



Universiteit
Leiden
The Netherlands

Functionele elementen in een cognitief perspectief

Verhagen, A.; Schilperoord, J.

Citation

Verhagen, A., & Schilperoord, J. (1997). Functionele elementen in een cognitief perspectief. Retrieved from <https://hdl.handle.net/1887/2411>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/2411>

Note: To cite this publication please use the final published version (if applicable).

* Functionele elementen in een cognitief perspectief

Evidentie uit taalproductie

Joost Schilperoord & Arie Verhagen*

Abstract

Today's most influential psycholinguistic models of language production assume that functional elements, such as articles, conjunctions and prepositions, are produced as a result of the computation of the syntactic structure of an utterance. These models are maximally structure-generating in that each non-terminal node in a syntactic tree is seen as an active procedure specialized in computing one particular type of phrase or constituent. The knowledge a speaker puts to use in the process of sentence production then consists predominantly of sets of rules or procedures that operate on minimal declarative input: the mental lexicon, containing the content words (with their syntactic properties) to be retrieved from memory under the control of the conceptualizing module, i.e. as a result of cognitive activities like thinking and planning. In this paper we argue that this idea of the division of labour between memory and computation cannot in general be correct, in view of the distribution of pauses with respect to functional elements. As an alternative, it is proposed that the cognitive status of many of the linguistic structures people produce, though they *may* in principle be computed, is that of established declarative *schemata*, with a functional element serving to retrieve such a schema, including its semantic properties, from memory.

The outline of this paper is, first, to present and discuss some of the data; second, to argue that they raise particular problems for a strictly computational theory of sentence production; third, to present proportional pausing data from a large corpus of text production processes; and finally, to explore how these data can be accounted for in a cognitive linguistic approach to the nature of linguistic knowledge. Special emphasis is given to the schematic representations of noun phrases and prepositional phrases.

1 Inleiding

In een van de afleveringen van de TV-serie "30 minuten" die de VPRO in het seizoen 1996/7 uitzond, werd onder de titel "Rondom Ons" het bekende gespreksprogramma "Rondom Tien" geparodieerd. Aan het begin van die aflevering sprak Arjan Ederveen de volgende woorden:

- (1) We gaan vanavond praten en de problemen die daarbij komen kijken.

* Adres van de auteurs: Utrechts instituut voor Linguïstiek OTS, Trans 10, 3512 JK Utrecht (Email: Joost.Schilperoord@let.ruu.nl, Arie.Verhagen@let.ruu.nl). Wij danken de redactieleden van *Nederlandse taalkunde* en een anonieme reviewer voor suggesties en commentaar.

Wat is er nu precies bijzonder aan deze uiting? Natuurlijk is zij direct *herkenbaar* als typerend voor het begin van een Rendom-Tien-aflevering. Maar Ederveen produceert niet zozeer een introductie *als* van Rendom-Tien, hij geeft a.h.w. *de* introductie. Hij presenteert het *sjabloon* van de introductie met weglating van het enige variabele deel, waarin gemeld wordt *waarover* in die specifieke aflevering gepraat gaat worden. Daarmee wordt dus indirect de boodschap overgebracht "Ze zijn allemaal hetzelfde". De herkenning van (1) als behorend bij een bepaald praatprogramma, en de herkenning van het schematische karakter zelf, berusten op de werking van het *geheugen*: de associatie van deze uiting met een bepaald soort TV-programma wordt door het sjabloon-karakter van het geheel geactiveerd, en niet berekend uit de samenstellende delen van de zin.

Taalgebruikers kunnen dus kennelijk *schema's* van taaluitingen als geheel uit hun geheugen opdiepen, met open plekken voor variabele delen, en een geval als (1) is zelfs op herkenning van dat schema-karakter afgestemd. Dit behoort dus tot de mogelijkheden, maar is het gebruik van het geheugen om een gestructureerde uiting te produceren of te begrijpen, nu erg uitzonderlijk of juist nogal gewoon? In dit artikel willen wij betogen dat kennis van gestructureerde taaluitingen in de vorm van schema's zo gewoon is, dat er consequenties uit voortvloeien voor fundamentele opvattingen over de organisatie van talige kennis en het gebruik ervan in (in ieder geval) taalproductie.

Gevallen als deze stellen de verhouding aan de orde tussen wat wel genoemd wordt het computationele aspect van menselijke cognitie, en het geheugen-aspect. Het idee is dat er, in algemene termen gesteld, twee fundamenteel verschillende manieren zijn waarmee een cognitief systeem bij een bepaalde input een bepaalde output kan bereiken (b.v. een interpretatie bij een waargenomen taaluiting, of juist een taalvorm bij een bedachte boodschap), d.w.z. vanuit de ene toestand in een volgende kan geraken. De ene is dat er met de elementen van de input 'gerekend' wordt volgens een bepaald algoritme, en dat het resultaat daarvan de output is, of die in ieder geval bepaalt. De andere is dat de input vergeleken wordt met een geheugenpatroon, en dat de informatie die gevonden wordt bij een patroon dat genoeg lijkt op de actuele input, geactiveerd wordt als output.

Iedereen is het er over eens dat mensen over beide soorten vermogens beschikken. Over de vraag hoe ze zich tot elkaar verhouden, bestaat daarentegen juist wel veel controverse. Is het typische van de manier waarop een expert een probleem oplost nu dat hij een complex programma volgt om allerlei, liefst zoveel mogelijk, keuzes en hun consequenties door te rekenen, of is de crux juist dat hij over een grote en gedifferentieerde database ('ervaring') beschikt waar hij nieuwe situaties mee kan vergelijken om tot een maximaal adequate diagnose te komen? In het eerste geval speelt procedurele kennis de hoofdrol, in het tweede declaratieve. Het zal duidelijk zijn dat dit soort controverses van groot theoretisch en praktisch belang zijn (men denke b.v. maar aan de consequenties voor onderwijs).

In de taalwetenschap is het algemeen cognitieve-wetenschappelijke verschil tussen berekening en geheugen geassocieerd met dat tussen grammatica en lexicon. De eerste is (idealiter) de verzameling van regels, en daarmee zowel de bron van regelmaat

als van structuur van taaluitingen.¹ Het lexicon is (idealiter) de verzameling van elementen die door de regels van de grammatica tot samenhangende (gestructureerde) en *regelmatige patronen* worden *samengevoegd*. Het normale geval is dus, in het kader van deze 'ideale' taakverdeling, dat als een taaluiting structuur vertoont, die structuur geproduceerd is door regels van het computationele systeem (de algoritmische grammatica), en niet uit het geheugen is gehaald. Het is dit ideaalbeeld dat problematisch wordt als taalproductie als die van Arjan Ederveen, en de manier waarop mensen die begrijpen, in feite behoorlijk normaal zouden zijn. In dit artikel willen we die visie dan ook aan een kritische beschouwing onderwerpen, en wel naar aanleiding van het verschijnsel dat (in ieder geval) in spontane taalproductie, pauzes nogal vaak optreden juist direct na *grammaticale* woordjes, de zgn. functiewoorden.

2 Pauzes in taalproductie

Het volgende transcript van een stukje dictaatproces is een goed voorbeeld van het soort pauzeringsgegevens dat we hier behandelen. Ons betoog zal uitsluitend betrekking hebben op zogeheten *stille* pauzes. Zulke pauzes zijn slechts één type van hesitatie-fenomenen waartoe ook gevulde pauzes ('eh') behoren. We beschikken echter niet over data met betrekking tot de plaatsing van zulke gevulde pauzes. Een extra reden om ze niet direct op een lijn te stellen met stille pauzes, is dat er in de literatuur voorstellen te vinden zijn over mogelijke specifieke functies van (verschillende soorten) pauze-vulsels (verg. Clark 1996). In transcript (2) geeft elke regel (afgesloten met een "/") een stukje taalproductie tot aan een pauze weer; we vragen in het bijzonder aandacht voor de regels (aangeduid met pijlen) die eindigen op een functiewoord.

- (2) (...)
- 1. Mijns inziens is het van belang om /
 - 2. de arbeidsovereenkomst /
 - 3. zodanig /
 - 4. in te kleden dat /
 - 5. duidelijk vaststaat dat /
 - 6. dit een arbeidsovereenkomst voor een bepaalde tijd is /
 - 7. dat wil zeggen dat /
 - 8. het eerste gedeelte van de overeenkomst. /
 - 9. Over de /
 - 10. overige onderdelen van /
 - 11. de arbeidsovereenkomst /
 - (...)

1 Dat wil zeggen: boven het niveau van woorden. Binnen het lexicon bestaat ook structuur, o.a. in de vorm van relaties tussen gelede woorden (het terrein van de morfologie). De gedachte is echter dat regelmaat op dat niveau veelal paradigmatisch van aard is, in tegenstelling tot de principiële syntagmatische relaties *tussen* woorden op het niveau van constituenten en zinnen; voor de grammatica van constituenten en zinnen zijn gelede en ongelede woorden allemaal elementen. Het is juist deze opvatting van het verschil tussen grammatica en lexicon die wij hier willen kritiseren vanuit het oogpunt van de analyse van taalproductieprocessen.

In de regels 1, 4, 5, 7, 9 en 10 treden pauzes op vlak nadat er een functiewoord is uitgesproken. Als we nu uitgaan van de globale 'hoofdregel voor pauzeren' (vgl. Schilperoord 1996, hoofdstuk 1), nl. dat tekstproducenten pauzeren op plaatsen waar daartoe een cognitieve noodzaak bestaat, dan is de vraag: *Waarom* wordt er precies op deze plaatsen gepauzeerd? Wat *is* die cognitieve noodzaak hier? Een voor de hand liggende hypothese is dat de spreker zich a.h.w. al gecommiteerd heeft aan bepaalde algemene aspecten van de frase die hij wil gaan uiten (hetgeen hem in staat stelt een bepaald functiewoord te realiseren), en dan nog een beetje moet nadenken over verdere invulling. In die richting ging ook al enigszins Boomer (1965), die op een vergelijkbaar fenomeen stuitte in wat hij "phonemic clauses" noemde; hij suggereerde het volgende:

The initial word [waar dus vaak een pauze na valt – JS/AV] in a phonemic clause sets certain constraints for the structure of what is to follow. The selection of a first word has in greater or lesser degree committed the speaker to a particular construction or at least a set of alternative constructions. (Boomer 1965:156)

Met andere woorden: door de productie van het eerste woord sluit de spreker allerlei mogelijkheden uit ten aanzien van de constructie die nu gerealiseerd gaat worden; dit eerste woord levert een lineaire modificatie (Bolinger 1965) van het deel van de taaluiting dat op dat moment nog in de toekomst ligt. Fragment (2) suggereert dat dit verschijnsel niet beperkt is tot beginwoorden van deelzinnen. 'Kleinere' eenheden zoals substantief- en prepositiegroepen vertonen het evenzeer (zie b.v. r. 9-11 van (2): [...] *over de / overige onderdelen van / de arbeidsovereenkomst*). Precies deze lineaire modificatie (de *globale* keuze van de soort woordgroep wordt vastgelegd) zou het mogelijk en nuttig kunnen maken functiewoorden reeds te realiseren terwijl er nog nagedacht (en dus gepauzeerd) moet worden over de verdere invulling van de eigenlijke inhoud.

Maar hoe voor de hand liggend wellicht ook, deze interpretatie stuit op een groot probleem in verband met een gangbare theorie over de productie van syntactische constructies: Kempen en Hoenkamp's *Incremental Procedural Grammar* (Kempen & Hoenkamp 1987); dat komt doordat de hypothese impliceert dat een syntactische deelstructuur vast kan liggen voordat de conceptuele inhoud ervan bepaald is, en dat is iets dat door de bedoelde theorie juist wordt uitgesloten. Omgekeerd zijn gegevens als die in (2) natuurlijk problematisch voor die theorie – aangenomen althans dat ze niet uitzonderlijk zijn. Deze vaststellingen nopen ons er dus toe eerst zowel de aard van het theoretische probleem precies vast te stellen als na te gaan hoe '(on)gewoon' data als in (2) nu precies zijn, om vervolgens (als het probleem intussen niet verdwenen is) een alternatief voor te stellen. Dat is dan ook wat wij nu achterenvolgens zullen doen. We zullen daarbij zien dat de achterliggende kwestie precies die is van de verhouding tussen de rollen van 'berekening' en van 'herinnering', tussen het computationele systeem en het geheugen, in feitelijke taalproductie.

3 Incrementele Procedurele Grammatica

Kempen en Hoenkamp's *Incrementele Procedurele Grammatica*, vanaf nu kortweg *IPG*, is een theorie over de aard en de rol van syntactische kennis bij het proces van

zinsproductie (Kempen & Hoenkamp 1987). Het model vormt de basis voor Levelts *Blueprint* model van het spreekproces, en is daarin geïncorporeerd als de *formulator*-component, d.w.z. de component die de omvorming beregelt van conceptuele structuren – de ‘boodschap’ die de spreker wil overbrengen – naar syntactische en fonologische representaties (Levelt 1989).

IPG is een *incrementeel* model van taalproductie. Dit houdt in dat de modelcomponenten die verantwoordelijk zijn voor de inhoud, vorm en articulatie (respectievelijk de *conceptualiseerder*, de *formulator* en de *articulator*) van een uiting goeddeels *parallel* opereren. Een zekere mate van serialiteit is onvermijdelijk, maar de formulator, bijvoorbeeld, kan in deze optiek zijn werk beginnen zodra de conceptualiseerder een ‘fragment of characteristic input’² aflevert. Zo’n conceptueel fragment representeert in de regel maar een deel van de volledige boodschap in wording. Het proces van grammaticale codering begint dus reeds voordat er een volledige boodschap gevormd is: de zinsinhoud groeit ‘mee’ met het verwoordings- en articulatieproces. Incrementele zinsproductie appelleert aan de alledaagse ervaring van sprekers dat er, op het moment waarop ze een uiting beginnen, vaak nog geen vastomlijnd idee is over hoe de uiting precies zal verlopen. Kempen en Hoenkamp benadrukken de psychologische motivatie voor incrementaliteit onder meer met erop te wijzen dat op die wijze het optreden van *pauzes* binnen uitingen/zinnen verantwoord kan worden. Met andere woorden, uitgaande van incrementaliteit zijn pauzes een normaal en natuurlijk verschijnsel bij zinsproductie.

Dit lijkt ons een alleszins redelijk uitgangspunt, en in ieder geval een krachtig argument voor het incrementele karakter van een taalproductiemodel. Maar nu komt ook de vraag op *waar* we die pauzes dan kunnen verwachten, en waar niet. Door middel van een bespreking van de werking van de procedurele grammatica zullen we laten zien dat IPG daarover specifieke voorspellingen impliceert. Voorspellingen, echter, die zich lastig laten rijmen met wat we daadwerkelijk in taalproductieprocessen aantreffen.

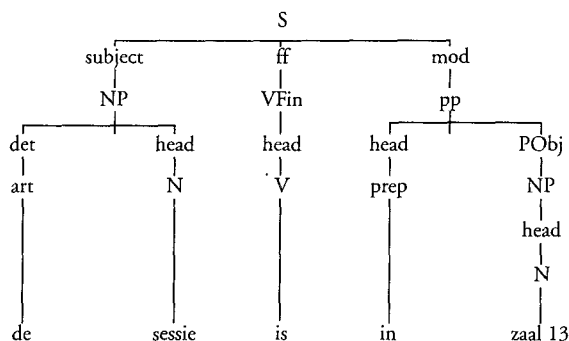
Als uitgangspunt voor de bespreking van de procedurele grammatica beschouwen we een eenvoudige declaratieve zin als (3)a en zijn syntactische constituenten.³

(3) a. De sessie is in zaal 13

2 Zie Levelt (1989:24). De vraag wat dan een ‘volledige’ boodschap, en dus ook wat een ‘fragment’ is, is niet eenvoudig te beantwoorden. De veiligste aanname lijkt te zijn dat een volledige conceptuele structuur die structureel is die correspondeert met de betekenis van een volledig opgetuigde clause. Levelt (1989:24) spreekt in algemene termen over een ‘characteristic unit of information’, ‘bits and pieces’ en fragmenten van ‘characteristic input’. Blijkens het voorbeeld op p.25 corresponderen de ‘bits and pieces’ met ontologische categorieën als ‘EVENT’, ‘PLACE’ en ‘TIME’.

3 De representatie in (3)b bevat zowel labels die categoriale informatie aanduiden (NP, PP, e.d.), als labels voor relationele informatie (subject, head, e.d.). Of dit vanuit een grammatica-theoretisch gezichtspunt een aantrekkelijke of plausibele voorstelling van zaken is, laten we hier in het midden. Voor de doelen van Kempen & Hoenkamp, evenals voor de onze, maakt dit niet uit. Tevens valt aan (3)b te zien dat Kempen & Hoenkamp voor het Nederlands geen VP aannemen (anders dan voor het Engels); ook dit punt doet voor ons betoog hier niet ter zake.

b.



Een boomstructuur als (3)b vat een aantal van de regels samen waaraan een (Nederlandse) zin moet voldoen om als grammaticaal aangemerkt te kunnen worden, en specificeert een aantal grammaticale functies zoals 'subject' en 'modificatie'. In een procedurele grammatica zoals IPG representeert (3)b een verzameling *routines, procedures of computationele specialisten* die gezamenlijk betrokken zijn bij de productie van zin (3)a. Dat wil zeggen dat iedere non-terminale knoop in de boom gezien wordt als een specialist, of een 'expert in assembling one type of syntactic constituent' (Kempen & Hoenkamp 1987:210). De procedure 'NP' bouwt derhalve nominale groepen, de procedure 'PP' voorzetselgroepen, enzovoorts. De vertakkingen in de boom representeren de hiërarchische verhoudingen tussen de verschillende specialisten. Iedere specialist is namelijk in staat subspecialisten aan te roepen die telkens specifiekere kenmerken van de boomstructuur beregelen. De procedure 'NP' is dus in staat de subprocedures 'DET' en 'NPhead' aan te roepen, en de procedure 'DET' kan op zijn beurt de subprocedure 'Art' aanroepen. Deze subroutines corresponderen met de lemma-informatie die bij het formuleerproces is betrokken.⁴

Het proces van grammaticale codering van een boodschap wordt in gang gezet op het moment dat er een intentie tot communiceren bestaat, en er een bijhorende conceptuele structuur is gevormd, dus een conceptuele representatie van de boodschap die de spreker wil communiceren. Deze representatie is de output van de *conceptualiseerder* (Levelt 1989) en vormt de input voor de formulator – de component die verantwoordelijk is voor het boom-constructieproces. Bij de interface tussen deze beide componenten is een cruciale rol weggelegd voor wat in Levelt (1989) het *mentale lexicon* genoemd wordt. We expliciteren die rol met behulp van een voorbeeld: de constructie van de NP 'de sessie' uit zin (1)a.

Laten we aannemen dat aan zin (3)a de conceptuele structuur (4) ten grondslag ligt:

⁴ Een procedure mag niet willekeurig welke andere procedures aanroepen. Welke (sub)procedures door een bepaalde procedure aangeroepen mogen worden, is beregeld door wat Kempen & Hoenkamp *appointment*-regels noemen. Tezamen constitueren deze regels de 'aanroep-hiërarchie' van de diverse specialisten. De claim is dat zo alle en alleen alle structurele configuraties die in een Nederlandse zin voor kunnen komen, gevormd kunnen worden. Appointment-regels vormen derhalve de kern van de procedurele grammatica.

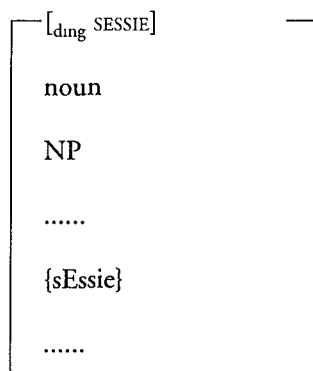
- (4) [_{state} ZIJN [_{ding} SESSIE][_{plaats} IN [_{ding} ZAAL 13]]]

Omdat IPG een incrementeel model van zinsproductie is, volstaat de aanwezigheid van de conceptuele deelstructuur [_{ding} SESSIE] om het boom-construcieproces in gang te zetten (de spreker behoeft de gehele zinsinhoud nog niet te kennen om met de productie van de subjects-NP te beginnen; de zinsinhoud groeit dus als het ware 'mee' met het spreekproces). De bedoeling is nu om de conceptuele structuur [_{ding} SESSIE] om te vormen naar de syntactische structuur [_{NP}[_{art}DE] [_nSESSIE]]. De cruciale rol die het mentale lexicon hierbij speelt komt tot uiting in de *lexicale hypothesis*, door Levelt als volgt geformuleerd:

(...) nothing in the speaker's message will *by itself* trigger a particular syntactic form
 (...). There must always be mediating lexical items, triggered by the message, which, by their grammatical properties (...) cause the Grammatical Encoder to generate a particular syntactic structure (Levelt 1989:181).

Het lexicon wordt in IPG gemodelleerd als het vaste kennisbestand van de spreker aangaande de *woorden* van zijn taal. Die kennis betreft de betekenis, de syntaxis (de categorie, subcategorisatie-informatie), en de fonologie van een bepaald woord. Zo'n kennispakketje wordt een *lemma* genoemd. In het proces van spreken betekent lexicalisatie dat een deel van de conceptuele structuur gekoppeld wordt aan een lemma met een overeenkomstige conceptuele specificatie: het 'adreskaartje' van het lemma (c.q. '(...) *triggered by the message* (...)'). Wanneer die koppeling tot stand is gebracht, komt de overige lemma-informatie beschikbaar voor het zinsbouwproces. Het lexicon implementeert derhalve de interface tussen het conceptuele systeem en het 'inhoudsloze' syntactische systeem. Dit is vereenvoudigd weergegeven in (5):

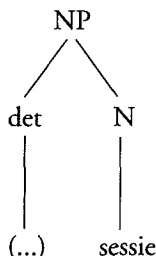
- (5) [_{ding} SESSIE]



Het lemma voor 'sessie' karakteriseert dit als een telbaar substantief (het kan een meervoud krijgen), en bevat informatie omtrent de woordklasse waartoe 'sessie' behoort: N, en informatie over het constituenttype waarvan het de kern kan vor-

men: NP⁵ Deze informatie komt dan ter beschikking van het bouwproces:⁶ de lemma-informatie 'sessie' roept de NP-routine of specialist aan die het verdere bouwproces beregelt.⁷ De N-positie (het lexicale hoofd) wordt vervuld door het substantief 'sessie', zodat de boom-in-aanbouw er op dat moment uitziet als (6):

(6)



Hoe vindt nu de insertie van het bepaalde lidwoord 'de' plaats? De NP-procedure inspecteert daartoe de conceptuele structuur en gaat, bijvoorbeeld, na wat de waarde is van de parameter die de 'toegankelijkheidsstatus' van het concept [_{ding} SESSIE] expliciteert. De procedure vindt als waarde +toegankelijk, en mede op basis van de lemma-specificatie voor het geslacht van het substantief 'sessie', komen de subroutines 'DET' en 'Art' uiteindelijk met het bepaalde lidwoord 'de' op de proppen, en niet met 'het' of 'een'. Dit proces van functiewoord-insertie wordt door Kempen en Hoenkamp aangeduid met de term *functorisatie*, met als doel '(...) refining the set of procedure calls contained by a lemma' (Kempen & Hoenkamp 1987:218, zie ook Levelt 1989:238-239). Procedures als "DET" en "Art" hebben dus een aparte status in het model omdat de aanwezigheid van functiewoorden '(...) [is] chiefly motivated on syntactic grounds, so they cannot be supposed to originate simply from lexicalization.' (ibid, 218). Hieruit volgt dat de conceptuele tegenhanger van het bepaalde lidwoord 'de' gevormd wordt door parameters die zijn geassocieerd met de conceptuele structuur 'sessie'. De motivatie voor (de keuze van) het lidwoord is dus niet onafhankelijk van het nominale concept in kwestie. Met betrekking tot de 'real-time'-organisatie van het formuleerproces betekent dit dat functiewoord-insertie *volgt* op lexicalisatie.

Uit de bovenstaande beschrijving van IPG kunnen de volgende algemene kenmerken van de theorie worden afgeleid. Allereerst maakt het model een strikte scheiding tussen *declaratieve* en *procedurele* taalkennis, een scheidslijn die bovendien precies langs het onderscheid tussen 'lexicon' en 'grammatica' loopt. De lexicale hypothese stelt dat het proces van grammaticale codering volledig lexicaal gestuurd

5 Met betrekking tot substantieven vormt deze lemma-informatie de procedurele tegenhanger van herschrijfgels: $N^2 \rightarrow \text{det}$, N^1 en $N^1 \rightarrow \dots N \dots$. Op deze wijze kunnen dus specificerders en complementen van nomina berekend worden.

6 Verder komt ook de fonologische lemma-informatie beschikbaar aangaande de uitspraak van het woord en het accentpatroon (aangegeven met een hoofdletter). We zullen deze informatie verder negeren.

7 Dit is een sterk vereenvoudigde beschrijving. Voor meer details, zie Kempen & Hoenkamp (1987), en Levelt (1989), vooral hoofdstuk 7.

is: boom(deel)structuren worden procedureel gevormd 'rond' de lexicale kern. Het model is bijgevolg *maximaal structuurgenererend*: iedere syntactische structuur die aan een uiting ten grondslag ligt, wordt gezien als de resultante van de procedurele grammatica. Hieruit volgt dat een structuur ((x)y), die op basis van de syntactische procedures 'berekend' kan worden vanuit de lexicale basiseenheden X en Y, *geen* deel kan, of zelfs mag uitmaken van het mentale lexicon. Kortom, wat in beginsel *kan worden* beregeld, wordt *niet* individueel gerepresenteerd.

Uit deze modellering van het taalproductiesysteem komt een uitgesproken, *generatieve* visie naar voren aangaande de aard van linguïstische kennis. De kennis die een spreker in staat stelt grammaticale uitingen te produceren, bestaat voornamelijk uit regelsystemen (de procedures, of routines) die opereren op een minimale declaratieve database (het mentale lexicon). Syntactische structuren *resulteren* dus uit de activiteiten van en de interacties tussen de procedures, en zijn, uiteindelijk, *geprojecteerd* vanuit geactiveerde lemmata. Merk op dat deze conceptie een *volledige analyseerbaarheid* van uitingen vereist: de betekenis van de uiting als geheel moet eenduidig af te leiden zijn uit de betekenis van de delen (de lexicale kernen), en de syntaxis zelf is betekenisloos.⁸

Gesteld nu dat we (onvrijwillige) pauzes kunnen opvatten als het gevolg van vertragingen in de werking van de bij taalproductie betrokken componenten, of van onderbrekingen in de informatiedoorgifte tussen de componenten, dan doet zich de vraag voor waar die hun oorsprong vinden in termen van het IPG-model. In beginsel kunnen we stellen dat iedere subcomponent pauzes kan veroorzaken, maar hier laten we pauzes die veroorzaakt worden door articulatoire problemen buiten beschouwing, evenals pauzes veroorzaakt door de prosodiegenerator.⁹ Pauzes kunnen dan nog veroorzaakt worden door problemen in de conceptualiseerder en in de formulator. Maar omdat IPG een input-output-model is, en omdat de werking van de formulator geacht wordt 'automatisch' te verlopen, zullen pauzes in de spraakstroom vooral voortvloeien uit haperingen in de activiteiten van de conceptualiseerder. Pauzes kunnen derhalve optreden wanneer de conceptuele specificaties (nog) niet geheel uitgewerkt zijn (geen conceptuele input), wanneer het lexicalisatieproces stagneert (wel conceptuele input, geen item in het lexicon gevonden), of wanneer uit meerdere lemmata een keuze gemaakt moet worden (wel conceptuele input, ook lemmata geactiveerd, maar nog geen keuze gemaakt).

In al deze gevallen, echter, volgt uit het model dat de subroutines die bij functorisatie zijn betrokken, nog niet aangeroepen kunnen worden. Immers, de mogelijkheid ze aan te roepen is volledig afhankelijk van de aanwezigheid van zowel een conceptuele input als een geactiveerd lemma (ze komen systematisch op de 'tweede' plaats in de hiërarchie van procedure-aanroepen). Hieruit volgt een heldere voorspelling over de plaatsen waar pauzes zullen optreden en waar niet: *wanneer de spreker bezig is de conceptuele input te specificeren, of een lemma te activeren, mogen we pauzes (als uiterlijke reflecties van deze cognitieve processen) verwachten vóór de articulatie van functiewoorden, en niet daarna*. Immers, zolang er geen lemma geacti-

8 Op dit punt wordt, door Levelt, alleen een uitzondering gemaakt voor idiomatische expressies (Levelt 1989: 186-187).

9 Zie ook Schilperoord (1996:274). Ook pauzes veroorzaakt door *monitor*-activiteiten laten we buiten beschouwing (zie Levelt 1989, hoofdstuk 12).

veerd is kunnen de relevante functorisatie-procedures nog niet aangeroepen worden, en als de articulator een functiewoord kan produceren, *moet* het lexicale hoofd geselecteerd zijn en kan de articulator dit dus ook direct produceren; tussen een functiewoord en een lexicaal hoofd zijn geen pauzes te verwachten.

Een en ander laat zich ook als een conceptuele kwestie formuleren. Stel dat er wel een pauze op een functiewoord *volgt*; als we die zoals gewoonlijk interpreteren als evidentie voor conceptuele activiteit, dan lijkt er sprake te zijn van 'structuur' zonder 'conceptuele input'. Dat is, zoals duidelijk zal zijn, flagrant in strijd met de lexicale hypothese. De aanwezigheid van zoiets als een lidwoord of een voegwoord in de spraakstroom impliceert volgens het IPG-model immers de selectie van, per definitie specifieke, lexicale kernen van de categorie N, V, of wat dies meer zij. Is er sprake van syntactische informatie zonder zulke conceptuele specificaties, en daarop zou de aanwezigheid van functiewoorden 'in isolatie' kunnen wijzen,¹⁰ dan kan deze informatie niet zijn geprojecteerd vanuit het geactiveerde lemma en de daarin gegeven categoriale informatie, en ziet de lexicale hypothese zich met een fundamenteel probleem geconfronteerd.

Kortom, het IPG-model impliceert problematische voorspellingen aangaande de plaats van pauzes ten opzichte van functiewoorden. Alvorens aan de orde te stellen hoe we de aanwezigheid van pauzes na functiewoorden wel kunnen verklaren, presenteren we in de volgende paragraaf eerst kwantitatieve informatie over de frequentie van het verschijnsel in kwestie.

4 Pauzepercentages

In de inleiding hebben wij het verschijnsel 'pauzes na functiewoorden' als zodanig aan de orde gesteld. In deze paragraaf buigen we ons over de vraag *hoe frequent* dit verschijnsel eigenlijk is. Gaat het om een ietwat exotische *faux pas* van de tekstproducent, of hebben we met een systematisch verschijnsel te maken? Hoewel het verschijnsel eerder is gerapporteerd (zie bijvoorbeeld Maclay & Osgood 1959, Boomer 1965, Clark & Clark 1977), ontbreekt een kwantitatief beeld. Om de genoemde vraag statistisch te kunnen beantwoorden hebben we gebruik gemaakt van een corpus van 120 tekstproductieprocessen.¹¹ Het corpus bestaat voor het overgrote deel

10 We bespreken nog apart (§ 4) in hoeverre zulke pauze-locaties een frequent verschijnsel zijn.

11 Voor een beschrijving van de wijze van verzamelen, de bewerking en de analyse van dit corpus, zie Schilperoord (1996), hoofdstukken 2 en 3.

uit (gedicteerde) routinebrieven van advocaten aan cliënten en collega's.¹² Iedere transitie tussen elk woordpaar in het corpus werd (o.a.) gescoord voor de lexicale categorie waartoe de flankerende woorden behoorden. Een algemeen onderscheid werd aangebracht tussen functiewoorden (determinatoren, preposities, complementeerders en conjuncties) enerzijds en 'inhouds'woorden (zoals de categorieën N, V en A) anderzijds. Verder werd iedere transitie gescoord voor het al dan niet voorkomen van een pauze op die transitie. Op basis van deze informatie kan de database geanalyseerd worden op pauze proporties in woordstrings zoals in (7): alle subscripten markeren een woordovergang en derhalve een potentiële pauzelocatie.

- (7) a. ₁een₂tuin₃
 b. ₁een₂Engelse₃tuin₄
 c. ₁in₂een₃Engelse₄tuin₅
 d. ₁in₂de₃tuin₄van₅Monet₆
 e. ₁dat₂ik₃een₄Engelse₅tuin₆heb₇bezocht₈

Doel van de analyse is een vergelijk maken tussen de proporties pauzes op, bijvoorbeeld, posities 1 en 2 in (7)a, 1 en 2 in (7)c en 1 en 2 in (7)e. De volgende stringtypen werden geselecteerd voor een dergelijke analyse (zie (8)).

- (8) a. det - (adjectief) - N
 b. prep - NP
 c. comp - ondergeschikte clause

De geschatte¹³ proporties en vergelijkingen zijn samengevat in tabel 1.

12 De hier gerapporteerde data hebben dus betrekking op tamelijk formele vormen van (schriftelijk) taalgebruik. Gedicteerde teksten verschijnen uiteindelijk 'op papier', en worden dan ook met het oog daarop ingesproken (zie Schilperoord 1996:19-23). Wij laten hier in het midden in hoeverre de gegevens gegeneraliseerd kunnen worden naar andere, meer informele taalgebruikcontexten. We verwijzen in dit kader naar de eerder aangehaalde publicaties van Boomer (1965) en Maclay & Osgood 1959, maar ook naar Blaauw (1995) waarin het zelfde verschijnsel in meer informele, mondelinge contexten gemeld wordt. Een van de reviewers suggereerde verder dat de plaatsing van pauzes na functiewoorden (ook) een strategische of retorische functie zou kunnen hebben, en niet zozeer de cognitieve oorzaak die wij eraan toeschrijven. Het is waar dat conversatie-analytische studies melding maken van het zelfde verschijnsel, en dat het daar in verband gebracht wordt met 'floor-keeping'-strategieën van sprekers. Echter, eenzelfde verschijnsel in verschillende contexten kan heel goed verschillende oorzaken hebben. Bovendien is niet op voorhand uit te maken dat we hier met een of-of-kwestie van doen hebben; een en-en-verklaring heeft niet op voorhand de mindere papieren. Als de pauzes in ons materiaal al een strategische achtergrond hebben, dan is het verband hoogstens indirect: de context is hier duidelijk monologisch zodat het strategische aspect strikt genomen zinloos is, hoewel niet geheel is uit te sluiten dat we hier van doen hebben met een conversationeel residu.

13 Voor de statistische data-bewerking is gebruik gemaakt van de multi-level-procedure voor variantie-analyse (Van den Bergh 1990, zie ook Schilperoord 1996:120). Deze procedure converteert steekproef-observaties direct naar populatie-kenmerken (dus kenmerken van de 'populatie NPs', de 'populatie PPs' enzovoorts). Omdat populatie-parameters principieel niet met zekerheid gekend kunnen worden, wordt gesproken van 'schattingen' in plaats van van 'berekeningen'. De proporties in tabel 1 geven dus aan wat we op basis van de steekproef-analyse mogen verwachten aan te treffen in de populatie. Het is ook om die reden dat de proportie-schattingen *getoetst* worden.

Tabel 1

Proporties en vergelijkingen van proporties van pauzes voor drie geopponeerde paren van pauzeerlocaties

stringtypen	proporties	χ^2
1. det - pauze	.53*	93.91*
2. pauze - det	.39*	
3. prep - pauze	.25*	14.24*
4. pauze - prep	.30*	
5. comp - pauze	.59*	25.89*
6. pauze - comp	.47*	

* significant bij $p = .05$

Het is niet goed mogelijk om de proportiedata in tabel 1 op zichzelf te interpreteren. We zouden daarvoor bijvoorbeeld moeten weten wat de kans is op het voorkomen van een pauze op de geanalyseerde locaties bij een willekeurige verdeling over alle mogelijke locaties. De significantie-indicaties wijzen daarom slechts uit dat pauzes op de geanalyseerde locaties als een populatiekenmerk beschouwd kunnen worden (zie voetnoot 14). Nu zou zo'n onafhankelijke interpretatie ook niet bijster informatief of interessant zijn. Waar het vooral om gaat is dat de proportie pauzes op de locatietypen 1, 3 en 5 opmerkelijk hoog is, *gegeven de op IPG gebaseerde voorspelling dat op die locaties geen pauzes verwacht worden*. Het aantal pauzes na de functionele elementen 'det' en 'comp' overtreft zelfs dat *voor* die elementen. Dat is niet het geval bij het element 'prep', maar ook daar is de proportie aanzienlijk, in het licht van de IPG-voorspelling.¹⁴ De chi-kwadraattoetsen voor de vergelijkingen tussen proporties per geanalyseerd paar wijzen bovendien uit dat alle verschillen reëel zijn. We kunnen derhalve concluderen dat we, zeker voor het hier onderzochte type discourse, met een regelmatig verschijnsel van doen te hebben, en dat de IPG-voorspelling aangaande de te verwachten locaties van pauzes ten opzichte van functiewoorden empirisch beschouwd inderdaad problematisch is.

5 Schematiciteit en de Usage-based conceptie van de Mentale Grammatica

Modellen als IPG stoelen op de aanname dat, wanneer een uiting gezien *kan* worden als het resultaat van algoritmische berekening door syntactische procedures, dat dan ook de praktijk is van de productie ervan: het model is maximaal structuurgenererend (zie § 3). De stelling die we in dit artikel verdedigen is dat de aard van de talige kennis die sprekers bij het produceren van uitingen inzetten, niet noodzakelijk beperkt behoeft te zijn tot de in § 3 beschreven conceptie van syntactische procedures die werken op minimale lexicale input. We zullen in het vervolg betogen

14 Een verklaring daarvoor zou gezocht kunnen worden in de wat ambigue status van voorzetsels als zowel functionele als lexicale elementen. Zie § 6.2 hierna en ook Schilperoord (1996:261 evv.).

dat de gepresenteerde empirische data suggereren dat de cognitieve status van de linguïstische structuren die sprekers produceren er één is van *gevestigde, declaratieve schema's*, die *als zodanig* deel uitmaken van het declaratieve taalkennisbestand,¹⁵ en ook *als zodanig* zijn betrokken bij het proces van taalproductie. Aan dergelijke schematische constellaties kan in beginsel een bepaalde constituent-structuur toegekend worden, en het is ook in beginsel mogelijk dat die structuur berekend wordt bij de productie (of interpretatie) ervan. De vraag is echter, of dat ook de praktijk is. In Bolingers woorden:

It is very likely that much more of the talk that goes on is idiomatic than just those utterances that can't be analyzed. The fact that we can analyze doesn't necessarily mean that we do. (Bolinger 1975:297)

In deze paragraaf stellen we aan de orde hoe de notie van declaratieve schema's taaltheoretisch begrepen kan worden. Het relevante verschil tussen de computationele visie en de schemavisie op de aard van taalkennis heeft met name betrekking op de modellering van het mentale lexicon. Aan de orde is dus de verhouding tussen wat 'berekend' wordt (het computationele aspect), en wat er 'kant-en-klaar' ligt aan taalkennis (het geheugen-aspect). Die verhouding valt adequaat te karakteriseren aan de hand van het door Langacker ingevoerde onderscheid tussen een *minimalistische*¹⁶ en een *maximalistische* conceptie van de 'Mental Grammar' (Langacker 1988, 1991a).¹⁷ Volgens de minimalistische conceptie is de Mentale Grammatica 'reduced to the smallest possible set of statements, with all redundancy avoided (...) anything which follows from general statements is omitted from the grammar, on the assumption that it is computed rather than being represented individually' (ibid. 129). Merk op dat dit een adequate karakteristiek is van een psycholinguïstisch zinsproductiemodel als IPG. De maximalistische conceptie daarentegen stelt:

It is plausible, psychologically, to suppose that speakers represent linguistic structures in different ways, with considerable redundancy built in. It is also reasonable to assume that many structures are learned as established units even when they also follow from general principles – the computability of a structure does not in principle

15 Wij zijn ons er uiteraard van bewust dat het type data dat we hier presenteren nooit exclusieve evidentie kan bieden voor het 'bestaan' van schematische constructies, of voor het gebruik ervan in feitelijke taalproductie of -verwerking. Naar ons oordeel leveren taalgebruiksdata als pauze-locaties *flankerende* evidentie voor gedachten die in recente linguïstische paradigma's, zoals de cognitieve grammatica en de constructiegrammatica, meer en meer gemeengoed aan het worden zijn (Fillmore, Kay & O'Connor 1988, Goldberg 1995, Jackendoff 1997). Dat gedachtengoed wordt binnen deze paradigma's ontwikkeld en van evidentie voorzien, bijvoorbeeld op basis van analyses van zogeheten 'constructional idioms', zoals resultatieve werkwoordconstructies ('He sneezed the napkin off the table'). Andere evidentie voor constructies is afkomstig van corpus-linguïstisch onderzoek naar collocaties, zoals bijvoorbeeld in Altenberg (1993).

16 Deze term moet niet worden verward met de notie 'minimalisme' waarmee het meest recente generatieve paradigma wordt aangeduid (cf. Chomsky 1995).

17 Het verschil tussen 'mentaal lexicon' en 'mentale grammatica' is terminologisch en niet conceptueel van aard. In beide gevallen wordt het declaratieve taalkennisbestand bedoeld.

preclude its learnability and inclusion as a distinct element in the cognitive representation of the linguistic system. (Langacker 1988:129)¹⁸

Een methodologische implicatie is dat taalgebruiksdata direct significant zijn met betrekking tot theorievorming aangaande de wijze waarop sprekers taalkennis representeren. De notie 'redundantie' houdt in dat procedures voor het 'berekenen' van, bijvoorbeeld, NP-structuren vrijelijk voor kunnen komen naast 'geleerde' *uitkomsten* van de toepassing van de procedures, en dat zulke geleerde uitkomsten niet beperkt hoeven te blijven tot het domein van idiomen. Of, in Bolingers termen: veel meer taalkennis is 'idiomatisch' gerepresenteerd dan alleen die uitingen waarvan de kennis als zodanig gekarakteriseerd *moet* worden. Merk op dat deze visie vergaande consequenties heeft voor een productiemodel als IPG. De strikte scheiding tussen een 'minimaal' mentaal lexicon en een 'maximaal' structuur-genererende procedurele grammatica wordt aanzienlijk genuanceerd. Bovendien wordt de mogelijkheid gesuggereerd dat taalproducenten, in termen van het IPG-model, a.h.w. een bocht afsnijden en van conceptualisatie direct overstappen op articulatie, dus de formuleer-stap overslaan, eenvoudig omdat dat mogelijk is doordat ze een 'established unit' kennen die direct met een bepaalde betekenisstructuur geassocieerd is.

Het theoretisch relevante kader waarbinnen ideeën als deze kunnen worden gesitueerd, is dat van de *Cognitieve Grammatica*, en meer in het bijzonder de *Usage-based* conceptie van de Mentale Grammatica (Deane 1992, Langacker 1988, 1991a). Uitgaande van de leerbaarheid van structuren, en hun potentie als eenheid te worden gerepresenteerd, omschrijft Langacker de mentale grammatica als een 'structured inventory of conventional linguistic units' (Langacker 1991b:548). De gestructureerdheid houdt in dat geleerde expressies bestaan naast, maar gerelateerd zijn aan abstracte schema's of procedures waarmee zulke expressies 'uitgerekend' kunnen worden. Veelvuldig gebruikte expressies verkrijgen de eenheidsstatus, terwijl hun overeenkomsten gerepresenteerd zijn als een abstract schema. Het voorbeeld dat in Langacker (1988) wordt gegeven betreft een veelvuldig gebruikte (Engelse) meervoudsvorm als 'cats', die is gerepresenteerd als het declaratieve betekenis/vorm-paar [[CAT][PL]/[cat][s]] en is gerelateerd aan een abstract schema voor meervoudsvormen:¹⁹ [[*_{THING}][PL]/[{phon.repr.}][s]]. Zeer frequente meervoudsvormen worden daarom geacht *gefixeerd* te zijn. Het al dan niet activeren van dat schema is in hoge mate een zaak van taalgebruiksomstandigheden. In Langackers visie ligt het ten grondslag aan de 'berekening' van 'nieuwe', dat wil zeggen *niet* gefixeerde

18 Merk op dat de tegenwerping als zou Langacker hier een 'performance'-grammatica beschrijven, en de minimalistische conceptie een 'competence'-grammatica, geen hout snijdt voorzover het de karakterisering betreft van verschillende taalproductiemodellen. Een model als IPG is per definitie een performance-grammatica.

19 We gebruiken hierin gemakshalve de aanduiding "*_{THING}" voor de abstracte notie "bounded region in some domain" waarmee Langacker de conceptuele inhoud van de categorie 'nomen' omschrijft (de asterisk geeft aan dat dit niet volledig samenvalt met de betekenis van het engelse woord thing), en "{phon.repr.}" voor de fonologische representatie van de enkelvoudsvorm van een willekeurig nomen.

meervoudsvormen als 'quagmires': $[[[*_{\text{THING}}]_{\text{PL}}]/[\{\text{phon.repr.}\}]_{\text{S}}] \rightarrow [\text{QUAGMIRE}]_{\text{PL}}/[\text{quagmire}]_{\text{S}}$.²⁰

De hierboven weergegeven cognitief-linguïstische ideeën over schematiciteit zijn niet beperkt tot meervoudsvormen – schema's kunnen eveneens kennis van constituent- en zinstypen representeren.²¹ Bovendien is de mate van schema-abstractie variabel: een schema kan bijvoorbeeld meer of minder lexicaal ingevuld zijn.²² Om dit te illustreren zullen we kort de twee productietranscripten bespreken die zijn weergegeven in (9)a en (9)b.

- (9) a 1. In antwoord op uw schrijven aan mij van /
 2. 14 oktober jongstleden
 (...)

 b (...)
 1. in /
 2. de procedure voor /
 3. het /
 4. beantwoorden van /
 5. de post van /
 6. cliënten
 (...)

Sequentie (9)a is een sjabloon waarmee zakelijke brieven vaak openen. De relevante observatie is dat de pauze in de sequentie direct valt voor de enige *variabele* 'invulplaats': de datum. Bezien we nu de voorzetselconstituent in wording in (9)b. Details daargelaten kan hieraan in beginsel dezelfde syntactische structuur toegewezen kan worden als aan (9)a. Het feitelijk tot stand komen ervan wijkt echter aanzienlijk af en is heel wat 'incrementeler'. De relevante observatie in dit geval is dat 4 van de 5 pauzes optreden *na* de voorzetsels.

20 Twee onmiddellijke consequenties hiervan zijn dat a) de mentale grammatica van spreker tot spreker kan verschillen, en b) dat de mentale grammatica geen stabiel, maar een inherent dynamisch systeem is: uitingen kunnen a.h.w. 'on-line' in een gesprek de status van eenheid verkrijgen, maar die, na verloop van tijd, ook weer verliezen.

21 Merk op dat de meer abstracte schema's, die het meest overeenkomen met grammaticale regels in de meer traditionele zin, beperkingen opleggen aan de concretere, en omgekeerd. De abstractere moeten afleidbaar zijn van de concretere, omdat ze anders niet meer tot dezelfde categorie behoren; ook vorm en betekenis van de concretere schema's zijn daarom in overeenstemming met 'the facts of language and the specific properties of syntactic representation' (Deane 1992:138). Bovendien is het natuurlijk zo dat het ook voor communicatieve doelen functioneel is dat de vormelijke en semantische structuur van gefixeerde eenheden in overeenstemming is met de algemene regels van de taal: dat maakt ze 'berekenbaar' – en dus in beginsel interpreteerbaar voor ook andere leden van de taalgemeenschap, die ze niet per se *als schema's* hoeven te kennen. Hiermee houdt verband dat de hier te behandelen schema's geen nieuwe grammaticale feiten aan het licht zullen brengen (daar gaat het ons hier ook niet om). Ze zijn weergegeven volgens de feiten van, in dit geval, de Nederlandse taal. Het gaat ons dan ook niet om de (syntactische) eigenschappen van de schema's als zodanig, maar om de wijze waarop ze door moedertaalsprekers gekend worden.

22 De multi-representationele aard van schema-informatie is een van de centrale aannames in de cognitieve linguïstiek (Langacker 1987, Harris 1992).

De overeenkomst tussen beide expressies kan binnen het *Usage-based* kader verantwoord worden door een abstract voorzetselschema aan te nemen waarmee beide expressies zijn verbonden (zie (10)).

(10) PP:[prep + np]

Expressie (9)a verhoudt zich tot (10) in veel opzichten zoals het meervoud 'cats' zich verhoudt tot Langackers algemene schema voor meervoud. Dat wil zeggen, (9)a is een (vrijwel) volledig *gefixeerde* expressie die in de mentale grammatica als declaratieve eenheid is gerepresenteerd (zie (11)).²³

(11) [[IN ANTWOORD OP UW SCHRIJVEN AAN MIJ VAN][DATUM]]/[*in antwoord op uw schrijven aan mij van*][{datumstring}]].²⁴

Maar terwijl schema (10) hoogstens afleidbaar is uit (9)a, lijkt het bij de productie van (9)b actief te zijn betrokken. We kunnen veronderstellen dat 'procedures voor het beantwoorden van de post van cliënten' een relatief 'nieuwe' expressie is, die aan de hand van een recursieve toepassing van het PP-schema gevormd wordt. Het schematische karakter van de bouwsteen [prep + np] wordt sterk gesuggereerd door de locatie van de pauzes: het lijkt hier te gaan om instanties van schema (10) die geassocieerd zijn met bepaalde preposities 'voor', 'van', enzovoorts. Het is bovendien denkbaar dat die instanties collocatieel gerelateerd zijn aan de substantieven: de advocaten in het corpus zijn het gewoon te communiceren over 'procedures voor...', of het 'beantwoorden van...'. Los daarvan zijn schema-extensies van (10) als volgt voorstelbaar.

(12) PP:[voor/op/van/... + np]

Het schematische karakter van (12) komt vooral tot uiting in de variabele 'np'. De *usage-based* status van erop gebaseerde expressies ligt derhalve ergens tussen die van Langackers 'cats' en 'quagmires'; het gaat hier om *deels gefixeerde* expressies (zie (13)).

(13) a. [PROCEDURES VOOR + *DING]/[*procedures voor* + np]
b. [BEANTWOORDEN VAN + *DING]/[*beantwoorden van* + np]

Op basis van de variabele 'abstractiegraad' kunnen dus drie typen schema's worden onderscheiden ((10), (11) en (12) vormen voorbeelden). Allereerst algemene, abstracte schema's die elk bepaalde eigenschappen representeren van (Nederlandse)

23 Dit geldt voor clichés in het algemeen (zie Jackendoff 1995, Kuiper 1995, Schilperoord & De Vet 1996). Clichés zoals 'In antwoord op uw schrijven aan mij van NP...', beschouwen we dus als onderdeel van de mentale grammatica, *hoewel* de betekenis is af te leiden uit de samenstellende delen, waardoor ze *in principe* analyseerbaar/berekenbaar zijn. Het open karakter van zulke constructies is, hoewel zeer begrensd, evident: gegeven de pragmatische gebruiksvoorwaarden kan het prepositionale object vrijelijk ingevuld worden.

24 Dus te vergelijken met het sjabloon [we gaan vanavond praten over [X] en de problemen die daarbij komen kijken].

syntactische categorieën (nominale groepen, prepositiegroepen, bijzinnen, enz.). Voor een constituent van de gegeven categorie legt het schema minimale beperkingen op, en is het dus maximaal productief. Ten tweede zijn er schema's die deels variabel zijn, maar ook deels lexicaal vast zijn ingevuld. Deze schema's leggen enige beperkingen op aan de invulling van de rest van de constituent, gegeven het lexicale element (zoals we zullen zien: het functiewoord), en zijn dus partieel productief. Tenslotte zijn er schema-instanties die representaties vormen van vrijwel geheel gefixeerde expressies, met niet meer dan één variabele, waarvan de waarde uit een beperkte set gekozen moet worden. Tegenover het invariante karakter van dergelijke schema's staat een hoog gebruiksgemak: een gefixeerde expressie als "in antwoord op uw schrijven aan mij van [datum]" is sterk begrensd in zijn gebruiksmogelijkheden, maar *daarbinnen* als eenheid direct oproep- en bruikbaar.

Wat hierbij vooral van belang is, is dat het verschil in abstractiegraad ook het enige relevante verschil is tussen de drie soorten schema's. De overeenkomst tussen de drie typen, en alle tussenliggende gevallen, is dat bepaalde vormelijk te specificeren informatie geassocieerd is met een bepaalde structuur waarin plaats is voor meer of minder variabele gegevens. Die associatie is een kwestie van herinnering, in alle gevallen; de verschillen in abstractheid impliceren niet noodzakelijk een verschil in cognitieve status: de schema's kunnen allemaal herinnerd worden (terwijl anderszids ook erkend-gefixeerde gevallen in principe 'berekenbaar' blijven).

We sluiten deze paragraaf af met het opsommen van een reeks (mogelijke) schema's in de hier begrepen zin. Uit de behandeling tot dusverre kan worden geconcludeerd dat schema's beschreven kunnen worden in termen van de volgende drie aspecten:

- (14) 1. het functiewoord (de categorie of een concrete instantie);
2. een positie voor een variabele die 'verplicht' ingevuld moet worden (doorgaans de lexicale kern);
3. optionele invulplaatsen.

De algemene gedaante van een schema is dan (15), waarin Z een bepaalde categorie is, X het daarbij horende functiewoord en Y de lexicale kern.

(15)

Z-categorie: $[[X\text{-categorie (X-instantiatie)}] + \dots [Y\text{-categorie (Y-instantiatie)}] \dots]_{Z\text{-instantiatie}}$

Het meest abstracte schema bestaat uit louter categorieën; deels gefixeerde schema's bestaan uit een X-instantiatie en een Y-categorie; volledig gefixeerde expressies vormen in hun geheel een instantiatie van het algemene schema. Voorbeelden van (15) zijn opgesomd in (16).²⁵

- (16) a. NP:[det + ... + N] (det = {*de, het, een, ...*})
- b. PP:[prep + (16)a] (prep = {*in, aan, op, onder, ...*})
- c. PP:[prep + (16)b] (voor een deelverzameling van prep)

25 Merk op dat ook cross-schematische recursie mogelijk is, zie bijvoorbeeld (16)c.

- d. S:[comp₁ + finiete clause]
 (comp₁ = {*dat*, *of*, *omdat*, *nadat*, *hoewel*, ...})
 S:[*dat* + ... + V_{fin}]
 e. S:[comp₂ + infinitie clause]
 (comp₂ = {*om*, *door*, ..., ?})
 S:[*om* + ... + (*te* + V_{inf})]

Schema's als in (16) zijn naar ons idee vergelijkbaar met 'woorden' uit het mentale lexicon, of, meer in het algemeen gesteld, met al die talige elementen waarvan gesteld wordt dat ze niet op basis van de regels te vormen zijn. Schema's lijken op zulke elementen in die zin dat ze als geheel uit het geheugen opgeroepen kunnen worden, en ook op die wijze in het proces van taalproductie zijn betrokken. Dat houdt intussen ook in dat er conceptuele/semantische omstandigheden expliciteerbaar (moeten) zijn die aan de activatie van schema's ten grondslag liggen. Kortom, aan de hier weergegeven schema's moet een bepaalde abstracte betekenis toegekend kunnen worden. In de slotparagraaf betogen we dat functiewoorden eenvoudig de elementen zijn die de abstracte betekenis van schema's oproepen: gegeven de associatie van functiewoorden met syntactische schema's kunnen de noties 'betekenis van een constructie' en 'betekenis van een functiewoord' in feite samenvallen. Dit betekent dat we ons vanaf nu beperken tot deels gefixeerde schema's.

6 De schema-status van functiewoorden

6.1 Lidwoorden

We bezien nogmaals het NP-schema, herhaald als (17):

(17) NP:[det + ... + N] (det = {*de*, *het*, *een*, ...})

Wanneer het nu niet (noodzakelijk) het concept uitgedrukt in het lexicale hoofd N is dat aan de constructie van NPs ten grondslag ligt, maar een andere conceptuele specificatie, wat zou dat laatste dan kunnen zijn? Naar ons oordeel zal die specificatie de door Langacker ingevoerde notie *grounding* omvatten. In Langacker (1991a) merkt hij daarover het volgende op:

[...] In the case of English nominals, grounding is effected by articles, demonstratives, and certain quantifiers. Whereas a simple noun (for example *cat*) merely names a "type" of thing, a full nominal (*this cat*; *some cat*; *any cat*; *a certain cat*) designates an "instance" of that type and gives some indication of whether and how the speech-act participants have succeeded in establishing mental contact with that particular instance (id.:321).

Een lidwoord bij een substantief heeft dus het effect dat het met dat substantief uitgedrukte concept een 'token'-specificatie wordt van het algemene 'type'. Het wordt dus op een bepaalde manier gerelateerd aan kenmerken van de taalgebruikssituatie (de "ground") en de participanten daarin ("Over wat voor geval van *cat* hebben wij het hier en nu?"). Merk op dat dit effect dus typerend is voor taalgebruik: sprekers hebben het in geval van 'things' over een 'thing-token', en 'mental contact' kan,

althans in communicatief relevante (manifeste) zin, alleen *door* taalgebruik, en *met* thing-tokens verkregen worden. Grounding is derhalve een elementair onderdeel van de (on-line) conceptualisatie van 'things'. In Langackers optiek wordt zo'n grounding-effect met name bewerkstelligd door 'functionele' elementen:

[...] We can reasonably suppose that only "grammaticized" (as opposed to "lexical") elements can serve as true grounding predications (id.:322).

Schema (17) lijkt derhalve een consequentie te zijn van het feit dat NPs de prototypische wijze voor het aanduiden van 'things' zijn, en dat 'grounding' een onvervreemdbaar onderdeel is van het spreken over 'things'; 'grounding' legt per definitie een relatie tussen een concept en iets anders, i.c. de taalgebruikssituatie, en een element met de functie van een 'grounding predication' roept dus noodzakelijkerwijze de notie van een ander, nader in te vullen concept op, dat wil zeggen een schema. Het variabele en productieve betekenis/vorm-paar voor 'things' heeft dan de gedaante (18) (vergelijk noot 19 voor de notatie).

(18) [[GROUNDED, *THING_i]/[{*de, het, een,...* }_i + N_i]]

Dit schema is dan een deels gefixeerde instantie van het algemene schema:

(19) [[[GROUNDING] + *THING]/[(det) + N]]

Gesteld nu dat dit een plausibele redenering is, dan volgt eruit dat het bepaalde of onbepaalde lidwoord (de 'grounding predication' in Langackers terminologie) in wezen een *schematische* NP is. En omdat functionele elementen uniek zijn voor het type constituent waarin ze voorkomen, kunnen we lidwoorden en andere functiewoorden zien als *schematische profiel-determinanten* van het betreffende schema.²⁶ Dat schema kan, in geval van een NP, los van de conceptuele specificatie van het [*THING/N]-deel geactiveerd worden, namelijk op basis van conceptuele grounding. Grounding is dan een conceptuele operatie die in de ware spreektijd *vooraf* kan gaan aan de THING-specificatie. Die laatste operatie kunnen we dan veronderstellen te zijn gereflecteerd in de *pauze* die na het lidwoord (zo dikwijls) volgt.

Nu kan met betrekking tot substantief-groepen tegen deze redenering worden ingebracht dat de keuze voor 'de' dan wel 'het' in het Nederlands wordt bepaald door het geslacht van het lexicale hoofd zodat lexicalisatie van de kern al plaats moet hebben gevonden om tot de juiste keuze te komen. Om zo'n dispuut te beslechten kunnen twee soorten gegevens in nadere beschouwing worden genomen, die we hier echter slechts kort aanstippen. Het eerste gegeven heeft te maken met de *gebruiksfrequentie* van de beide lidwoorden (Uit den Boogaard 1975). Op basis daarvan valt te veronderstellen dat het '*de*-NP-schema' een default-prioriteit heeft boven het '*het*-schema'. Die aanname wordt ondersteund door een specifiek type gebruiksdata in ons corpus – te weten *reparaties* van lidwoorden die 'on the fly' aangebracht worden. Twee voorbeelden hiervan staan in transcript (20).

26 Voor de notie 'profiel-determinantie', zie Langacker (1987, 288evv).

(20) a. *de* → *het*

- (...)
 1. Het is jammer /
 2. dat de rechtbank niet /
 3. **de** /
 1-2*. ((jammer dat de rechtbank niet)) /
 4. **het** argument /
 (...)

b. *het* → *de*

- (...)
 1. welke terreinen /
 2. **het** meeste /
 1*. ((welke terreinen)) /
 3. de meeste mogelijkheden openen /
 (...)

In (20)a wordt het ongeschikte '*de*' vervangen door '*het*'; in (20)b gebeurt het omgekeerde. Een lexicaal gestuurd productiemodel zou gevallen als deze kunnen verantwoord worden door aan te nemen dat de '*ware*' trouble spot het substantief is, en niet het lidwoord. De tekstproducent in (20)a repareert na regel 3 dus een nog niet gerealiseerd substantief, en zet daarvoor '*argument*' in de plaats, of hij besluit on-line de meervoudsvorm '*argumenten*' te vervangen door het enkelvoudige '*argument*'. Wanneer dit het mechanisme is, wordt voorspeld dat beide typen reparaties even vaak voor zullen komen.²⁷ De schema-voorspelling, met het '*de*-schema' als default, voorspelt echter dat de reparatie *de* → *het* vaker voorkomt dan de omgekeerde, vanwege het default-karakter van het *de*-schema. Nu bevat het corpus in totaal slechts vijf van deze lidwoord-reparaties (waaronder de twee bovenstaande), maar vier daarvan zijn inderdaad van het type *de* → *het*.

Het tweede type gegevens heeft betrekking op een bepaalde regelmaat in de distributie van '*de*' en '*het*' die samenhangt met de semantiek van het substantief: '*het*-woorden' duiden vaker algemene, of abstracte zaken aan dan '*de*-woorden' (o.a. Zubin & Köpcke 1986), wat in ieder geval historisch overeenstemt met het 'kenmerkloze' karakter van het neutrum. Wanneer lidwoorden, zoals in deze paragraaf is betoogd, inderdaad grounding-predicaties zijn, dan lijkt het niet onaannemelijk

27 De reden is de volgende. Stel dat er 80% '*de*'-woorden, en 20% '*het*'-woorden zijn. De kans dat een '*de*'-woord geproduceerd dan wel gerepareerd wordt, is dan vier maal zo groot als de kans op productie/reparatie van een '*het*'-woord. De kans dat vervolgens het reparandum van hetzelfde geslacht is, is dan voor '*de*'-woorden $.8 \times .8 = .64$, en voor de '*het*'-woorden $.2 \times .2 = .04$. De kans op een switch van een '*de*' naar een '*het*'-woord is $.8 \times .2 = .16$. De kans echter op een switch van een '*het*'-woord naar een '*de*'-woord is precies even groot: $.2 \times .8 = .16$. Kortom, een model van productie dat bekendheid van het geslacht van een nomen vereist om het lidwoord te bepalen, voorspelt dat een '*de*' → '*het*'-reparatie even vaak voorkomt als een '*het*' → '*de*'-reparatie.

dat dit zeer algemene betekenis-onderscheid als het ware 'in' het betreffende lidwoord gebakken zit. Het onderscheid manifesteert zich vrij duidelijk op het punt van nominalisaties. De *het*-nominalisatie van een werkwoord als 'beslissen' duidt eerder op het algemene *proces* van beslissen ('het beslissen van de rechter van Assen'), terwijl de *de*-nominalisatie ervan eerder de gedachte oproept aan een *concrete* instantie daarvan ('de beslissing van de rechter van Assen').²⁸ Denkbaar is dat in bepaalde contexten een dergelijk conceptueel onderscheid bij de productie van nominale groepen een rol speelt reeds *voor* de lexicalisatie van het kern-concept (die dan wellicht juist mede gestuurd wordt door de geslachtskeuze). Hoe dit zij, het is in ieder geval logisch evenzeer mogelijk dat de keuze van een lidwoord vervolgens de selectie van de daarbij passende nomina beperkt als andersom, en dat er daardoor betrekkelijk weinig reparaties op hoeven treden. Het is overigens ook duidelijk dat het pragmatische en psycho-linguïstische onderzoek naar de mogelijke functie(s) van grammaticaal geslacht in ieder geval voorlopig een nogal gecompliceerde aangelegenheid is (vgl. Van Berkum 1996).

6.2 Voorzetsels (en voegwoorden)

De functie van voorzetsels lijkt niet die van 'grounding' te zijn – ze relateren niet een bepaald concept aan abstracte kenmerken van de taalgebruikssituatie. Eerder relateren ze een instantiatie van het ene concept aan die van een ander. Typische 'voorzetsel'-functies zijn daarbij het aangeven van temporele of ruimtelijke relaties van allerlei aard, en metaforische toepassingen daarvan. Doordat soms het grammaticaal-relaterende aspect van voorzetsels de overhand lijkt te hebben (in gevallen als *de overdracht van de olifanten aan het stamhoofd door de stropers*), en soms het conceptueel-inhoudelijke (*het verblijf onder water gedurende de dag*), krijgen voorzetsels vaak een enigszins aparte, of ook wel ambivalente behandeling (vgl. Sturm 1986). Ook in IPG is zoiets het geval.

Kempen & Hoenkamp (1987) delen de voorzetsels in in twee groepen. De ene wordt gevormd door de korte en meest frequente voorzetsels (door hen 'clitics' genoemd), dus *van, in, op* e.d.; deze hebben volgens hen uitsluitend een grammaticale functie, en daarom worden ze geproduceerd door de functorisatie-procedures, die dus na lexicalisatie werken (vgl. Kempen & Hoenkamp 1987:218/9). De tweede groep wordt gevormd door de langere en minder frequente voorzetsels (typische gevallen: *gedurende, ondanks*, e.d.); deze rekenen Kempen & Hoenkamp wel tot de inhoudswoorden.

In het licht van onze eerdere uiteenzetting (§ 3) kunnen we dus zeggen dat in deze vorm IPG systematisch verschillen voorspelt over pauzeer-patronen bij voorzetsels. Na de korte, grammaticale voorzetsels vallen, omdat ze voortgebracht worden door de functorisatieprocedures die *volgen* op lexicalisatie, *geen* pauzes te verwachten, na de lange, lexicale voorzetsels zou dat daarentegen heel wel mogelijk zijn. In ons corpus valt van zo'n verschil echter weinig te bespeuren. Een blik op de in dit artikel gegeven voorbeelden (2) en (9) laat ook al voldoende zien dat er na 'grammaticale' voorzetsels makkelijk gepauzeerd kan worden (*om, in, voor, van* (4x)). De empirie van pauzeerpatronen biedt dus geen steun voor de verdeling van voorzet-

28 Merk op dat de '*de*-nominalisatie' ook veel ruimere predicatie-mogelijkheden biedt.

sels over enerzijds het lexicon en anderzijds de procedurele grammatica.²⁹

Dat neemt niet weg dat preposities wel degelijk een 'tussenpositie' zouden kunnen innemen. In tabel 1 is een aanwijzing daarvoor ook dat er *voor* voorzetsels vaker gepauzeerd wordt dan *erna*, dit in tegenstelling tot de andere onderscheiden functiewoorden. Onze benadering biedt ook een heel direct aanknopingspunt voor zo'n gedachte. De essentie van het schema-idee is dat een schema een gedeeltelijk gefixeerde structuur kent, met daarin plaats voor variabele informatie. Daarin stemmen het NP-schema en het PP-schema overeen. Een verschil is echter dat voorzetsels, zoals gezegd, typisch relaties tussen instantiaties van concepten leggen, en daarmee bijdragen aan de conceptuele inhoud van de uiting, en niet zozeer aan de relatie ervan met de taalgebruikssituatie (zoals determinatoren doen). In die zin zijn voorzetsels dus zowel functiewoorden (d.w.z. profiel-determinanten van schema's) als inhoudswoorden (bijdragen aan de conceptuele inhoud); in onze opvatting van die termen is daar niets tegenstrijdigs in.

Een aparte kwestie in dit verband is het verschijnsel van de zgn. vaste voorzetsels (m.n. bij voorzetselvoorwerpen, maar ook elders). Gegeven het idee dat schema's in verschillende gradaties van abstractheid kunnen voorkomen, inclusief varianten die gedeeltelijk lexicaal gefixeerd zijn, is het verschijnsel weinig meer dan een illustratie hiervan; (vrijwel) alle volwassen sprekers van het Nederlands beschikken over het schema "antwoord(en) op X" (en "praten over X"). In een model als IPG, met een principiële scheiding tussen lexicon en grammatica, is de status van dergelijke uitdrukkingen veel meer een probleem; Kempen & Hoenkamp (1987) zijn er ook niet erg expliciet over; ze behandelen impliciet wel de combinatie *listen to* alsof het één werkwoord is, maar doen over dit type verschijnselen geen algemene uitspraak. Van belang is allereerst uiteraard dat naast de lexicale 'eenheid-status' van dergelijke uitdrukkingen, ook de ontleedbaarheid ervan in de theorie gehandhaafd blijft (*antwoord* in *antwoord op* betekent nog steeds hetzelfde als *antwoord* op zichzelf, of in *antwoord van*, etc.) - iets dat in de maximalistische conceptie (zie § 5) vanzelf spreekt. Hiermee hangt, als tweede punt, samen dat substantieven en werkwoorden, zoals *antwoord(en)*, *besluit(en)*, *praten*, wel krachtige beperkingen op kunnen leggen aan de voorzetsels die erop kunnen volgen, maar dat die nog niet 100% voorspelbaar zijn; naast *besluiten om* hebben we b.v. ook nog *besluiten tot* en *besluiten over*. Dat betekent dat de feitelijke keuze van een voorzetsel, ook daar waar die keuze beperkt is, toch altijd een nadere keuze voor het vervolg met zich meebrengt, dus een projectie van specifiekere structuur: bij *besluiten om* legt de producent zich vast op een infinitief-constructie, bij *besluiten tot* op een nominale groep (met bepaalde semantische eigenschappen). Dit soort gevallen is gemakkelijk inpasbaar in de schema-benadering voor de productie van grammaticale constituenten: de gedachte is dat de taalgebruiker zich met de keuze van een functiewoord (in de zin van schematische profiel-determinant) vast kan leggen op een bepaald *type* relatie (in de vorm van een bepaald syntactisch schema) terwijl de conceptuele specificatie daarvan nog moet plaats vinden, hetgeen de oorzaak kan zijn van een pauze. In een IPG-

29 Naast dit empirische probleem geldt natuurlijk ook nog de vraag of het theoretisch wel verantwoord is de leden van één grammaticale klasse, i.c. die van de preposities (die zich syntactisch toch erg gelijksoortig gedragen), zo over 'lexicon' en 'syntaxis' te verdelen.

achtige benadering, waarin syntactische informatie niet beschikbaar kan komen zonder dat er conceptuele inhoud gespecificeerd is, leveren dergelijke verschijnselen stelselmatig een probleem op.³⁰

De generalisatie in het semantische vlak over lidwoorden en voorzetsels is dus dat die een wezenlijk relationeel karakter hebben (zij het met betrekking tot verschillende domeinen), hetgeen direct samenhangt met het feit dat ze syntactische schema's specificeren. Ook onderschikkende voegwoorden kunnen vermoedelijk goed in deze analyse ondergebracht worden. Wellicht moet daarbij een onderscheid gemaakt worden tussen enerzijds onderwerps- en voorwerpszinnen en anderzijds bijwoordelijke bijzinnen, waarbij de functie van de eerste meer in de richting van 'grounding' gaat, terwijl de tweede soort functioneel overeenkomt met voorzetselgroepen. Mogelijk zou dit een verklaring kunnen vormen voor het feit dat pauze-proporties na voorzetsels minder groot zijn dan na complementeerdere; omdat we op dit punt in de data niet de relevante onderscheidingen kunnen maken, laten we het hier echter bij deze speculatieve opmerkingen.

7 Tot slot

In dit artikel hebben we betoogd dat syntactische schema's (constituenten, zinnen) in uiteenlopende graden van abstractie deel uitmaken van het declaratieve lexicon, en als zodanig betrokken zijn bij het proces van taalproductie. Daarnaast hebben we de prevalentie van functie-woorden bij schema-activatie en -gebruik aan de orde gesteld. Wij zijn, zoveel moge duidelijk zijn, niet de eersten om de mogelijkheid van schema's te onderkennen en te onderzoeken (zie bijvoorbeeld Bolinger 1975, Langacker 1987, Deane 1992). Wel nieuw is, naar ons weten, het type data – pauzes, in spontane taalproductie, na functiewoorden – dat we in dit artikel aan de orde hebben gesteld als mogelijk relevant in het onderzoek naar de verhouding tussen 'geheugen' en 'berekening'. Verder onderzoek is nodig naar de mogelijkheid dat het gebruik van gefixeerde uitdrukkingen en schematische constellaties (mede) afhankelijk is van bepaalde spreek-psychologische factoren, zoals de vereiste *snelheid* van spreken,³¹ of de reikwijdte van het semantische domein dat aan de taalproductie ten grondslag ligt (zie bijvoorbeeld Blaauw 1995). Hoe het zij, wanneer de aannemelijkheid van de schema-benadering verder onderbouwd kan worden heeft dat tamelijk verstrekkende gevolgen voor de modellering van het productieve taalvermogen. Die gevolgen zouden allereerst betrekking kunnen hebben op de verhouding 'lexicon-grammatica', dus op de rol van het 'onthouden' en die van het 'berekenen' in taalgebruik. De tijd lijkt rijp voor zo'n heroverweging, zoals ook moge blijken uit de conclusie die Jackendoff (1995, zie ook Jackendoff 1997) trekt op basis van een analyse van idiomen, clichés en grammaticale constructies:

[...] the boundary between lexicon and rules of grammar begins to blur. Along with Fillmore & Kay (1993) and in some sense Langacker (1987), one might even want to

30 Stel bijvoorbeeld dat *om* (of *te*) opgevat wordt als de grammaticale markering van een infinitief-constructie, dan moet het geïntroduceerd worden door een functorisatie-procedure, zodat het infinitie werkwoord als hoofd van de constructie eerst gelexicaliseerd moet zijn, waardoor voorspeld wordt dat er na *om* (of *te*) geen pauzes zullen optreden – in tegenstelling tot wat de feiten laten zien.

31 Een suggestie die gedaan wordt door Kuiper (1995).

view the core rules of phrase structure for a language as maximally underspecified constructional idioms. (...) A strict separation of lexicon and grammar (...) may prove to be but a methodological prejudice. (Jackendoff 1995:155/6)

Bibliografie

- Altenberg, B. (1993). Recurrent verb-complement constructions in the London-Lund Corpus. In: J. Aarts, P. de Haan & N. Oostdijk (eds.), *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam/Atlanta, GA: Rodopi, 227-245.
- Bergh, H. van den (1990). *Unilevel versus multilevel analysis*. Paper gepresenteerd op de statistische dagen van de Nederlandse Vereniging voor Statistiek.
- Berkum, van, J.J.A. (1996). *The Psycholinguistics of Grammatical Gender. Studies in language comprehension and production*. Proefschrift K.U.N.
- Blaauw, E. (1995). *On the Perceptual Classification of Spontaneous and Read Speech*. Utrecht: OTS dissertation series.
- Bolinger, D. (1965). Linear modification. In: Bolinger, D., *Forms of English: accent, morpheme, order*. Cambridge MA: MIT Press, 279-307.
- Bolinger, D. (1975). *Aspects of Language*. New York: Harcourt, Brace, Jovanovich.
- Boomer, D.S. (1965). Hesitation and grammatical encoding. *Language and Speech* 8, 148-158.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge MA, London England: The MIT Press.
- Clark, H.H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H.H. & E. Clark (1977). *Psychology and Language. An introduction to psycholinguistics*. New York: Harcourt Brace.
- Deane, P.D. (1992). *Grammar in Mind and Brain. Explorations in cognitive syntax*. Berlin: Mouton de Gruyter.
- Fillmore, C.F., P. Kay & M.C. O'Connor (1988). Regularity and idiomacity in grammatical constructions. The case of *let alone*. *Language* 64, 501-538.
- Goldberg, A.E. (1995). *Constructions. A construction grammar approach to argument structure*. Chicago/London: The University of Chicago Press.
- Harris, C.L. (1992). Connectionism and cognitive linguistics. In: N.E. Sharkey, (ed.) *Connectionist Natural Language Processing. Readings from Connection Science*. Dordrecht: Kluwer Academic Publishers, 1-27.
- Jackendoff, R. (1995). The boundaries of the lexicon. In: M. Everaert, E. van der Linden, A. Schenk & R. Schreuder (eds.) *Idioms: Structural and Psychological Perspectives*. Hillsdale NJ: Lawrence Erlbaum Ass. Publishers, 133-166.
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge MA, London England: The MIT Press.
- Kempen, G. & E. Hoenkamp (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science* 11, 201-257.
- Kuiper, K. (1995). *Smooth Talkers. The linguistic performance of auctioneers and sportcasters*. Mahwah NJ: Lawrence Erlbaum Ass. Publishers.
- Langacker, R.W. (1987). *Foundations of Cognitive Grammar. Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R.W. (1988). A Usage-based model. In: B. Rudzka-Ostyn (ed.) *Topics in Cognitive Linguistics*. Amsterdam: Benjamins, 127-161.

- Langacker, R.W. (1991a). *Concept, Image, and Symbol. The cognitive basis of grammar*. Berlin/New York: Mouton de Gruyter. Cognitive Linguistic Research I.
- Langacker, R.W. (1991b). *Foundations of Cognitive Grammar. Vol. 2. Descriptive application*. Stanford: Stanford University Press.
- Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge MA: MIT Press.
- Maclay, H. & C.E. Osgood (1959). Hesitation phenomena in spontaneous English speech. *Word* 15, 19-44.
- Schilperoord, J. (1996). *It's about Time. Temporal aspects of cognitive processes in text production*. Amsterdam/Atlanta: Rodopi.
- Schilperoord, J. & D. de Vet (1996). Over clichés. *Tekst[blad]* 2, 5-9.
- Sturm, A. (1986). *Primaire syntactische structuren in het Nederlands*. Proefschrift Utrecht.
- Uit den Boogaart, P.C. (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht: Oosthoek, Scheltema en Holkema.
- Zubin, D.A. & K.M. Köpcke (1986). Gender and folk taxonomy: the indexical relation between grammatical and lexical categorization. In: C. Craig (ed.) *Noun Classes and Categorization*. Amsterdam: Benjamins, 139-180.