie</i>	MOD	bijwoordelijke bepaling
	PREDM	bepaling van gesteldheid ‘tijdens de handeling’

Figuur 1: *d*-labels: verbale domeinen

**Werkwoordelijk hoofd** We onderscheiden finiete en niet-finiete verbale domeinen. Het werkwoord (finiete dan wel niet-finiete vorm) krijgt in beide gevallen het *d*-label HD. De combinatie van het *d*-label en de kenmerk-informatie in het POS-label differentieert tussen de finiete en de niet-finiete domeinen – in het voorbeeld hieronder T301: WW(pv,tgw,ev) versus T314: WW(inf,vrij,zonder).

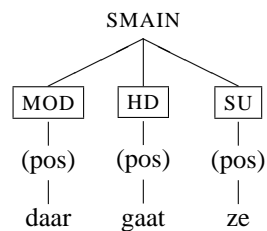
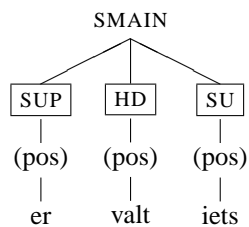
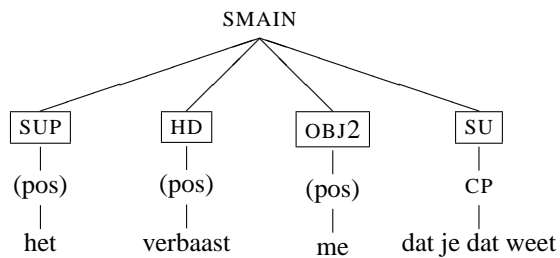
(14) je zal wel moe zijn





**Onderwerp, voorlopig onderwerp** De dependentierol SU (onderwerp) komt alleen voor in combinatie met een tempusdragend werkwoordelijk hoofd (de persoonsvormen T301–T309). De dummies ‘het’ en ‘er’ kunnen als voorlopig onderwerp (SUP) fungeren als het inhoudelijke onderwerp niet zijn kanonieke zinsinitiële positie inneemt. De voorlopige onderwerpen zijn gereduceerde, semantisch lege vormen, anders dan bijvoorbeeld het niet-gereduceerde, betekenisdragende *daar* in (18).

- (15) je verbaast me (SU)
- (16) het verbaast me dat je dat weet (SUP+SU)
- (17) er valt iets (SUP+SU)
- (18) daar gaat ze (MOD+SU)

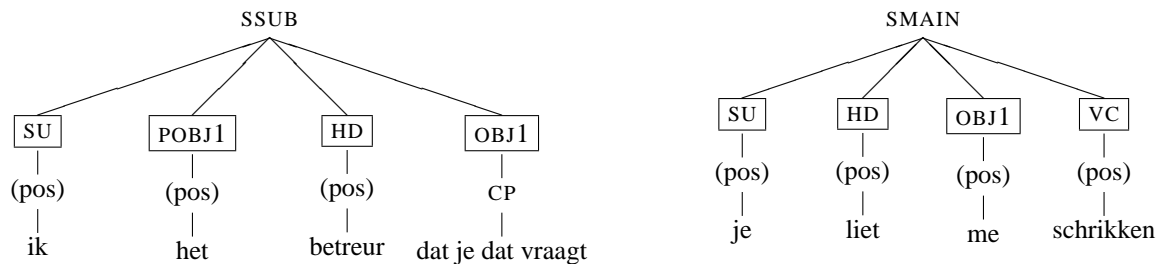


**Direct object, voorlopig direct object** In werkwoordelijke domeinen wordt het *d*-label OBJ1 toegekend aan het complement dat in het passief in de SU rol verschijnt, en aan het accusatief object in de ‘accusativus cum infinitivo’-constructie ACI-constructie (bij causatief *doen*, *laten*, en perceptiewerkwoorden *zien*, *horen*, *voelen*, ...), waar passivisering niet mogelijk is. Net als het onderwerp, kan ook de OBJ1-positie door een voorlopig lijdend voorwerp (POBJ1) worden ingenomen.

(19) ik betreur dit (OBJ1)

(20) (omdat) ik het betreur dat je dat vraagt (POBJ1+ OBJ1)

(21) je liet me schrikken (OBJ1)



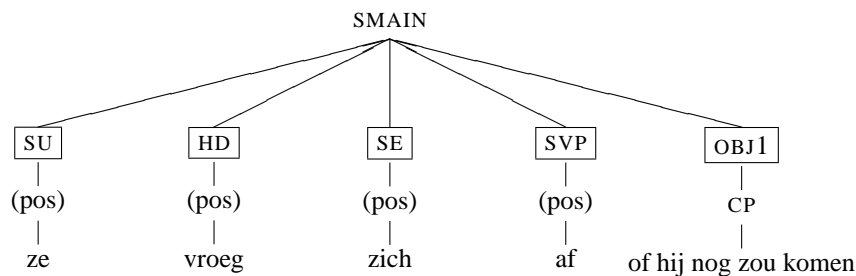
**Verplicht reflexief object** Criterium voor toekenning van het label SE: het is niet mogelijk om, met behoud van betekenis voor het werkwoordelijke hoofd, het SE gelabelde element te vervangen door een volledige NP. Vergelijk:

(22) ze vroeg zich | \*de jongen af of hij nog zou komen

(23) we schaamden ons | \*de jongens

(24) hij voelt zich | \*hem tevreden

(25) hij wast zich | zijn voeten (geen SE)



OPMERKING. We gebruiken het label SE ook bij niet-verbale hoofden, zoals in

(26) hij houdt het voor zich | \*de jongens

(27) hij is zich | \*de jongens bewust van de consequenties

**Secundair object: meewerkend, belanghebbend, ondervindend voorwerp** Het is niet altijd gemakkelijk meewerkend (IO), belanghebbend (BO) en ondervindend (EO) voorwerp van elkaar te onderscheiden. De ANS (die zelf meewerkend en belanghebbend voorwerp in één sectie behandelt) zegt hierover het volgende.

MEEWERKEND VW zinsdeel dat bij bepaalde w.w. en n.w. gezegdes kan of moet optreden en de zelfstandigheid aanduidt waarop de door dat gezegde en het lijdend of oorzakelijk v.w. uitgedrukte werking gericht is.

BELANGHEBBEND VW geeft aan in het belang van wie of wat de door een w.w. gezegde uitgedrukte handeling verricht wordt. I.t.t. meewerkend v.w. vrijwel nooit verplicht.

ONDERVINDEND VW zinsdeel dat bij bepaalde w.w. en n.w. gezegdes kan of moet voorkomen en de zelfstandigheid aanduidt die de door het gezegde uitgedrukte werking ondervindt. Nooit in de vorm van een voorzetselconstituent.

(28) hij geeft haar alles (IO)

(29) hij heeft alles aan haar gegeven (IO)

(30) ze schenkt hem een borrel in (BO)

(31) het verbaast me dat je nog aan het werk bent (EO)

(32) hij rookt me te veel (EO)

Om de annotatie werkbaar te houden, gebruiken we een overkoepelend label OBJ2 (secundair object), dat dan alleen van direct object (OBJ1) onderscheiden moet worden.

**Predicatief complement** Dit is een overkoepelend label voor subject- en object-georiënteerde predicatieve complementen. Motivatie voor het samennemen: de semantische oriëntatie kan worden vastgelegd in de lexicale ingang van de hoofden die deze complementen selecteren. In de traditionele terminologie (cf. ANS) verschijnt PREDC als:

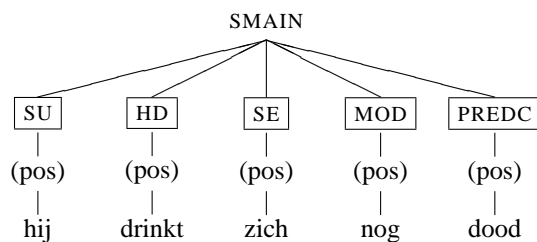
- naamwoordelijk deel van gezegde bij koppelwerkwoorden;
- bepaling van gesteldheid volgens de handeling ('tijdens de handeling' – zie het modificeerder-label PREDM in §2.1.1);
- resultatieve werkwoordsbepaling (of 'bepaling van gesteldheid ten gevolge van de handeling').

(33) ze is moe

(34) ze vindt hem een schat

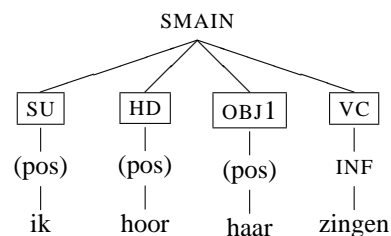
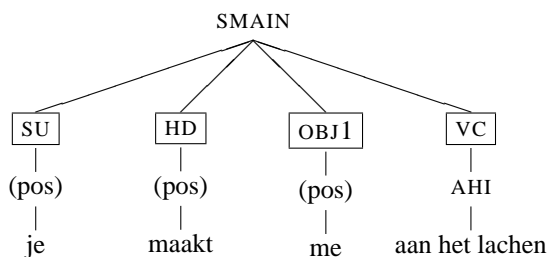
(35) hij drinkt zich nog dood

(36) zij verfde het kastje rood



**Verbale complementen** Voor niet-finiete verbale complementen beschikken we over het *d*-label VC. De dependentierol VC wordt toegekend aan werkwoordgroepen binnen de werkwoordelijke constituent, zoals behandeld in hoofdstuk 18 van de ANS (pp 937–1056). Het gaat om complementen met *c*-labels INF, PPART, TI, AHI, OTI (zie de tabellen 2 en 3). Het betreft hier de verbale complementen van verplicht groepsvormende hulpwerkwoorden van tijd en modaliteit, en van causatieve/permisieve (*doen, laten*) en waarnemingswerkwoorden (*horen, zien, voelen, ...*). Voorbeelden:

- (37) het blijft tobben (INF als VC)
- (38) hij heeft veel gelezen (PPART als VC)
- (39) zij bleek heel intelligent te zijn (TI als VC)
- (40) je maakt me aan het lachen (AHI als VC)
- (41) ik hoor haar zingen (kale infinitief als VC)



Merk op dat de accusatief in ACI-constructies (*haar* in (41)) wordt aangemerkt als OBJ1 van het matrixwerkwoord *horen* en niet als (logisch of anderszins) SU van het ingebedde werkwoord *zingen*.

De dependentierol VC wordt ook toegekend aan de beknopte bijzin (*c*-label TI of OTI) in complementatiepatronen waarin ook een nominale constituent als lijdend voorwerp (OBJ1) optreedt. Er is bij deze complementen geen sprake van groepsvorming.

- (42) je dwingt me (om) strenger op te treden (TI | OTI als VC)

In het geval van nominale OBJ2- of SE-complementen is er geen bezwaar om de beknopte bijzin de rol van lijdend-voorwerpszin toe te kennen.

- (43) ik vroeg hun om hun mond te houden (OTI als OBJ1)
- (44) hij verbood hun om nog langer te praten (OTI als OBJ1)

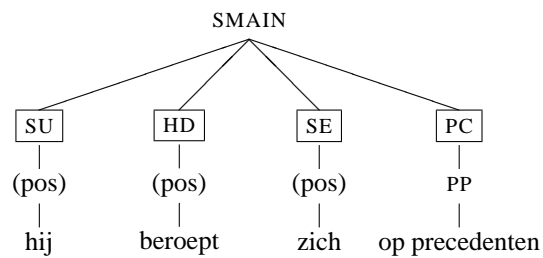
**Voorzetselvoorwerp** Het label PC wordt gebruikt voor PP-complementen met een ‘vast’ voorzetsel als hoofd. Eigenschappen:

- het ‘vaste’ voorzetsel laat zich niet door een ander voorzetsel vervangen zonder een betekenisverandering te veroorzaken voor het verbale hoofd met het PC-complement;
- het ‘vaste’ voorzetsel draagt niet compositioneel bij tot de betekenisopbouw van dit complementatiepatroon (in tegenstelling tot de voorzetsels in modificeerders en LD-complementen).

- (45) hij wacht op haar (PC)

(46) hij wacht op het perron

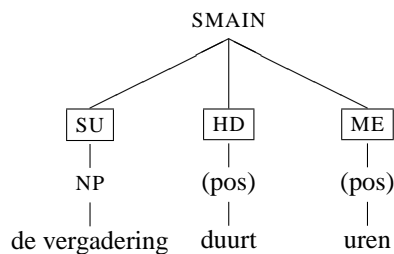
(MOD)



**Maatcomplement** Het *d*-label ME is voor complementen die een maat (gewicht, duur, prijs, ...) van het onderwerp uitdrukken. In tegenstelling tot het lijdend voorwerp (OBJ1) zijn deze ME-complementen niet passiviseerbaar.

(47) het kost 20 gulden

(48) de vergadering duurt uren



Optionele maataanduidingen, als in

(49) we hebben uren gefietst

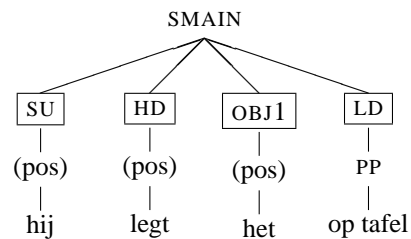
worden overeenkomstig de traditie aangemerkt als modificeerder (MOD).

**Locatief of directioneel complement** We gebruiken LD als overkoepelend label voor de plaats- en richtingscomplementen. De lexicale semantiek van het werkwoord en van (het hoofd van) het complement (het bijwoord en de voorzetsels in de voorbeelden hieronder) kan gebruikt worden voor nadere specificatie.

(50) het ligt boven

(51) hij legt het op tafel

(52) ze ging naar huis

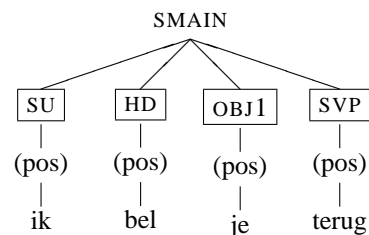


**Scheidbaar deel van samengesteld werkwoord** Het ‘partikel’-deel van een scheidbaar samengesteld werkwoord (*d*-label SVP) beschouwen we als het innigst met het werkwoordelijke hoofd verbonden element van een complementatiepatroon: het kan (en moet soms) weliswaar van het werkwoordelijke deel gescheiden worden, maar lang niet altijd.

- (53) ik bel je morgen terug (verplicht gescheiden)
- (54) of je hem morgen terugbelt (verplicht bij elkaar)
- (55) ik zal je morgen terugbellen (verplicht bij elkaar)
- (56) ik probeer je morgen terug te bellen (verplicht gescheiden)
- (57) of je hem morgen terug wil bellen (optioneel gescheiden)

De SVP-rol kan gerealiseerd worden door de categorieën voorzetsel, bijwoord, adjectief en zelfstandige naamwoord.

- (58) of je haar door wil laten (POS: voorzetsel)
- (59) of je het weg wil brengen (POS: bijwoord)
- (60) of je het klaar wil maken (POS: adjectief)
- (61) dat je hier moeilijk adem kan halen (POS: zelfst. naamw.)



Het is te voorzien dat er bij de annotatie problemen zullen zijn wat betreft het onderscheid tussen SVP en PREDC dan wel LD in die gevallen waar

- de semantiek van de combinatie volledig compositioneel is;

- de SVP bij een bepaalde werkwoordelijke stam behalve door een lexicaal element ook gerealiseerd kan worden door een volledige frase.

Voorbeelden:

(62) hij stuurt haar weg | naar huis (LD)

(63) de deur viel dicht | in het slot (PREDC)

(64) ze steekt haar handen omhoog | in de lucht (LD)

Hieronder enkele criteria die bij de keuze voor SVP kunnen helpen.

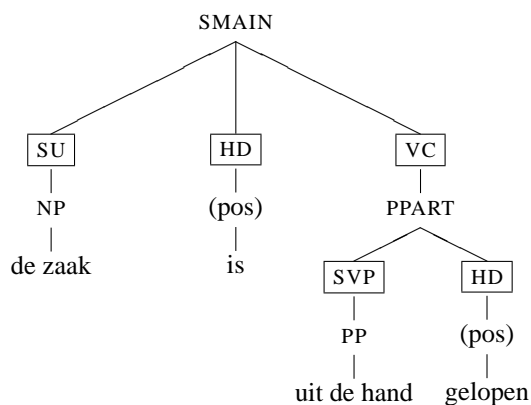
- De combinatie SVP + werkwoordelijk deel kan de input vormen voor morfologische derivatie (bijvoorbeeld: *vrijlating*, *bekendmaking*, *ademhaling*, maar *\*groenverving*) – voldoende maar niet noodzakelijke voorwaarde.
- De combinatie SVP + werkwoordelijk deel kan bij elkaar blijven binnen de werkwoordelijke cluster (bijvoorbeeld: (*of je haar*) *wil wegbrengen* naast *weg wil brengen*) – noodzakelijke maar niet voldoende voorwaarde (Vlaamse clusters!).

Niet-compositionele gevallen worden wel benoemd als SVP:

(65) de zaak is uit de hand gelopen

(66) hoe een natuurlijke taal precies in elkaar zit

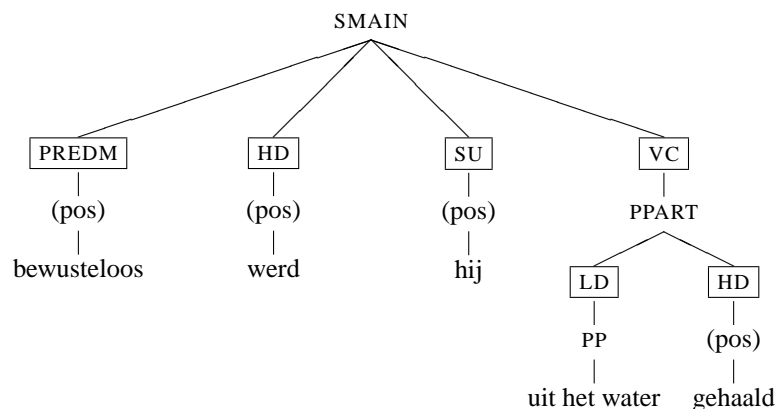
(67) om nieuwe samenwerkingen tot stand te brengen



**Modificerende elementen in het verbale domein** Bij de bepalingen binnen het verbale domein maken we een onderscheid tussen PREDM en MOD.

Het *d*-label PREDM gebruiken we voor wat wel de bepaling van gesteldheid ‘tijdens de handeling’ wordt genoemd. Dit type bepaling van gesteldheid speelt een modificerende rol, en moet dus onderscheiden worden van de bepaling van gesteldheid ‘volgens/ten gevolge van de handeling’ die tot het complementatiepatroon behoort, zoals hierboven gezien in de discussie van zin (33) en volgende.

- (68) hij kwam dronken thuis (‘tijdens de handeling’, PREDM)
- (69) je maakt je nog vuil (‘ten gevolge van handeling’, PREDC)
- (70) dat smaakt lekker (‘volgens de handeling’, PREDC)



Wat betreft de traditionele bijwoordelijke bepalingen maken we geen verdere semantische onderscheidingen tussen bepalingen van plaats, tijd, doel, gevolg, modaliteit, ... (in de ANS een totaal van 21 soorten): ze krijgen alle het default modificeerderlabel MOD. Dat het om een adverbiale bepaling gaat, blijkt uit het verbale *c*-label van de moederknoop in het lokale dependentiedomein.

De *door*-bepaling in passieve zinnen (“handelend voorwerp”)

- (71) de winnaar werd gekust door de rondemiss
- (72) de winnaar werd door de rondemiss gekust
- (73) door de rondemiss werd de winnaar gekust
- (74) de overtreders kregen door de politierechter een rijverbod opgelegd.

wordt eveneens benoemd als MOD. Deze bepaling wordt aangehecht op het niveau van het lexicale werkwoord en niet op het niveau van de HD, want bij nominalisering treffen we eenzelfde soort bepalingen aan:

- (75) het kussen van de winnaar door de rondemiss
- (76) het door minderjarigen roken van sigaren
- (77) de verovering van Carthago door de Romeinen

Speciale vermelding bij de bepalingen binnen het verbale domein verdienen de zogenaamde ‘modale partikels’ of ‘oordeelspartikels’: deze elementen vormen een cluster waarbinnen geen hoofd aan te wijzen is en vereisen daarom de invoering van een *c*-label dat niet vanuit één van de hoofdwoordsoorten wordt geprojecteerd.

De ANS onderscheidt oordeelspartikels van de klassieke bijwoorden van modaliteit (*misschien, wellicht, allicht, weliswaar, inderdaad; godweet, wieweet* (ANS: 456)) als elementen “die betrekking hebben op de gehele zinsinhoud en deze op subtiële wijze schakeren” (ANS: 457). Hieronder ter illustratie een inventaris van de functies die *maar* als modaal partikel kan vervullen:<sup>3</sup>

<sup>3</sup>Let wel: behalve als oordeelspartikel heeft *maar* natuurlijk nog vele andere functies!



- afzwakker van de stelligheid van beweringen en bevelen: *ik ga maar weg; kom maar; ik doe je heus geen kwaad*
- uitdagend: *kom maar; ik lust je rauw*
- duratief 1: *hij zit maar te zeuren de hele dag*
- duratief 2: *en maar zeuren, de hele dag door*
- in wenszinnen 1: counterfactual: *was ik maar bij moeder thuis gebleven*
- in wenszinnen 2: *als hij me maar niet vergeet*

De oordeelspartikels hebben een vaste positie in de zin, op de grens van ‘oude’ en ‘nieuwe’ informatie.<sup>4</sup> Op die positie vormen ze met elkaar een cluster, die soms een lengte van zes elementen kan bereiken.

(78) ik heb haar het boek nu toch eindelijk maar eens teruggegeven

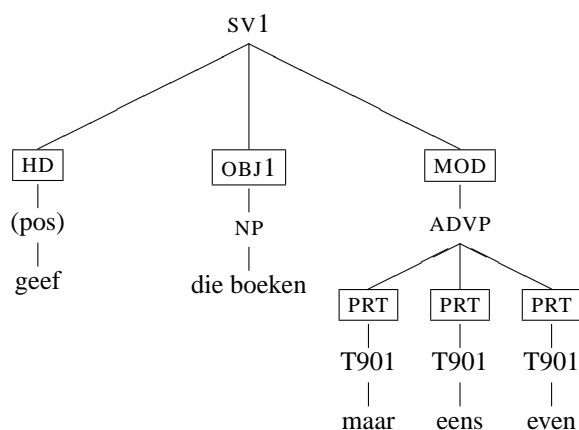
(79) ik heb haar nu toch eindelijk maar eens een boek gegeven

(80) geef de boeken dan nu toch maar eens even hier

Kleinere clusters hebben de neiging te verstenen; soms komt dat tot uitdrukking in de spelling (*weleens* naast *wel eens*, *allang* naast *al lang*), maar daar valt geen peil op te trekken, en we willen er hier niet van uitgaan dat de transcribeurs precies de regels van “het Groene Boekje” zullen eerbiedigen. Dikwijls is het zinvol om ze als eenheid te behandelen, omdat de combinatie ook semantisch als eenheid fungeert. Voorbeelden van zulke ‘vaste’ combinaties zijn:

*al eens, dan ook, maar eens, wel eens, weer eens*

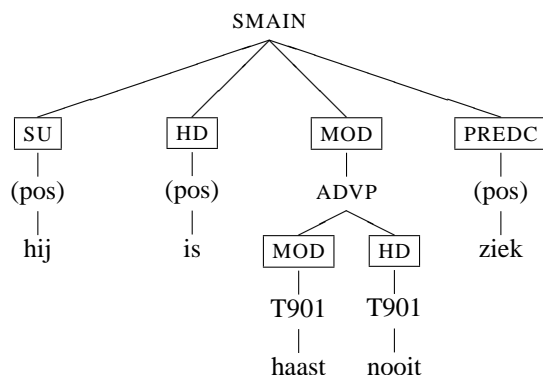
In het geval van clusters van oordeelspartikels groeperen we de samenstellende delen onder een *c*-label ADVP. Deze groep vervult dan als geheel een MOD-rol binnen het werkwoordelijke domein. Omdat we ons niet willen committeren aan de keuze van één van de partikels als hoofd van de ADVP-groep, krijgen de onderdelen als *d*-label PRT.



Overigens kunnen we het *c*-label ADVP natuurlijk ook gebruiken voor frases waarbinnen we wel degelijk een MOD+HD-articulatie kunnen aanbrenge.

<sup>4</sup>Of: rhema en thema.

(81) hij is haast nooit ziek



Ook kan een ADVP (syntactisch) ambigu zijn tussen een partikelgroep en een 'gewone' ADVP. Of er van het één dan wel het ander sprake is, valt vaak op te maken uit het zins- of tekstverband. Ook het accentpatroon van de zin geeft vaak uitsluitsel.

(82) ik krijg nog 'ns een PUNThoofd van jou (partikelgroep)

(83) zeg dat NOG 'ns als je durft (MOD+HD)

## AANTEKENINGEN

**SVP** In de volgende voorbeelden zijn respectievelijk *af*, *toe*, *rekening* en *naar voren* geassocieerd als SVP.

(84) we vroegen ons af of of ...

(85) we zijn dan toe aan agendapunt negen de mededelingen

(86) we zullen er rekening mee houden

(87) ik denk dat het belangrijk is dat u zo'n punt van zorg naar voren brengt

Het eerste geval, *af*, is oncontroversieel: het is deel van het werkwoord *zich afvragen* dat geheel andere eigenschappen, en een andere betekenis, heeft dan het (niet-transitieve) grondwoord *vragen*; *zich afvragen* staat dan ook apart vermeld in de woordenboeken (als de woordenboeken elkaar tegenspreken beslist in principe het CGN-lexicon). De redenering bij *toe* is analoog, hoewel er geen apart werkwoord *toezijn* (met vast voorzetsel *aan*) in het woordenboek te vinden is.

Het derde geval, *rekening* is een twijfelgeval: de combinatie SVP en werkwoordelijk hoofd kan niet door bijwoordelijke bepalingen doorbroken worden, maar het woord gedraagt zich (in elk geval voor sommige sprekers en schrijvers) als een 'gewoon' lijdend voorwerp voor de passief-test (*rekening moet worden gehouden met de volgende punten ...*). Aanvullende argumenten: het kan geen lidwoord krijgen (in andere constructies wel), het kan niet gemodificeerd worden (*idem*) en er kan niet naar terugverwezen worden (*idem*).

*Naar voren* in voorbeeld (87) is oorspronkelijk een LD, maar anderzijds wordt voldaan aan het criterium dat de combinatie *naar voren* en werkwoordelijk hoofd *brengt* niet door bijwoordelijke bepalingen doorbroken kan worden – maar dat geldt in het algemeen ook voor richtingsbepalingen. Maar er zijn meer argumenten om *naar voren* hier toch als SVP te benoemen: de letterlijke directionele betekenis is verdwenen, en modificatie is onmogelijk. Vergelijk:

- (88) het punt van zorg dat u (\*recht) naar voren brengt (SVP)  
 (89) de ballen die de keeper (recht) naar voren schiet (LD)

**Loos *het en loos er*** Behalve als voorlopig onderwerp, voorlopig lijdend voorwerp enzovoort kent *het* ook een aantal loze gebruiken, bijvoorbeeld als onderwerp (Haeseryn *et al.* 1997:1133).

- (90) *het* vriest dat *het* kraakt  
 (91) hoe gaat *het* met u?  
 (92) *het* ontbreekt hem aan moed  
 (93) *het* werkt prettig op deze kamer

“In deze gevallen heeft *het* geen eigen betekenis en geen verwijzende, maar alleen een syntactische functie” (Haeseryn *et al.* 1997:1133). *Het* kan ook loos lijdend voorwerp zijn (Haeseryn *et al.* 1997:1154):

- (94) de presidente had *het* warm  
 (95) samen zullen we *het* wel rooien  
 (96) hij zette *het* op een lopen  
 (97) de veldwachter had *het* op mij gemunt  
 (98) weer een keer dat we *het* erover hebben

Gevalen van *het* als loos meewerkend voorwerp kunnen we niet verzinnen. De ANS bespreekt wel het gebruik van het voornaamwoordelijk bijwoord als loos voorzetselvoorwerp (Haeseryn *et al.* 1997:1176):

- (99) nu ben je *erbij*  
 (100) we trokken *erop* uit

Er lijkt niets op tegen deze gevallen binnen CGN te benoemen als, respectievelijk, onderwerp, lijdend voorwerp en voorzetselvoorwerp – syntactisch vervullen de elementen de genoemde rollen, en dan is het voor onze doelen niet belangrijk dat daar semantisch niet zo veel van te merken is. Ingewikkelder zit het met *er*, dat de eerste zinspositie kan innemen in zinnen zonder onderwerp (Haeseryn *et al.* 1997:471):

- (101) *er* wordt hard gereden op de snelweg  
 (102) *er* is niets te zien  
 (103) *er* valt niets te veranderen

Als we een ander zinsdeel voorop zetten, dan wordt dit *er* soms optioneel, maar soms ook niet – en dat kan ook nog wel per spreker (of dialect) verschillen (Bennis 1986):

- (104) op de snelweg wordt (er) hard gereden  
 (105) hard gereden wordt \*(er) op de snelweg  
 (106) te zien is \*(er) niets  
 (107) niets is ?(er) te zien  
 (108) te veranderen valt er niets  
 (109) niets valt (er) te veranderen

Volgens Stef Grondelaers (hij heeft er een proefschrift ([Grondelaers 2000](#)) aan gewijd, maar we citeren uit zijn abstract voor de partikelconferentie in december 2000 in Brussel) is *er* in dit soort zinnen niet loos: hij denkt “that *er* is a presentative particle the speaker inserts when he expects that the hearer will invest too much energy in the processing of a subject which, although it is located in the sentence-final position restricted to maximally ‘newsworthy’ items, is of limited communicative importance. By using *er*, the speaker warns the hearer for the subject, and advises him to deactivate the inferences generated with respect to the subject, in order to guarantee as cost-free a processing of the subject as possible”, en dat zou dan weer samenhangen met de volgende eigenschappen:

- that presentative sentences with *er* rarely contain informationally rich subjects, but often contain negated subjects;
- that the subject in presentative sentences with *er* rarely “persists” in discourse after it has been introduced ([Givón 1983:14 ff.](#));
- that *er* significantly facilitates the processing of the subject entity in presentative sentences in which the preposed adjunct refers to an abstract or contextually non-activated entity;
- that *er* significantly facilitates the processing of subject entities which are incongruent with the inferences generated by the preposed adjunct and the main verb.

Dat alles zou ervoor pleiten *er* niet als puur grammaticaal op te vatten, maar als iets met een betekenis: een MOD dus. Dat doen we dus. In gevallen als

(110) waar zit 'm dat in

maakt 'm geen deel uit van het complementatiepatroon van *zitten*. Het lijkt er eerder op dat we hier met een relict van de datus ethicus te maken hebben, die fungeert als een modaal partikel (vergelijk bijvoorbeeld [Wegener \(1989\)](#)), dus als MOD. (Zie ook ‘*Me en daar* als modaal partikel’ op p. [23](#))

**PREDC** De traditie (Den Hertog) volgend, noemen we *weg* in *het boek is weg* en *uit de mode* in *die hoed is uit de mode* naamwoordelijk deel van het gezegde (PREDC), en wel omdat het bijwoord en de voorzetselconstituenten de functie van bijvoeglijk naamwoord vervullen (ANS).

In

(111) wie is deze meneer?

(112) wat wil jij later worden?

(113) wat zijn de kenmerken van de verbrandingsmotor?

wordt de rol van naamwoordelijk deel van gezegde door vragende voornaamwoorden. Het laatste voorbeeld laat zien dat het werkwoord *zijn* congrueert met *de kenmerken* . . . , dus die woordgroep moet wel onderwerp zijn. Volgens Den Hertog moeten we bij *die japon staat goed* niet aan een koppelwerkwoordconstructie denken, met andere woorden, *goed* is daar geen PREDM, maar een MOD.

Bij zinnen van het type *hij is dertien* volgen we de traditie, i.c. de ANS (435, 1204): *dertien* wordt beschouwd als een zelfstandig gebruikt hoofdtelwoord dat fungeert als een naamwoordelijk deel van het gezegde (elliptisch voor *dertien jaar oud*), dat wil zeggen, PREDC.

Zo ook:

(114) die toren is 100 meter (hoog)

(115) het water is 21 graden (warm)

**MOD bij verzwegen werkwoord** *Hij moet echt op de hand* (gewassen (worden)) (sfn000740,18). We zijn gewend bij zinnen als *we moeten naar huis* te zeggen: *gaan* is weggelaten, maar we laten *naar huis* toch maar de rol vervullen die

het bij *gaan* zou doen, dan maar bij *moeten*, dus krijgt *moeten* hier een LD. Volgens een vergelijkbare redenering heeft *moeten* ook de mogelijkheid van een MOD voor zinnen als *de auto is kapot dus we moeten op de fiets*. Als we dat eenmaal toelaten, en de parser al gewend is aan *moeten* met een MOD, dan kunnen we die aanpak ook wel uitbreiden naar minder ‘lichte’ werkwoorden als *wassen*: slechts een zeer beperkt aantal werkwoorden kunnen op deze manier argumenten en modificeerders erven, en die doen dat dan kennelijk bij allerlei soorten werkwoordelijke complementen.

**Kijken/zoeken naar/in/achter/onder/...** Voorzetselcomplementen bij *kijken* en *zoeken* krijgen het label LD. (Voor *kijken/zoeken of ...*, zie de AANTEKENINGEN na 2.1.3)

**Werken met complementen en modificeerders** Bij *werken aan* hebben we te maken met een PC, maar bij *werken met* (*je handen, twee collega's*), bij *werken in* (*het onderwijs, een inrichting voor zwakzinnigen*), *werken bij* (*de bakker, het spoor, je tante in de huishouding*) en *werken voor* (*een hongerloontje, een Griekse oliemaganaat*) met een MOD.

**Vinden** *Vinden* in de betekenis ‘van mening zijn’ heeft een PREDC en een OBJ1 (ik vind dat stom) of een PC (*wat vindt u van Brahms?*)

**Komen met** *Hij kwam met een heel bijzondere verklaring voor zijn afwezigheid. Met ...*: PC.

**Beginnen met, bij** PC.

**Praten met, over; denken over** PC.

**Werkwoorden met mee** Het is in het algemeen voorspelbaar dat een eventueel complement van een samengesteld werkwoord dat met *mee* begint een voorzetselgroep met *met* is:

- (116) meelopen met de optocht
- (117) meegaan met de tendens
- (118) twintig minuten laten meekoken met de aardappels
- (119) meekomen met de meester
- (120) meespelen met de Lotto

We benoemen die *mee*-bepalingen als PC.

Er zijn ook uitzonderingen op deze regel:

- (121) we weigeren mee te werken *aan* deze schending van de internationale rechtsorde
- (122) ik zou graag willen meedenken *over* een oplossing

maar ook hier benoemen we de voorzetselbepalingen als PC.

**Doen aan, met** De PP krijgt PC, ongeacht of het prepositioneel object wel of niet ‘affected’ is:

- (123) zij doet aan karate
- (124) zij doet iets aan de kapotte kraan
- (125) wat doen we met moeder met de feestdagen

De eerste *met*-constituent in (125) is een PC, de tweede is natuurlijk een tijdsbepaling, dus een MOD.

Let op: ook in *moeder doet niet mee aan/met de rituele vleesverbranding*, waar we te maken hebben met het werkwoord *meedoen*, is *aan/met* ... een PC.

**Aan ... gaan** In *aan de gang gaan, aan de wandel gaan, aan het bier gaan* enz. benoemen we *aan* ... als PC.

**Krijgen** De *van*-constituent bij *krijgen* (*dat heb ik van Sinterklaas gekregen*) is een PC.

**Maken** In *zij maakte van boter een toverheks* is *van boter* PC. Aan semantische haarkloverijen of het object wel of niet ‘affected’ is doen we niet, dus van *van je kamer in je maakt een troep van je kamer* maken we ook een PC, en zelfs van (het tweede) *werk in werk maken van werk*.

**We komen er niet uit** *Er ... uit* is een LD als het ‘letterlijk’ is (*we komen uit het bos*), en een PC als het figuurlijk is (*we komen vandaag nog uit het probleem*). Twee tests: de letterlijke variant laat modificatie met *regelrecht* toe, en de letterlijke variant laat toe dat *uit* een achterzetsel wordt (*we komen het bos uit* vs. *\*we komen het probleem uit*).

**We zijn het eens (over ...)** *Eens* is hoofd van een AP die als PREDC fungeert, met *het* als (loos) OBJ1 en de optionele *over*-bepaling als PC.

**Het werkwoord zijn** Bij het werkwoord *zijn* is het soms moeilijk uit te maken met welk soort complement we te maken hebben. Voorbeelden:

(126) het weegt twee kilo	(ME)
(127) het is zwaar	(PREDC)
(128) het is twee kilo	(ME?   PREDC?)
(129) hij was in Amsterdam	(LD)
(130) het was in 1993	(MOD?)
(131) het was in Amsterdam	(LD?   MOD?)

In (131) kan *in Amsterdam* betekenen ‘in de stad A’dam’ of ‘toen we in A’dam waren | woonden’. In het laatste geval gaat het om een tijdsbepaling en zou *in Amsterdam* het label MOD moeten krijgen, net als in (130), maar dat is dan wel een verplichte modificeerder, want het is niet weglaatbaar (*\*het was*). In het eerste geval is *het was* ambigu tussen ‘het bevond zich’ (LD) en ‘het vond plaats’ (LD? | MOD?). We hakken de knoop als volgt door:

(132) het is twee kilo   tien euro   ...	PREDC
(133) het was (= bevond zich) in Amsterdam	LD
(134) het was (= vond plaats) in Amsterdam	MOD

**Dat heeft niks met syntaxis te maken** Een lastige constructie. We hebben gekozen voor een TI die de rol van SVP bij *heeft* vervult, en een PC *met syntaxis* en een OBJ1 *niets*, ook bij *heeft*. Het werkwoord is (afhankelijk van de spreker en de inhoud van de zin) te vervangen door *schaften, doen, uitstaan*.

**Erbij, ervoor staan** Meestal LD: *ik stond erbij en ik keek ernaar*, maar soms SVP: *hoe staan de tuinbonen erbij dit jaar?* Zo ook *ervoor*: meestal LD, soms SVP: *hoe staan we d’rvoor?*

**Even** Volgens onze intuïtie hoort een woordje als *even* (of *effen*, *eventjes*, *efkes*) in onderstaande voorbeelden eerder bij het hoofdwerkwoord dan op zinsniveau:

(135) ik zal even kijken

(136) ga even zitten

*Even* komt echter ook in langere partikelclusters voor:

(137) ik zal er wel nog wel eens even naar kijken

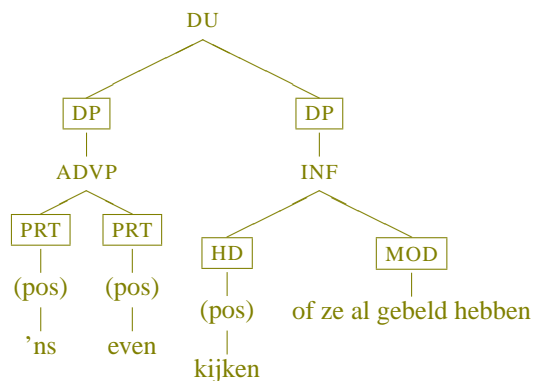
(138) ga maar eerst even rustig zitten

We hebben de neiging die lange reeksen kleine woordjes als een verder ongeanalyseerd cluster (ADVP) te benoemen, een cluster dat direct aan de SMAIN-knoop aangehecht is. Nu kunnen we zeggen: *even* laten we daar wel buiten, en hechten we in voorkomende gevallen toch lager aan, want het staat toch vrijwel altijd achteraan het cluster (de Vriendt *et al.* 1991). Dat is echter tegen de geest van het hele (nihilistische) idee van het partikelcluster, dat immers een manier om een heel moeilijk probleem (dat van de syntactische beschrijving van de modale partikels in het middenveld) niet nu even te hoeven oplossen.

Gegeven dat de eerder genoemde intuïtie over waar *even* bij hoort niets te maken hoeft te hebben met de syntactische positie van dat woord, behandelen we *even* daarom als zinsmodificerder, en nemen het in voorkomende gevallen op in een partikelcluster ADVP.

Nota Bene: als uit de context of de uitspraak of uit modificeerbaarheid met *heel* duidelijk is dat *even* letterlijk moet worden geïnterpreteerd ('een korte spanne tijds'), dan kan het wel geanalyseerd worden als behorend bij het hoofdwerkwoord.

Nota Bene 2: dit heeft als consequentie dat een veel voorkomend constructie-type als '*ns even kijken of ze al gebeld hebben*' wordt geanalyseerd als een DU (die later misschien wel wordt gespecificeerd als een adhortatief), met *eens even* samen in een ADVP onder een DP:



**Wat ... nog meer?** *Nog meer* is een PREDM, en *nog* en *meer* zijn PRTs.

**Me en daar als modaal partikel** Bijvoorbeeld in *toen heb ik ze me daar toch een pak rammel gegeven. Me daar toch samen ADVP, drie keer PRT.* (Zie ook '*m*' in (110))

## 2.1.2 Categorielabels

In Figuur 2 geven we de *c*-labels die geprojecteerd worden vanuit (finitie dan wel niet-finitie) werkwoordelijke hoofden. Bij de *c*-labels voor finiete zinnen (SSUB, SV1, SMAIN) beperken we ons in deze fase van de annotatie tot zuiver vormelijke

onderscheidingen in termen van de positie van de persoonsvorm (respectievelijk: finaal, initieel of op de tweede plaats). Wanneer ooit ook hogere discourse-niveaus in de annotatie worden betrokken, wordt het mogelijk deze hoofden te voorzien van functionele informatie met betrekking tot hun dependentierol (vraag, bevel, ...).

c-LABEL	OMSCHRIJVING
SMAIN	declaratieve zin (werkwoord op tweede plaats)
SSUB	bijzin (V finaal)
SV1	werkwoordsinitieële zin (ja/nee-vraag, bevel)
INF	kale infinitiefgroep
PPART	voltooid/passief-deelwoordgroep
PPRES	tegenwoordig-deelwoordgroep

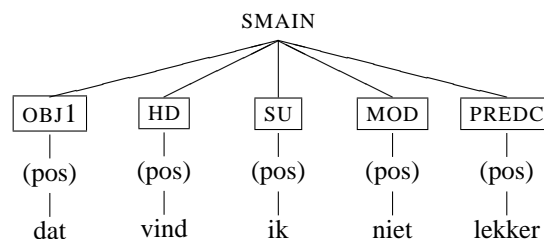
Figuur 2: c-labels: moederknopen van verbale domeinen

**Declaratieve zin** Hoofd is de persoonsvorm (met label HD), die de ‘tweede positie’ inneemt.

(139) het waait

(140) soms regent het

(141) dat vind ik niet lekker



Merk op dat constituentvragen als *wanneer vertrekt de trein naar Brussel* een ander label krijgen: zie §3.3.

We zetten ‘tweede positie’ tussen aanhalingstekens omdat sommige bijwoordachtige elementen tussen eerste zinspositie en verbogen werkwoord kunnen staan:

(142) Wim echter was allang ingevroren

In sommige exclamatieven staat het verbogen werkwoord eveneens op de tweede plaats:

(143) hij heeft me toch een mooie boeken gekocht!

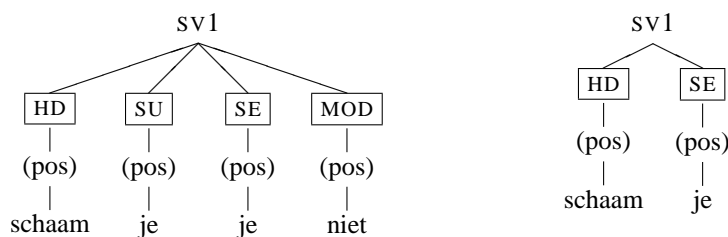
(144) in die kast staan toch een mooie boeken!

Syntactisch onderscheiden deze uitroepen zich in geen enkel opzicht van mededelingen; in de syntactische annotatie krijgen ze dan ook geen bijzondere behandeling.



**Werkwoordsinitieële zin** Het label SV1 is voor finiete werkwoordelijke domeinen waarbij de persoonsvorm (HD) op de eerste plaats komt. Het betreft hier vooral polaire vragen (ja/nee-vragen), directieven (bevelen met een imperatieve vorm van het werkwoord) en ‘voegwoordloze voorwaardelijke bijzinnen’.

- (145) schaam je je niet (polaire vraag)  
 (146) schaam je (bevel)  
 (147) val dood (bevel)  
 (148) houd je van vlees, braad je in Croma (voegwoordloze voorwaardelijke bijzin)



Ook topic-drop-zinnen als

- (149) doen we  
 (150) heb ik al gedaan

worden, op grond van de vorm, gecategoriseerd als SV1. In een later stadium kan, met een meer verfijnde verzameling discourse-labels, functioneel onderscheid gemaakt worden tussen de verschillende zinstypen met de persoonsvorm op de eerste plaats.

**Bijzin: werkwoordsfinaal** Voor de behandeling van SSUB verwijzen we naar §2.1.3. Daar worden onderschikkende voegwoorden behandeld, die SSUB-constituenten als complement vereisen.

### 2.1.3 Extensies van verbale domeinen: complementeerderelementen

De besproken finiete en niet-finiete verbale projecties kunnen worden uitgebreid met complementeerder-elementen. In Figuur 3 geven we de *c*-labels die door complementeerderelementen worden geprojecteerd. We gebruiken het *d*-label CMP voor de complementeerder en BODY voor de romp van de verbale projectie. (Voor de behandeling van extensies van het verbale domein met (vragende, betrekkelijke) voornaamwoorden met labels REL, WHREL, WHQ, WHSUB, verwijzen we naar §3, waar meervoudige dependentierelaties worden geïntroduceerd.)

CP-frases kunnen een complementsrol of een modificeerder-rol vervullen. Hieronder enkele voorbeelden van CP-complementen. De keuze van de specifieke complementeerder (*dat*, *of*) kan in het lexicon worden vastgelegd.

- (151) hij vroeg zich af of ze nog zou komen (CP in de rol van OBJ1)  
 (152) ze vermoedde dat hij niet meer zou komen (CP in de rol van OBJ1)

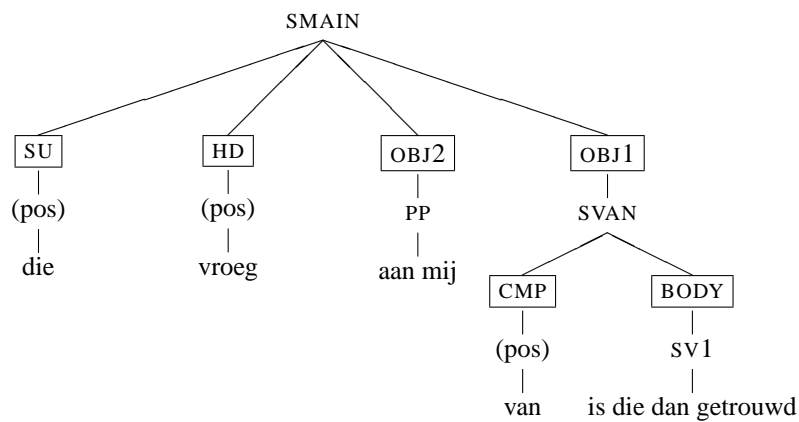
We beschouwen *van* in het voorbeeld hieronder als complementeerder ter inleiding van een complementszin in ‘directe rede’ (SV1 of SMAIN).

<i>d</i> -LABEL	OMSCHRIJVING
CMP	complementeerder
BODY	romp
<i>c</i> -LABEL	OMSCHRIJVING
CP	frase ingeleid door onderschikkend voegwoord
SVAN	<i>van</i> -zin (complement in directe rede)
TI	<i>te</i> -infinitief-groep
OTI	<i>om te</i> -infinitief-groep
AHI	<i>aan het</i> -infinitief-groep

Figuur 3: Complementeerder-extensies van verbale domeinen

(153) die vroeg aan mij van: is die dan getrouwd?

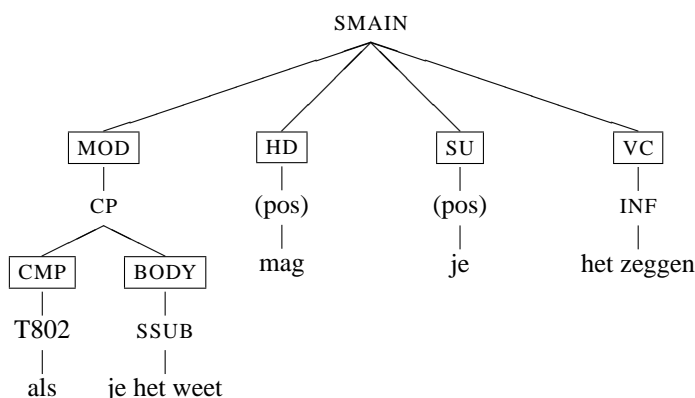
(154) dan heb ik zoets van: laat maar . . .



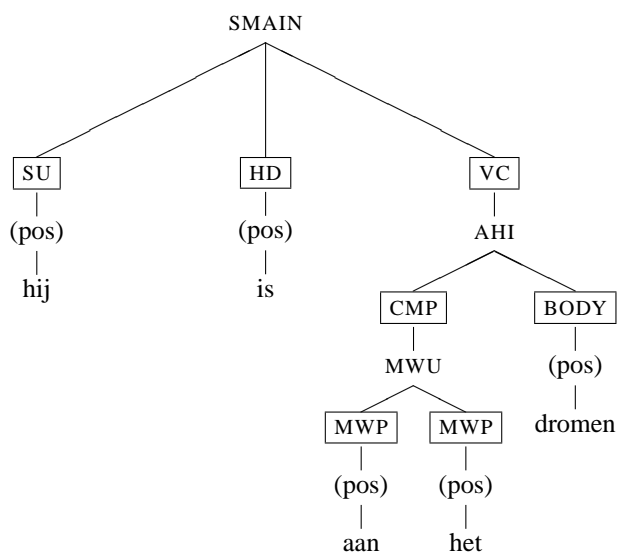
In de voorbeelden hieronder speelt de CP een modificerende rol.

(155) als je het weet, mag je het zeggen (MOD)

(156) omdat je het vraagt, zal ik het je zeggen (MOD)



OPMERKING In het geval van AHI behandelen we de combinatie *aan het* als complexe (meer-woord)complementeerder: het element *het* speelt hier een zuiver grammaticale rol. Dat betekent dat de elementen *aan* en *het*, die op het POS-annotatieniveau elk hun individuele tag krijgen, in de syntactische annotatiegraaf worden samengevoegd. Een dergelijke versteende, niet nader te analyseren woordgroep krijgt het label MWU (voor 'multiple word unit'), met gelijkwaardige dochters MWP. In §2.2.1 kiezen we dezelfde behandeling voor *het* + adjectief in vergelijkingsconstructies (*het meest*, *ik vind hem het aardigst*, ...).



**Complexe voegwoorden** Het is een klassiek gegeven dat de hoofdpositie van ingebedde Nederlandse zinnen door meer dan een element tegelijk kan worden vervuld.

(157) *zijn er meer mogelijkheden dan wat of je nu hebt?*

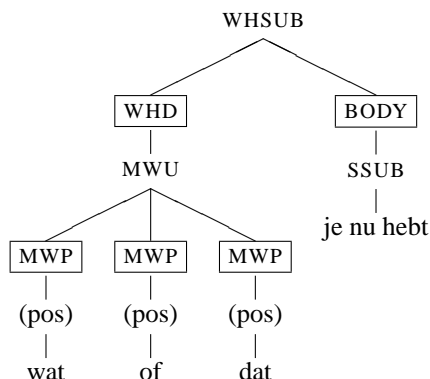
(158) ik zie niet in *wie dat* dat allemaal gaat kunnen realiseren

Varianten met *of* schijnen in België niet of nauwelijks voor te komen, vandaar dat we varianten met drie elementen alleen in Nederland verwachten:

(159) zijn er meer mogelijkheden dan *wat of dat* je nu hebt?

(160) ik zie niet in *wie of dat* dat allemaal gaat kunnen realiseren

De combinaties behandelen we als een MWU, die de rol van WHD in een WHSUB vervult:



met in dit geval een secondary edge OBJ1 van de MWU naar de SSUB.

Het komt ook voor dat een voegwoord schijnbaar ontbreekt. In beknopte bijzinnen weten we dat *om* regelmatig weg kan blijven: het label van de bijzin wordt OTI of TI, al naar gelang de aan- of afwezigheid van *om*:

(161) je dwingt me (om) strenger op te treden (TI | OTI als VC)

(162) de kans (om) een gouden medaille te winnen (TI | OTI als VC)

Er zijn evenwel ook gevallen waarin het voegwoord *dat* ontbreekt, zowel in Nederland als in België:

(163) voor het geval er iemand wordt gepakt of doorslaat

(164) voor het geval ik in de verleiding kwam

De betekenis van deze constructie (*voor*) *het geval* is ongeveer 'indien', en *dat* kan alleen wegblijven als *geval* niet gemodificeerd wordt (*\*voor het zeldzame geval ik in de verleiding kwam*); *het* kan niet vervangen worden door iets anders (*\*voor een geval ik in de verleiding kwam*, *\*voor dat geval ik in de verleiding kwam*).

Syntactisch fungeert *het geval* hier als een voegwoord (CMP), en zo analyseren we de combinatie ook. Andere enigszins vergelijkbare voegwoorden met een frasale herkomst zijn onder meer *indien*, *voor zover* (*voorzover*) en het nauw verwante *ingeval*.

## AANTEKENINGEN

**Of-zinnen** In een zin als

(165) ik zal (hem) eens bellen of hij komt

noemen we het *of hij komt*-‘complement’, dat een doel aangeeft, MOD. Het object is immers de degene die gebeld wordt en het label VC is gereserveerd voor ‘echte’ (d.w.z. niet-finale) werkwoordelijke complementen. (Hoewel uitgerekend dit type *of*-zinnen een argument zouden kunnen zijn voor het toelaten van finiete VC’s!)

Dit is dus anders dan

(166) ik zal (hem) eens vragen of hij komt

waarin de *of*-zin de rol van OBJ1 vervult: *of hij komt* kan worden vervangen door een pronominaal lijdend voorwerp (*dat*) en als degene aan wie iets gevraagd wordt, uitgedrukt wordt, dan gebeurt dat in een indirect object IO (*ik zal hun eens vragen of hij komt, ik zal het aan hem vragen*).

*Kijken of ...* en *zoeken of ...* is weer een ander verhaal. Allebei de werkwoorden kunnen een OBJ1 krijgen:

(167) we hebben gisteravond een film gekeken

(168) ik zoek m’n pen

en allebei de werkwoorden kunnen een *of*-complement krijgen:

(169) kijk | zoek ’ns even of m’n pen daar ergens ligt

maar niet tegelijk. In deze gevallen krijgt het *of*-complement ‘gewoon’ het label OBJ1.

**Ik denk van wel/niet** *Van wel/niet* is een SVAN met de rol van OBJ1.

**Nevenschiktelijke voorwaardelijke bijzinnen** Als twee voorwaarden samengetrokken worden dan komt het voor dat je maar één *als*-zin met bijzinsvolgorde krijgt; het tweede lid van de coördinatie heeft hoofdzinsvolgorde:

(170) Als je gaat fietsen en je komt langs de bakker, wil je dan brood kopen?

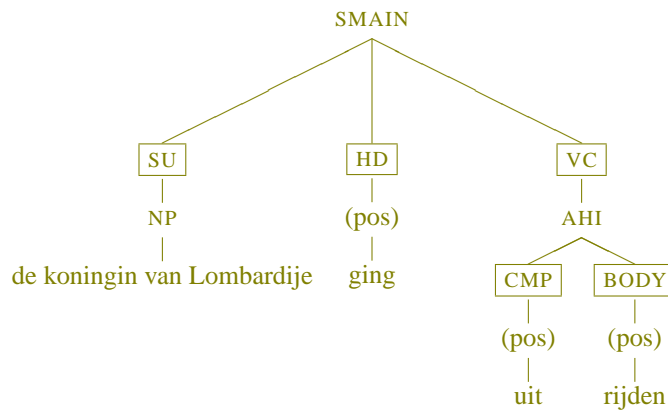
We analyseren dit zinstype als volgt: het geheel is een CP met *als* als CMP en als body een CONJ van een SSUB en een SMAIN.

**Als het ware** MWU.

**Uit + infinitief** In zinnen met een absentief (in de zin van [de Groot \(1995\)](#)) met *uit* (ANS: 1054–56): krijgt de ‘*uit*-infinitief’ het label AHL.

(171) de koningin van Lombardije die ging uit rijden

(172) pa is uit vissen



**Op klappen staan, op sterven liggen** De ANS (1056–7) suggereert dat dit een soort van verbaal complement is, en wij volgen die suggestie; *op klappen* enz. benoemen we als TI met *op* als complementizer.

**Voor + te-infinitief** In veel gesproken varianten van het Nederlands kun je *voor* als voegwoord gebruiken: *humor voor te lachen, iets voor te drinken, ben je geselecteerd voor te dansen* (sfn000771,141) enz. Syntactisch is er geen verschil met OTI, dus benoemen we de constructie als OTI.

**Dat is goed te doen** Het predikaat (TI) van zinnen als *dat is (goed, nauwelijks, ...) te doen* (doenbaar/doenlijk) en *dat is (goed, nauwelijks, ...) te zien* (zichtbaar) wordt geanalyseerd als een PREDC met de interne opbouw *goed*: MOD, *te*: CMP, rest: BODY.

**Laat staan (dat)** Krijgt het label MWU. *Laat staan* fungeert doorgaans als CRD in een nevenschikking (zie 4), *laat staan dat* als een voegwoord (CMP).

## 2.2 Het adjectiefdomein

In Figuur 4 geven we de *d*-labels voor adjectiefgroepen. Het *c*-label voor de moederknoop is AP.

	<i>d</i> -LABEL	OMSCHRIJVING
<i>hoofd</i>	HD	hoofd van AP
<i>compl</i>	OBJ1 POBJ1 PC VC OBCOMP	oorzakelijk voorwerp voorlopig oorzakelijk voorwerp voorzetselvoorwerp verbaal complement vergelijkingscomplement
<i>mod</i>	MOD	bepaling

Figuur 4: *d*-labels: adjectiefdomein

## 2.2.1 Complementatiepatronen

In de adjectiefgroep treffen we de *d*-labels PC, VC aan voor complementen die we ook in het verbale domein hebben ontmoet.

(173) bereid tot samenwerking (PC)

(174) op wraak belust (PC)

(175) bereid (om) te helpen (VC, vergelijk *bereid tot*)

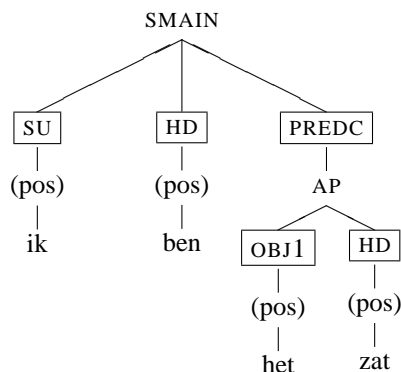
(176) bang dat hij zou opvallen (VC, vergelijk *bang voor*)

Het label OBJ1 is voor het ‘kale’ naamwoordelijk complement dat traditioneel ‘oorzakelijk voorwerp’ wordt genoemd. De ANS neemt hier geen duidelijke positie in: het oorzakelijk voorwerp wordt op pagina 901 behandeld als complement binnen de adjectivische constituent, maar op pagina 1179 als ‘hoger’ complement binnen het verbale domein. We kiezen voor de eerste optie, omdat weglating van het oorzakelijk voorwerp van invloed is op de interpretatie van het adjectivische hoofd. Vergelijk: *hij is het leven zat* en *hij is zat*. Merk op dat, zoals in het verbale domein, een voorlopige realisering (POBJ1) mogelijk is in combinatie met een (beknopte) bijzinscomplement.

(177) het gezeur beu (OBJ1)

(178) het zat (OBJ1)

(179) ik ben het zat om telkens op jullie te moeten wachten (POBJ1+OBJ1)

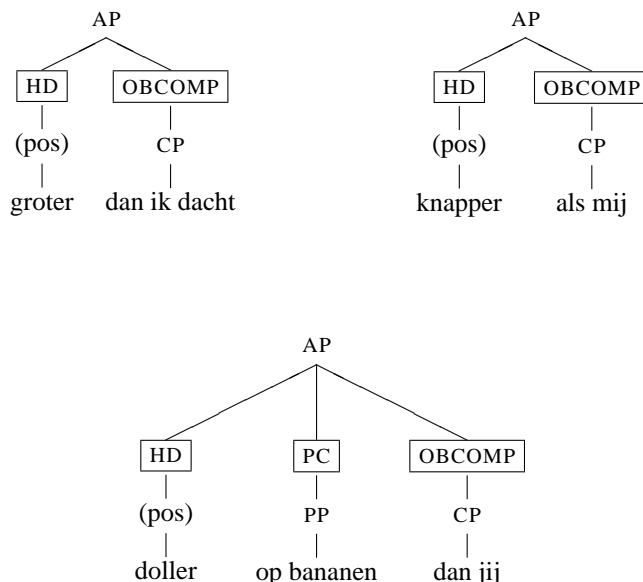


Het label OBCOMP is voor het complement van comparatieven. Een OBCOMP-complement kan ook voorkomen in het gezelschap van andere complementen, zoals in (182).

(180) groter dan ik dacht

(181) knapper als mij

(182) doller op bananen dan jij



Comparatieven en superlatieven kunnen soms ook vergezeld worden door een (optionele) partitief, maar als dat kan, dan kan het ook met de andere vormen van hetzelfde adjectief, en dan is het dus geen eigenschap van de comparatief-morfologie of -semantiek, maar van het adjectief. Deze partitief benoemen we dus als een modificeerder.

(183) dat is het dikkere boek (*van de twee*)

(184) dat is het dikste boek (*van de twee*)

(185) dat is het dikke boek (*van de twee*)

De annotatie van superlatiefconstructies verloopt geheel parallel aan wat hierboven is gezegd. We volstaan hier met de opmerking dat we het element *het* in de voorbeelden hieronder een puur grammaticale functie toedichten. We maken daarom een MWU van *het* met de superlatief, analoog aan de behandeling van de complementeerdercombinatie *aan het*.

(186) [het meest] gefrustreerd van allemaal

(187) [het dolst] op bananen van die twee

Tenslotte behandelen we hier de vergelijkingsconstructie *even | zo...als...*, hoewel het hoofd van die constructie (*even | zo*) natuurlijk een bijwoord is. De *als*-frase fungeert hier als OBCOMP van het bijwoord *even | zo*, en de constructie als geheel als modificeerder van een bijvoeglijk naamwoord.

(188) *even* groot als ik

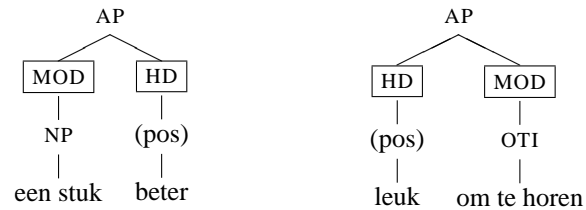
(189) *zo* gek als een deur

### 2.2.2 Modificatie

Het *d*-label voor bepalingen binnen de adjectivische constituent is MOD. De bepalingen hebben verschillende syntactische realiseringen. Hieronder enkele voorbeelden.



- (190) heel erg druk (AP als MOD)
- (191) vers van de pers (PP als MOD)
- (192) een stuk beter (NP als MOD)
- (193) leuk om te horen (OTI als MOD)
- (194) het meest gefrustreerd (MWU als MOD)



#### AANTEKENINGEN

's *Morgens vroeg* 's *Morgens vroeg* is een AP met *vroeg* als hoofd, en 's *morgens* als modificeerder. De argumentatie is een beetje indirect. Laten we deze constructie vergelijken met de semantische equivalente

- (195) Vroeg in de morgen

waarvoor je dezelfde vraag kunt stellen. Zeg je: het is een PP met een modificeerder, dan zeg je eigenlijk dat het te vergelijken is met *pal voor het huis* of *recht tegenover de bank*. Uit het proefschrift van Van Riemsdijk weten we dat je daarin R-pronomen kunt krijgen:

- (196) pal ervoor
- (197) er pal voor
- (198) waar woont hij recht tegenover?

maar bij *vroeg in de morgen* kan dat niet:

- (199) \*vroeg erin
- (200) \*er vroeg in

(die constructies zijn wel mogelijk, maar niet met een temporele lezing). Dus, zouden we zeggen, *vroeg in de morgen* heeft een andere structuur. Het enige redelijke alternatief is dan *vroeg* als hoofd en *in de morgen* als PP-modificeerder, en het geheel een PP. Na analogie is dan *vroeg* ook in 's *morgens vroeg* het hoofd, en het geheel een AP.

**Te + adjectief** *Te* modificeert het adjectief, eventueel met complement (OBCOMP) (*te dom om te helpen onweten*), eventueel gemodificeerd (*een jaar te jong om naar de universiteit te mogen*).

#### Heel oud versus twee jaar oud

- (201) ze is (al | \*pas) heel oud (MOD)

(202) ze is (al | pas) twee jaar oud (ME)

(203) het is (\*maar) heel lang (MOD)

(204) het is (maar) twee centimeter lang (ME)

In (201) en (203) fungeert *heel* als MOD van het adjectief, maar in (202) en (204) is er iets anders aan de hand. In dit soort constructies betekenen *oud*, *lang*, *breed*, *hoog*, *diep*, ... niet ‘ouder, langer, groter, dieper, ... dan gemiddeld’, maar zijn het ‘neutrale’ adjectieven, d.w.z. betekenen ze ‘met een leeftijd, lengte, breedte, hoogte, diepte, ... van ...’. M.a.w. *hij is x jaar oud* betekent niet ‘hij is oud, en wel x jaar’, in welk geval *x jaar* een MOD zou zijn geweest, maar ‘hij heeft een leeftijd van x jaar’, waarin *x jaar* niet kan worden weggelaten en dus een argument is, i.c. een ME.

## 2.3 Het prepositionele domein

In Figuur 5 geven we de *d*-labels voor voorzetselgroepen. Het *c*-label voor de moederknoop is PP.

	<i>d</i> -LABEL	OMSCHRIJVING
<i>hoofd</i>	HD	hoofd van PP
	HDF	afsluitend element van circumpositie
<i>compl</i>	OBJ1	object van prepositie
	POBJ1	voorlopig object van prepositie
	PREDC	predicerend complement
	LD	locatief/directioneel complement
<i>mod</i>	MOD	modificeerders binnen PP

Figuur 5: *d*-labels: prepositionele domeinen

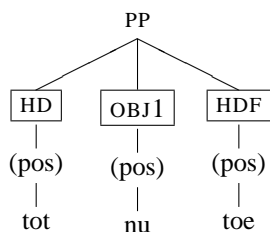
### 2.3.1 Prepositionele hoofden

Het voorzetsel is het hoofd van de PP. We onderscheiden twee hoofdlabels: HD en HDF. Het label HD wordt gebruikt voor de groepen waar er maar één hoofdkandidaat is. Het label HDF gebruiken we voor het afsluitende element van circumpositieconstructies (zie ANS, pp 526 e.v. voor een opsomming). Het label HDF komt dus alleen voor in combinatie met HD. Door de combinatie HD+HDF krijgen de circumpositieconstructies een niet-geneste annotatie. De behandeling van de discontinue circumposities is daarmee analoog aan die van continue samengestelde voorzetsels (*onderuit*, *bovenin*, ...).

(205) van het dak af (vergelijk: vanaf ...)

(206) tot nu toe

(207) van hier uit (vergelijk: vanuit ...)



OPMERKING. De elementen *om*, *te*, *aan+het* met POS-tag ‘voorzetsel’ krijgen een *d*-label als complementeerders, zoals gezien in §2.1.3.

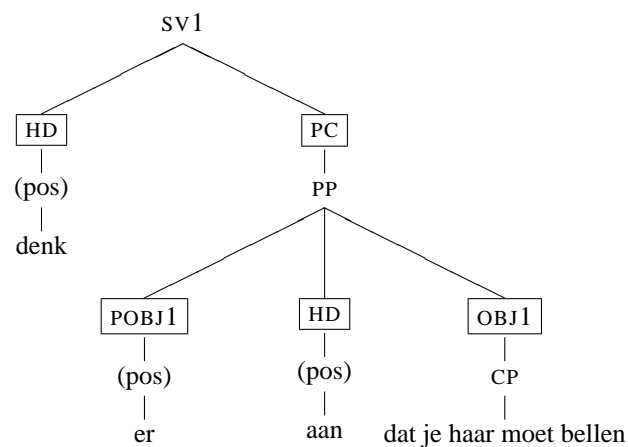
### 2.3.2 (Voorlopig) object van een prepositie

Net als in het verbale domein maken we een onderscheid tussen OBJ1, voor het ‘echte’ object van een prepositie, en POBJ1 voor *er* als ‘voorlopig’ object. Het label POBJ1 kan dus alleen voorkomen als ook OBJ1 in het betreffende domein is toegekend. De dochterknoop van de OBJ1-tak is in dat geval een (finit of niet-finit) verbaal complement. Vergelijk:

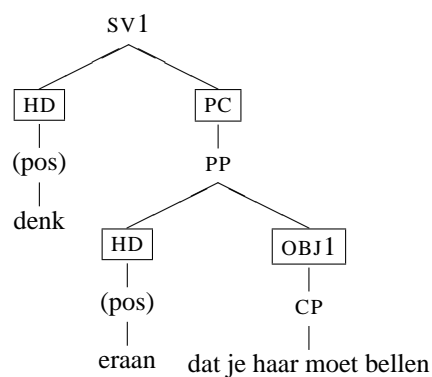
(208) je moet er niet aan denken (OBJ1)

(209) denk \*(er) aan dat je haar nog moet bellen (POBJ1+OBJ1)

(210) denk \*(er) aan haar vanavond nog te bellen (POBJ1+OBJ1)



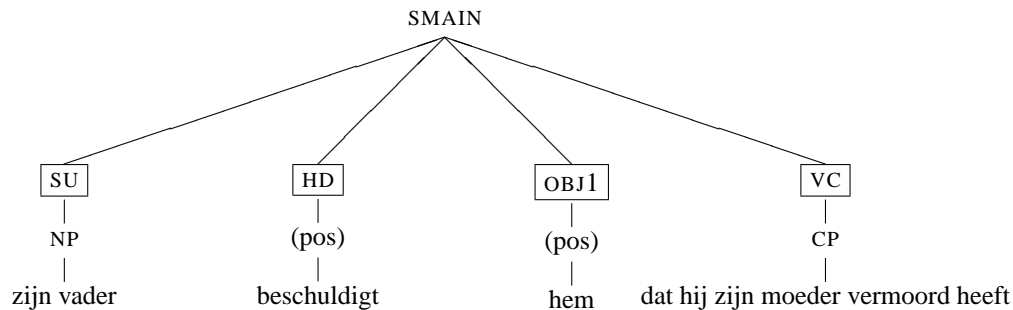
De Nederlandse orthografie schrijft evenwel voor dat voornaamwoordelijke bijwoorden dikwijls aan elkaar geschreven moeten worden (*eraan* in plaats van *er aan*). Het is niet aan de syntactische annotatie om de orthografische transcripties van het CGN-corpus aan te tasten, vandaar dat de volgende ad-hoc-oplossing wordt gehanteerd:



Met name in spreektaal komt het wel voor dat het voorlopig voorzetselvoorwerp niet wordt uitgedrukt:

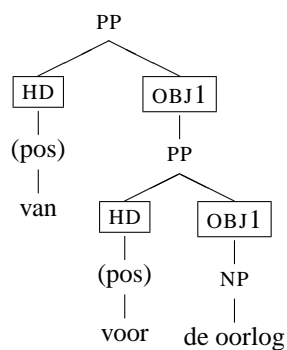
(211) hij zorgt dat het goed komt

(212) zijn vader beschuldigt hem dat hij zijn moeder vermoord heeft



In die gevallen ontbreekt dus het “vaste” voorzetsel. In het gewone geval (211) krijgt de CP het functielabel OBJ1, bevat de zin nog een eerste complement (212) dan wordt dat OBJ1 en krijgt de CP het label VC.

De functie OBJ1 heeft verschillende syntactische realisaties: NP, PP, finiete dan wel niet-finiete verbale complementen.



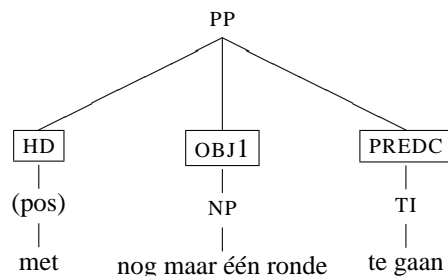
### 2.3.3 Absolute *met*-constructies

De absolute *met*-constructie is een PP-constituent die de rol van bijwoordelijke bepaling (MOD) of bepaling van gesteldheid (PREDM) vervult. Behalve het prepositional object treedt in deze constructie nog een tweede complement op met een LD- of PREDC-functie.

(213) met Haider in de regering (LD)

(214) met de helft van de ploeg dronken (PREDC)

(215) zonder wapens in huis (LD)



### 2.3.4 Modifierende elementen

Zoals in het verbale domein, gebruiken we een enkel *d*-label (MOD) voor de bepalingen van hoeveelheid, plaats, tijd, ... binnen de voorzetselconstituent.

(216) een uur voor de wedstrijd

(217) pal achter het voorzetsel

(218) bij hem thuis

(219) zes graden onder nul

### 2.3.5 Voorzetseluitdrukkingen

Met de ANS (510–11) beschouwen we voorzetseluitdrukkingen als lexicaal gefixeerde groepen (ingeleid door een voorzetsel) die in hun geheel de functie van een voorzetsel vervullen. Voorbeelden zijn *op grond van* in (220) en *aan de hand van* in (221), omdat ze respectievelijk gelijkgesteld kunnen worden met de voorzetsels *wegens* en *met*.<sup>5</sup>

(220) op grond van haar grote verdiensten werd zij in de adelstand verheven

(221) aan de hand van deze plattegrond zul je je wel kunnen oriënteren

Groepen als deze kunnen vaak ook een letterlijke betekenis hebben, zoals in (222) en (223); het zijn dan geen voorzetseluitdrukkingen:

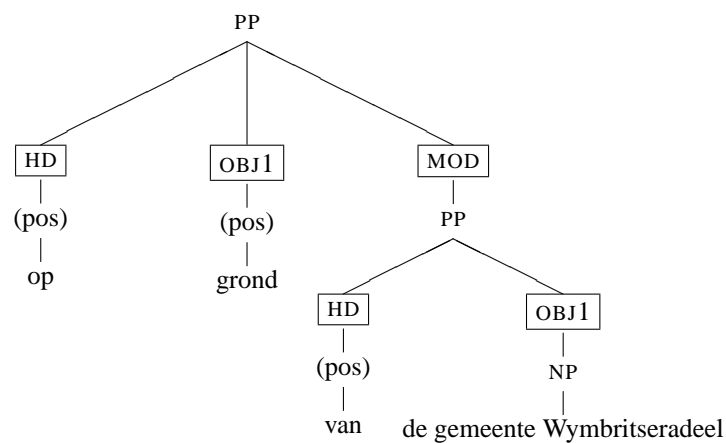
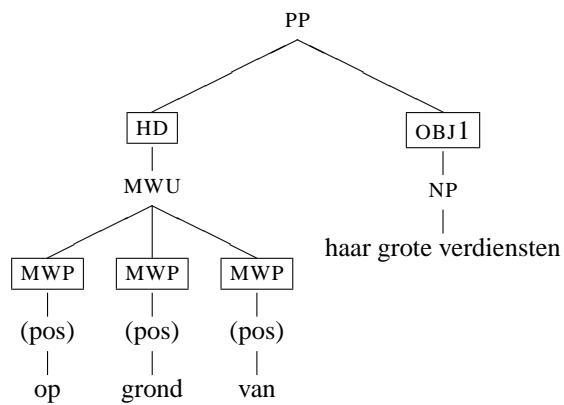
(222) we bouwen hier op grond van de gemeente Wymbritseradeel

(223) het kindje liep aan de hand van zijn moeder

<sup>5</sup>De ANS volgt hieraan toe: “Voorzetseluitdrukkingen zijn (enkele uitzonderingen daargelaten) niet varieerbaar (bijv. door toevoeging van bepalingen). [...] Toch is de categorie van de voorzetseluitdrukkingen op deze manier niet in alle gevallen duidelijk af te bakenen. Niet altijd is een groep die gewoonlijk als voorzetseluitdrukking beschouwd wordt, door een voorzetsel te vervangen (bijv. niet *in antwoord op*, *bij monde van* of *met medeweten van*). Bovendien is in een enkel geval een voorzetseluitdrukking ook wel uitbreidbaar. Zo is naast (*het Concertgebouw*) *onder leiding van*... bijvoorbeeld ook mogelijk: (*het Concertgebouw*) *onder de eminente leiding van*...”

Voorzetseluitdrukkingen komen ook hierin met voorzetsels overeen, dat ze vrijwel altijd vóór de constituent of bijzin staan waar ze bij horen. Slechts in een enkel geval komt achterplaatsing voor. Zo bijv. *ten spijt* in: *Het mooie weer ten spijt zijn we de hele dag binnen gebleven.*”

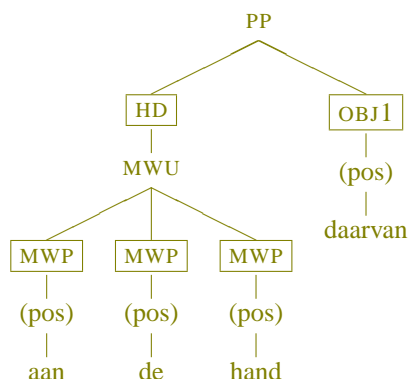
De CGN-annotatie volgt in dezen de traditionele grammatica, in die zin dat de voorzetseluitdrukking niet nader geanalyseerd wordt, maar als MWU als hoofd van een voorzetselconstituent fungeert.



De ANS (530–538) geeft een lange, maar niet uitputtende lijst van voorzetseluitdrukkingen.

#### AANTEKENINGEN

**Vorzetseluitdrukkingen** Je kunt ook nog voornaamwoordelijke bijwoorden krijgen bij voorzetseluitdrukkingen: dat leidt dan tot constructies als *aan de hand daarvan*. Die analyseren we als volgt:



Dus een PP met als HD een MWU *aan de hand* en een OBJ1 *daarvan*.

**Absolute *met*-constructies** In 2.3.3 zijn we ervan uitgegaan dat er alleen absolute *met*-constructies met een LD (*met Cruyff in het doel*) of een PREDC (*met Kitty Courbois als Hedda Gabler*) voorkomen, maar er lijkt meer te kunnen, zoals een tijdsbepaling: *met af en toe een interruptie van Rouvoet*.

***Vroeg in de morgen*** Als je zegt: *vroeg in de morgen* is een PP met een modificeerder, dan zeg je eigenlijk dat het te vergelijken is met *pal voor het huis* of *recht tegenover de rechtbank*. Uit het proefschrift van Van Riemsdijk weten we dat je daarin R-pronomen kunt krijgen:

(224) pal ervoor

(225) er pal voor

(226) waar woont hij recht tegenover?

maar bij *vroeg in de morgen* kan dat niet:

(227) \*vroeg erin

(228) \*er vroeg in

(die constructies zijn wel mogelijk, maar niet met een temporele lezing). Dus, zouden we zeggen, *vroeg in de morgen* heeft een andere structuur. Het enige redelijke alternatief is dan *vroeg* als hoofd en *in de morgen* als PP-modificeerder, en het geheel een PP. (Zie ook de discussie over *'s morgens vroeg* op p. 33).

## 2.4 Het nominale domein

De globale opbouw van de naamwoordelijke constituent (*c*-label van de moederknoop: NP) kan aan de hand van het volgende schema (zie ANS, hoofdstuk 14) inzichtelijk worden gemaakt:

NAAMWOORDELIJKE CONSTITUENT			
Determinator	Voorbepalingen	Kern	Complementen/Nabepalingen

In Figuur 6 geven we de *d*-labels binnen het nominale domein. Over de vraag wat het hoofd is van een naamwoordelijke constituent bestaan verschillende meningen, afhankelijk van de interpretatie van het begrip 'hoofd'. De voorgestelde CGN-annotatie stelt zich in deze kwestie theorie-neutraal op.

<i>d</i> -LABEL	OMSCHRIJVING
HD	naamwoordelijke kern
DET	determinator
PC	voorzetselvoorwerp
VC	verbaal complement
MOD	voor- of nabepaling

Figuur 6: *d*-labels: nominale domeinen

### 2.4.1 De nominale kern

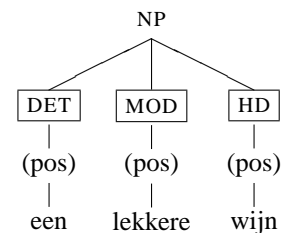
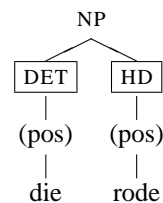
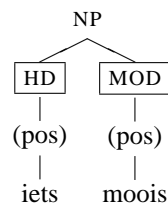
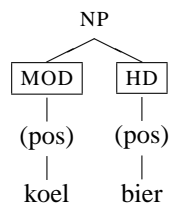
We hanteren het label HD voor de nominale kern van de NP.

(229) een lekkere wijn

(230) koel bier

(231) iets moois

(232) die rode



### 2.4.2 De determinator

De ANS onderscheidt verschillende soorten determinator-elementen, waarvoor een aantal volgorde restricties gelden, die gehoorzamen aan de volgende systematiek:

DETERMINERENDE ELEMENTEN			
POSITIE 1	POSITIE 2	POSITIE 3	POSITIE 4
kwantiteits-aanduidende elementen	determinerende elementen	adjectivische elementen	kwantiteits-aanduidende elementen

De verschillende determinator-elementen hebben in deze visie van rechts naar links een toenemend bereik: [de [laatste [twee [opties]]]]. In het geval van *de twee laatste opties* maakt *laatste* volgens de ANS geen deel uit van de determinator, maar fungeert het als bepaling bij *opties*.

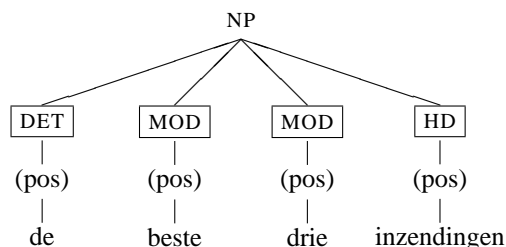
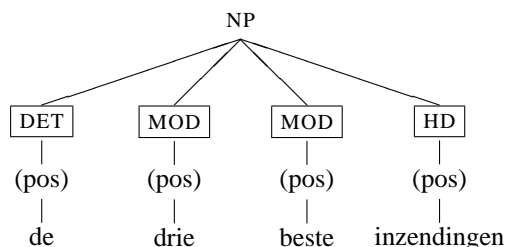
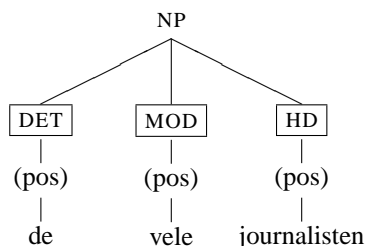
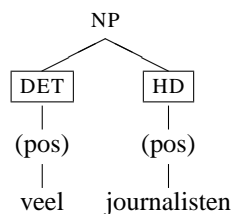
Wat ook de verdiensten zijn van dit schema voor determinatoren, het is incompatibel met de CGN-visie dat een NP direct de categorie van de naamwoordelijke kern projecteert. Wij houden het er dan ook op dat zowel in het geval van *de twee laatste opties* als in het geval van *de laatste twee opties*, *de* de determinator is en dat *laatste* en *twee* beide fungeren als bepaling bij *opties*. Wij baseren ons hierbij op data als de volgende:



(233) veel | vele journalisten belaagden het jonge paar

(234) de vele | \*veel journalisten belaagden het jonge paar

waarin het maar de vraag is of *vele* in (234) als determinator-element kan worden opgevat, gegeven dat het hier, anders dan in (233), verplicht verbogen is. Verder is *de vele journalisten* equivalent met *de journalisten, die vele | met velen waren*, maar *veel | vele journalisten* niet met *journalisten(,) die vele | met velen waren*.



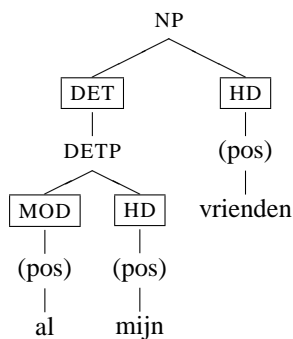
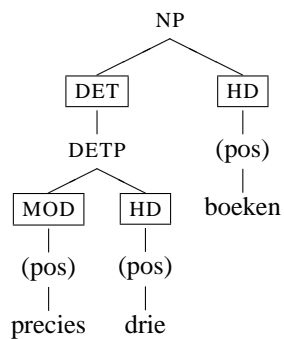
**Complexe determinatoren** Determinatoren kunnen natuurlijk ook gemodificeerd worden:

(235) precies drie boeken

(236) juist drie kolibrie

(237) al mijn vrienden

Daarvoor introduceren we een nieuw label, DETP (determinator-frase)

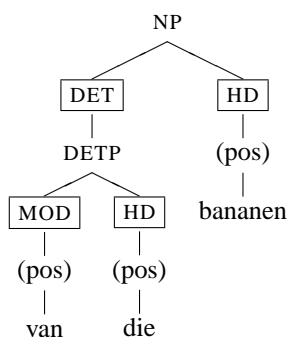


Eenzelfde structuur kennen we toe aan de complexe determinatoren *van die*, *van dat*, *van zulke* enz., waarin het paradigma-loze *van* een modifier-rol krijgt toegekend:

(238) hebt u nog van die bananen voor me?

(239) je kunt daar van dat heerlijke gebak kopen

(240) ze kan daar van zulke diepe dingen over zeggen



De schijnbaar analoge uitdrukking *van alles* in zinnen als

(241) in de Winkel van Sinkel is van alles te koop

wordt niet nader geanalyseerd: die wordt, als vaste uitdrukking, aangemerkt als een MWU.

Woordcombinaties als *een heleboel (flessen bier)* en *een stuk of wat (pandjes aan de Oudegracht)* fungeren eveneens als complexe determinator, en hetzelfde geldt voor de discontinue constructie *een N of Telwoord*, als in *een boek of zes*.

Wat voor structuur moet er aan bijvoorbeeld *een boek of zes* worden toegekend? Tegen analyse als een coördinatie (van ongelijke zinsdelen)

(242) [een boek] [of] [zes]

pleiten de semantiek (de uitdrukking betekent immers ‘ongeveer zes boeken’, en niet ‘een of zes boeken’) en de intonatie. Liever zien we de constructie dan ook als een NP met *boek* als HD en een complexe determinatorfrase DETP *een ... of zes*, waarin *zes* als achteropgeplaatst HD fungeert, en *een ... of* een MWU is die een MOD-rol vervult.

*Wat voor (een)* is een vragende NP-determinator:

(243) wat voor boeken heb je gelezen?

(244) wat voor een auto wil je?

*Wat* en *voor* zijn in de terminologie van Paardekooper allebei paradigmaloos, en over de constructie schrijft hij “*Wat voor 'n* en zelfs *wat voor* zijn [...] niet helemaal eenwoordig, maar het scheelt niet zo heel veel.” Daarom wordt ook deze woordgroep aangemerkt als een MWU.

Zoals andere vraagwoorden staat dit *wat (voor)* niet altijd op de plaats waar we het zouden verwachten, maar we benoemen het desondanks als een determinator van het nominale hoofd en storen ons niet aan de discontinue afhankelijkheden.

*Wat* uit de *wat voor (een)*-constructie kan ook “vanuit deze constructie verplaatst worden” en helemaal vooraan de zin staan, los van de NP waarin het thuishoort (Honcoop 1998):

(245) wat heb je voor boeken gelezen?

(246) wat zeg je dat je voor een auto wilt?

zonder nominale kern voorkomen (net zoals *welke* enz.):

(247) A: ik wil ook graag soep. B: wat voor?

(248) C: ik ga een nieuwe auto kopen. D: wat voor één?

We vinden dezelfde constructie ook met sommige onbepaalde voornaamwoorden, met name *iemand* en *iets* (*wat voor iemand is dat eigenlijk?*) en met genominaliseerde adjectieven (*wat voor moois heb je nu weer gekocht?*). In het laatste geval fungeert *wat voor* als hoofd van de NP *wat voor moois* (vgl. *iets moois* in 2.4.1).

In gesproken taal vinden we ook het vergelijkbare *hoe een*, maar hier lijkt het vraagwoord in de standaardtaal niet naar voren verplaatst te kunnen worden en lijkt het vraagwoord *hoe een* eerder als modificeerder van het (verplichte) adjectief te fungeren dan als NP-determinator:

(249) kijk eens pappa hoe een grote vissen ik gevangen heb

(250) \*kijk eens pappa hoe ik een grote vissen gevangen heb

(251) kijk eens pappa wat een vissen ik gevangen heb

(252) \*kijk eens pappa hoe een vissen ik gevangen heb

Toch beschouwen we ook *hoe een* als een NP-determinator, omdat het de functie overneemt van 'gewone' determinatoren op posities waar een determinator verplicht is:

(253) een grote vis

(254) hoe een grote vis

(255) \*grote vis

Vergelijk verder de combinatie *wat een* die we in exclamatieven aantreffen, ook met meervoudige zelfstandige naamwoorden, die normaliter onverenigbaar zijn met het enkelvoudige lidwoord *een*:

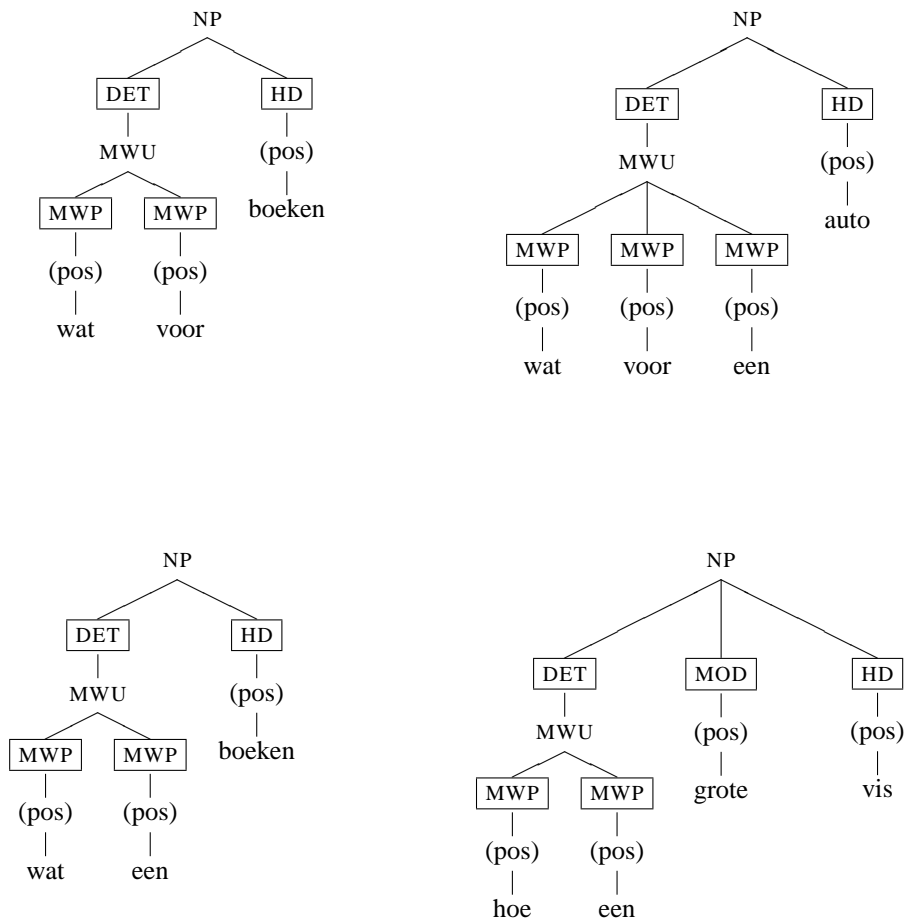
(256) wat een feest was dat zeg!

(257) wat een boeken heeft die man gelezen!

Ook hier is 'extractie' van het WH-element mogelijk:

(258) wat heeft die man een boeken gelezen!

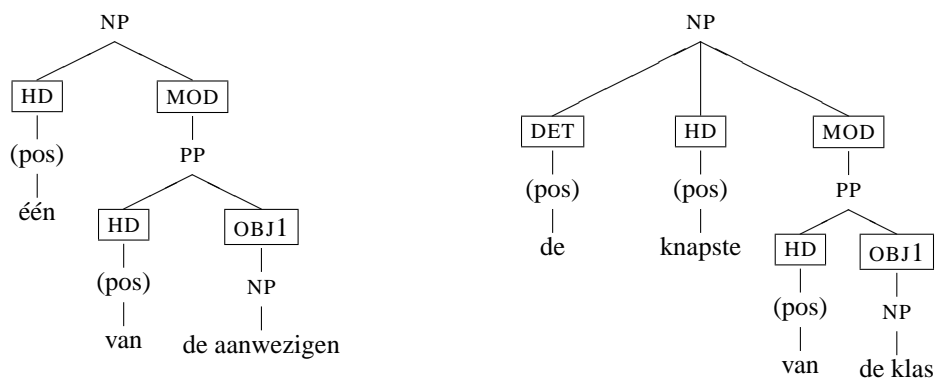
Ook in deze gevallen krijgt de complexe determinator het label MWU.



**De partitieve *van*-constructie** In afwijking van de ANS (p 835) beschouwen we de nominale kern van de NP in een partitieve *van*-frase niet als het hoofd van de volledige partitieve constructie. In een zin als

(259) één van de aanwezigen nam het woord

is er immers getalsovereenkomst tussen *nam* en *één*, niet tussen *nam* en *aanwezigen*.



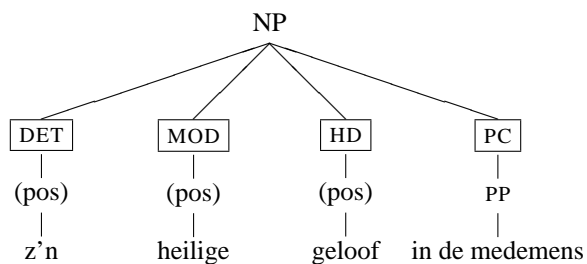
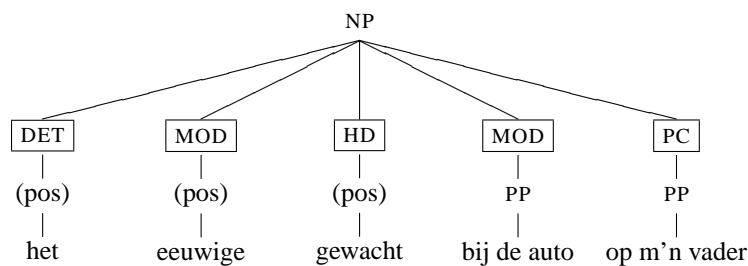
### 2.4.3 Complementen binnen de nominale constituent

De *d*-labels PC, VC en hun syntactische realisaties zijn reeds besproken bij de behandeling van het verbale domein. Binnen het nominale domein betreft het hier dan ook in de regel complementatie bij deverbale nominale kernen (vergelijk sectie 2.4.6).

(260) het eeuwige gewacht bij de auto op m'n vader (PC)

(261) z'n heilige geloof in de medemens (PC)

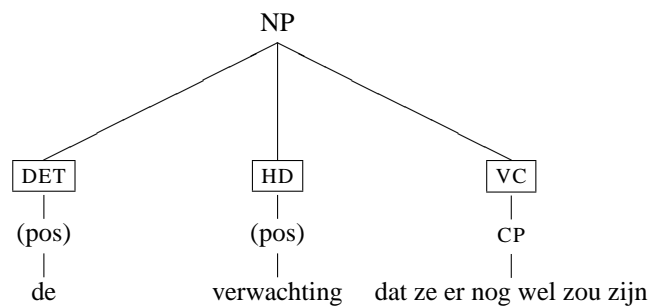
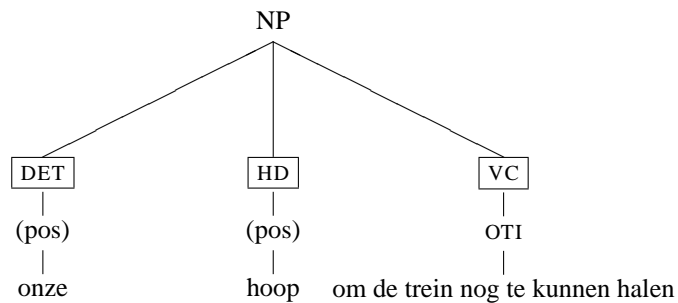
(262) twijfels aan de oprechtheid (PC)



(263) onze hoop (om) de trein nog te kunnen halen (VC)

(264) de verwachting dat ze er nog wel zou zijn (VC)

(265) de vraag of alle moeite voor niets was geweest (VC)



Maar ook niet-deverbale nomina kunnen een complement met een vast voorzetsel krijgen. Ook die worden benoemd als PC:

(266) de kans op winst

(267) het gebrek aan mensen bij de politie

(268) gelegenheid tot bloot zwemmen

Neemt het complement een verbale vorm aan (wat niet altijd kan), dan wordt dat benoemd als VC:

(269) de kans (om) te winnen

(270) gelegenheid (om) bloot te zwemmen

In Vlaanderen zal men hier, afhankelijk van het nomen, soms *van* in plaats van *om* aantreffen (Van Craenenbroeck 2000).

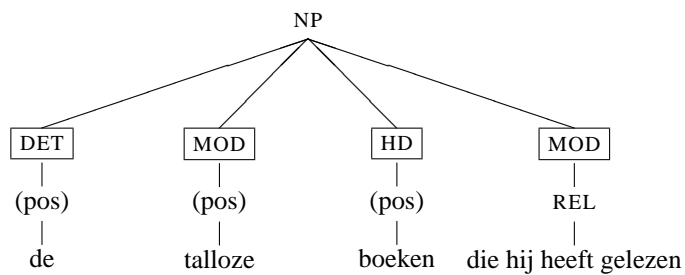
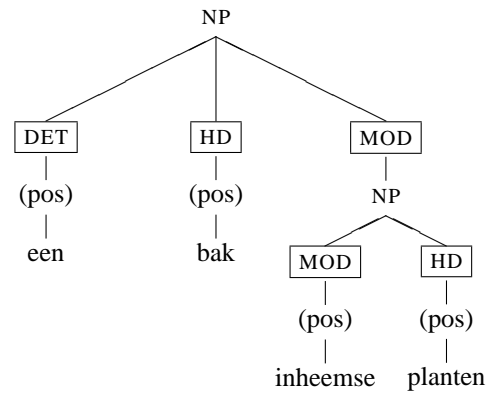
#### 2.4.4 Modifierende elementen

Als modificeerders kunnen tal van syntactische categorieën fungeren. Hieronder een paar voorbeelden.

(271) een bak inheemse planten (NP)

(272) heerlijk helder Heineken (adjectieven)

(273) de talloze boeken die hij heeft gelezen (adjectief, REL)



Met de ANS (844) beschouwen we de geursiveerde zinsdelen in de volgende voorbeelden als modificatoren in het nominale domein:

(274) flink wat

(275) dezelfde ongeveer

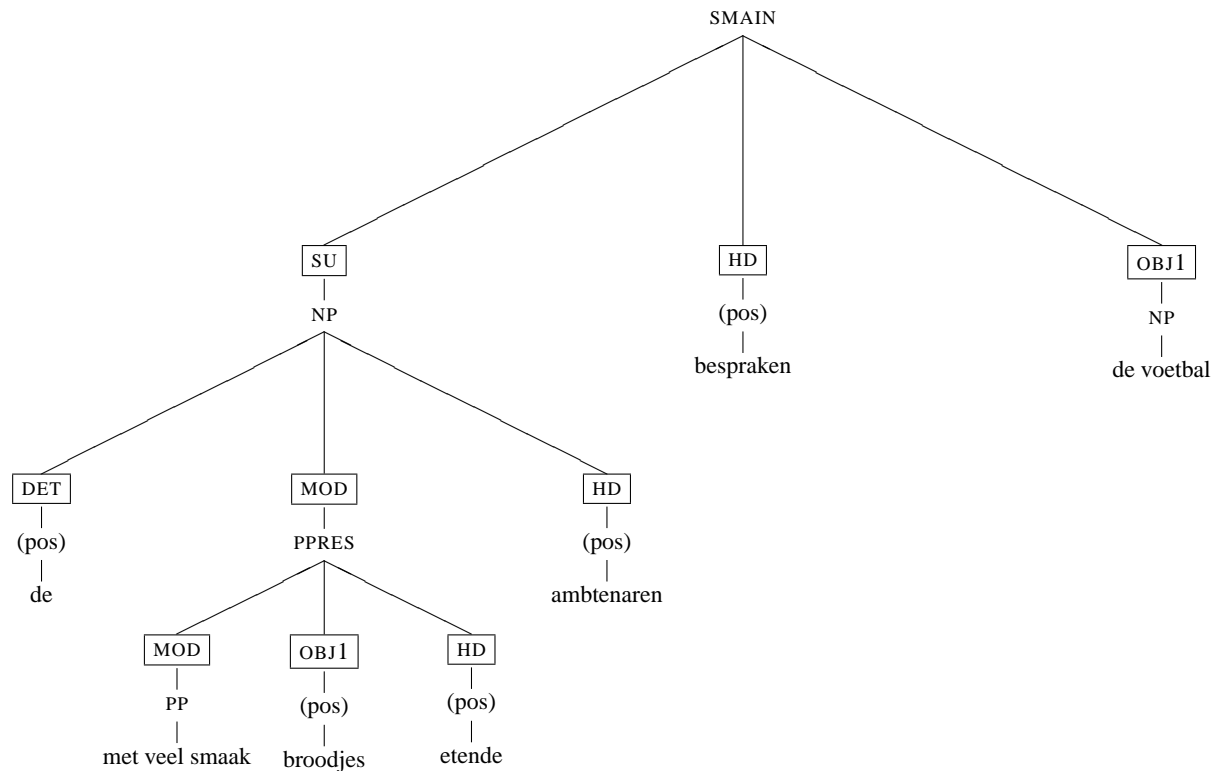
(276) in het geheel niets

(277) om het even wie

(278) al wie (dat hoorde wou bij hem zijn)

De modificerende bepaling tussen lidwoord en nominale kern kan ook verbaal materiaal bevatten, dat dan ook als verbaal materiaal behandeld wordt:

(279) De met veel smaak broodjes etende ambtenaren bespraken de voetbal



In theorie is eindeloze recursie mogelijk, en de praktijk van ambtenaren kan een eind komen:

(280) Het interventiebureau houdt bij de verkoop van het magere-melkpoeder rekening met de datum waarop het product is ingeslagen en slaat van de totale beschikbare hoeveelheid of, in voorkomend geval, van de in de door de marktdeelnemer aangewezen opslagplaats of opslagplaatsen beschikbare hoeveelheid telkens eerst de oudste producten uit.

De ANS (p.844) wijst er ook nog op dat gelijkkluidende adjacente lidwoorden soms kunnen wegvallen: vergelijk het volgende contrast:

(281) het moeilijkst te onderhouden gebouw

(282) de het moeilijkst te beklimmen berg

In Chomskyaanse termen zou men hier kunnen spreken van een PF-effect, wat niet wegneemt dat deze zinnestjes wel een syntactisch boompje moeten krijgen.

Nabepalingen (MOD) bij nominale constituenten kunnen allerlei vormen aannemen; hun aantal is theoretisch onbeperkt (cf. 291):

(283) het feest deze middag (NP)

(284) de plek des onheils (NP)



- (285) wie uwer (NP)
- (286) die herrie hier ook altijd (ADVPS)
- (287) koekjes voor bij de koffie (PP)
- (288) kennis van de natuur (PP)
- (289) jongens zonder verantwoordelijkheidsbesef (PP)
- (290) een dag als vandaag (PP)
- (291) de tocht te voet van Karel naar Rome (PPS)
- (292) een kind om te zoenen (OTI)
- (293) iets te eten (kan ik je niet aanbieden, maar we hebben drank genoeg in huis voor een heel peloton) (TI)
- (294) het dorpje dat wij allemaal zo goed kennen (relative bijzin)
- (295) God almachtig (Adj)
- (296) zaterdag aanstaande (Adj)
- (297) iets heel raars (AP)
- (298) wij drieën (volgens de ANS een telwoord)
- (299) wij drietjes (telwoord)
- (300) wij allemaal (bijwoord)
- (301) wij als redacteurs van de ANS (PP)

Met name voorzetselbepalingen hoeven niet direct naast de nominale kern te staan, maar dat geldt niet alleen daarvoor:

- (302) Jan heeft een boek gekocht over reptielen
- (303) ik vind dat je een kind hebt om door een ringetje te halen

In (300) en (301) fungeren *allemaal* en *als redacteurs van de ANS* als NP-modificeerders. Dit is echter alleen het geval als er sprake is van adjacentie en de NP de eerste positie in een hoofdzin inneemt. In alle andere gevallen fungeren deze constituenten als PREDM op zinsniveau. Er zijn twee redenen waarom het geval adjacentie + NP-op-de-eerste-positie-in-de-hoofdzin in dezen een aparte behandeling krijgt. De eerste is dat de constituent die 'normaal' als PREDM op zinsniveau fungeert in dit geval wel deel moet uitmaken van de NP, omdat er in de hoofdzin immers maar één constituent voor de persoonsvorm kan staan. De tweede reden is dat een dergelijke NP bijvoorbeeld geen gereduceerd pronomen als kern toestaat (*wij* versus *\*we* – vgl. ook *arme jij* versus *\*arme je* en *zij die dat willen* versus *\*ze die dat willen*), een beperking die alleen geldt in (304), maar niet in (305–308).

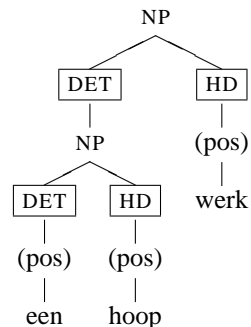
- (304) wij | \*we allemaal gaan naar oma morgen (MOD bij *wij*)
- (305) we gaan allemaal naar oma morgen (PREDM)
- (306) omdat we allemaal naar oma gaan morgen (PREDM)

(307) omdat we morgen naar oma gaan met z'n allen (PREDM)

(308) omdat ze dat als redacteurs van de ANS belangrijk vinden (PREDM)

In een aantal constructies van het type (271) is een zelfstandig naamwoord zo vaak opgetreden als kwantiteitsaanduiding dat de afgesloten combinatie functioneert als een (complexe) determinator:

(309) een hoop werk



Frases als (309) zijn in principe ambigu: de *hoop* kan, als de nominale kern iets is dat zich op laat hopen, ook letterlijk ('berg', 'bult') bedoeld zijn. De context zal in het algemeen helpen desambigueren.

NB: in een aantal noordelijke Nederlandse dialecten (cf. werk van Eric Hoekstra) is *veel* negatief-polair en zijn in positieve zinnen constructies à la *een hoop*, *een berg* de enige mogelijkheid; dat zal wel zijn weerslag hebben in het Nederlands van sprekers van die dialecten.

#### 2.4.5 Bijstellingen enz.

De ANS (846) verstaat onder bijstelling "elke naamwoordelijke constituent die als aanvullende informatie achter een andere naamwoordelijke constituent geplaatst wordt", en suggereert (847) dat het zinvol is "bijstellingen [te] beschouwen als gereduceerde zinnen". Dus een zin als

(310) De schipper, een voorzichtig man, hield zijn schip aan de kade

moeten we volgens de ANS interpreteren als, "de schipper, die een voorzichtig man was/is, hield zijn schip aan de kade". Daaraan wordt nog toegevoegd dat die nominale constituent van verbale origine kan zijn (311) en dat er allerlei zinsdelen in geïncorporeerd kunnen zijn (311, 312):

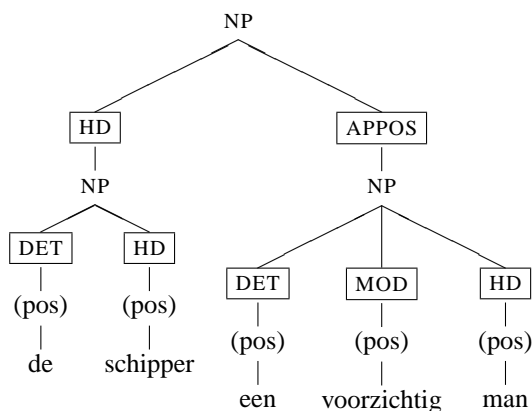
(311) de dagelijkse activiteit van een auteur, stukjes schrijven, wordt niet altijd gewaardeerd

(312) die man daar, tot voor kort gymnasiumleraar hier ter stede, is onlangs tot hoogleraar benoemd

Niet altijd staat de bijstelling onmiddellijk na de bijbehorende zelfstandig-naamwoordgroep:

(313) we hebben Pollini ontmoet, de beste pianist ter wereld

Dick Klein heeft in zijn dissertatie betoogd dat het geen goed idee is bijstellingen te beschouwen als gereduceerde zinnen. Arie Sturm stelt in zijn proefschrift voor ‘klassieke’ bijstellingen van het type (313) te analyseren als asyndetische coördinaties, en wel verplicht asyndetisch (*We hebben Pollini ontmoet en de beste pianist ter wereld* is niet ongrammaticaal, maar betekent iets anders). Semantisch is dat in dit geval niet onredelijk, omdat het gaat om twee predicaties bij een referent. Syntactisch is die analyse echter bevreemdend, want verder zijn coördinaties nooit verplicht asyndetisch. Voor deze ‘klassieke’ bijstellingen gebruiken we het label APPOS.



Het Nederlands kent nog een aantal constructies waarbij een nomen een soort nabepaling heeft, een naam of een telwoord. Denk daarbij aan de volgende types (naar Paardekooper):

(314) hoofdstuk vier

(315) Karel één

(316) de maand maart

(317) de stad Groningen

(318) de afstand Mechelen-Brussel

(319) oom Wim

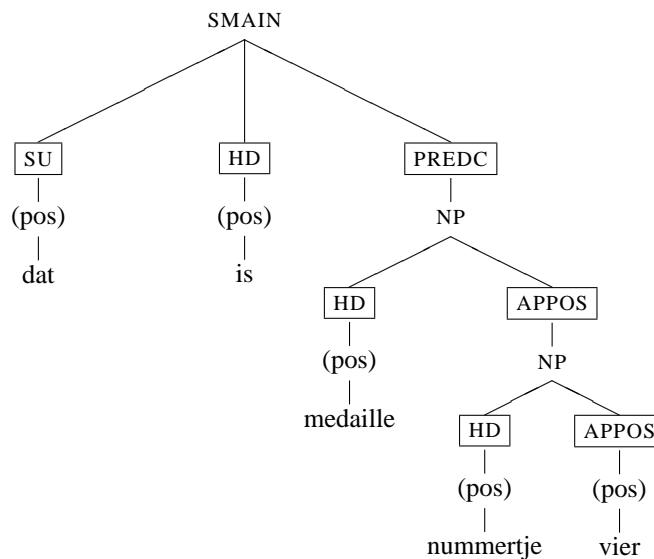
(320) het voorstel-Kollewijn<sup>6</sup>

Traditioneel (Den Hertog) heet een aantal van deze constructies “bijstelling”, maar zoals Paardekooper schrijft “Vanuit het ABN is er [...] geen mogelijkheid om die types onder één formele noemer te brengen”.

Paardekooper laat zien dat de constructies in (314–320) syntactisch allemaal verschillend zijn. De analyse van Sturm van bijstelling als asyndetische coördinatie is niet over te dragen op de constructies in (314–320): het is niet zo dat in *het voorstel-Kollewijn voorstel* en *Kollewijn* van een en dezelfde referent worden geprediceerd. De analyse van de ANS, als gereduceerde (bijvoeglijke) bijzinnen, is ook contra-intuïtief: *we lezen hoofdstuk vier* – *we lezen het hoofdstuk dat vier is?* *de afstand Mechelen-Brussel kan op een dag te voet worden afgelegd* – *de afstand die Mechelen-Brussel is kan op een dag te voet worden afgelegd?*, tenzij we ad-hoc analyses aannemen als *we lezen het hoofdstuk dat het vierde van het boek is en de afstand die Mechelen en Brussel scheidt kan op een dag te voet worden afgelegd*.

<sup>6</sup>In theorie krijgt *voorstel-Kollewijn* in CGN één POS-tag, dus die constructie hoeft niet syntactisch geanalyseerd te worden.

Omdat geen taalkundige vooralsnog een oplossing voor alle soorten gevallen weet, laat staan dat er zelfs maar het begin van een consensus is, stellen we ons agnostisch op. Ook voor al dit soort gevallen gebruiken we het (recursieve) label APPOS.



NB: We onderscheiden dus twee typen bijstellingen: enerzijds de 'klassieke' (zie 310–313), met een duidelijke 'komma-intonatie', waarvan het hoofd een NP is (en dus niet-lexicaal!) en anderzijds het type nabepalingen dat hierboven besproken is (314–320), met een lexicaal hoofd.

## 2.4.6 Deverbale nominalisaties

In de sfeer van de deverbale nominalisatie kunnen we een aantal gevallen onderscheiden:

### 1. morfologische derivatie

afleiding op *-ing* (*vernedering*, *verdediging*)

afleiding op *-atie* (*diskwalificatie*, *hospitalisatie*)

afleiding met *ge-* (*getreuzel*, *gedachte*)

conversie van werkwoordelijke stam (*hoop*, *vlucht*)

### 2. de constructie *het* + infinitief, al dan niet met frasale conversie (zie (a) en (b) hieronder)

### 3. frasale conversie van een infinitiefgroep

Voorbeelden:

(321) de verdediging van de stad (type 1)

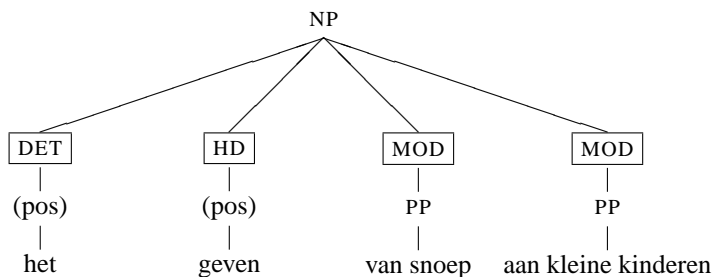
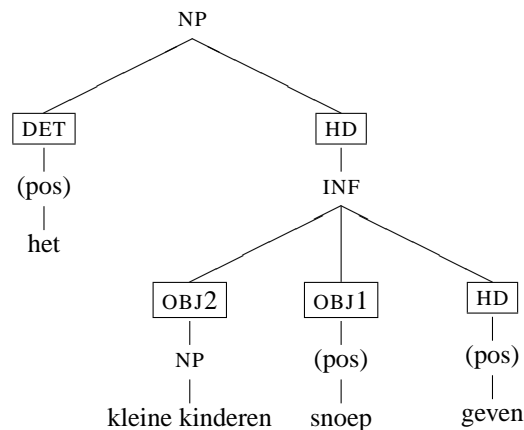
(322) het kleine kinderen snoep geven (wordt afgeraden) (type 2a)

(323) het geven van snoep aan kleine kinderen (type 2b)

(324) cognac aanbieden aan tante Fien (vindt hij zonde)

(type 3)

In de gevallen waar het complementatiepatroon van de infinitief aan de verbale bouwprincipes beantwoordt (type 2a en 3 in de voorbeelden) labelen we de moederknoop van dit verbale domein INF (merk op dat er in dit geval dus sprake is van een niet-lexicaal hoofd van de NP); in de gevallen waar de nominale realisering van het complementpatroon optreedt, hebben we te maken met een nominale groep met (prepositionele) postmodificerende.



Wat betreft de NP-complementen van een werkwoord (SU, OBJ1, OBJ2) merken we op dat deze als PP worden gerealiseerd in het nominale domein. Ze zijn dan altijd weglaatbaar, en worden als MOD gelabeld. We kunnen overwegen in een latere fase naast het default label MOD enkele meer specifieke labels te introduceren, die de 'onderliggende' verbale complementsrol duidelijk maken. De 'casus'-aanduidende hoofden (*van, aan, door*) kunnen daarbij een sturende rol spelen.

(325) het getreuzel van Piet(326) het zingen van het volkslied(327) de gift door Han van een tientje aan Jantje Beton

#### 2.4.7 Vocationen enz.

Vocationen of aanspreekvormen kunnen verschillende vormen aannemen:

- (328) ha, Kees
- (329) bakker, mag ik een brood?
- (330) die rommel, Jan, die moet je toch eens opruimen
- (331) goedemorgen Nijmegen
- (332) dat is het juist, jongen
- (333) geef maar hier, liefste
- (334) maar honneponnetje, we hoeven toch niks?
- (335) kijk uit, hufter

In voorbeelden als deze zijn ze nooit echt deel van de zin: we benoemen ze dan ook als TAG.

Namen kunnen in dit soort gebruik voorafgegaan worden door iets dat lijkt op een aanwijzend voornaamwoord:<sup>7</sup>

- (336) ha, die Kees
- (337) ha die bakker, mag ik een brood?

In de standaardtaal is alleen *die* mogelijk, *dat* en *deze* zijn uitgesloten. Toch ligt analyse van *die Kees* als NP met *die* als DET en *Kees* als HD het meest voor de hand. Het geheel is weer een TAG.

Ook sommige ‘volle’ NPs (met modificatoren en determinatoren) kunnen als aanspreekvorm gebruikt worden; de determinator, voor zover aanwezig, zal meestal *mijn* zijn:

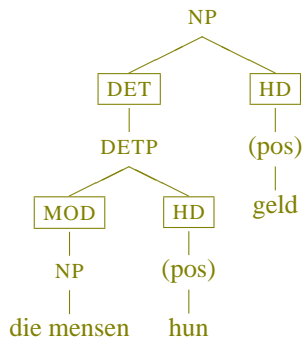
- (338) kom maar hier, m'n jongen
- (339) ik zal altijd bij je blijven, mijn allerliefste lieveling

Ook deze vocatieven benoemen we als TAG.

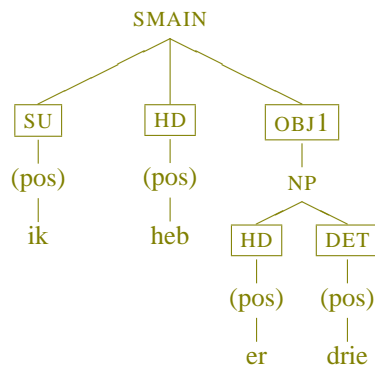
## AANTEKENINGEN

**Jan zijn fiets** *Jan zijn* vormt een complexe determinator met *zijn* als HD en *Jan* als MOD. Zo ook met *die mensen hun geld enz.*

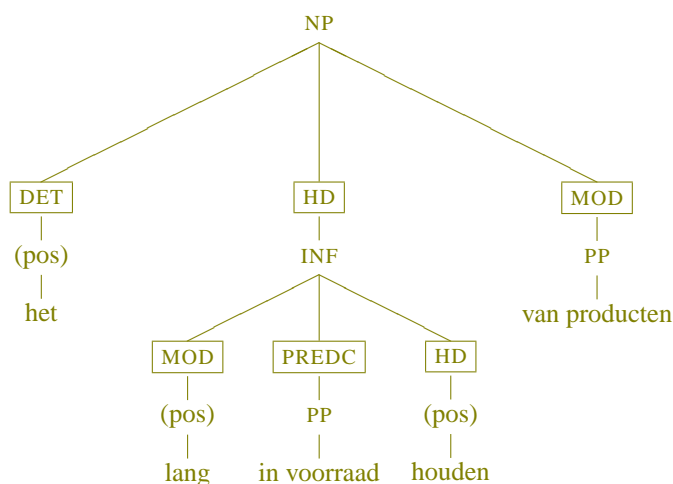
<sup>7</sup>Paardekooper merkt op (3.3.1): “Een zeker verband is per se vereist; *die Jan* komt bijna nooit totaal zelfstandig voor. Het is m.i. wel zelfstandiger als de aanspreking en onderscheidt zich bovendien daarvan doordat vervanging door een pers. vn hier uitg. is: *ha, jij* (uitg.)”. Maar ondertussen is er weer niets mis met *hè, jij* (ook altijd met je flauwe grapjes) en *hè, jij* (met je roze broekje).



**Kwantitatief *er*** We kiezen ervoor om geen nieuw label te introduceren, maar om *er* + telwoord te analyseren als een gewone NP:



**Hybride deverbale nominalisering** Soms moet je wel een hybride structuur bouwen, bijvoorbeeld bij *het lang in voorraad houden van producten* (sfn000920,72): *lang in voorraad* vraagt om een verbale structuur, *van producten* om een nominale:



### **Problemen met, problemen ermee**

(340) we hebben problemen met de koeling

(341) daar heb ik problemen mee

De voorzetselbepaling is een PC.

**Van alles en nog wat** In 2.4.2 zagen we al dat *van alles* in zinnen als *in de Winkel van Sinkel is van.alles te koop* niet nader geanalyseerd wordt: die woordgroep wordt, als vaste uitdrukking, aangemerkt als een MWU. Datzelfde geldt voor de langere variant *van alles en nog wat*.

**Weet ik veel** Frases als *weet ik (veel)* en *ik weet niet* in zinnen als

(342) weet ik veel

(343) weet ik (veel) wat ze allemaal doet

(344) hij heeft ik weet niet hoeveel boeken over katten

krijgen het label MWU. Ze fungeren als MOD bij het woord waar ze bij horen (in de voorbeelden respectievelijk bij *wat* en bij *hoeveel*).

**Vragend voornaamwoord + dan ook** Een constructie als *wie/wat/hoer/waar/welke ... dan ook* krijgt de volgende behandeling: *dan ook* is een ADVP: PRT PRT, die fungeert als MOD bij het vragend voornaamwoord.

**Projectie van voornaamwoorden** Sommige voornaamwoorden zijn nogal 'multi-categoriaal': ze kunnen fungeren als hoofd van verschillende soorten constituenten. Een voorbeeld is *veel*:

(345) heel veel mensen (DETP)

(346) hij eet nogal veel (NP)

(347) je zeurt te veel (AP)



*Iets van, rond de, meer dan* + telwoord Voorbeelden:

(348) iets van | rond de vijfhonderd kinderen doen d'raan mee

(349) meer dan één deelnemer(\*s) had(\*den) alle vragen goed

In (349) is goed te zien dat *meer dan* een modificeerderrol vervult, aangezien er agreement is tussen *één* en het zelfstandig naamwoord, niet tussen *meer* en het zelfstandig naamwoord, en ook het werkwoord de enkelvoudvorm heeft. Zowel *meer dan* als *iets van* en *rond de* modificeren dus het telwoord. *Iets van* en *rond de* krijgen het categorielabel MWU, *meer dan* krijgt ADVP: HD + OBCOMP.

## 2.5 Restgevallen

Er zijn constructietypes waar de syntaxis niet veel mee kan, omdat niet uit te maken is hoe de dependentieverhoudingen liggen. Voorbeelden zijn getallen, data, tijdstippen en complexe eigennamen.

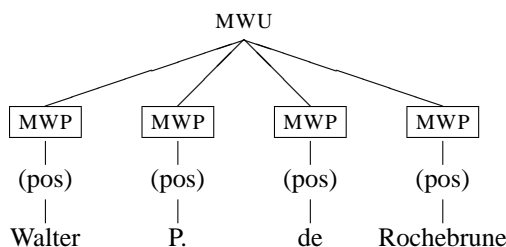
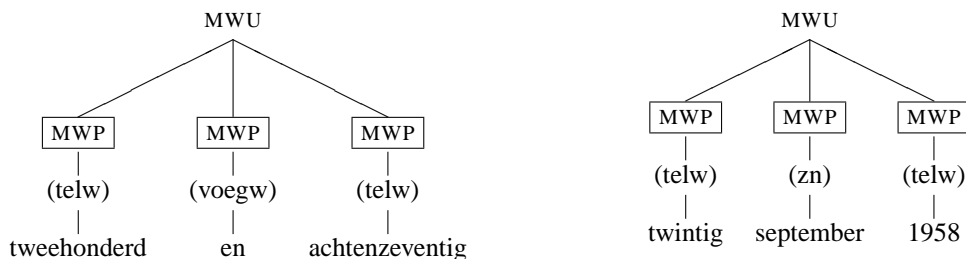
(350) tweehonderd (en) achtenzeventig

(351) 20 september 1958

(352) half drie

(353) Walter P. de Rochebrune

We kiezen in dit geval voorlopig voor een pragmatische, agnostische aanpak: we labelen de knoop MWU:



### AANTEKENINGEN

**Een uur of elf** Als tijdstip: de MWU *een ... of* modificeert de MWU *elf uur*; als tijdsduur: NP met *uur* als HD en een complexe determinatorfrase *DETP een ... of elf*, waarin *elf* als achteropgeplaatst HD fungeert en *een ... of* een MWU is die een MOD-rol vervult (zie ook 2.4.2).

### 3 Niet-lokale dependenties

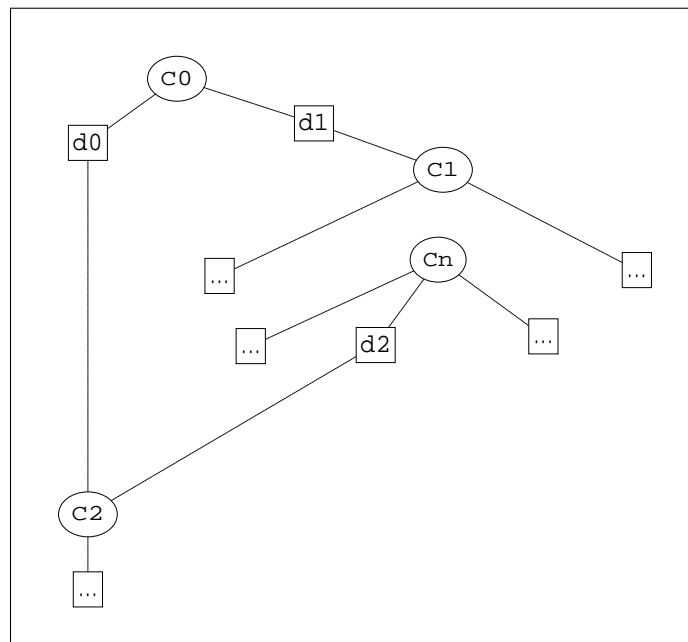
In deze sectie behandelen we de *d*- en *c*-labels in niet-lokale dependenties, zoals die zich voordoen bij:

- relatieve zinnen (met expliciet dan wel met ingesloten antecedent, cf. ANS, 329 ev);
- constituentvragen, op hoofdzins- en op bijzinsniveau

Wat betreft de uitgangspunten van de CGN-annotatie zien we in deze gevallen een spanning:

- de elementen die deze configuraties inleiden (constituenten met een vragend of betrekkelijk voornaamwoord) zijn bepalend voor het *c*-label van de moederknoop (relatieve zin *vs* vraagzin, ...); dat bestempelt ze tot hoofden;
- tegelijk willen we aangeven welke rol deze inleidende elementen vervullen in een dependentiedomein ergens in het vervolgstuk; het relevante dependentiedomein kan arbitrair diep in dat vervolgedeel zijn ingebed.

We kunnen deze spanning oplossen door gebruik te maken van het feit dat de definitie van de CGN-annotatiestructuren het mogelijk maakt om een element van verschillende moederknopen afhankelijk te maken (en dus meerdere dependentierollen toe te kennen).<sup>8</sup> We geven de situatie schematisch weer in Figuur 7.  $C_2$  is de ‘inleider’ van de niet-lokale dependentieconfiguratie. De primaire rol van  $C_2$  is  $d_0$ : de rol van hoofd, die bepalend is voor de moederknoop  $C_0$ . Het vervolgedeel  $C_1$  vervult de rol van dependent  $d_1$  ten opzichte van  $d_0$ . Ingebed in  $C_1$  bevindt zich een domein  $C_n$ ;  $C_2$  vervult een secundaire dependentiefunctie  $d_2$  in  $C_n$ .



Figuur 7: Niet-lokale dependenties: schematisch.

In Figuur 8 geven we een overzicht van de gebruikte labels.

Opmerking: relatieve zinnen en constituentvragen worden zo veel mogelijk parallel behandeld. Ze hebben dan ook, in de lijn van de filosofie van de CGN-annotatie, een hoofd. Dat leidt tot de paradoxale notie “hoofd van een hoofdloze relatieve zin”. Daarvoor kan men evenwel net zo goed lezen: “hoofd van een betrekkelijke bijzin met ingesloten antecedent”.

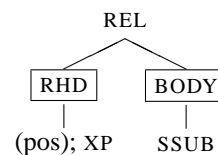
<sup>8</sup>In de terminologie van de ANNOTATE software betreft het hier *secondary edges*.

c-LABEL	OMSCHRIJVING
REL	relatieve zin
WHREL	'hoofdloze' relatieve zin
WHQ	constituentvraag: hoofdzin
WHSUB	constituentvraag: ondergeschikte zin
d-LABEL	OMSCHRIJVING
WHD	hoofd van WHQ, WHSUB
RHD	hoofd van ('hoofdloze') relatieve zin
BODY	romp

Figuur 8: Labels voor niet-lokale dependenties

### 3.1 Relatieve zinnen

De primaire dependentierelaties worden in de betreffende bijzin als volgt gerealiseerd. Het inleidend element is ofwel een betrekkelijk voornaamwoord op POS-niveau, ofwel een constituent XP die een betrekkelijk voornaamwoord bevat. Het daarop volgende rompedeel is een SSUB. De volledige relatieve zin REL vervult de functie van MOD in het nominale domein. We maken geen onderscheid tussen beperkende en uitbreidende relatieve zinnen.

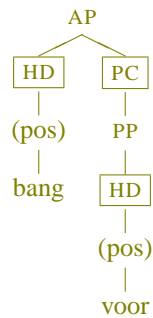


Hieronder enkele voorbeelden van secundaire dependentierelaties die het RHD-element onderhoudt binnen SSUB.

- (354) (de jongen) die hem wil plagen (SU bij 'plagen')
- (355) (de jongen) die hij wil plagen (OBJ1 bij 'plagen')
- (356) (de jongen) op wie hij zeker dacht te kunnen rekenen (PC bij 'rekenen')
- (357) (de muziek) waar hij het meest van houdt (OBJ1 bij 'van')

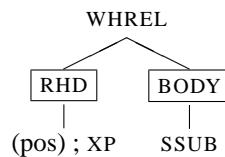
#### AANTEKENINGEN

*Niks om bang voor te zijn* OTI als MOD bij *niks*; *voor* een PP als PC bij *bang*, met een secondary edge (OBJ1) van *niks* naar de PP.



### 3.2 'Hoofdloze' relatieve zin

Anders dan de relatieve zin, die een modifierende functie heeft, vervult de betreffende bijzin met ingesloten antecedent ('hoofdloze' relatieve zin) de functie van een naamwoordelijke constituent.



De voorbeelden hieronder illustreren mogelijke secundaire dependentierelaties binnen SSUB.

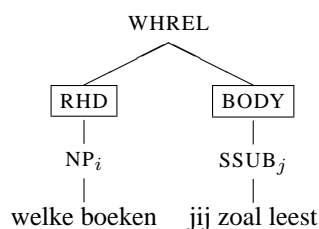
- (358) wie het weet, mag het zeggen (SU bij 'weet')
- (359) wat je zegt, ben je zelf (OBJ1 bij 'zegt')
- (360) waar jij mee bezig bent, dat laat me koud. (OBJ1 bij 'mee')

De RHD-elementen in een WHREL zijn op POS-niveau verankerd: een grotere XP-constituent waarin het betreffende voor-naamwoord is ingebed, is als inleider van WHREL uitgesloten. In de voorbeelden die volgen maakt het voorzetsel dus geen deel uit van WHREL; WHREL speelt de rol van OBJ1 bij dat voorzetsel.

- (361) nog een aanvulling bij wat ik je gisteren vertelde (OBJ1 bij 'bij')
- (362) geef het aan wie het verdiend heeft (OBJ1 bij 'aan')

Wat wel kan als XP-inleider, is een NP met DET 'welk(e)'.

- (363) welke boeken jij zoal leest, laat me koud (OBJ1 bij 'leest')



(*i, j*) boog: OBJ1 Het verschil tussen WHSUB en WHREL in termen van de

traditionele grammatica: WHSUB is voor afhankelijke vraagzinnen; een WHSUB hangt dan ook af van een werkwoord, bijv. *weten* in:

(364) ik weet niet wat hem bezielt

WHREL is voor relatieve bijzinnen met ingesloten antecedent, hetgeen inhoudt dat het relativum vervangen moet kunnen worden door een aanwijzend voornaamwoord + relativum. In het bovenstaande zinnetje kan dat niet:

(365) \*ik weet niet datgene wat hem bezielt

In de zin

(366) ik geloof nog steeds in wat ik gezegd heb

kan dat wel:

(367) ik geloof nog steeds in datgene wat ik gezegd heb

Een WHREL fungeert als, en kan dus vervangen worden door, een ‘echte’ NP:

(368) ik geloof nog steeds in jou | de quantumtheorie

Een WHSUB kan alleen vervangen worden door een voornaamwoord in de derde persoon, oftewel een ‘zinsvervangende’ NP, vergelijk

(369) ik weet het niet

(370) ik weet \*jou | \*de quantumtheorie niet

### 3.3 Constituentvraag: hoofd- en bijzin

Constituentvragen op het hoofdzins- en bijzinsniveau hebben dezelfde inleiders: vragende voornaamwoorden, dan wel grotere XP-constituenten waarin deze voornaamwoorden (of de DET ‘welk(e)’) zijn ingebed. De twee typen constituentvragen verschillen in de syntactische realisering van het rompdeel: een SV1 in het geval van WHQ, en een SSUB in het geval van WHSUB.



Zoals in het geval van de relatieve zinnen, vervult het WHD-element een secundaire dependentiefunctie binnen de BODY-component. We geven hieronder enkele voorbeelden.

- (371) WHQ: wat wil hij zeggen? (OBJ1 bij ‘zeggen’)
- (372) WHSUB: (ik weet) wat hij wil zeggen (OBJ1 bij ‘zeggen’)
- (373) WHQ: hoeveel denk je dat hij weegt? (ME bij ‘weegt’)
- (374) WHSUB: (ik vroeg) hoeveel je denkt dat hij weegt (ME bij ‘weegt’)

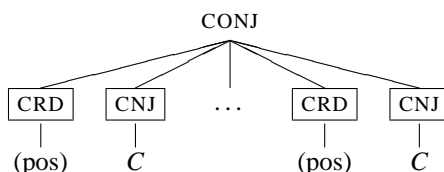
## 4 Nevenschikking

<i>d</i> -LABEL	OMSCHRIJVING
CRD	nevenschikker (voegwoord of reeksvormer)
CNJ	lid van nevenschikking
LP	elk der delen van een LIST
<i>c</i> -LABEL	OMSCHRIJVING
CONJ	nevenschikking (polymorf)
LIST	asyndetische conjunctie

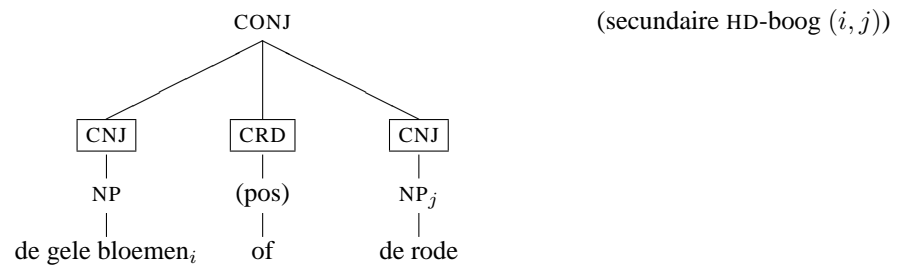
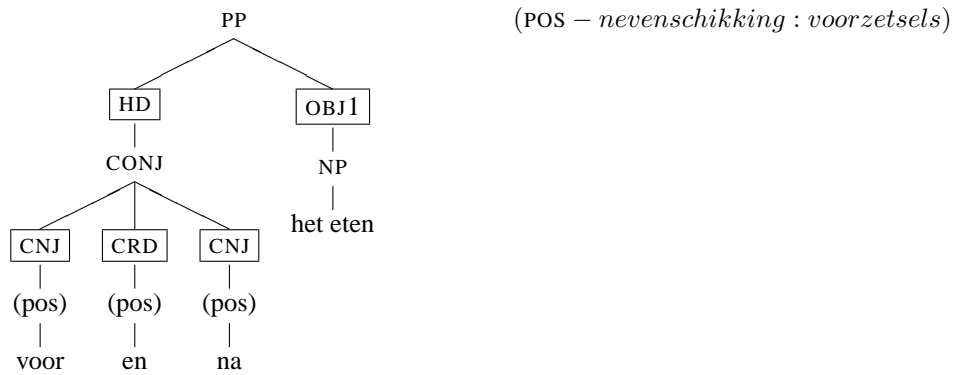
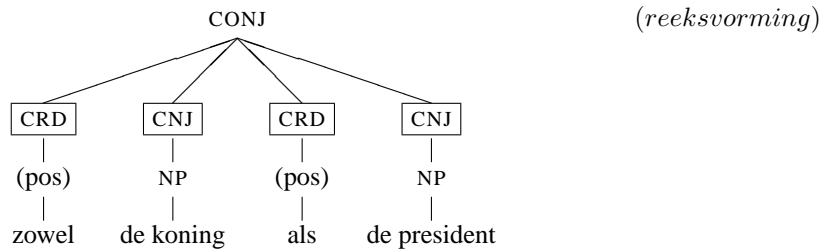
Figuur 9: Labels voor nevenschikking

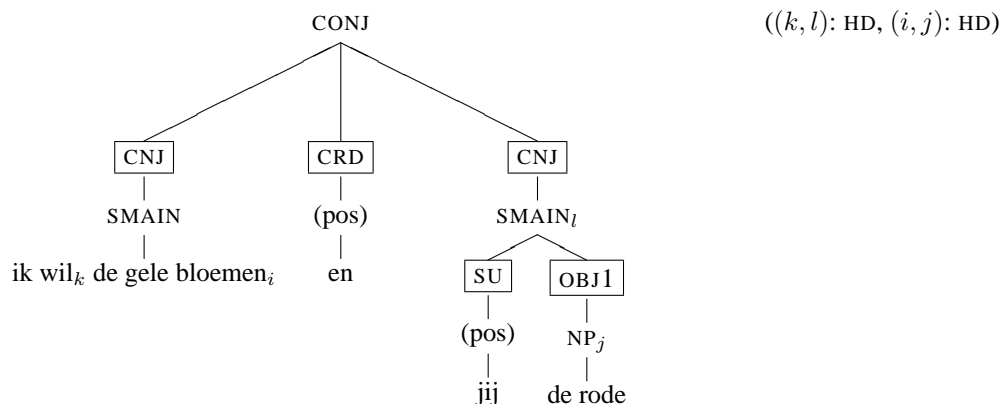
De *d*- en *c*-labels voor de annotatie van nevenschikkingen zijn te vinden in Figuur 9. Schematisch ziet het structurele patroon voor een nevenschikking eruit als in (375). We lichten de componenten van dit patroon toe (onder verwijzing naar ANS, hoofdstukken 24–26) en geven een aantal voorbeelden.

(375)



- In een nevenschikking worden twee of meer ‘gelijkwaardige taalelementen’ – de leden van de nevenschikking, met *d*-label CNJ – naast elkaar geplaatst, verbonden door nevenschikkende voegwoorden of reeksvormers met *d*-label CRD.
- Voor *C*, de categorie van de leden, geldt  $C \in \text{CAT}$ . Met andere woorden: we kunnen zowel op het POS-niveau als op het niveau van de frasale labels nevenschikkingen vormen.
- De moederknoop krijgt een *polymorf* label CONJ. De categoriale identiteit van CONJ wordt impliciet gegeven door de categorie *C* van de leden van de nevenschikking.
- Als het aantal verbindende elementen (CRD) kleiner is dan het aantal verbonden elementen (CNJ), spreken we van nevenschikkende voegwoorden; als het aantal verbindende elementen even groot is als het aantal verbonden elementen, van reeksvormers.
- In het geval van nevenschikking met samentrekking (ANS, hoofdstuk 27) vullen we de restantleden aan door middel van secundaire dependentierelaties met betrekking tot de volledige leden.



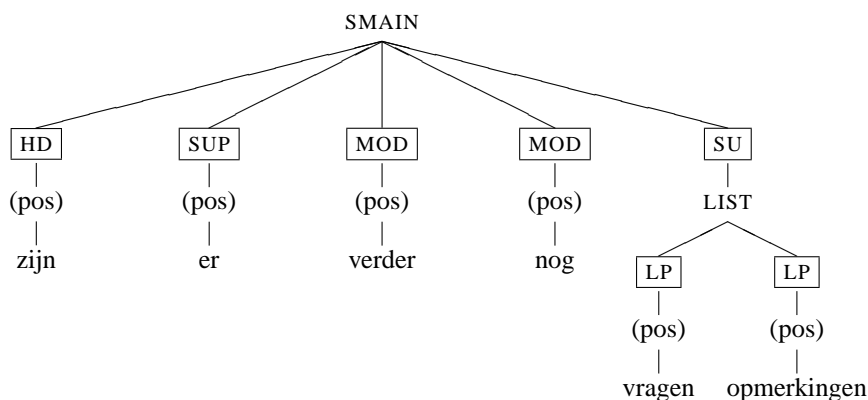


## 4.1 Asyndetische nevenschikking

Asyndetische nevenschikkingen, zoals

(376) zijn er verder nog uh vragen, opmerkingen?

passen niet goed in de hoofd-gebaseerde aanpak van de CGN-annotatie; het postuleren van een verzwegen, onderliggende coördinator *en* of *of* ligt niet in onze lijn. Daarom roepen we voor deze structuren twee nieuwe labels in het leven, een moederlabel LIST, met als dochters LP (voor 'list-part'). Het gebruik van deze labels wordt geïllustreerd in het volgende voorbeeld.



### AANTEKENINGEN

*En zo, of zo* Beide krijgen het label MWU en fungeren als MOD:

(377) ik ga m'n eigen bedrijfje beginnen of zo

(op INF-niveau)



- (378) je moet dan wel wat regen en zo op de koop toe nemen (op NP-niveau)
- (379) 't liep allemaal heel slecht en zo (op zinsniveau)
- (380) die was pas getrouwd of die zou gaan trouwen of zo (op conjunctieniveau)

## 5 Discourse-structuren

De syntactische annotatie binnen CGN is gericht op het aanbrengen van zinsinterne dependentierelaties. Toch worden we ook nu al geconfronteerd met dependenties die behoren tot het domein van de discourse. Het betreft enerzijds uitbreidingen van de kernzin, anderzijds asyndetische constructies waarbinnen de samenhang alleen in een wijder discourse-verband vast te stellen valt. Figuur 5 geeft een overzicht van de *d*-labels en *c*-labels die in deze sectie aan bod komen.

<i>d</i> -LABEL	OMSCHRIJVING
NUCL	kernzin
DLINK	discourse-link
SAT	satelliet: aan- of uitloop
TAG	aanhangsel, tussenvoegsel
DP	discourse-part
<i>c</i> -LABEL	OMSCHRIJVING
DU	discourse-unit

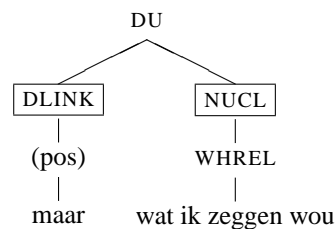
Figuur 10: Labels voor discourse-structuren

### 5.1 Discourse-markeerders

In samenhangende tekst komt het vaak voor dat een spreker een zin begint of eindigt met iets dat op een voegwoord lijkt, maar geen syntactisch verband heeft:

- (381) maar wat ik zeggen wou ...
- (382) dus je bent zeker wel blij dat je niet naar Brussel bent gegaan?
- (383) is dat nou gewoon alleen een broek of ...

Deze elementen krijgen het *d*-label DLINK. De kernzin krijgt het *d*-label NUCL. De moederknoop krijgt het *c*-label DU (discourse eenheid).



Merk op dat woorden als *dus* verschillende analyses krijgen, afhankelijk van de zinspositie waarin ze voorkomen:

- (384) Jan gaat *dus* naar huis (na de persoonsvorm: MOD)

(385) dus gaat Jan naar huis

(eerste zinspositie: MOD)

(386) dus Jan gaat naar huis

(buiten zinsverband: DLINK)

## 5.2 Uitbreidingen van de zin

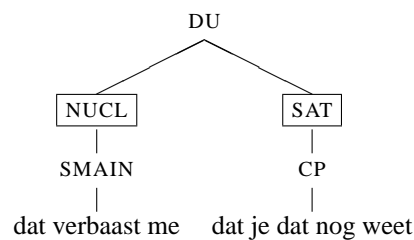
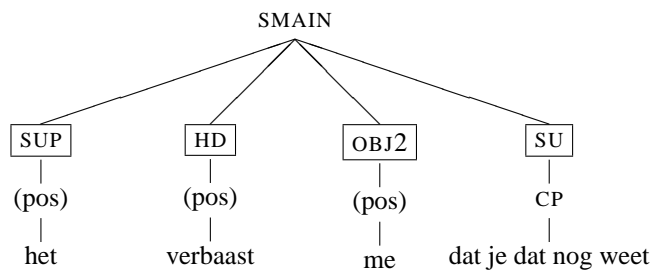
Onder de bedoelde uitbreidingen vallen ‘aanloop’- en ‘uitloop’-elementen die een anaforische relatie onderhouden met verwijswaarden in de kernzin. Zie ANS, §21.8 en 21.9. De aanloop/uitloop-constructies kunnen een kenmerkend intonatiepatroon hebben (‘komma-intonatie’) waardoor ze van extrapositie te onderscheiden zijn: vergelijk (387) en (388). Maar ook zonder die intonatie-informatie is (388) als NUCL+SAT te identificeren: immers, het element *dat* in de kernzin kwalificeert niet als dummy-SUP.

(387) het verbaast me dat je dat nog weet

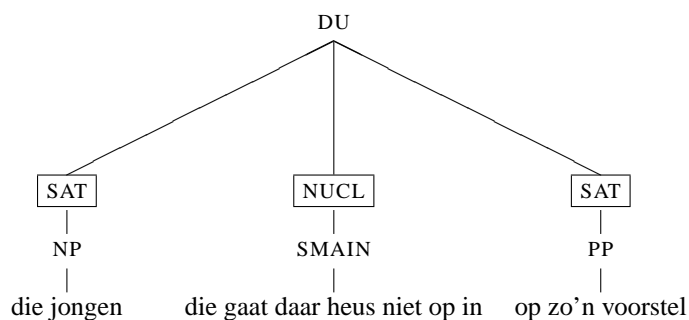
(SUP+ SU)

(388) dat verbaast me, dat je dat nog weet

(NUCL+ SAT)



We gebruiken het *d*-label SAT (satelliet) zowel voor de aanloop- als voor de uitloopelementen.



Ook *als-dan*-constructies worden op deze manier behandeld: vergelijk (389) en (390). *Dan* fungeert hier als het verwijswoord in de kernzin waar de *als*-zin een relatie mee onderhoudt.

(389) als je wint heb je vrienden (eerste zinspositie: MOD)

(390) als je wint dan heb je vrienden (buiten zinsverband: SAT)

Een tweede type uitbreiding van een NUCL-component wordt gevormd door een heterogene groep van elementen die, op een andere manier dan de hierboven beschreven DLINK en SAT, buiten de kernzin staan. We geven deze uitbreidingen het *d*-label TAG. Een eerste groep van TAG-uitbreidingen zijn elementen die op POS-niveau als tussenwerpsels zijn benoemd, dan wel die een tussenwerpsel-achtige functie vervullen, ook al hebben ze een andere POS-waarde gekregen (bijwoord, zelfstandig naamwoord, werkwoord, ...). Het functielabel TAG maakt in die gevallen duidelijk dat de elementen niet hun 'normale' rol vervullen van bijwoord (MOD binnen de zin), zelfstandig naamwoord (HD van een NP), etc.

(391) hè, wat zitten we hier toch weer gezellig (tussenwerpsel)

(392) 't was gisteren weer ouderwets gezellig, hè (tussenwerpsel)

(393) je gelooft me niet, wel? (bijwoord)

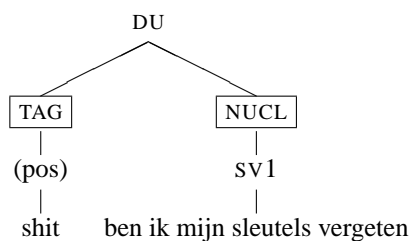
(394) ze kwam vanavond later, toch? (bijwoord)

(395) verrek, da's waar ook (werkwoord)

(396) shit, ben ik mijn sleutels vergeten (zelfst. naamw.)

(397) God, wat was ik misselijk (eigenaam)

(398) Jan, doe jij de deur even dicht? (eigenaam)



Een tweede groep van TAG-elementen wordt gevormd door (gereduceerde) zinnen die als inleider, aanhangsel of tussenvoegsel gerealiseerd worden. Ook hier maakt het TAG-label duidelijk dat het niet gaat om imperatieven of vragen in de gebruikelijke zin.

(399) dat is lang geleden, zeg

(400) zeg eens, was je nu nog van plan iets te vertellen

(401) kom, ik ga er maar weer eens vandoor

(402) kijk, daar heb je haar ook weer

(403) moet je horen, ik ben hier niet voor niks naartoe gekomen

(404) dan heb ik zoiets van: laat maar, weet je

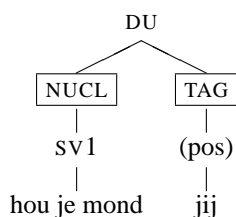
(405) hij is daar nog niet klaar voor, denk ik

(406) hij is daar, denk ik, nog niet klaar voor

We gebruiken de combinatie NUCL+TAG tenslotte ook voor gevallen waar de kern een imperatiefzin is (cf. sectie 7). Het TAG-element onderhoudt in deze gevallen een relatie met het niet uitgedrukte subject van de kernzin.

(407) hou je mond, jij (TAG)

(408) jongens, opgepast (TAG)



### 5.3 Asyndetische constructies

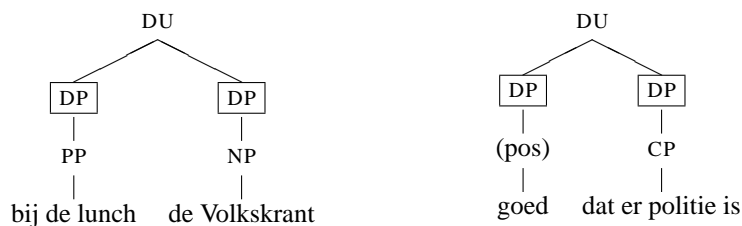
Labels: DU, DP

Met asyndetische constructies doelen we hier op reeksen van woordgroepen waartussen geen expliciete dependentierelaties zijn uitgedrukt (vergelijk ook sectie § 4.1 over asyndetische nevenschikking), maar die als elliptische zinnen kunnen worden opgevat. Bijvoorbeeld:

(409) (welke kranten lees jij?) bij de lunch de Volkskrant; 's avonds de NRC

(410) goed dat er politie is

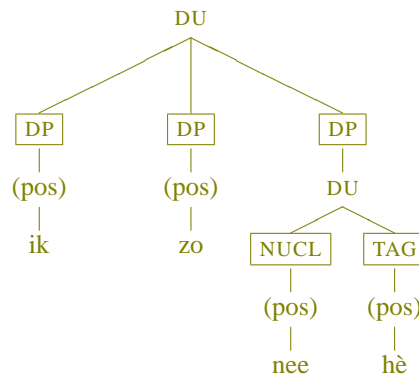
We geven de coherentie in dergelijke reeksen weer door de frases onder een DU-knoop samen te voegen. We zien er evenwel van af om expliciete dependentierelaties te reconstrueren: we beschouwen dit als een inferentie-taak, niet als een taak van de basisannotatie zelf. De dochters van DU krijgen in deze gevallen een uniform *d*-label DP ('discourse-part').



## AANTEKENINGEN

**Aldus premier Kok** Wordt geanalyseerd als een asyndetische DU (DP DP).

**Ik zo nee hè** Wordt geanalyseerd als een asyndetische DU (DP DP DP), hoewel er misschien best iets voor te zeggen zou zijn om *zo* HD van een SMAIN te maken.



**Hoe meer, hoe beter** en andere constructies van dat type: *hoe meer vragen je goed hebt, hoe meer je verdient* (sfn000841,129). Volgens de ANS fungeert de eerste *hoe*-vraag als een bepaling bij de tweede, maar dan is het toch gek dat de tweede ook de vorm van een ondergeschikte vraag heeft. Zowel het elliptische *hoe meer hoe beter* als de 'complete' gevallen met twee werkwoorden (*hoe meer vragen je goed hebt, hoe meer je verdient*) analyseren we als een asyndetische DU (DP DP), in het eerste geval bestaande uit twee APs en in het tweede geval uit twee WHQs.

## 5.4 Anakoloeten en dergelijke

Niet zeldzaam in gesproken taal zijn verschillende soorten anakoloeten, zinnen die syntactisch min of meer lijken te corresponderen.<sup>9</sup> Aangezien de grondhouding van de syntactische annotatie zo veel mogelijk descriptief en niet normatief is, wordt, op duidelijke gevallen van kennelijke versprekingen en correcties na, zo veel mogelijk getracht zulke structuren in te passen in de standaard syntactische configuraties (vergelijk ook het gezegde in §1.2). In sommige gevallen leidt dat tot onorthodoxe structuren. In de volgende paragrafen geven we enige voorbeelden.

### 5.4.1 Spiegelconstructies

Behoorlijk frequent is het volgende fenomeen: spreker begint een zin in de standaardvolgorde subject-persoonsvorm, dan volgt een bepaling, en vervolgens is het alsof de spreker het bouwplan van de zin herziet en de bepaling herinterpreteert als topic op de eerste zinspositie, waar de persoonsvorm en het onderwerp op moeten volgen (411–413). Ook bij zinnen met inversie, waarin een ander zinsdeel dan het onderwerp op de eerste zinsplaats staat, is dit mogelijk (414–415).

(411) je moet heel snel moet je weer wisselen

(412) ik ben eigenlijk ben ik docente Frans

<sup>9</sup>de Vries (1910:6) schrijft: "Soms wordt in het vervolg van een zin niet de noodige rekening gehouden met een reeds geuit bestanddeel." Toch lijkt het niet correct deze zinnen te beschouwen als performance-fouten; ze lijken eerder onderdeel van de spreektaalgrammatica. Sommige sprekers gebruiken bijvoorbeeld veel van de in de volgende paragraaf besproken spiegelzinnen, en naar het ons toeschijnt min of meer bewust, omdat ze een duidelijk retorisch effect kunnen hebben (vergelijk Hoekstra *et al.* (2001)).

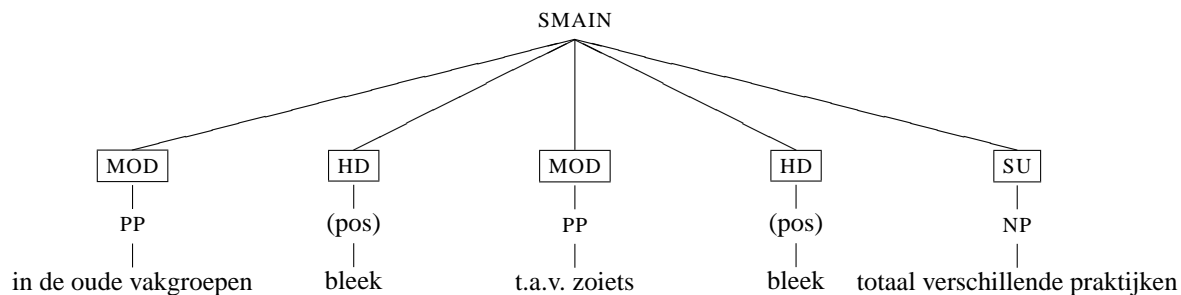
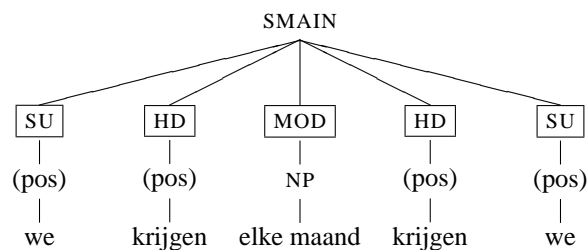
(413) we krijgen elke maand krijgen we dat soort mailtjes

(414) in de oude vakgroepen bleek ten aanzien van zoiets als het maken van dienstreizen bleek totaal verschillende praktijken te bestaan

(415) dan is toch vrijwel altijd is dan cruciaal of men met een gebied te maken heeft waar laat ik 't maar zo zeggen uh sprake is van eenduidigheid

De bepaling lijkt steeds een soort van spiegelpunt te vormen (vergelijk Huesken (2001), Jansen (1981:h.7) en de daar vermelde literatuur).<sup>10</sup>

De analyse is puur descriptief: de zin krijgt eenvoudigweg twee verbale hoofden (en eventueel twee subjecten).<sup>11</sup>



#### 5.4.2 Ingebedde dislocaties

Beschouw een zin als de volgende:

(416) wat vindt u d'r eigenlijk van dat 't vak literatuur of literatuurgeschiedenis dat dat niet meer bij alle vakken gegeven wordt?

De schrijftaalvariant zou, vermoedelijk, zijn

(417) wat vindt u d'r eigenlijk van dat 't vak literatuur of literatuurgeschiedenis niet meer bij alle vakken gegeven wordt?

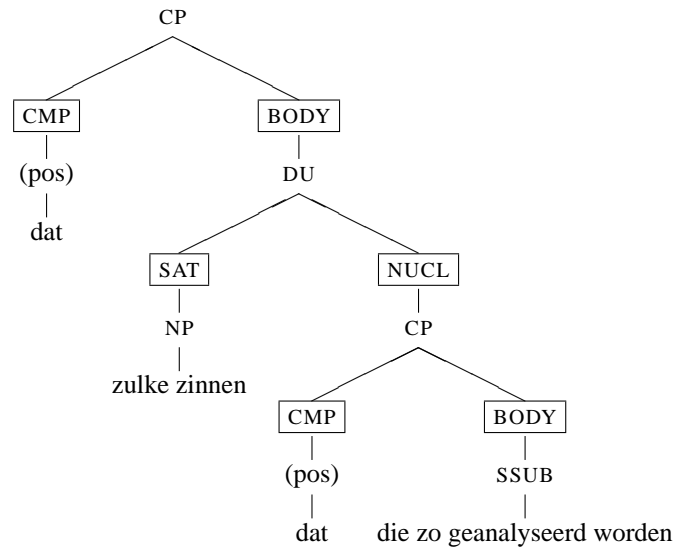
maar dat wordt niet zo gezegd. Het is ook niet correct om te denken dat de sequentie *dat dat* hier aarzeling uitdrukt. Veranderen we namelijk het zelfstandig naamwoord in het eerste gedeelte van een *het*-woord in een *de*-woord, dan zien we dat het tweede *dat* verandert in *die*:

<sup>10</sup>Dikwijls is het zaak het spraaksignaal goed te beluisteren om te bepalen dat het inderdaad om een spiegelzin gaat en niet om een hapering, verspreking of zelfcorrectie.

<sup>11</sup>Er kunnen niet meer dan vijf takjes aan een knoop (L<sup>A</sup>T<sub>E</sub>X).

(418) wat vindt u d'r eigenlijk van dat zulke zinnen dat die zo geanalyseerd worden?

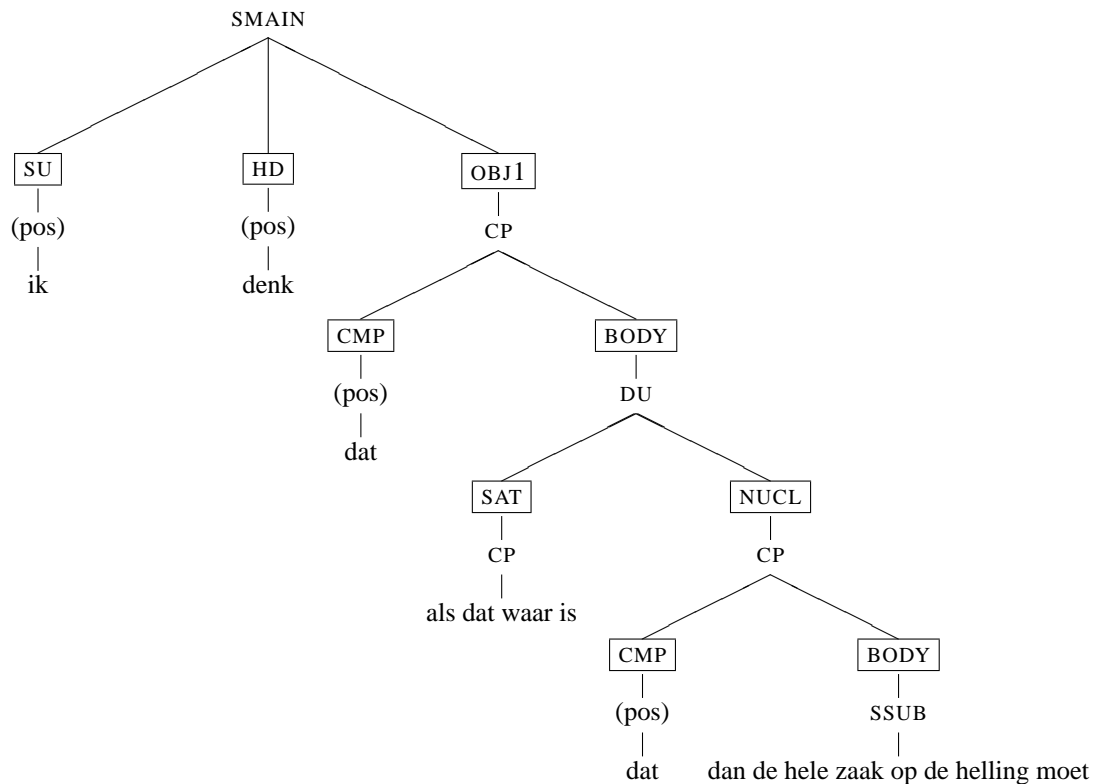
We analyseren zinnen als (417) en (418) als ingebedde links-dislocaties: *dat die zo geanalyseerd worden* is een CP die als NUCL fungeert binnen een DU waarin *zulke zinnen* een SAT is. Deze DU is de BODY van de hoogste CP.



Een ander voorbeeld van een ingebedde discourse-eenheid is de volgende (vergelijk [Jansen \(1981:h.7\)](#))

(419) ik denk dat als dat waar is dat dan de hele zaak op de helling moet

(Schrijftaalvariant: *ik denk dat als dat waar is de hele zaak op de helling moet.*) In dit geval fungeert de *als*-frase als SAT van de ingebedde DU:



Deze analyses laten zien dat het hierboven gemaakte onderscheid tussen zins-interne structuren enerzijds en boven-syntactische discourse-structuren anderzijds niet zo absoluut is als daar wel gesuggereerd wordt. We zijn gewend de relatie tussen de twee leden van een voorwaardelijke bijzin tot de laatste categorie te rekenen, maar zinnen als de hier behandelde laten zien dat dat discourse-geheel op zijn beurt weer kan fungeren als onderdeel van een groter syntactisch geheel.

#### AANTEKENINGEN

Naast ‘spiegelzinnen’ als de volgende (spiegelpunt: het lijdend voorwerp):

(420) ik vind fietsen in de bergen vind ik altijd wel heel lekker

en zinnen met (al of niet ingebedde) dislocatie:

(421) fietsen in de bergen en dan ook wel een beetje op tempo dat vind ik altijd wel heel lekker

komen er ook mengvormen voor, zoals:

(422) ik vind fietsen in de bergen en dan ook wel een beetje op tempo dat vind ik altijd wel heel lekker



De zin is begonnen als:

(423) ik vind fietsen in de bergen en dan ook wel een beetje op tempo (vind ik) altijd wel heel lekker

maar wordt op zeker moment hernomen als een zin met links-dislocatie:

(424) ik vind fietsen in de bergen en dan ook wel een beetje op tempo dat vind ik altijd wel heel lekker

Andere voorbeelden van hybride zinnen:

(425) ... in die zin dat ik in elk geval als we wat gaan doen dat ik jou vrijhou (sfn000887,7)

(426) ... dat ie instelt dat ie om de minuut hetgene wat je getypt hebt dat ie dat even wegschrijft (sfn000400,163)

(427) ik denk als we m\*a mijn vrouw gaan vragen dat ze zal zeggen dat ... (sfn007391,142)

In al dergelijke gevallen moet zo veel mogelijk in de geest van de CGN-annotatie gehandeld worden. In (427), bijvoorbeeld, is geen sprake van ingebedde links-dislocatie als besproken in 5.4.2 (er is immers geen ‘anker’ voor de *als*-zin in de *dat*-zin), maar van pure verplaatsing uit de *dat*-zin. Deze zin wordt geanalyseerd als een ‘gewone’ hoofdzin met een discontinue bijzin: de *als*-zin wordt aangehecht onder de BODY van de *dat*-zin.

Onnodig te zeggen dat dergelijke zinnen altijd zorgvuldig moeten worden beluisterd, om te kunnen vaststellen of er sprake is van een echte spreektaalconstructie of van (een) performance-fout(en)!

## Index

- ADVP, 17, 18, 23, 49, 56, 57  
AHI, 12, 26, 27, 29, 30  
ANS, 11, 12, 16, 19, 20, 29–31, 34, 37–40, 44, 47–51, 58, 62, 66, 69  
APPOS, 51, 52  
AP, 8, 22, 30–33, 56, 60, 69  
BODY, 5, 25–28, 30, 59–61, 71–73  
CMP, 5, 25–30, 71, 72  
CNJ, 62–64  
CONJ, 29, 62–64  
CP, 9, 10, 25–27, 29, 32, 35, 36, 46, 66, 68, 71, 72  
CRD, 30, 62–64  
DETP, 41, 42, 55–57  
DET, 7, 40–42, 44–48, 50, 51, 53–56, 60, 61  
DLINK, 65–67  
DP, 23, 65, 68, 69  
DU, 23, 65–69, 71, 72  
HDF, 34  
HD, 5, 6, 8–18, 23–27, 30–42, 44–48, 50–57, 60, 63, 64, 66, 67, 69, 70, 72  
INF, 9, 12, 23, 24, 27, 53, 56, 64  
IO, 11, 29  
LD, 8, 12–16, 18, 19, 21, 22, 34, 36, 39  
LIST, 7, 62, 64  
LP, 62, 64  
ME, 8, 13, 22, 34, 62  
MOD, 8, 9, 11, 13, 15–18, 20–27, 29, 30, 32–34, 36–38, 40–42, 44, 45, 47–49, 51, 53–57, 59, 64–67, 70  
MWP, 27, 28, 38, 39, 44, 57  
MWU, 27–30, 32, 33, 38, 39, 42–44, 56, 57, 64  
NP, 7, 8, 10, 13, 15, 17, 30, 33, 36–57, 60, 61, 63–68, 70, 71  
NUCL, 65–69, 71, 72  
OBCOMP, 30–33, 57  
OBJ1, 5, 6, 8, 10–14, 17, 21, 22, 24–26, 28–31, 34–39, 45, 48, 53, 55, 59–64, 72  
OBJ2, 8, 9, 11, 12, 26, 53, 66  
OTI, 12, 26, 28, 30, 33, 46, 49, 59  
PC, 8, 12, 13, 21, 22, 30–32, 35, 40, 45, 46, 56, 59, 60  
POBJ1, 8, 10, 30, 31, 34, 35  
POS, 4, 5, 8, 14, 27, 34, 51, 59, 60, 62, 63, 67  
PPART, 12, 15, 16, 24  
PPRES, 24, 48  
PP, 5, 8, 12–16, 21, 26, 32–39, 45, 48, 49, 53, 56, 59, 60, 63, 66, 68, 70  
PREDC1, 8, 9, 11, 14–16, 18, 20–22, 24, 30, 31, 34, 36, 37, 39, 52, 56  
PREDM, 8, 11, 15, 16, 20, 23, 36, 49, 50  
PRT, 17, 23, 56  
REL, 25, 47, 59  
RHD, 59–61  
SAT, 65–67, 71, 72  
SE, 8, 10–13, 25  
SMAIN, 5, 6, 9–16, 18, 23–27, 29–31, 36, 48, 52, 55, 64, 66, 69, 70, 72  
SSUB, 10, 23–25, 27–29, 59–61, 71, 72  
SUP, 8, 9, 64, 66  
SU, 5, 6, 8–16, 18, 24–27, 30, 31, 36, 48, 52, 53, 55, 59, 60, 64, 66, 70, 72  
SV1, 17, 23–26, 35, 61, 67, 68  
SVAN, 26, 29  
SVP, 8, 10, 14, 15, 18, 19, 22  
TAG, 54, 65, 67–69  
TI, 5, 12, 22, 26, 28, 30, 37, 49  
VC, 5, 8–10, 12, 15, 16, 27–31, 36, 40, 45, 46  
WHD, 28, 59, 61  
WHQ, 25, 59, 61, 62, 69  
WHREL, 25, 59–61, 65  
WHSUB, 25, 28, 59, 61, 62

## Lijst van figuren

1	<i>d</i> -labels: verbale domeinen . . . . .	8
2	<i>c</i> -labels: moederknopen van verbale domeinen . . . . .	24
3	Complementeerder-extensies van verbale domeinen . . . . .	26
4	<i>d</i> -labels: adjectiefdomein . . . . .	30
5	<i>d</i> -labels: prepositionele domeinen . . . . .	34
6	<i>d</i> -labels: nominale domeinen . . . . .	40
7	Niet-lokale dependenties: schematisch. . . . .	58
8	Labels voor niet-lokale dependenties . . . . .	59
9	Labels voor nevenschikking . . . . .	62
10	Labels voor discourse-structuren . . . . .	65

## References

- BENNIS, HANS. 1986. *Gaps and dummies*. Dordrecht: Tilburg dissertation. (published by I.C.G. Printing and Foris, Dordrecht). 19
- DE GROOT, CASPER. 1995. De absentief in het Nederlands. *Forum der Letteren* 36, 1–18. 29
- DE VRIENDT, SERA, WILLY VANDEWEGHE, & PIET VAN DE CRAEN. 1991. Combinatorial aspects of modal particles in Dutch. *Multilingua* 10, 43–59. 23
- DE VRIES, JELLE. 2001. *Onze Nederlandse Spreektaal*. Den Haag: SDU Uitgevers. 6
- DE VRIES, WOBBE, 1910. Dymelie. Opmerkingen over syntaxis. Verhandeling behorende bij het programma van het gymnasium der gemeente Groningen voor het jaar 1910–1911. 69
- , 1911. Dymelie. Opmerkingen over syntaxis (vervolg). Verhandeling behorende bij het programma van het gymnasium der gemeente Groningen voor het jaar 1911–1912. 6
- DEN HERTOOG, C.H. 1974–1979. *Nederlandsche spraakkunst: handleiding ten dienste van aanstaande (taal)onderwijzers*. Amsterdam: Versluys. 4e bew. dr. ingel. en bew. door H. Hulshof [1e dr. 1892–1895].
- GIVÓN, TALMY (ed.). 1983. *Topic continuity in discourse: a quantitative cross-language study*. Amsterdam [etc.]: Benjamin. 20
- GRONDELAERS, S. 2000. *De distributie van niet-anaforisch er buiten de eerste zinsplaats. Sociolexicologische, functionele en psycholinguïstische aspecten van er's status als presentatief signaal*. K.U. Leuven dissertation. 20
- HAESERYN, WALTER, & OTHERS (eds.). 1997. *Algemene Nederlandse Spraakkunst*. Groningen and Deurne: Martinus Nijhoff and Wolters Plantijn. 2e, geheel herz. dr. 19
- HOEKSTRA, HELEEN, MICHAEL MOORTGAT, BRAM RENMANS, INEKE SCHURMAN, & TON VAN DER WOUDE. 2001. On certain syntactic properties of spoken Dutch. Paper delivered at Computational Linguistics in the Netherlands, Enschede, November 2001. 69
- HONCOOP, MARTIN. 1998. *Dynamic excursions on weak islands*. Leiden dissertation. (The Hague: Holland Academic Graphics). 42
- HOOGVLIET, J.M. 1903. *Lingua: een beknopt leer- en handboek van Algemeene en Nederlandsche taalkennis, meer bepaaldelijk bestemd voor leeraren en onderwijzenden in moderne en oude talen*. Amsterdam: S.L. van Looy.
- HUESKEN, NICOLE. 2001. Mirrorsentences. Repetition of inflected verb and subject in Spoken Dutch. Master's thesis, Utrecht University, General Linguistics. [www.let.uu.nl/~Nicole.Huesken/personal/scriptie/scriptie.pdf](http://www.let.uu.nl/~Nicole.Huesken/personal/scriptie/scriptie.pdf). 70
- HULSHOF, HANS. 1987. Geef de boeken dan nu toch maar 'es even hier. *Forum der Letteren* 28, 85–89. ANS-nummer.
- JANSEN, FRANK. 1981. *Syntactische constructies in gesproken taal*. Leiden dissertation. 6, 70, 71
- KLEIN, DICK. 1977. *Appositionele constructies in het Nederlands*. Nijmegen dissertation.
- KRIVONOSOV, ALEKSEJ T. 1963. *Die modalen Partikeln in der Deutschen Gegenwartssprache*. Humboldt Universität Berlin dissertation. published Göttingen (1977): Kümmerle.
- PAARDEKOOPER, P.C. [n.d.]. *Beknopte ABN-syntaxis*. Eindhoven: Uitgave in eigen beheer. Zevende druk.
- RENKEMA, JAN. 1989. *Schrijfwijzer*. 's-Gravenhage: SDU uitgeverij. Volledig herz. ed. (Oorspr. uitg.: 1979).
- STURM, ARIE. 1986. *Primaire syntactische structuren in het Nederlands*. Nijmegen dissertation. (published by Martinus Nijhoff, Leiden).
- VAN CRAENENBROECK, JEROEN. 2000. Complementerend van. Een voorbeeld van syntactische variatie in het Nederlands. *Nederlandse Taalkunde* 5, 133–163. 46

VAN RIEMSDIJK, HENK. 1978. *A Case Study in Syntactic Markedness: The binding nature of prepositional phrases*. Foris Publications, Dordrecht.

WEGENER, HEIDE. 1989. Eine modalpartikel besonderer art: Der dativus ethicus. In *Sprechen mit Partikeln*, ed. by Harald Weydt, 56–73. Berlin: Walter de Gruyter. 20

Woordenlijst. 1995. *Woordenlijst Nederlandse Taal*. Den Haag, Antwerpen: Sdu, Standaard. samengest. door het Instituut voor Nederlandse Lexicologie; met een leidraad door Jan Renkema.