

Een web van woorden

Rede uitgesproken door

Prof dr. Geert E. Booij

bij de aanvaarding van het ambt van hoogleraar
op het gebied van de morfologie en het lexicon
aan de Universiteit Leiden
op 9 september 2005.

Mijnheer de Rector Magnificus, zeer gewaardeerde toehoorders,

De term ‘morfologie’ die voorkomt in de naam van mijn leerstoel ‘De morfologie en het lexicon’, zal niet direct iedereen het hart sneller doen kloppen. Sommigen van U denken bij dat woord aan een saai corvee, het leren van rijtjes woordvormen in het voortgezet onderwijs, bij talen als het Latijn en het Duits. Velen associëren het leren van vreemde talen met het leren van zulke rijtjes voor de vervoeging van werkwoorden en de verbuiging van naamwoorden. Zo begon mijn kennismaking met het Latijn in 1960 met de *Vormleer* van dr. K. van der Heyde, deel 1 van zijn *Latijnse Grammatica*, en pas na een jaar of twee kwam deel 2 over *Syntaxis* aan de orde, min of meer als een toegift. Morfologie, oftewel vormleer, had dus een centrale plaats in het taalonderwijs, leek belangrijker te zijn dan syntaxis, maar maakte tegelijkertijd niet bepaald de indruk een opwindend terrein van wetenschap te zijn, maar eerder een soort noodzakelijk kwaad, nodig om een taal goed te beheersen.

Morfologie is overigens niet beperkt tot de verbuiging of vervoeging van woorden, doorgaans aangeduid als het domein van de inflectie. Wat er ook bij hoort, is de vorming van nieuwe woorden, woordvorming dus.

De term ‘morfologie’ werd voor het eerst gebruikt in de biologie, en het Duitse genie Goethe was een van de eerste gebruikers ervan. In de taalwetenschap werd deze term voor het eerst gebruikt in 1859 door de Duitse taalwetenschapper August Schleicher (Salmon 2000). Begrippen en termen uit het vakgebied van de morfologie nemen van oudsher een belangrijke plaats in bij het taalonderwijs en bij de beschrijving van talen.

Bij morfologie gaat het, anders dan de soms gebruikte vertaling ‘vormleer’ suggereert, niet alleen om de vorm, maar breder, om de systematiek in het verband tussen de klank en de betekenis van woorden, en het gaat daarom altijd om gelede woorden, dat wil zeggen woorden met een interne structuur. Een woord als *werker* is een geled woord met een interne structuur, omdat de delen *werk-* en *-er* beide een eigen bijdrage leveren aan de betekenis van dit woord. Maar een woord als *erker* is geen geled woord, omdat het niet kan worden gesplitst in twee delen *erk-* en *-er*, die ieder een eigen betekenis hebben. Er is immers geen werkwoord *erken* is. Het woord *werker* heeft een interne structuur omdat het deel uitmaakt van een rijtje woordparen zoals:

(1)	<i>werkwoord</i>	<i>zelfstandig naamwoord</i>
	eet	eter
	fiets	fietser
	koop	koper
	werk	werker

waarin er een systematische correspondentie is tussen het vormverschil tussen de werkwoorden en de zelfstandige naamwoorden op *-er*, en het betekenisverschil tussen die rijtjes woorden: de zelfstandige

naamwoorden op *-er* noemen doorgaans personen of dingen die de door het werkwoord aangeduide handeling verrichten. Zo'n rijtje overeenkomstige woordparen noemen we een paradigma, een term die velen van U zullen kennen met een wat andere betekenis, namelijk als begrip uit de wetenschapsfilosofie. Als het in Leiden gaat over paradigmastrijd, dan is er geen sprake van een morfologisch debat.

Morfologie is dus een subdiscipline van de taalwetenschap, maar niet de meest populaire. Dat is wetenschappelijk gezien niet zonder meer gerechtvaardigd. Terecht stellen Spencer en Zwicky in de inleiding van hun recente handboek van de morfologie uit 1998 dat morfologie conceptueel gezien centraal moet staan in de taalwetenschap:

“Morphology is at the conceptual centre of linguistics. This is not because it is the dominant subdiscipline, but because morphology is the study of word structure, and words are at the interface between phonology, syntax and semantics.” (Spencer and Zwicky 1998: 1).

Dit is een belangrijke uitspraak, omdat bij sommige theoretische taalkundigen de neiging heeft bestaan, en nog bestaat, om morfologie als aparte subdiscipline weg te redeneren, en te verdelen tussen fonologie en syntaxis. Een recente vorm van die opvatting staat bekend onder de naam ‘distributed morphology’. Vanwege die neiging is de morfologie wel het Polen van de linguïstiek genoemd. Polen heeft immers het lot moeten dragen dat het ruim een eeuw van de kaart van Europa weggevaagd was. Deze vorm van reductionisme in de taalwetenschap heeft als groot gevaar dat er geen oog is voor de eigenwettelijkheid van het niveau van woordstructuur tegenover dat van zinsstructuur. Gelukkig staat Polen weer op de kaart, en dat geldt ook voor de morfologie, zoals ik U vandaag hoop duidelijk te maken. Ik zal dus aandacht vragen voor het onreducerebare van morfologische verschijnselen.

Spencer en Zwicky stellen in het hiervoor gegeven citaat het begrip ‘interface’ centraal als definiërend kenmerk van de morfologie. Met ‘interface’ wordt bedoeld dat de verschillende soorten informatie over taalbouwsels elkaar kunnen zien, dat er een systematisch verband tussen die soorten informatie bestaat. Laat ik dat begrip ‘interface’ wat precieser met U bekijken, en daarna een voorbeeld geven van die eigenwettelijkheid van de morfologie.

De eigenschappen van een geleed woord als *werker* moeten op drie niveaus worden gerepresenteerd die elkaar kunnen ‘zien’: de klankvorm, de morfologische structuur, en de betekenis. In de woorden van de linguïst Uhlenbeck in zijn diesrede voor deze universiteit in het jaar 1976: “woorden zijn eenheden waaraan drie dimensies zijn te onderkennen. Zij vertonen een hoorbare vorm - dit is hun fonische dimensie -, zij leveren in het gebruik een kennisbijdrage tot het geheel waarvan zij deel uitmaken - dit is hun semantische dimensie -, en tenslotte hebben zij een grammatische dimensie, waaronder allereerst moet worden verstaan dat zij over systematische verbindingsmogelijkheden beschikken ten opzichte van andere woorden” (Uhlenbeck 1976: 9). Elk van deze dimensies heeft zijn

eigen systematiek, zoals ook wordt benadrukt in het recente werk van de Amerikaanse taalkundige Jackendoff die in dit verband spreekt over de drievoudige parallelle architectuur van de grammatica, onder andere in zijn boek *Foundations of Language* (Jackendoff 2002). Dat het in de morfologie gaat om meer dan alleen de interne structuur van woorden wordt ook uitgedrukt door de titel van het tekstboek over morfologie dat ik begin van dit jaar publiceerde: het heet *The Grammar of Words*.

Het is van groot belang in te zien dat de drie niveaus die Uhlenbeck noemde in zijn Leidse diesrede, systematisch met elkaar zijn verbonden. Zo is er een systematische relatie tussen de structuur en de klankvorm van het woord *werker*. Het bestaat uit twee morfemen, *werk* en *-er*, en fonologisch uit twee lettergrepen, (wɛr) en (kər). U hoort dus dat de verdeling van de klanken van dit woord in lettergrepen zich niets aantrekt van de interne morfologische bouw. Anders zouden we de lettergreepverdeling [wɛrk.ər] gekregen hebben.

Dit is een voorbeeld van de asymmetrie tussen fonologische en morfologische structuur, en het lijkt er dus even op dat fonologie en morfologie elkaar niet kunnen zien, en dat de interface dus negatief van aard is. We noemen het Nederlandse suffix *-er* in dit verband een ‘cohering’ suffix. De term ‘cohering suffix’ is geïntroduceerd door de linguïst Dixon in zijn beschrijving van de Australische taal Yidin^y. Voor zo’n suffix geldt dat het met de stam waaraan het is toegevoegd samensmelt tot één prosodisch woord. Binnen dat prosodisch woord vindt de lettergreepverdeling plaats op louter fonologische gronden, zonder dat de morfologische bouw een rol speelt. Dank zij dit gedrag van het suffix *-er* wordt voorkomen dat de tweede lettergreep van het woord *werker* met een klinker begint. Voor alle talen van de wereld geldt namelijk dat een lettergreep bij voorkeur begint met een medeklinker, en de prosodische vorm van het woord *werker* is in overeenstemming met deze universele tendentie: zowel de eerste als de tweede lettergreep begint met een medeklinker.

Er is ook een interface tussen het niveau van de morfologische bouw en het niveau van de betekenis. Het systematische verband tussen de bouw van het woord *werker* en de betekenis ervan zal duidelijk zijn: de betekenis van dit woord is een combinatie van de betekenis van de werkwoordsstam *werk* en de betekenisbijdrage van het suffix *-er*.

Met deze observaties over één geleed woord, het Nederlandse woord *werker*, zien we al hoe de bestudering van de eigenschappen van woorden een rol speelt in de beantwoording van belangrijke onderzoeksvragen van de taalwetenschap: Wat is de architectuur van de grammatica’s van natuurlijke talen? Wat is wel mogelijk in natuurlijke taal, en wat niet? Hoe verhouden de drie soorten eigenschappen (klank, vormeigenschappen, en betekenis) van woorden en zinnen zich tot elkaar?

Het voorbeeld *werker* zou U, zo zagen we, op de gedachte kunnen brengen dat het een algemene wetmatigheid van taal is dat de morfologische structuur van een woord geen invloed heeft op de uitspraak van dat woord. Dat is zeker niet het geval. Voor veel talen geldt bijvoorbeeld dat bij samenstellingen (woorden die zelf weer combinaties zijn van woorden of stammen) de interne bouw wel degelijk een rol speelt. U spreekt de samenstelling *handappel* niet uit als *han.dap.pel* maar als *hant.ap.pel*, en aan die slot-[t] van het eerste gedeelte *hand* kunnen we horen dat U daar een

lettergreepgrens legt, die dus samenvalt met het einde van het woord *hand*. Dat komt omdat er een fonologische wet geldt voor het Nederlands: aan het eind van een lettergreep zijn plof- en wrijfklanken altijd stemloos. Vandaar dat we het woord *handen* uitspreken met een [d], maar het woord *hand* met een [t]. Er moet dus gezien de uitspraak van het woord *handappel* wel een lettergreepgrens zijn direct voor de [ɑ] van *appel*.

We zeggen daarom dat met het ene grammaticale woord *handappel* twee prosodische woorden, *hand* en *appel* corresponderen. Een prosodisch woord bestaat uit een of meer lettergrepen. Deze asymmetrie tussen grammaticaal en prosodisch woord is typisch voor samenstellingen, woorden die zelf weer uit meer dan één woord bestaan.

Woorden kunnen echter hun morfologische structuur en dus hun doorzichtigheid verliezen, en dat kunnen we dan soms horen aan hoe het woord wordt uitgesproken. Aan de uitspraak van het woord *aardappel* als [a:r.dɑ.pəl] horen we dat taalgebruikers die dit zeggen dit woord niet opvatten als de naam voor een bepaald soort appel, het woord niet als een samenstelling van *aard-* en *appel* ervaren, maar als een ongeleed woord. Immers, als er geen interne morfologische bouw meer is, zal de lettergreepverdeling plaats vinden over het woord in zijn geheel. De [d] van *aard* komt dan terecht aan het begin van de tweede lettergreep, en zal dus stemhebbend worden uitgesproken, anders dan de /d/ van *handappel*.

Zo'n terminologisch onderscheid tussen grammaticaal woord en prosodisch woord is buitengewoon nuttig en verhelderend bij de beschrijving van talen. Termen zijn kleine stukjes gestolde theorie, beproefd gebleken inzichten die we gebruiken bij de verdere bestudering van een bepaald aspect van de werkelijkheid. Het gebruik van termen laat zien dat theorie en descriptie in de wetenschap onlosmakelijk met elkaar verbonden zijn. Of, zoals ik het formuleerde in mijn oratie in 1982 aan de Vrije Universiteit: "theorie zonder descriptie is leeg, maar descriptie zonder theorie is blind" (Booij 1982: 24).

Laat ik nog een voorbeeld geven van het nut van dit theoretisch onderscheid tussen de begrippen 'grammaticaal woord' en 'prosodisch woord'. In talen als het Fins en het Hongaars vinden we het verschijnsel van klinkerharmonie: de klinkers binnen één woord komen ten aanzien van bepaalde kenmerken overeen. Binnen een woord zijn alle klinkers of voorklinkers (klinkers die voor in de mond gevormd worden zoals [i], [e], en [y]), of achterklinkers, klinkers die achter in de mond gevormd worden, zoals [u], [o], en [a]. Uitzonderingen hierop in het Hongaars vormen de zogenaamde neutrale klinkers [i], [i:] en [e:]. Het woord *Katalin* in de titel van het boek van Magda Szabó, *Katalin Utca* (vertaald als *Katalinstraat*) is dus welgevoemd, want de achterklinkers [a] in de eerste en tweede lettergreep worden gevolgd door de neutrale klinker [i].

Samenstellingen vormen een ogenschijnlijke uitzondering op deze regel. Zo bevat het woord *Budapest*, de naam van de hoofdstad van Hongarije, eerst twee achterklinkers (/u/ en /ɑ/), en daarna een voorklinker /ɛ/. Dit woord is een samenstelling van *Buda* (de stad op de rechteroever van de Donau), en *Pest* (de stad op de linkeroever). We moeten dus concluderen dat het domein van de regel

van klinkerharmonie niet het grammaticaal woord, maar het prosodisch woord is. Zo kon ik ook, toen ik met mijn gezin een aantal jaren geleden voor vakantie aankwam in het Finse gehucht *Urajärvi* meteen zien dat deze naam een samenstelling is omdat het een ogenschijnlijk disharmonisch woord is: eerst de achterklinkers [u] en [ɑ], en daarna de voorklinkers [ε] en [i]. Inderdaad, het woord *Urajärvi* is een samenstelling, en betekent Ura-meer.¹

Er is, zo zagen we, bij gelede woorden ook een systematische samenhang tussen hun betekenis en hun syntactische verbindingsmogelijkheden, en dit is dus een ander aspect van de ‘interface’. Zoals ik (voor een deel samen met collega Ton van Haaften) aan de hand van Nederlandse woordvormingsprocessen heb laten zien (cf. Booij en Van Haaften 1988), is het de functie van veel morfologische processen om woorden van een bepaalde semantische klasse te vormen, waarbij die betekenis een bepaalde syntactische valentie impliceert. Een voorbeeld: nieuw gevormde Nederlandse werkwoorden die met het prefix *be-* beginnen drukken – anders dan sommige bestaande werkwoorden met dat prefix zoals *beginnen* - doorgaans een handeling uit die gericht is op een bepaald object. Denk aan werkwoorden als *beargumenteren*, *bespreken*, en *bestoken*. Het gevolg daarvan is dat het verplicht overgankelijke werkwoorden zijn, die altijd een lijdend voorwerp bij zich moeten hebben. Je kunt niet zeggen **Laten we bespreken*, er moet een lijdend voorwerp zoals *dit probleem* bij. Vanwege die effecten op syntactische valentie neemt daarom in veel beschrijvende grammatica’s de morfologie van werkwoorden een belangrijke plaats in.

Na deze globale schets van de notie ‘interface’ vraag ik kort uw aandacht voor de in het begin al genoemde eigenwettelijkheid van de morfologie. Er zijn regelmatigheden in de vorm van gelede woorden die niet terug te voeren zijn tot fonologische of syntactische principes. Ik zal dit met een voorbeeld uit het Frans illustreren. In het Frans kunnen we bijwoorden maken van bijvoeglijke naamwoorden door toevoeging van het suffix *-ment*. Zo heeft het bijvoeglijk naamwoord *lent* ‘langzaam’ als bijbehorend bijwoord *lentement*. De vorm van het bijvoeglijk naamwoord waaraan we dit suffix moeten hechten is de vrouwelijke vorm, in dit voorbeeld *lente*, hoewel in de betekenis van het bijwoord het begrip ‘vrouwelijk’ geen enkele rol speelt. Je kunt het verwijzen naar de vrouwelijke vorm niet vermijden door te zeggen dat het suffix eigenlijk *-ement* is. Als namelijk de vrouwelijke vorm van het adjectief onregelmatig is, dan geldt dat ook voor het corresponderende bijwoord: *beau* heeft als vrouwelijke vorm *belle*, en het bijwoord is dan ook *bellement*. De vrouwelijke vorm van *fou* is *folle*, en dus is het bijwoord *follement*, etc. In een synchrone analyse van het Frans moeten we dus zeggen dat de vorm van de stam die we moeten gebruiken voor de vorming van bijwoorden identiek is aan die van de vrouwelijke vorm van het adjectief. Om de juiste vorm van het bijwoord uit te rekenen moeten we dus verwijzen naar een ander woord dan de stam zelf. We maken dus, klassiek gezegd, gebruik van paradigmatische betrekkingen tussen woorden. Deze vormregelmatigheid heeft niets van doen met fonologie of syntaxis, is een zuiver morfologische regelmaat, en dus een voorbeeld van de relatieve autonomie van de morfologie.

Historisch is deze regelmaat overigens wel verklaarbaar. Het suffix *-ment* stamt af van een naamvalsvorm van het Latijnse vrouwelijke woord *mens* ‘geest’. Een Latijnse woordgroep met de ablatief-naamval zoals *clara mente* ‘met een heldere geest, op heldere wijze’ heeft een vrouwelijke vorm van het adjectief, *clara*, omdat het woord *mens* ‘geest’ waar het bij hoort, vrouwelijk is. Deze woordgroep is geherinterpreteerd als één woord, een bijwoord *claramente*, en zo vinden we dit suffix als *-mente* of *-ment* terug in de verschillende Romaanse talen. Als residu van deze herinterpretatie van *-mente* tot een suffix vinden we de keus voor een speciale stam van het adjectief, die gelijk is aan de vrouwelijke vorm.

Het ontstaan van het Romaanse suffix *-mente* of *-ment* is een mooi voorbeeld van wat taalkundigen ‘grammaticalisatie’ noemen: het ontstaan van grammaticale elementen zoals suffixen uit woorden of woordgroepen. Dit type taalverandering staat de laatste jaren sterk in de belangstelling, en biedt een fascinerend perspectief op hoe taalsystemen zich ontwikkelen. Dit perspectief kan kort worden samengevat in de slagzin ‘Today’s morphology is yesterday’s syntax’. Het helpt ons zeker om de synchrone eigenaardigheden van de morfologie beter te begrijpen.

De kennis van een taalsysteem die we aanduiden als morfologische kennis heeft twee functies. Deze stelt ons in staat de juiste vorm van woorden in hun syntactische context te bepalen (traditioneel inflectie genoemd), en daarnaast om nieuwe gelede woorden te maken en zo de woordvoorraad van een taal uit te breiden. Morfologische kennis maakt dus deel uit van het creatieve vermogen dat eigen is aan het beheersen van een natuurlijke taal.

Maar hoe verwerven we die kennis? Laten we ons bij de beantwoording van die vraag voor dit moment beperken tot ons vermogen om nieuwe woorden te maken, woordvorming dus. Er is daarbij een centrale rol weggelegd voor het lexicon. De term ‘lexicon’ wordt in de taalwetenschap gebruikt om de kennis van de lexicale eenheden van een bepaalde taal aan te duiden. Het is een theoretische grootheid, te onderscheiden van de term ‘woordenboek’ die verwijst naar praktische naslagwerken met betrekking tot de woordenschat van een of meer talen. Er zijn bovendien twee perspectieven op dat lexicon. Vanuit psychologisch gezichtspunt zit het in het hoofd van de taalgebruiker, en is het dus een mentale entiteit, het mentale lexicon. Anderzijds is er ook een boven-individueel perspectief (taal als het bezit van een taalgemeenschap), met een boven-individueel lexicon. In dat geval fungeert het lexicon als een norm die bepaalt welke woorden gelden als gevestigd in een taalgemeenschap, en zo mede het taalgebruik van ieder individu bepaalt. We weten hoe dingen conventioneel aangeduid worden, ook al had het ook anders gekund.² We noemen bij voorbeeld het stukje papier dat we op een brief plakken ter frankering een *postzegel*, maar het had ook een *briefbon* kunnen heten.

Kennis van woordvormingspatronen kan alleen verworven worden doordat we eerst afzonderlijke gelede woorden verwerven. Pas als we voldoende zelfstandige naamwoorden van het type ‘werkwoord + *-er*’ hebben gehoord en opgeslagen in ons lexicaal geheugen zijn we in staat daaruit een abstract patroon te destilleren dat ons in staat stelt ook nieuwe woorden volgens dat

patroon te vormen. Zulke patronen of regels zijn dus enerzijds redundantieregels met betrekking tot al opgeslagen woorden van dat type, en anderzijds recepten om er nieuwe van dat type bij te maken.

De formuleringen die ik tot nu toe gekozen heb om het begrip ‘morfologie’ toe te lichten veronderstellen dat woorden in het lexicon betrekkingen met elkaar onderhouden, en dus een web van woorden vormen. De web-metafoor is inderdaad heel geschikt om het begrip ‘lexicon’ inzichtelijk te maken, en nieuwe onderzoeksvragen op te roepen.

Door middel van metaforen proberen we het onbekende te vatten in termen van het bekende, het is een vorm van transport van begrippen van het ene, bekende domein naar het andere, onbekende domein. Het Griekse woord *metaforá* betekent inderdaad nog steeds letterlijk ‘transport’; je kunt dat woord bijvoorbeeld in Griekenland zien staan op vrachtwagens van transportbedrijven, en de abstracte betekenis van het woord *metafoor* is dus zelf metaforisch van aard.

Zoals Douwe Draaisma prachtig laat zien in zijn boek *De metaforenmachine*, dat gaat over het wetenschappelijk onderzoek van het menselijk geheugen, worden metaforen continu gebruikt in de wetenschap, en hebben ze daar zeer belangrijke functies. Een van de metaforen voor het menselijk taalvermogen is die van de computer, die een geheugen bevat met representaties van informatie, en daarnaast software om nieuwe informatie uit te rekenen. Zeker voor de studie van het mentaal lexicon is dat nog steeds een dominante metafoor (Aitchison 1987: 32).

De computer-metafoor heeft geleid tot een belangrijke onderzoeksvraag in de taalwetenschap: de verhouding tussen opslag (‘storage’) en berekening (‘computation’) in onze taalkennis. Er zijn taalkundigen die een redundantievrij lexicon bepleiten: alles wat met regels kan worden gegenereerd, moet niet worden opgeslagen in het lexicon. Het lexicon bevat alleen die informatie over lexicale eenheden die niet voorspelbaar is. Deze visie kan gezien worden als de weerspiegeling van de beginperiode van de computer, waarin het geheugen beperkt was, en de reken capaciteit de belangrijkste component voor het verwerken van informatie.

Deze zuinige opvatting van het lexicon van natuurlijke talen is wel gekarakteriseerd als de ‘rule/list fallacy’, het misverstand dat ‘regel’ en ‘lijst’ elkaar uitsluitende domeinen van informatie zijn. Dat een taalbouwsel regelmatig is, en dus volgens regels te vormen, sluit immers niet uit dat het desondanks met zijn voorspelbare eigenschappen wordt opgeslagen in ons geheugen. Als we als taalgebruikers woordvormingsregels verwerven op grond van patroonherkenning in een verzameling ons bekende, verwante gelede woorden, dan zou het raar zijn om al die gelede woorden weer uit ons lexicale geheugen te verwijderen zodra we die regel ontdekt hebben.

Dit geldt ook voor het domein van de inflectie. Verbogen en vervoegde vormen van woorden kunnen we vaak gemakkelijk uitrekenen met behulp van regels, maar we kunnen ze ook direct uit ons geheugen opvissen. Uit taalpsychologisch onderzoek is gebleken dat we volstrekt regelmatige meervoudsvormen van zelfstandige naamwoorden in talen als het Italiaans en het Nederlands opslaan in ons lexicale geheugen als ze een zekere frequentie hebben (Baayen, Burani en Schreuder 1997, Baayen, Dijkstra en Schreuder 1997). We kunnen ze iedere keer ter plekke uitrekenen, maar in plaats

daarvan kunnen we zulke vormen ook direct uit ons geheugen halen, en dat zal de snelheid van taalherkenning en taalproductie zeker ten goede komen.

Wat we uit dit gebruik van de computer-metafoor kunnen leren is, dat metaforen in de wetenschap een belangrijke heuristische functie hebben, maar ons ook onterechte conceptuele beperkingen kunnen opleggen. Het belang van de reflectie op het gebruik van metaforen in wetenschappelijk onderzoek is een kleine illustratie van het belang van de wetenschapsleer, een vakgebied dat ook in het onderwijs aan een letterenfaculteit van grote waarde kan zijn.

De opslag van woorden in het geheugen hangt op een intrigerende manier samen met de gebruiksfrequentie van die woorden. Hoewel frequentie een statistische eigenschap is van woorden in corpora van feitelijk taalgebruik, heeft het een correspondentie in het individuele lexicon in de vorm van de sterkte van geheugensporen. Met frequentie correspondeert het niveau van activatie, waardoor woorden met een relatief hoge frequentie sneller herkend of opgeroepen worden dan woorden met een relatief lage frequentie.

Naast taalpsychologische experimenten zijn er andere bronnen van evidentie voor hoe het mentale lexicon in elkaar zit. Zo is ook taalverandering een venster op de aard van ons lexicaal geheugen. Dit kan geïllustreerd worden aan het Fries. Het Fries kent een verschijnsel dat bekend staat als breking, een afwisseling tussen een ongebroken, vallende tweeklank in het enkelvoud en een gebroken, stijgende tweeklank, onder andere in het meervoud van zelfstandige naamwoorden. Zo is het meervoud van het Friese woord [stiən] ‘steen’ het woord [stjɪnən]. Dit verschil in klankvorm tussen het enkelvoud en het meervoud is echter aan het verdwijnen, met als gevolg dat dan ook in de meervoudsvorm de vallende diftong optreedt van de enkelvoudsvorm optreedt. In plaats van de meervoudsvorm [stjɪnən] krijgen we dan de meervoudsvorm [stiənən]. Maar wanneer de meervoudsvorm frequenter is dan de enkelvoudsvorm, zoals bij de woorden *armen* en *tranen* het geval is, dan is het juist de minder frequente enkelvoudsvorm die zich qua klankvorm aanpast aan de meervoudsvorm. Zo krijgen we in sommige vormen van het Fries blijkens een onderzoek van de Amerikaanse taalkundige Tiersma voor het woord *traan* [triən] in plaats van het woordpaar [triən] - [trjɪnən] het paar [trjɪn] - [trjɪnən] (Tiersma 1982: 834). Het is duidelijk dat dit proces van taalverandering waarin frequentie een rol speelt, de opslag van althans de meer frequente meervoudsvormen vooronderstelt.

Een andere illustratie van het frequentie-effect is het Nederlandse woord *natuurlijk*. Dat kan gebruikt worden als bijvoeglijk naamwoord, als in *natuurlijke behoeften*, maar ook als een bijwoord, met de betekenis ‘vanzelfsprekend’. Dit tweede gebruik van *natuurlijk* is de meest frequente. In dat laatste geval is dit woord onderhevig aan fonetische erosie. Het kan worden uitgesproken als *natuurlijk*, maar ook als *tuurlijk*, en op nog wel twaalf andere manieren, zoals [tylək] en [tyk] (cf. Ernestus 2000: 141). Maar we zullen niet spreken van *tuurlijke behoeften*. Het bijwoord *natuurlijk* blijkt zijn morfologische doorzichtigheid verloren te hebben, wordt niet meer verbonden met het

grondwoord *natuur*, en schuift dan qua uitspraak op in de richting van de prosodische vorm van ongelede woorden, die doorgaans twee- of een-lettergrepig zijn.

We keren nu terug naar onze web-metafoer en proberen daar wat meer inhoud aan te geven. De relatie tussen een woordvormingspatroon en de volgens dat patroon gevormde woorden kan gemodelleerd worden als een omgekeerde boom, waarin het patroon de top vormt, verbonden met de individuele woorden die de takken van de boom vormen. Zo kunnen we ons voorstellen dat alle bestaande samenstellingen van het type zelfstandig naamwoord + zelfstandig naamwoord, zoals *bureaustoel* en *kaasschaaf* gedomineerd worden door een knoop met dat patroon ‘zelfstandig naamwoord + zelfstandig naamwoord’ in de top van de boom. Daarmee geven we aan dat deze woorden een overeenkomstige bouw hebben, en zijn gevormd volgens een in het Nederlands productief patroon van woordvorming. Tegelijkertijd zal het woord *bureaustoel* ook verbonden zijn met de woorden *bureau* en *stoel*, en het woord *kaasschaaf* met de woorden *kaas* en *schaaf*. Kortom, hier ontvouwt zich het idee van het lexicon als een web van niet alleen woorden, maar ook van patronen, met een veelvoud aan verbindingen, zoals ook het internet een web vormt met velerlei verbindingen tussen stukken informatie.

Een algemene eigenschap van dit productieve samenstellingspatroon ‘zelfstandig naamwoord + zelfstandig naamwoord’ is dat het rechterdeel het hoofd vormt. Immers, een bureaustoel is een bepaald type stoel, niet een bepaald type bureau. Dit zal dus als algemene eigenschap van deze morfologische constructie van het Nederlands moeten worden gespecificeerd. Er zijn ook talen waar dat precies omgekeerd is, en het linkerdeel het hoofd is. Zo is het woord voor ‘computer’ in het Maori, gesproken in Nieuw-Zeeland, *roro hiko*, letterlijk ‘hersenen electriciteit’, dus zoiets als elektronische hersenen, met het hoofdwoord *roro* ‘hersenen’ in de linkerpositie. De positie van het hoofd is dus taalspecifiek bepaald.

Tussen het abstracte patroon en de afzonderlijke woorden kunnen er ook tussenpatronen zijn, subgeneralisaties over deelverzamelingen van de relevante groep woorden. Een voorbeeld uit het Nederlands is de verzameling van samenstellingen die beginnen met het woord *hoofd* zoals *hoofdonderwijzer* en *hoofdingang*. Dit soort woorden zijn opnieuw combinaties van twee zelfstandige naamwoorden. Hoewel het woord *hoofd* allerlei betekenissen heeft, krijgt het in samenstellingen doorgaans de specifieke betekenis ‘voornaamste’ (daarnaast zijn er natuurlijk ook samenstellingen als *hoofdpijn* en *hoofddekseel* waarin *hoofd* een andere betekenis heeft). Zo is er dus een subpatroon *hoofd* + zelfstandig naamwoord’ ontstaan met de betekenis ‘voornaamste x’ dat we vinden in woorden zoals *hoofdlijn*, *hoofdzaak*, *hoofdgerecht*, *hoofdedacteur*, *hoofdkantoor*, *hoofdvak*, *hoofdindex*, *hoofdvraag* en *hoofdedachte*.

Er zijn talloze voorbeelden van dergelijke contextueel beperkte interpretaties van woorden binnen samenstellingen. Zo heeft *oud* in samenstellingen de betekenis ‘voormalig’, en een oud-decaan hoeft dus niet oud te zijn. Dergelijke patronen met specifieke betekenissen waarin sommige posities

lexicaal vastliggen, en andere nog geheel variabel zijn, worden wel ‘constructional idioms’ of idiomatische patronen genoemd (cf. Jackendoff 2002, Booij 2002). Algemeen kunnen we patronen voor de vorming van gelede woorden kwalificeren als ‘constructional idioms op woordniveau’. Dat geldt bijvoorbeeld voor de constructie van zelfstandige naamwoorden van het type [werkwoord + *er*] die ten grondslag ligt aan een woord als *werker*.

Het begrip ‘constructie’ heeft een lange staat van dienst in de taalwetenschap. In de Nederlandse taalkunde speelt het een belangrijke rol in het werk van Paardekooper die niet zo lang geleden nog enkele jaren aan deze universiteit heeft gedoopt. Tegenwoordig staat het begrip ‘constructie’ centraal in de theorie van de Constructie-Grammatica.

Waar ik in dit verband kort bij wil stil staan, is de verwevenheid van morfologische en syntactische constructies. In een simplistisch model van de grammatica gaat de vorming van zinnen in twee opeenvolgende stappen: morfemen worden gecombineerd tot woorden, en vervolgens woorden tot zinnen. Bij nadere beschouwing blijkt echter dat syntaxis en morfologie vaak een wederzijdse afhankelijkheid vertonen. Dit kunnen we mooi zien in het gebruik van woordgroepen zoals *met zijn drieën*.

In het Nederlands gebruiken we de meervoudsvormen van telwoorden in een aantal verwante constructies, zoals

- (2) in drie-en
- wij / ons drie-en
- met ons / jullie / hun drie-en
- met zijn drie-en

Al deze woordgroepen zijn voorbeelden van constructies. Het voorbeeld *met zijn drieën* is een voorzetselgroep geconstrueerd volgens de algemene regels voor de bouw van zulke woordgroepen in het Nederlands, maar met een specifieke groeps-betekenis. De woorden *met* en *zijn* vormen de vaste bezetters van de eerste twee posities van deze constructie, terwijl op de plaats van het zelfstandig naamwoord alle telwoorden (behalve, om buitentalige redenen, *nul* en *één*) kunnen voorkomen mits gevolgd door het suffix *-en*. Het is dus een typisch geval van een ‘constructional idiom’. De groeps-betekenis is gebonden aan deze constructie, en kan niet zonder meer worden afgeleid uit de betekenis van de samenstellende woorden.

Zoals uit de zin

- (3) We komen morgen met zijn drie-en

blijkt, is er geen overeenstemming in persoon en getal tussen het onderwerp (hier eerste persoon meervoud) en het bezittelijk voornaamwoord *zijn* [zɛn] dat derde persoon enkelvoud is. Bijzonder is

dat altijd het suffix *-en* gekozen moet worden, ook al zou het reguliere meervoud op een *-s* eindigen. Zo hebben we:

(4) met zijn zeven-en, met zijn negen-en

Maar als deze telwoorden gebruikt worden als zelfstandige naamwoorden, krijgen ze het regelmatige meervoud op *-s*, zoals in:

(5) Ik heb drie zeven-s en twee negen-s op mijn rapport

Dit laat zien dat we bij het beschrijven van deze constructies de aanwezigheid van een specifiek suffix *-en* als noodzakelijk aanwezig moeten vermelden. Omgekeerd komen woorden als *tweeën* en *zevenen* alleen voor in de genoemde constructies. Je kunt bijvoorbeeld het woord *tweeën* niet gebruiken als onderwerp in de zin:

(7) *Tweeën gingen naar huis / Twee gingen naar huis

Zulke meervouden zijn alleen mogelijk bij woorden als *honderd*, *duizend*, *miljoen*, etc. Je kunt wel zeggen: *Honderden gingen naar huis*. Er is dus sprake van onderlinge afhankelijkheid van morfologische en syntactische constructies. Deze verwevenheid leidt tot de conclusie dat het web van woorden van een taal verbonden is met het web van morfologische en syntactische constructies. Mijn leeropdracht had dus ook omschreven kunnen worden als ‘morfologie en het constructicon’, omdat ook bouwsels groter dan een woord of een morfologische constructie deel uitmaken van het web. Het is dus de vraag of lexicon en grammatica wel als twee volledig aparte grootheden van het taalvermogen kunnen worden beschouwd.

Deze laatste conclusie wordt ondersteund door het verschijnsel dat in de Nederlandse taalkunde wel het ‘scheidbaar samengestelde werkwoord’ wordt genoemd, woordcombinaties zoals *opbellen* en *opvrolijken*, die in enkelvoudige zinnen gesplitst worden, getuige een zin als:

(7) Deze dag vrolijkt me op

Een woordje als *op* wordt doorgaans een partikel genoemd, en *opvrolijken* is dus een partikelwerkwoord. Dat er sprake is van twee woorden blijkt ook uit de vorm van het voltooid deelwoord: *opgevrolijkt*, waarin het voorvoegsel *ge-* na het partikel komt, en uit een zinnetje als *Ik ben er zeker van dat deze dag me op zal vrolijken*, met het hulpwerkwoord *zal* tussen het partikel *op* en het werkwoord *vrolijken* in. We staan er niet altijd bij stil dat hier sprake is van woordcombinaties, en niet van gelede werkwoorden met een prefix, mede omdat de Nederlandse spellingvoorschriften zeggen dat

we deze woordreeksen als één woord moeten schrijven als ze naast elkaar staan. Deze combinaties van een werkwoord met een ander woord, die in heel veel talen voorkomen (cf. Booij en Van Kemenade 2003), zijn natuurlijk wel lexicale eenheden, opgeslagen in het mentaal lexicon, maar geen woorden in strikte zin, waarvoor immers het criterium van de onscheidbaarheid geldt.

Sommige van deze partikels combineren met een werkwoord dat niet zelfstandig bestaat. Zo bestaat het werkwoord *vrolijken* niet op zich, maar alleen in combinatie met het woordje *op*. Denk naast *opvrolijken* ook aan partikelwerkwoorden als *opleuken* en *opfrissen*. De werkwoorden *vrolijken*, *leuken*, en *frissen* bestaan niet op zich naast de ermee corresponderende bijvoeglijke naamwoorden *vrolijk*, *leuk* en *fris* hoewel de omzetting van bijvoeglijk naamwoord naar werkwoord wel structureel mogelijk is in het Nederlands, zoals blijkt uit werkwoorden als *bleken* en *drogen*. Dat betekent dat dit gebruik van bijvoeglijke naamwoorden als *leuk*, *fris* en *vrolijk* als werkwoorden alleen voorkomt in combinatie met een partikel.

Een ander actueel voorbeeld is het werkwoord *ophokken* dat op dit moment veel gebruikt wordt in relatie tot pluimvee en vogelgriep. Denk maar aan de veel bediscussieerde *ophokplicht*. Het werkwoord *hokken* dat gevormd is op basis van het zelfstandig naamwoord *hok*, komt met de betekenis ‘in een hok opsluiten’ alleen maar voor in combinatie met *op* (*hokken* als zelfstandig werkwoord heeft andere betekenissen zoals ‘ongehuwd samenwonen’). Het proces van vorming van werkwoorden uit zelfstandige naamwoorden kan dus in gang gezet worden door een partikelwerkwoord-constructie met het partikel *op*. Er is dus een wederzijdse afhankelijkheid tussen woordvormingspatronen en woordcombinatiepatronen. Dit is een ander voorbeeld van de webgedachte: patronen op woord- en woordgroepsniveau zijn met elkaar verweven, zoals dat ook al bleek uit constructies als *met zijn drieën*. De grammatica bevat niet alleen een web van woordpatronen, maar ook een web van constructies.

Het wordt tijd voor een conclusie. In het begin van deze rede heb ik gesproken over de mogelijke saaiheid van de morfologie. Ik hoop u er vanmiddag van overtuigd te hebben dat niets minder waar is. Door de studie van de klank, de structuur, en de betekenis van woorden in hun onderlinge samenhang hebben we een blik kunnen werpen op de architectuur van natuurlijke talen, de verhouding tussen regels en lexicon, en het belang van de notie ‘constructie’. Wat daarbij steeds weer opvalt, is hoe complex de grammatica’s van talen zijn georganiseerd, als een duizelingwekkend geheel van patronen en subpatronen. Het is U hopelijk duidelijk geworden dat morfologie een uitstekend venster is op die intrigerende organisatie van taalsystemen, en ons helpt bij het verkrijgen van inzicht in het fenomeen van natuurlijke taal, met zijn vele verschijningsvormen in de talen van de wereld.

Terloops heb ik daarbij ook Uw aandacht willen vragen voor een aantal methodische uitgangspunten die van belang zijn voor goed taalkundig onderzoek, te weten:

- de nauwe verwevenheid van theorievorming en descriptie;

- het belang van discussie tussen taalkundigen die werken in verschillende onderzoekstradities, en het open staan voor alternatieve concepties;
- het belang van het betrekken van zoveel mogelijk verschillende soorten feiten bij de beschrijving van talen en de modellering van het menselijk taalvermogen: intuïties, corpora, informanten, experimenten, taalverandering, etc.;
- het waardevolle van reflectie op metaniveau, zoals het onderkennen van de rol van metaforen bij het stellen van wetenschappelijke vragen en het modelleren van menselijke kennis.

Aan het eind van deze oratie gekomen dank ik het College van Bestuur van de Universiteit Leiden van harte voor mijn benoeming en het in mij gestelde vertrouwen. Deze dank geldt eveneens voor het Bestuur van de Faculteit der Letteren. De wijze waarop beide besturen mij zijn tegemoet getreden was en is buitengewoon plezierig. Ook de persoonlijke reacties van veel collega's binnen de faculteit die ik ontving na mijn benoeming waren zeer bemoedigend. Met vreugde aanvaard ik daarom de taken die ik de komende jaren te vervullen heb.

Mijn taak voor de eerstkomende vijf jaren is primair die van het decanaat van de Faculteit der Letteren. Om die reden heeft deze oratie niet het karakter van het ontvouwen van een onderzoeksprogramma, maar van een kennismaking met een taalwetenschapper die van buiten als decaan deze faculteit en universiteit binnentreedt. Wel ben ik zeker van plan de hierboven genoemde methodische overwegingen te betrekken bij mijn werk als decaan dat erop gericht zal zijn optimale omstandigheden te scheppen voor uitstekend wetenschappelijk onderwijs en onderzoek.

Ik zal er mijn best voor doen dat de prachtige Faculteit der Letteren van de Universiteit Leiden zal blijven bloeien en groeien. Ik reken daarbij op de steun van de collega's in de faculteit en de universiteit, van wie ik vele al in andere verbanden heb leren kennen als mensen met wie het goed samenwerken is.

Heel bijzonder is het voor mij, en een eer, om als decaan de opvolger te mogen zijn van Ton van Haften, die ik al zoveel jaren ken, als student, daarna als promovendus en collega, en nu dan als voorganger en boven mij gestelde. Ook met de andere leden van het faculteitsbestuur, Jef Jacobs en Jelle Kaldewaij, was het vanaf het begin goed werken. Daarvoor ben ik ze zeer erkentelijk.

Tenslotte: deze rede draag ik op aan mijn vrouw Herry. Ik ben haar dankbaar voor de steun die ze mij bij deze stap in mijn loopbaan heeft gegeven, en voor nog veel meer.

Ik heb gezegd.

Noten

1. Er zijn ook suffixen die de status van prosodisch woord hebben. Een bekend voorbeeld uit het Nederlands is het suffix *-achtig* dat we vinden in *roodachtig*. We noemen dit suffix ‘non-cohering’ omdat de stam *rood* en het suffix *-achtig* ieder op zich een prosodisch woord vormen. Dit valt te horen omdat we dit woord uitspreken als [rot.ɑx.təx], dat kan worden gecontrasteerd met het woord *rodig* [rodəx] dat wel één prosodisch woord vormt.

2. Conventies zijn overigens niet erg strikt, en de creativiteit van de taalgebruiker kan deze ook negeren. Zo kent het Amerikaanse Engels als conventioneel woord voor het Nederlandse begrip *geldautomaat* het woord *ATM*, een acroniem for *automatic* (of *automated*) *teller machine*. Omdat blijkbaar dit acroniem niet meer als zodanig herkend wordt, vindt men ook wel *ATM machine*, een redundante constructie derhalve omdat ‘machine’ tweemaal wordt uitgedrukt. De samenstelling *money machine* zou natuurlijk een prachtige doorzichtige samenstelling vormen voor dit begrip, ideaal als het Engels als lingua franca moet fungeren. Dit woord blijkt wel te bestaan maar primair gebruikt te worden voor de betekenis ‘iets dat veel geld opbrengt’. Zo wordt McDonalds als ‘money machine’ gekwalificeerd. Toch vinden we ook al advertenties waarin het woord *money machine* wordt gebruikt, bijvoorbeeld door een firma die graag zulke automaten in winkels wil plaatsen. Lexicale conventies hebben dus geen absoluut blokkerend effect, zoals deze feiten, verzameld via een simpele Google search, laten zien.

Verwijzingen

- Aitchison, J. 1987. *Words in the Mind. An Introduction to the Mental Lexicon*. Oxford: Blackwell [2002³]
- Baayen, R. H., Burani, C., en Schreuder, R. 1997. 'Effects of Semantic Markedness in the Processing of Regular Nominal Singulars and Plurals in Italian', in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1996*, Dordrecht: Kluwer, 13-34.
- Baayen, R.H., Dijkstra, T., en Schreuder, R. 1997. 'Singulars and Plurals in Dutch: Evidence for a Parallel Dual Route Model', *Journal of Memory and Language* 36, 94-117.
- Booij, G. E. 1982. 'Hiërarchische fonologie'. Rede uitgesproken bij de aanvaarding van het ambt van gewoon hoogleraar in de algemene taalwetenschap aan de Faculteit der Letteren van de Vrije Universiteit op 12 februari 1982.
- Booij, G. E. en T. van Haaften 1988. 'The external syntax of derived words: evidence from Dutch'. *Yearbook of Morphology 1988*, 29-44.
- Booij, G. E. 2002. 'Constructional Idioms, Morphology, and the Dutch Lexicon', *Journal of Germanic Linguistics* 14, 301-327.
- Booij, G. E. en A. van Kemenade 2003. 'Preverbs: An Introduction.' In G. Booij and J. van Marle (eds.) *Yearbook of Morphology 2003*. Dordrecht: Kluwer, 1-12.
- Dixon, R. M. W. 1977. *A Grammar of Yidin'*. Cambridge: Cambridge University Press.
- Draaisma, D. 1995. *De metaforenmachine. Een geschiedenis van het geheugen*. Groningen: Historische Uitgeverij.
- Ernestus, M. 2000. *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-based Study of the Phonology-Phonetics Interface*. Dissertatie Vrije Universiteit Amsterdam.
- Jackendoff, R. S. 2002. *Foundations of Language*. Oxford: Oxford University Press.
- Salmon, P. 2000. 'The Term 'Morphology''. In G. Booij, Ch. Lehmann, Ch., en J. Mugdan (eds.), *Morphology. An International Handbook of Inflection and Word Formation*. Berlin: De Gruyter, Vol. 1, 15-21.
- Spencer, A. and A. Zwicky (eds.) 1998. *The Handbook of Morphology*. Oxford: Blackwell
- Uhlenbeck, E. M. (1976) 'Taal en taalwetenschap'. Diesrede Universiteit Leiden, uitgesproken op 9 februari 1976.
- Tiersma, P. 1982. 'Local and General Markedness', *Language* 58, 832-49.