

Op weg naar W(E)TEN

**Een nieuw wetenschappelijk woordenboek van het
twintigste- en eenentwintigste-eeuwse Nederlands**

P.G.J. van Sterkenburg

bron

P.G.J. van Sterkenburg, *Op weg naar W(E)TEN. Een nieuw wetenschappelijk woordenboek van het twintigste- en eenentwintigste-eeuwse Nederlands*. Sdu Uitgevers, Den Haag 1997

Zie voor verantwoording: http://www.dbnl.org/tekst/ster007opwe01_01/colofon.htm

© 2009 dbnl / P.G.J. van Sterkenburg



Mijnheer de Rector Magnificus, waarde toehoorders,

1. Inleiding

De Australische theologe Barbara Thiering (1996) ontwikkelde na 20 jaar studie van de Dodezeerollen een revolutionaire nieuwe theorie die het christendom opnieuw uitdaagt. Jezus, betoogt zij, was de leider van een radicale factie van Esseense priesters. Hij was niet geboren uit een maagd en stierf niet aan het kruis. Hij trouwde met Maria Magdalena, stichtte een gezin en scheidde later. Hij stierf vlak na het jaar 64.

Het is bijna een blasfemie om sprekend over woorden en woordenboeken de indruk te wekken, dat er veranderingen zijn die de bestaande theoretische lexicografie of metalexigrafie op vergelijkbare wijze op haar kop zetten als Barbara Thiering de axioma's van het christendom. Vergelijkingen gaan altijd mank, ook deze. Maar toch.

Aan de ene kant heeft de automatisering het traditionele handwerk van de lexicograaf revolutionair gewijzigd, en in zoverre gaat de vergelijking dus wel op, maar aan de andere kant moet daar onmiddellijk aan toegevoegd worden dat de basisbeginselen van het lexicografische ambacht, zoals wij die vinden in Mesopotamië in de eerste Sumerische eentalige woordenboeken uit de tijd van Uruk IV, d.w.z. omstreeks 3300 vóór Christus (Cavigneaux 1976, 1989) en dus in een periode dat er helemaal nog geen theorie over het maken van woordenboeken bestond, niet zo zijn aangetast of op de proef gesteld als het christendom door Barbara Thiering. Dat betekent natuurlijk niet dat er geen spectaculaire veranderingen in het vak zijn waar te nemen.

Over veranderingen bij het maken van wetenschappelijke woordenboeken van de Nederlandse standaardtaal gaat deze oratie. Ik meen een en ander het

best te kunnen verduidelijken door nauw aan te sluiten bij de werkzaamheden van het Instituut voor Nederlandse Lexicologie, waar zeer binnenkort het monumentale *Woordenboek der Nederlandsche Taal*, het WNT, voltooid wordt en waar een begin gemaakt moet worden met een wetenschappelijk woordenboek van het twintigste- en eenentwintigste-eeuws Nederlands. In ons jargon heet dat laatste al W(E)TEN. Voor dat nieuw op te zetten woordenboek reserveer ik in het vervolg een ereplaats.

Het karakter, de identiteit van dat W(E)TEN is nog niet bekend. Welk concept eraan ten grondslag zal liggen, is dat evenmin. Wat wel bekend is, is dat zo'n nieuw woordenboek niet meer gemaakt kan worden zoals in de negentiende eeuw, de eeuw van het taalhistorisch comparatisme, de eeuw ook waarin de Romantiek tot grote woordenboekkathedralen als *Oxford English Dictionary*, *Deutsches Wörterbuch* en *Woordenboek der Nederlandsche Taal* inspireerde.

Zo'n nieuw woordenboek zal niet alleen de computer tot bondgenoot moeten maken, maar ook over de omheining van de traditionele lexicografie heen moeten kijken en, zeker waar het de beschrijving en presentatie van betekenissen betreft, gebruik moeten maken van de verworvenheden van de theoretische en computationele linguïstiek, de computationele lexicografie en van de kennistechnologie (Swanepoel 1994: 11; Verkuyl 1994: 5-20).

Tot de opvallende veranderingen in de hedendaagse lexicografie behoren beschouwingen over de identiteit en functies van een woordenboek. Maar ook zaken als de aard van het corpus dat eraan ten grondslag moet liggen, de wijze waarop omgesprongen moet worden met de voor beschrijving beschikbare ruimte, de verbetering van de beschrijving van de betekenissen en gebruiksmogelijkheden van woorden en woordgroepen, van stoplappen, van verraderlijke steunwoorden zijn aan veranderingen onderhevig, om niet te spreken over de dynamische opslag van de informatie en over de kortere bewerkingstermijn. Deze en nog andere veranderingen zullen ook het

W(E)TEN niet onberoerd laten. Slechts iets van dat alles zal nu de revue passeren met de nadrukkelijke kanttekening dat ik vanmiddag natuurlijk geen lexicologische catechismus aanbied die alle antwoorden omvat.

2. Identiteit

Wetenschappelijke woordenboeken binnen ons taalgebied hebben altijd met hun identiteit geworsteld. Wie het *Middelnederlandsch Woordenboek* van Verwijs en Verdam regelmatig gebruikt, weet dat dat product niet uitsluitend de taal uit de periode 1200-1500 beschrijft. Ook de taal van Kiliaan, Plantijn, Hooft, Huygens en Bredero wordt er regelmatig geciteerd. Nog duidelijker in dit verband is de hoofdredacteur van het *WNT*: ‘Vragen naar de identiteit van het *Woordenboek der Nederlandsche taal* ... is vragen om moeilijkheden. Eén korte en ondubbelzinnige beantwoording is namelijk niet mogelijk’ (Moerdijk 1994: 5).

Ook woordenboeken die bestemd waren voor het onderwijs en de algemene publieksmarkt, wat dat ook moge zijn, bekommerden zich tot het laatste kwart van onze eeuw weinig of niet om de uitdrukkelijke formulering van hun identiteit. Bij de verklarende woordenboeken had men geen oog voor de taalproducerende moedertaalspreker en bij de vertaalwoordenboeken vroeg men zich nauwelijks af voor wie de opgenomen woordenschat geselecteerd was, laat staan dat men voor het deel Nederlands-vreemde taal bewust rekening hield met taalproductie. Beter, maar riskanter gezegd: ‘Nergens wordt gesproken over de gebruiker van het woordenboek’. En het concept ‘leerwoordenboek’, zo vitaal voor hen die Nederlands als tweede taal leren, leek bij de lexicografen van het Nederlands, in tegenstelling tot bij hun Engelse collega's nog nauwelijks bekend. Dar alles veranderde in de jaren zeventig en tachtig toen het huis Van Dale Lexicografie een groot assortiment van nieuwe woordenboeken op de markt ging brengen. Alleen al om uit hetzelfde fonds drie verschillende woordenboeken van het Nederlands te

positioneren, moesten de redacties hun concept wel zodanig inrichten en aanscherpen dat er van een unieke identiteit en niet van een fletse kloon gesproken kon worden.

Van het WNT en van de commerciële woordenboeken hebben wij geleerd dat een kruising van identiteiten wetenschappelijk nauwelijks te verdedigen is. Daarom dient een W(E)TEN van meet af aan te voldoen aan hoge identiteitsstandaarden. In mijn optiek wil dat zeggen dat het een alfabetisch, wetenschappelijk woordenboek van de hedendaags Nederlandse standaardtaal moet zijn, dat niet puristisch, normatief of pedagogisch is en dat bovenal een verklarende, relationele en documenterende functie dient te hebben. Op deze eisen ga ik nu wat nader in.

Alfabetisch

Hoe vanzelfsprekend het alfabetiseringsbeginsel voor het verschaffen van kennis en inzicht ook moge lijken, het is dat, en ik zeg dat zonder onderton van melancholie, niet of beter misschien nog, niet meer. De opbouw van complexe elektronische bestanden heeft ons geleerd dat de computer juist een ideaal instrument is, mits natuurlijk goed geïnstrueerd, om zonder dat alfabet uit een verzameling de gegevens op te halen die de gebruiker vraagt. Als wij dan toch praten over alfabetisering komt dat omdat wij van een nieuw woordenboek een bestand maken dat voor de gebruikers niet alleen in elektronische vorm, maar desgewenst ook als folio beschikbaar moet komen. Ik kom daar aan het slot van mijn beschouwingen nog op terug.

Wetenschappelijk

Waar een woordenboek als de *Grote van Dale* gemaakt kan worden volgens wetenschappelijke beginselen, moet aan een wetenschappelijk woordenboek een wetenschappelijke theorie, bijvoorbeeld met betrekking tot de beschrijving en presentatie van betekenissen, grammaticale en lexicale verbindingen enz., ten grondslag liggen. Bovendien dient zo'n woordenboek uitputtend te

zijn. Ik bedoel met dit laatste dat een wetenschappelijk woordenboek uit puristische of morele overwegingen geen woorden de toegang mag ontzeggen. Ook niet als die woorden volgens de thans geldende normen bijvoorbeeld zaken in de werkelijkheid een naam geven die als niet ethisch of esthetisch gelden. Ik denk daarbij natuurlijk aan allerlei woorden die verwijzen naar emoties, zoals vloeken, verwensingen en scheldwoorden, dysfemismen, racismen en etnofaulismen. Wetenschappelijk impliceert immers objectief selecteren, interpreteren en beschrijven. In een wetenschappelijk woordenboek dient de lexicograaf descriptief te werk te gaan. De hedendaagse lexicograaf kent niet meer de emoties die de vermaarde en o zo menselijke achttiende-eeuwse Engelse lexicograaf Samuel Johnson (1709-1784) zo sympathiek, maar tevens subjectief maakten. Ik breng u daarvoor de beruchte definitie van het zelfstandig naamwoord *oats* ‘haver’ in herinnering: ‘a grain, which in England is generally given to horses, but in Scotland supports the people’.

Op deze plaats in mijn betoog heb ik er behoefte aan om het zojuist genoemde aspect ‘wetenschappelijk’ onmiddellijk te relativiseren. Ik doe dat met een citaat van de cardioloog Dunning, die schrijvend over de zestiende-eeuwer Michel de Montaigne, opmerkt: ‘Wetenschap is een product van menselijke interpretatie en creativiteit, wisselend in de tijd en zonder absolute of objectieve waarheid. Wetenschappelijke kennis is afhankelijk van de waarnemer, zijn theoretische vooronderstellingen, de aanvaarde verklaringsmodellen en de wereld waarin ze ontstaan en worden afgelost door nieuwe’ (Dunning 1997: 36).

Hedendaags Nederlands

Grosso modo is het hedendaags Nederlands het communicatiemiddel van drie generaties Nederlandssprekenden. Betekent dit nu dat de taal van mijn moeder, die dit jaar 84 jaar is geworden, in het W(E)TEN wordt beschreven? Anders geformuleerd betekent dit dat de materiaalverzameling die ten grondslag zal liggen aan genoemd woordenboek bronnen zal omvatten vanaf

1900? Ik bepleit dat niet en bevind mij in dezen in goed gezelschap (Verkuyl 1994: 3). Ik wijs die tijdsspanne af, niet omdat ik zo gemakkelijk kan definiëren wat hedendaags Nederlands is, maar omdat ik voor de beschrijving van wat daarvoor doorgaat het hedendaags Nederlands vereng tot de taal van de hedendaagse maatschappij. Die hedendaagse maatschappij definieer ik dan als de maatschappij vanaf 1960 tot 2010. Nederlandse teksten die geproduceerd zijn in en over die maatschappij worden daarbij gekozen als basismateriaal omdat lexicografen denken dat zij de woordenschat bevatten die wij willen beschrijven.

Is het voorstel om het hedendaags Nederlands te begrenzen met 1960 en 2010 meer dan een elegante en praktische hypothese?

In de jaren zestig en later gaan de veranderingen zich aftekenen die onze samenleving en cultuur werkelijk gaan veranderen en in het voetspoor daarvan ook de taal. De media winnen aan invloed en verruimen het wereldbeeld van iedere Nederlander. De generatie die na 1945 geboren is en die de volwassenheid bereikt in de late jaren zestig en in het decennium van 1970 tot 1980 gaat zich afzetten tegen het gezag. Ik noem bijvoorbeeld de verlangens naar democratie op de universiteiten, ingeleid door de bekende Maagdenhuisbezetting in Amsterdam van 1969, de kraakacties, de Kabouters, het optreden van de Dolle-Mina-beweging, het jeugdtoerisme, de Damslapers, de strijd voor biologisch-dynamische voeding, de grote antikernwapenbetogingen, de toenemende ontkerkelijking, de veranderde bedrijfscultuur (met inspraak en ondernemingsraden). Het betreft hier allemaal exponenten van diepgaande veranderingen in de samenleving en ze vinden hun weerspiegeling in veranderingen in de Nederlandse taal. Uiteraard manifesteert zoiets zich in de woordenschat, maar niet alleen daar (Van den Toorn 1997).

De eindgrens van 2010 is wel ingegeven door praktische en opportunistische overwegingen. Voorspellende gaven heb ik niet dus een antwoord geven op

de vraag welk jaartal in het begin van de eenentwintigste eeuw een periode markeert, kan ik niet. De keuze van 2010 berust m.a.w. op willekeur.

De keuze voor een vaste begin- en eindgrens betekent in elk geval dat de wetenschappelijke lexicografie van het Nederlands ook vanuit het perspectief van kosten beter beheersbaar is. Zij is dat eens te meer als de subsidiërende overheid zich engageert voor een bekostigingstermijn van 20 jaar. Ik teken daarbij overigens onmiddellijk aan dat een woordenboek van het hedendaags Nederlands natuurlijk nooit af is omdat menselijkerwijze gesproken de hedendaagse maatschappij nooit af is. Door te werken met een vaste begin- en eindgrens krijgt iedere generatie de gelegenheid om haar Nederlands volgens de in haar tijd geldende wetenschappelijke normen en opvattingen te beschrijven.

Standaardtaal

Met een definitie van taal van de hedendaagse maatschappij zijn wij er echter niet. In die maatschappij worden ook dialecten, regiolecten, sociolecten, stadsdialecten en groepstalen gebezigd. W(E)TEN moet een woordenboek zijn van de bovengewestelijke standaardtaal. Deze term, en ik vind dat dit niet vaak genoeg gezegd kan worden, heeft het gewonnen van de aanduiding Algemeen Beschaafd Nederlands (ABN), omdat men die aanduiding te discriminerend vond. Immers, wat is precies beschaafd? Onder bovengewestelijke standaardtaal zou men het ideaal van Van Haeringen (1925) kunnen verstaan, die daaronder de taal begreep waaraan je niet kon horen waar iemand vandaan kwam. Later is dat omschreven in termen als verzorgde schrijftaal of niet regionaal noch sociaal gemarkeerd Nederlands. Maar het ideaal van Van Haeringen heeft schipbreuk geleden. De ontwikkelingsgang van het Nederlands laat een andere ontknoping zien.

De gedemocratiseerde maatschappij heeft er sinds de jaren zestig voor gezorgd dat het Nederlands informeler werd en dat er steeds minder woorden van het etiket onverzorgd of slecht Nederlands werden voorzien. Die

informalisering betekent dat de standaardtaal gaat divergeren, d.w.z. in plaats van een na te streven norm ontstaan er steeds meer variëteiten van die standaardtaal. Omdat dialecten zich ontwikkelen tot talen die niet langer een lokaal karakter hebben, maar in een steeds groter gebied gebruikt worden, kruipen standaardtaal en dialect steeds dichterbij elkaar toe. En niet alleen op het gebied van de uitspraak. Lexicaal, grammaticaal, morfologisch en syntactisch brengt die ontwikkeling problemen met zich mee voor de lexicograaf. De taalgemeenschap verwacht juist van hem immers dat hij kleur bekent met betrekking tot de juistheid van die toenemende variatie. Niet dat hij niet opgewassen zou zijn tegen het taalgebruik van Leidenaar Rubberen Robbie of tegen dat van Haagse Harry. In teksten van die vertolkers van het lokale levenslied gaat het om primaire dialectkenmerken, m.a.w. om dialectverschijnselen waardoor het Leids, het Hoge of Lage Rijndijks, zich onderscheidt van bijvoorbeeld het Amsterdams of Delfts, en dat heeft niets met de divergentie van de standaardtaal te maken.

De problemen met betrekking tot de niet-regionale gemarkeerdheid van het hedendaags Nederlands liggen meer in het Zuiden van ons taalgebied, waar de vraag naar de taal die gesproken en geschreven wordt door alle Nederlandstalige Belgen nog beantwoord moet worden. De lexicograaf van 1997 heeft nog geen empirische basis tot zijn beschikking op grond waarvan hij kan besluiten wat algemeen Vlaams taalgoed is en dus opgenomen dient te worden in een woordenboek van de hedendaagse Nederlandse standaardtaal. Die empirische basis moet overigens gemaakt worden door Vlamingen en dient rekening te houden met de aanvaardbaarheidsoordelen van de Vlaamse taalgebruikers zelf (Taeldeman 1992: 40; Deygers 1996). Het betreft daarbij niet alleen woorden ‘die officiële of eigenaardige functies in of instellingen van de Belgische staat weergeven’, maar ook ‘woorden, uitdrukkingen of betekenisnuances waarvan het gebruik zeker tot de standaardtaal in België beperkt is, terwijl het daar dan zo ruim verspreid is dat het niet als gewestelijk beschouwd kan worden’ (Claes, 1985:22). Een paar voorbeelden kunnen de zaak verduidelijken. Moeten *Ardeens*, *badstad* ‘badplaats’,

beenhouwer ‘slager’, *behangpapier* ‘behang, behangselpapier’, *bibbergeld* ‘gevaarpremie’, *bijzit* ‘man die ongehuwd met een vrouw samenleeft’, *bobijn* ‘spoel, garenklos’, *broekkousen* ‘panty’, *dieetwinkel* ‘reformhuis’, *drinkgeld* ‘fooi, zakgeld, service’, *fruitsap* ‘vruchtensap’, *herpakken* ‘zich herstellen, weer moed vatten’, *onderpastoor* ‘kapelaan’, *studentenkot* ‘kast’, *zonnslag* ‘zonnesteek’ enz. in een W(E)TEN opgenomen worden of niet? Het antwoord is ‘ja’, als de genoemde empirische basis daarvoor voldoende evidentie bevat en het aanvaardbaarheidsoordeel van Vlamingen daarover positief is.

Wat ik opmerkte over de standaardtaal in Vlaanderen geldt ook voor het Nederlands dat typisch mag heten voor Suriname en de Nederlandse Antillen. Waarom zou een volbloed Nederlands woord als *gasbom* ‘met gas gevulde cilinder, gebruikt om te koken’, dat uitsluitend algemeen gebruikt wordt in Suriname, niet in W(E)TEN beschreven moeten worden? Of een woord als *ordeteken* dat wij kennen als *ereteken* (De Bies 1994)? Ook op het niveau van de woordenschat van de standaardtaal kenmerkt het Nederlands zich als een eenheid in verscheidenheid.

Vaktermen

Bij het begrenzen van wat tot de verzorgde, niet-gewestelijk gebonden standaardtaal behoort, doen zich nog andere problemen voor. Een van die problemen is het beantwoorden van de vraag welke vaktermen een plaats in het woordenboek verdienen. Moet men zich wellicht beperken tot alleen die vaktermen die in de algemene taal een rol spelen? En hoe bepaal je dat dan? Moet je bij de selectie van bronnen het aandeel van de specifieke vakliteratuur in de materiaalverzameling beperken? Voor zover mijn kennis reikt, bestaat er nog geen enkel criterium op grond waarvan men kan uitmaken wat tot de algemene woordenschat behoort en dus ook niet wat tot het vakexterne jargon. Wel bieden teksttypologische studies aanknopingspunten om op basis van interne kenmerken van een tekst tot genuanceerder uitspraken te komen. Maar, als vaktermen als *glycoproteïnetransferrine*, *acetylglucosamine*, *mannose*, *galactose*, *glutamyltranspeptidase*, *transferrine*, *malariaprofylaxe* of

erytromycine voorkomen in de materiaalverzameling waarop wij W(E)TEN baseren, dan dienen zij als sprekende getuigen van het hedendaags Nederlands beschouwd en in de basisnomenclatuur opgenomen te worden. Die verzameling is toch in feite het enige objectieve criterium dat voor selectie gehanteerd kan worden. Alle overige criteria, inclusief het inschakelen van specialisten, blijven op onderdelen aanvechtbaar. Maar hiermee loop ik op de zaak vooruit.

Groepstalen

Wat ik zo-even zei over de opname van vaktermen, geldt in gelijke mate voor allerlei groepstalen: de vluchtige en nauwelijks betrapbare jeugdtaal, de taal van Jan Soldaat, de taal van studenten, van katholieken of van lesbische vrouwen. Zij maken slechts deel uit van de algemene woordenschat als zij zich verschansen in een materiaalverzameling die pretendeert in de mate van het mogelijke representatief te zijn voor die algemene woordenschat.

Leenwoorden

En wat te doen met al die nieuwe leenwoorden die het Nederlands overstelpen? Een strenge kanttekening is hier op haar plaats. In geschreven vorm wordt het Nederlands niet door lexicale vreemdelingen overstelpt, in tegenstelling tot de gesproken taal misschien, waarin men zich, wellicht om anderen vol bewondering naar zich te laten kijken, nog wel eens zinnnetjes laat ontvallen als: ‘Deze urban villa's, krijgen ten onrechte een Engelse naam. Althans volgens mijn judgement.’ Volgens mijn observatie is het met de zondvloed van uitheemse woorden in het Nederlands helemaal niet zo droevig gesteld als velen ons willen doen geloven. Weliswaar laat de internationalisering Nederland en dus het Nederlands niet onberoerd, maar emotionele woorden als overwoekering, naasting en bedreiging zijn daarop niet van toepassing, zoals mijn onderzoek naar de actualisering van de negende druk van de Grote van Dale (Kruyskamp 1970) leert. Die wijzigingen voor de negende druk zijn eerst als afzonderlijk supplement verschenen. Het is dat supplement, dat in totaal uit 16.164 trefwoorden bestaat, dat ik per trefwoord

gecodeerd heb met categorienamen als: samenstelling, afleiding, eigennaam, letterwoord, nuance, ongeleed, leenwoord, verkorting en uitbreiding. Het aantal leenwoorden dat aan de negende druk werd toegevoegd, bedroeg niet meer dan 350 of wel 2,17%. Dezelfde exercitie herhaalde ik op het nieuwe materiaal voor de nieuwe druk van de woordenboeken van hedendaags Nederlands die onder mijn hoofdredactionele verantwoordelijkheid bij Van Dale uitkomen (Van Sterkenburg 1991a, 1995). Die updating omvat de periode 1985-1995. Op een totaalbestand van 4.574 trefwoorden werden er daar 517 als leenwoord gecodeerd. Dat is 11,30%. Deze toename in 25 jaar met meer dan 9% vind ik niet verontrustend. Zeker niet als men bedenkt hoe klein de moderne media de wereld gemaakt hebben. Die leenwoorden, vaak internationalismen, bedreigen het Nederlands niet. Dat laatste gebeurt eerst als kinderen het Nederlands niet meer als moedertaal leren, er geen wetgeving meer in het Nederlands verschijnt of het onderwijs niet meer in het Nederlands gegeven wordt. Ook hier geldt mijns inziens slechts één criterium voor selectie: het voorkomen in de als uitgangspunt gekozen materiaalverzameling. Die selectie is, zoals straks zal blijken, wel onderworpen aan het vetorecht van de lexicograaf.

Verklarend, receptief en productief, relationeel

Sprekend over de identiteit van W(E)TEN hebben wij opsommenderwijs ook gezegd dat het een verklarend woordenboek moest zijn. Dat wil zeggen dat van ieder woord aangegeven wordt welke betekenissen het heeft. Met behulp van een definitie wordt geprobeerd de betekenis die in het hoofd van de taalgebruiker zit in andere woorden op heterdaad te betrappen. Je kunt ook zeggen dat de definitie de betekenis van het woord verklaart. Maar niet alleen de hoofdbetekeningen, ook de betekenisnuances. Verklarend wil eveneens zeggen dat inlichtingen worden gegeven over de gebruiksmogelijkheden van het woord. Beter is wellicht te spreken van beperkingen in gebruik, omdat dit soort informatie altijd betrekking heeft op het voorkomen in een bepaalde tijd, in een bepaalde stijlslag, in een bepaalde groep, in een bepaald domein of in een bepaalde regelmaat. De karaktereigenschap verklarend

voert ons van de vorm van het woord naar het begrip en onderstreept daarmee de receptieve functie van zo'n woordenboek. Taalkundig geformuleerd wil dat zeggen dat het een semasiologisch woordenboek is. Maar een wetenschappelijk woordenboek als het onderhavige moet ook de relaties van woorden met andere woorden beschrijven, moet m.a.w. bij iedere betekenis, bij ieder begrip de betrekkingen daarvan met andere woorden in kaart brengen. Het moet daarom ook onomasiologisch zijn. De gebruiker moet kunnen vinden tot welke bovenklasse een bepaald woord behoort, maar ook tot welke onderklasse. Hij moet ook kunnen zien welke netwerken de afzonderlijke betekenissen van de woorden vormen. En als hij gelooft in synoniemen, moet hij bij de desbetreffende betekenis de min of meer gelijkwaardige woorden en woordgroepen aangeboden krijgen. Hiermee wordt W(E)TEN ook tot een productiewoordenboek, althans ten dele. Het wordt dat eerst ten volle als het ons van begrip naar woord voert.

Documentaire functie

In samenstellingen beschikt de moedertaalspreker van het Nederlands over een geraffineerd en op maat gesneden instrument om nieuwe dingen in de werkelijkheid een naam te geven. Benoemt hij iets nieuws, dan doet hij dat bij voorkeur met een samenstelling. Een enkel voorbeeld.

Als er een instrument wordt uitgevonden dat de temperatuur berekent als een functie van de hoeveelheid door het trommelvlies afgegeven infraroodstraling, is de samenstelling *trommelvliesinfraroodthermometer* snel geboren. Samenstellingen houden ons een spiegel voor van maatschappelijke veranderingen in de hedendaagse maatschappij. Als er tumult ontstaat rond het affiche dat het Groninger Museum ter promotie van een tentoonstelling van de Amerikaanse fotograaf Andres Serrano wilde verspreiden, ontstaan er samenstellingen als *plasposter*, *plasseks*, *plasseksaffiche*, *plasseksfoto* en *plasseksposter*.

In de negende druk van de Grote van Dale werden 8.443 nieuwe samenstellingen opgenomen op een totaal van 16.164 nieuwe trefwoorden. Dat is

52,23%. Voor de Hedendaagse 1985-1995 luiden die getallen 2.850 op 4.574, of wel 62,31%.

Als men het standpunt huldigt dat ieder woordenboek van de algemene taal ook een documentatiefunctie heeft, dan moet het ook de woordenschat van die taal registreren en vastleggen zoals deze op een bepaald moment in de tijd is samengesteld, en daardoor weerspiegelen wat voor die samenleving van dat moment kenmerkende omstandigheden en ontwikkelingen zijn. Onder die invalshoek zijn samenstellingen, zowel de doorzichtige als de ondoorzichtige, van uitzonderlijk belang (Moerdijk 1994: 14).

Voor een uitputtend wetenschappelijk woordenboek van het hedendaags Nederlands vormen de samenstellingen niet langer een probleemcategorie. Het zal niemand meer verbazen dat in mijn opvatting het corpus opnieuw dicteert welke samenstellingen in W(E)TEN worden opgenomen. Niet alleen omdat zij ons, zoals gezegd, maatschappelijke ontwikkelingen laten zien, maar ook omdat zij nieuwe woordvormingspatronen aan het licht kunnen brengen. Zo maakte Matthias Hüning mij attent op de eendagsvlieg *Irangate*. Maar blijkens zijn verzameling heeft het Nederlands in de laatste jaren een hele reeks woorden op *-gate* ontleend: *Irakgate*, *Saddamgate*, *Dianagate*, *Camillagate*, *Tamilgate*. Het gaat hierbij om een internationaal gangbare naam voor een bepaald schandaal, waarbij het eerste deel van het woord duidelijk maakt om welk schandaal het gaat. Natuurlijk stond *Watergate* hier model. Wij vinden intussen evenwel ook enkele ondubbelzinnige nieuwvormingen, zoals *stadiongate* ‘renovatie van het Amsterdamse Olympische Stadion’ en *klusjesgate* ‘bijbaantje van hoogleraar dat tot een schandaal leidt’. Aan laatstgenoemde schandaal danken wij overigens ook de samenstelling *snabbelprof*.

Een documentaire functie hadden WNT en MNW in extremis. Zij namen immers een representatieve selectie op van de citaten, van de woorden in context, die de bouwstoffen vormden voor de redacteur ter bewerking van

een bepaald woord. Die selectie diende de interpretatie van de lexicograaf te staven en nader toe te lichten. Dat deel van de documentaire functie verandert in W(E)TEN. Met citaten en illustratieve voorbeelden zou ik de afzonderlijke trefwoorden niet willen bemannen. Zij dienen wel op een of ander elektronisch medium ter beschikking te staan van de gebruiker. Een cd-rom zou in dat verband al prachtig zijn.

3. Materiaalverzameling

Als men vastgesteld heeft wat de identiteit van het te compileren woordenboek is, rijst de vraag: ‘Hoe kom ik aan de noodzakelijke data die voor beschrijving in zo'n woordenboek in aanmerking komen?’

J.B. Drewes (1976), de bewerker van het woordenboek van M.J. Koenen, schreef in het voorwoord van zijn handwoordenboek: ‘Voor zover een taalbeschrijver daarvan nog niet doordrongen mocht zijn, heeft het gebruik van de computer in de taalkunde hem kunnen leren, dat hij slechts een zeer geringe taalervaring heeft en weinig taalfeiten kent’. Alle lexicografen roepen dan ook in koor om een corpus, een verzameling teksten. Maar er bestaan verschillende soorten corpora en welk type verschaft aan W(E)TEN het beste lexicale cement? Voor W(E)TEN, dat zoals gezegd een periode van 1960-2010 beschrijft, pleit ik voor een zeer groot corpus, in termen van tientallen miljoenen woorden. Het corpus moet zo groot mogelijk zijn en het moet representatief zijn. Helaas wordt het begrip representatief gekenmerkt door de gladheid van een epistel. Al in 1988 toonde mijn Amsterdamse collega Willy Martin aan, hoe moeilijk je het statistische begrip representativiteit kunt toepassen op de samenstelling van een tekstencollectie die als representatieve steekproef van ‘het’ Nederlands van een bepaalde periode moet dienen. De meeste deskundigen op dit gebied erkennen tegenwoordig dat representativiteit onmogelijk is. Het streven naar een zo groot en gevarieerd mogelijke verzameling, zonder de claim van statistische representati-

teit, wordt momenteel dan ook de meest haalbare optie geacht (Vgl. ook Kruyt 1995).

Voor W(E)TEN zou ik allereerst een referentiecorpus willen hebben. Een dergelijk corpus staat het dichtst bij het verlaten concept van het representatieve corpus en wordt daarom ook wel *corpus met gebalanceerde samenstelling* genoemd. Een referentiecorpus wordt samengesteld op basis van relevante variabelen waarover binnen de gemeenschap van linguïsten overeenstemming bestaat (Sinclair 1991; Teubert 1995: 120; Kruyt 1996). Dit corpus zou ik het basiscorpus willen noemen. Ik teken daarbij tevens aan dat het uit verschillende deelverzamelingen kan bestaan. Bijvoorbeeld uit een verzameling van literaire teksten, een verzameling van krantentaal en een verzameling van technisch Nederlands. Van zo'n basiscorpus moeten alle elementen op de een of andere wijze gehonoreerd worden voor het lexicon. Die honorering kan ook betekenen een plaats in een elektronische vergaarbak. Er is evenwel een tweede corpus nodig voor W(E)TEN en wel voor het voortdurend actualiseren van het materiaal. Wij starten immers met een woordenboek dat de periode bestrijkt van jaren die voor een deel nog moeten aanvangen. Ik pleit daarom ook voor een monitorcorpus. 'Een monitorcorpus is een corpus dat vanuit een welomschreven uitgangssituatie veranderingen bijhoudt ("monitort") binnen bepaalde goed te traceren taalvarianten - krantentaal bijvoorbeeld. Boven een zekere lexicografisch bepaalde verzadigingsgrens komen nieuwe voorkomens van een woordvorm (met hun directe context) niet bovenop, maar in plaats van, eerder geregistreeerde voorkomens van dezelfde woordvorm (waarbij overigens wel de "totaalscore" van alle voorkomens per woordvorm wordt bijgehouden). Daarnaast geldt een bepaalde, automatische "vervalwaarde" voor woordvormen waarvoor geen nieuwe voorkomens worden geregistreeerd. Op deze manier wordt het bestand a.h.w. voortdurend ververst, en kan worden bijgehouden welke woordvormen om zich heen grijpen en welke er in onbruik raken' (Meys 1991: 321). Via zo'n monitorcorpus betrappen wij niet alleen nieuwe woorden, maar krijgen wij ook aanwijzingen over betekenisveranderingen van bestaande woorden op

basis van veranderingen van de context waarin die woorden voorkomen. Bijvoorbeeld doordat andere woorden, zoals *verpleging*, *ziekenhuis*, *patiënt* in de omgeving van *versterven* frequenter worden.

Voor het registreren van contemporaine ontwikkelingen is alleen een monitorcorpus niet voldoende. Als dr. L.G.M. Stevens, hoogleraar fiscale economie aan de Erasmus Universiteit Rotterdam, in *NRC/H.* van 19 maart dit jaar als cri de coeur laat afdrukken: ‘Rigide onderwijsregels frustreren scharrelstudenten’, dan worden die *scharrelstudenten* wel door een monitorcorpus betrappt. Anders is het waar het gaat om betekenisveranderingen bij bestaande woorden. Ik geef een paar voorbeelden: *koffieshop* betekent ‘openbare gelegenheid waar men koffie kan drinken’. Hoe betrap ik via een monitorcorpus de hedendaagse betekenis ‘ruimte, gelegenheid waarin men (soft)drugs kan kopen en gebruiken’. Dezelfde vraag moet gesteld worden voor *gotspe* ‘banaliteit’, dat ten onrechte steeds vaker ook ‘godswonder’ gaat betekenen. Hoe attendeert mijn genoemde corpus mij op het feit dat *surfen* niet langer alleen maar betekent ‘zich staande op een plank voortbewegen op de branding van de zee’, maar ook ‘zappen op internet’. Hoe kom ik er achter dat *jungle* tegenwoordig ook verwijst naar ‘omgeving waar het recht van de sterkste geldt’? Om van het werkwoord *mekken* met de betekenissen ‘zaniken’ en ‘het voor geiten kenmerkende geluid maken’, maar te zwijgen. Wie van u weet al dat het ook ‘Mac Donald’s bezoeken’ betekent? En dat het echt met een *e* gespeld wordt? Voor het opsporen van dat soort betekenisontwikkelingen schiet de computer nog te vaak te kort. Een netwerk van goed getrainde correspondenten is, naast dat monitorcorpus, voor de lexicograaf onontbeerlijk om dit soort vernieuwingen te signaleren, zoals bijvoorbeeld de redactie van *Ordbog over Dansk Nutidssprog* met *Spordhund* laat zien.

De twee corpora voor W(E)TEN, het referentie- en het monitorcorpus, omvatten corpora die bestaan uit volledige teksten en losse zinnestjes die door de zojuist genoemde zegslieden aangeleverd zijn. Die teksten zijn door de redactie zelf geselecteerd voor dit speciale lexicaal product. Ieder taalkun-

dig product stelt immers zijn eigen eisen aan de materiaalverzameling. Daarbij kan W(E)TEN een heleboel leren van de materiaalverzameling van het WNT (Van der Voort van der Kleij 1975; Moerdijk 1994), maar ook van de Europese expertise, zoals die ontwikkeld is in een aantal projecten als Network of European Corpora (Nerc) en Parole (Kruyt 1996). Een zo groot mogelijke variatie aan geschreven en gesproken teksttypen in de standaardtaal is daarbij vereist. Die typen kan men ondermeer indelen naar medium of naar topic. Met ‘naar medium’ bedoelen wij een verdeling van de verschillende teksten over bijvoorbeeld de categorieën woordenboek, boek, brief, krant, magazine, brochure, advertentie en voor de gesproken taal over conversatie, discussie en debat. Topic staat dan voor domeinen als godsdienst, techniek, wetenschap, sport, kunst, politiek, geschiedenis, medicijnen, filosofie enz. Hoe de verhoudingen moeten liggen in de verdeling over medium, literair genre, topic, medium, e.d. wordt door de samenwerkende taalgebieden binnen Parole in ondergrenzen en bovengrenzen berekend.

Alleen een elektronisch referentie- en monitorcorpus zijn voor de lexicograaf niet genoeg. Voor hem is meer nodig dan de charme van een expositiehal. Die corpora moeten aan nog andere strenge eisen voldoen. Alle teksten en zinnen die er deel van uitmaken, moeten verrijkt worden. Met tekstverrijking bedoelen wij het toevoegen van allerlei soorten informatie aan elektronische teksten of aan taalkundige eenheden in die teksten. Daarom hebben wij een computerprogramma ontwikkeld dat woordvormen in een corpus automatisch verrijkt met woordsoortcodes en trefwoordvormen. Zo wordt bijvoorbeeld bij *zagen* automatisch vermeld, dat die vorm hoort bij het trefwoord *zien* en bij het zelfstandig naamwoord *zaag*. Door aan iedere woordvorm een woordsoortcode toe te kennen, is het mogelijk om bijvoorbeeld bij *erg* onmiddellijk te zien of wij met een zelfstandig naamwoord (*geen erg in iets hebben*) of met een bijvoegelijk naamwoord (*het is meer dan erg*) te maken hebben. Het hulpmiddel dat die lemmatisering alsmede de morfosyntactische codes per woordvorm aanbrengt, wordt in de vakliteratuur tagger-lemmatiseerder genoemd.

Naast automatische taalkundige verrijgingsprogramma's moet er ook een opvraagstelsel ontwikkeld worden waarmee men kan zoeken naar afzonderlijke woorden of naar woordpatronen. Het resultaat van een zoekopdracht is altijd een lijst van woorden met frequentie- of spreidingsgegevens of woorden en woordpatronen in hun context. De analyse van de resultaten wordt ondersteund door diverse sorteermogelijkheden van de concordanties. Binnen het INL zijn al die verrijgingsprogramma's en retrievalsystelen ontwikkeld en operationeel. Dat is plezierig voor de lexicograaf, maar wat hij vooral nodig heeft alvorens hij met zijn beschrijving per woord kan beginnen, is de volledige trefwoordenlijst, de nomenclatuur, uit zijn corpus. En de belangrijkste vraag die nieuwe woordenboekplannen een grote spanning geeft, is in dat verband dan ook: 'Kun je volledig langs automatische weg een nomenclatuur voor W(E)TEN afleiden?' Het antwoord is ontkennend.

Altijd blijft de lexicograaf nodig om de door de lemmatiseerder niet gecodeerde woordvormen te beoordelen. Hij bepaalt of zich daartussen materiaal bevindt dat tot de nomenclatuur gerekend moet worden en neemt maatregelen om de vereiste morfosyntactische codes aan te doen brengen. Hij bepaalt ook welke vormen de toegang naar het lexicon ontzegd moeten worden, omdat zij uniek specificerend zijn of evident fout. Formules, dialect, vreemde talen enz. wijst hij af en in feite bepaalt hij ook wat er met hapaxen gebeurt (Vgl. Van Sterkenburg 1997a). Als er van de volgende vier voorbeelden maar één bewijsplaats in het hele corpus voorkomt, bepaalt de lexicograaf of hij deze woorden beschrijft of op een lexicale parkeerplaats achterlaat. De voorbeelden:

- (1) De vossenstand in de duinen is zo groot, dat het natuurlijke voedsel opraakt en de dieren steeds vaker hun kostje in de stad bij elkaar scharrelen. Het gezelschap van jagers en een enkele natuurliefhebber hoefde gisteravond niet eens zo lang te wachten tot de 'patatvossen' - zoals ze al worden genoemd - zich lieten zien.' *Haagsche Courant* 10 oktober 1996. Het gaat hier natuurlijk om de status van het woord *patatvos*.

- (2) ‘De Schotse BBC houdt vol dat de kerk als geheel sinds een jaar of vijf een “sekseverzekering” heeft. Het zou om een speciaal fonds gaan, ondergebracht bij een maatschappij in belastingparadijs Guernsey, waaruit zowel de slachtoffers van ongewenste pastorale intimiteiten als priesters die hiervan valselijk beschuldigd zijn schadeloos kunnen worden gesteld.’ *Haagsche Courant* 5 oktober 1996. Hier verdient *sekseverzekering* de aandacht van de lexicograaf.
- (3) ‘Bij zijn oratie als bezetter van de Cleveringastoel aan de Leidse universiteit heeft mr. B. Asscher, de oud-president van de rechtbank in Amsterdam, het integratiebeleid vanmiddag sterk bekritiseerd. “Het integratiebeleid van de Nederlandse overheid faalt. De overheid probeert diverse problemen met tientallen minderheidsgroepen in één samenhangend beleid aan te pakken. Dit “stampotbeleid” is echter tot mislukking gedoemd”, zei hij.’ *Haagsche Courant* 26 november 1996. In de stad van de *ollapodrida* of gewoon hutspot, twijfelt men wellicht minder aan het bestaansrecht van ‘*stampotbeleid*’, maar de doorsnee-lexicograaf doet dat onmiskenbaar.
- (4) ‘Minister Sorgdrager van Justitie heeft haar bedenkingen bij het moderne “onderhandelingshuishouden”, waarin kinderen en ouders op gelijke voet afspraken maken. Overleg is goed, maar de ouders zijn uiteindelijk verantwoordelijk voor de uitkomst. Een kind dat niet af en toe op ouderlijke grenzen stuit, zal het moeilijk krijgen: met zichzelf, met anderen en uiteindelijk met de samenleving, aldus Sorgdrager bij de opening van de studiedag van het Nederlandse Studiecentrum Criminaliteit en Rechtshandhaving op 16 oktober in Amsterdam.’ *Staatscourant* 210 17 oktober 1996. De vraag die de woordenboekmaker zich hier stelt, is: ‘Wat doe ik met *onderhandelingshuishouden*?’

Persoonlijke voorkeuren mogen geen rol spelen bij het opnemen van woorden. Altijd moet er een argumentatie aan de selectie ten grondslag liggen (Vgl. daarvoor Van Sterkenburg 1997a).

De definitieve nomenclatuur van W(E)TEN zal, als wij het beknopt moeten samenvatten, berusten op: (a) de taalkundig verrijkte populatie van een referentie- en monitorcorpus, (b) materiaal van correspondenten dat actualiteitswaarde heeft, (c) een afgesproken frequentie en spreiding waarmee de woorden moeten voorkomen en (d) de expertise van de lexicograaf, die bij zijn beslissingen natuurlijk ook gebruik maakt van wat andere woordenboeken van het eigentijdse Nederlands gedaan hebben.

Nu wij weten hoe wij aan onze hedendaagse Nederlandse woorden kunnen komen, kunnen wij overgaan tot zaken die te maken hebben met de wetenschappelijke interpretatie en beschrijving daarvan.

4. Beschrijving en theorie

Het blijft onmogelijk om in het bestek van deze openbare les alle onderdelen te bespreken die onlosmakelijk verbonden zijn met de lexicale beschrijving van woorden. Ik heb niet meer te bieden dan steentjes uit een mozaïek en omdat de lokroep van de beschrijving van betekenissen en gebruiksmogelijkheden en hun combinatorisch vermogen zo betoverend is, zal ik mij daartoe beperken, wetend dat ik de lexicografische werkelijkheid versimpel. Maar voor ik daartoe overga eerst nog dit.

De vraag of lexicografie een kunst is of een wetenschap is wat mij betreft meer een kwestie voor wetenschapsfilosofen. Persoonlijk ben ik van mening dat er geen praktijk is zonder theorie. Meer dan duizenden jaren hebben lexicografen lexicale triomfbogen en wolkenkrabbers opgetrokken en zoiets is onmogelijk zonder theoretische lexicografie, al was het alleen maar omdat iemand die een woordenboek schrijft zich bij alles wat hij doet laat inspireren door beginselen en beweringen van linguïstische aard. De lexicograaf is meestal een door de praktijk beproefde veldwerker die telkens opnieuw betere woordenboeken wil schrijven. Dat is evenwel eerst mogelijk als hij zich op

de hoogte houdt van de theoretische ontwikkelingen in de domeinen die relevant zijn voor lexicografie. Dat is dan o.a. de theoretische semantiek (Swanepoel 1994).

Het is het mal du siècle en tevens de uitdaging van ons vak dat er nog geen enkele samenhangende theorie bestaat om alle woordtypen te analyseren en de semantische structuur ervan systematisch te representeren. Melčuk met zijn lexicale functies, het semantische frame van Fillmore, de functionele grammatica van de veel te vroeg overleden Simon Dik, het generatieve lexicon van Pustojevski, de prototype semantiek van Lakoff en de conceptuele semantiek van Jackendoff, zij alle schieten tekort en noodzaken tot aanpassingen en aanvullingen (Verkuyl 1994: 4). Maar de lexicograaf kan daarmee leven. Hij gebruikt van de linguïstische theorievorming wat hij gebruiken kan en hij stoort zich niet aan een ongenueanceerde uitspraak als zou hij zich daardoor kwalificeren als de hyena van de linguïstiek.

De lexicograaf uit onze tijd is een zeer zelfbewuste taalkundige. Hij is pragmatisch, want hij maakt producten die door vele collega's gebruikt worden. Hij is goed geschoold en bereid om over de omheining van zijn eigen discipline heen te kijken. Hij is allesbehalve wereldvreemd en niet langer een 'harmless drudge' of dorre boekhouder van een taal. De redacteurs van W(E)TEN hebben hun training en opleiding gehad bij het WNT of VMNW. Daar hebben zij geleerd hoe zij de betekenissenmerken van verschillende woordsoorten kunnen identificeren. Daar ook hebben zij in praktijk gebracht op welke wijze die kenmerken het best beschreven kunnen worden. Als geen ander weten zij te onderscheiden tussen linguïstische betekenis en contextbetekenis. Dat alles betekent natuurlijk niet dat de beschrijving van de betekenis en de gebruiksmogelijkheden van de woorden op een identieke wijze moet gebeuren als in een historisch woordenboek van het type WNT.

De eerste vraag die de redactie van W(E)TEN zal stellen, is dan ook: ‘Wat is de geschiktste betekenisstructuur voor ons woordenboek? Is dat een platte structuur waarin alle betekenissen dezelfde status hebben, of is dat een hiërarchische?’ Met dat laatste bedoel ik dan: loopt de presentatie van de betekenis van algemeen naar bijzonder, van letterlijk naar figuurlijk, van frequent naar minder frequent, van prototypisch naar perifeer enz.

Aan het begin van een nieuw woordenboek zoekt de lexicograaf vooral ook naar nog betere antwoorden op vragen als: wat telt nu precies als een woordenboekbetekenis; op grond waarvan moeten de betekenissen van een woord of van een woordgroep onderscheiden worden; hoeveel en welke betekenissen moeten wij voor een specifiek woord of een specifieke woordgroep uittekenen; in welke gevallen moet een betekenis ondergebracht worden bij een bestaande betekenis en in welke gevallen moet er sprake zijn van een zelfstandige betekenis; wanneer breng je nuances onder een algemene noemer samen en wanneer splits je in een meer specifieke betekenis? (Vgl. ook Swanepoel 1994). En als iedere betekenis onderscheiden wordt in termen van bovenklasse en specificerende kenmerken geef je dan ook de relaties aan die de afzonderlijke betekenissen onderhouden met andere en hoeveel lagen onderscheid je daarbij? In het synoniemenwoordenboek dat ik voor Van Dale mocht maken (Van Sterkenburg 1991b), structureerden wij in de folio-uitgave slechts twee niveaus, omdat de typografie ons parten speelde. Die niveaus zijn dan de ingang (het hyperoniem) met alle directe hyponiemen. In voorkomende gevallen zijn er altijd doorverwijzingen naar hogere of lagere niveaus. Zeker in een elektronische uitgave is die beperking niet meer aanwezig.

En hoe zit het met aard en opbouw van de definitie? Niemand hoeft de redacteuren van het WNT te leren wat de beste lexicografische definitie is. Wat zij wel als wens hebben, is dat er voor bepaalde categorieën werkwoorden, zoals die welke een emotie uitdrukken, of voor werkwoorden van geluid vaste sjablonen ontwikkeld worden, die zij bij hun beschrijving kun-

nen gebruiken. Nemen wij een groep werkwoorden als de volgende: *bakken, barbecuen, blancheren, braden, frituren, fruiten, gratineren, grilleren, koken, pocheren, poffen, roken, roosteren, smoren, stomen, stoven*. Deze groep is niet uitputtend zult u zeggen. En dat klopt, want *fonduen* en *gourmetten* ontbreken bijvoorbeeld. Via een analytische definitie zal de beroepslexicograaf deze werkwoorden definiëren als ‘voedsel bereiden door verhitting’. Maar uit welke set kenmerken kan hij kiezen om al deze werkwoorden in voldoende en noodzakelijke kenmerken van elkaar te onderscheiden? Een sjabloon voor deze groep zou kunnen zijn:

- intensiteit en gradatie van de verhitting: lang, kort, snel, traag, heet, koel
- ingrediënt dat bij de bereiding gebruikt wordt: vocht, vet, as, olie, rook, kokend water
- instrument of apparaat dat gebruikt wordt: oven, bakplaat, braadpan, koekenpan, kookpan, rooster enz.
- selectierestricties op het onderwerp in termen van: bezielde, menselijk
- aard van de complementen bij het werkwoord: aardappels, kastanjes, spijzen, vlees enz.

Ik ben ervan overtuigd dat de lexicograaf voor het maken van adequate definities een bondgenootschap moet sluiten met de formele semantiek. Graag zeg ik Verkuyl na dat verbetering van de definities bewerkstelligd kan worden door een set van theoretische algebraïsche notaties te gebruiken. Niet in de definities zelf, maar als betrouwbare controle op die definities (Verkuyl 1994: 5 en 1996: 109-117, 125-128).

Sprekend over sjablonen en frames van samenhangende woorden is ook belangrijk wat Moerdijk opmerkt over het semantisch potentieel dat zo'n samenhangende groep van woorden heeft. Hij houdt een vurig pleidooi voor een frame voor de woorden die een handeling aanduiden en baseert zich bij de uitwerking daarvan op empirisch taal materiaal. Aan zo'n frame heeft de lexicograaf behoefte, omdat hij in zijn materiaal allerlei verschijnselen van

taalgebruik vindt die meestal als contextbetekenis of individuele afwijkingen werden afgedaan en waarvoor bij een polyseem woord zelden een structurele verklaring werd gegeven, laat staan dat de relaties tussen de conventionele betekenis en deze contextbetekenis tot een systematische beschrijving leidden. In het handelingsframe van Moerdijk is m.a.w. ook plaats voor gebruiksbetekeningen waartussen een aantoonbare samenhang bestaat, voor metoniemen dus (Moerdijk 1989; 1990; 1993).

Een mooi voorbeeld van zo'n geval van bijzonder taalgebruik is het woord *rilsnip*, dat in de koude winter van 1996-1997 ineens opdook. Het ging hier niet om 'een van kou rillende loopvogel met zeer lange buigzame snavel', maar om 'een biljet van honderd gulden dat bijstandsgerechtigden kregen als tegemoetkoming ter bestrijding van de hoge stookkosten ten gevolge van de extreme kou'. Tussen *snip* 'loopvogel' en de gebruiksbetekenis 'bankbiljet van honderd gulden' bestaat een betrekking die wij aanduiden met metonymie. Daarbij gaat het altijd om bijzondere betrekkingen: in ons geval om die tussen vogel en het voorwerp waarop die vogel is afgebeeld. Metonymische betrekkingen, zegt Moerdijk (1993: 26-27), 'zijn niet gebonden aan de afzonderlijke specifieke woordbetekenis maar vormen een soort van onderliggend stramien waarop de metonymische uitbreiding geborduurd kan worden'.

Wetenschappelijke woordenboeken mogen zich niet beperken tot de beschrijving van uitsluitend de conventionele betekenis van woorden. Ook de individuele afwijkingen moeten er in worden verantwoord. Moerdijk heeft de weg gewezen voor de handelingswoorden. Er moeten meer vergelijkbare stramieneen komen. En deze hartenkreet heeft niets met waan van de dag of met een nieuwe lexicografische hype te maken.

Wij hebben iets gezegd over een nog betere beschrijving van de betekenissen. Waar ik zelf buitengewoon aan hecht, is dat wij met W(E)TEN meer doen dan er woordvormen in alfabetische volgorde in opnemen, waarvan de

betekenis kan worden opgezocht. Ik wil een dynamischer product. Ik wil dat iedere betekenis die in W(E)TEN voorkomt, gerelateerd is aan het concept of de concepten waarmee die betekenis in een conceptuele structuur verenigd is. In een elektronisch woordenboek kan dat via wat hyperlinking heet. In eenvoudig Nederlands wil dat het volgende zeggen: een woordenboekbestand bestaat uit een eindig aantal knopen, waarin onze kennis is ondergebracht. Die knopen zijn door speciale computertalen (Standard General Markup Language, Hyper Text Markup Language) onderling te verbinden in een netwerk en te manipuleren. Het netwerk waarvan zij deel uitmaken is bovendien grafisch te presenteren, zoals Landman (1997) onlangs aantoonde. Maar ik wil nog meer; ik wil in de nomenclatuur naast alfabetische woordvormen, ook alfabetisch toegankelijke concepten, waarbij de hiërarchische structuur van de woorden die tot dat concept behoren, moeten worden aangeboden.

Wie zelf een synoniemenwoordenboek gemaakt heeft, is zich meer dan wie ook bewust van de problemen rond het iken van concepten. Een concept is immers geen vast gegeven, is ook niet constant en kan van gebruiker tot gebruiker verschillen (Uhlenbeck 1981: 15). Ik verwacht dat bestaande woordenboeken in dit verband goede diensten kunnen bewijzen. In het bijzonder verwacht ik veel van de omkering van het WNT, d.w.z. het alfabetiseren op betekenisomschrijvingen en in een eindfase ook van een omkering van W(E)TEN zelf. Het moge duidelijk zijn dat bij omkering alle betekenissen alfabetisch geordend worden met daarachter het trefwoord waaronder de desbetreffende betekenis is gevonden. Dat analytische en typologische definitieën een belangrijke rol kunnen spelen bij het opbouwen van de hiërarchie binnen concepten wordt duidelijk als men bedenkt dat dat soort definitieën meestal inzet met het vermelden van de bovenklasse waartoe het gedefinieerde woord behoort.

Als wij spreken over een nog betere descriptie in W(E)TEN dan kan men natuurlijk ook denken aan de verbetering van de beschrijving van de

gebruiksmogelijkheden van woorden. Ik doel daarbij niet zo zeer op rolpatronen bij werkwoorden of op wat in de traditionele literatuur bekend staat als selectierestricties, maar op domeinmarkeringen en stijlregisters, op labels die de aandacht vestigen op een bepaalde attitude van de spreker, op connotaties dus.

Laten wij er geen doekjes omwinden, in alle bestaande woordenboeken, dus niet alleen in die van het Nederlands, is de beschrijving van de gebruiksmogelijkheden het meest betwistbaar. Bij de domeinmarkering is, en ik weet dat ik door dat te herhalen fungeer als een soort gebedsmolen, vaak niet uit te maken of een vakterm nog vakintern is of reeds vakextern, d.w.z. ook voorkomt in een andere vaktaal dan uitsluitend die waarin hij geijkt is of misschien zelfs al in de algemene woordenschat. Een bijkomend probleem is de wijze waarop de markering wordt aangebracht. Wanneer wordt met een label gewerkt en wanneer mag uit de parafrase of definitie blijken dat het gaat om een woord of een betekenis, dat respectievelijk die specifiek is voor een groeps- of beroepstaal. Waar W(E)TEN wetenschappelijke objectiviteit predikt, moet hier naar een objectief criterium gezocht worden. Dat criterium kan welhaast geen ander zijn dan de spreiding. Komt het woord alleen voor in een bepaald domein, dan is de markering evident, komt het voor in verschillende bronnen die alle deel uitmaken van het subcorpus vaktaal, dan is het ook evident welke keuzes men moet maken bij zijn markeringen. Een actueel woord als *technolease* krijgt dan zeker geen label omdat er in alle dagbladen en verder over geschreven en gesproken is.

Veel moeilijker ligt het bij de stijlregisters en connotaties. Iedere moedertaalspreker van het hedendaags Nederlands weet, zonder dat de stoppen bij hem doorslaan, dat er met de volgende woorden wel iets aan de hand is. Ik noem hier *beneuzelen*, *je stinkende best doen*, *draaikonten*, *de grasdokter van de Amsterdamse ArenA*, *iemand een kloot afdraaien*, *lullo*, *modemmeid* of *webjuf*, *politiemeervoud* (u wordt aangehouden met de vraag: ‘Waar gaan wij naar toe?’), *spiekprof*, *snabbelprof* en *spookprof*, *uitbuiken*, *shit* en *fuck*.

Het instrument dat de taalkunde de lexicograaf voor de beschrijving van dit soort woorden adviseert, is het zogenaamde taalattitudeonderzoek. Men vraagt aan de taalgebruikers naar hun emotionele, negatieve of positieve waardering van woorden en woordgroepen als de genoemde. Zelfs als het om een bescheiden hoeveelheid woorden zou gaan, zou ik het de lexicograaf niet willen aandoen zijn lexicale descriptie te baseren op de resultaten van een onderzoek naar woordattitudes. In een recent door mij uitgevoerd grootschalig onderzoek naar vloeken (Van Sterkenburg 1997b), heb ik kunnen vaststellen dat de waarderingen van taalgebruikers te veel verschillen om ze met voldoende gezag onder een noemer in een woordenboek te kunnen brengen. Te veel variabelen beïnvloeden dat oordeel. De gemeten waarde per woord verschilt per leeftijdscohort, per regio, per geslacht zelfs, en is tevens afhankelijk van de vraag of men zich gelovig of niet-gelovig noemt. In een vragenlijst waarin gevraagd werd de taboewaarde op te geven van een aantal verschijnselen, gedragingen, scheldwoorden en vloeken bleek er tussen Nederlanders en Vlamingen een zeer grote consensus met betrekking tot o.a. kindermishandeling, verkrachting, moord, getuige zijn van extreem geweld, een hakenkruis, abortus en euthanasie en ook over een scheldwoord als *jood* ‘al te handige zakenman, bedrieger’ was men het roerend eens, maar de negatieve lading van vloeken als *shit*, *fuck*, *klote* en *godverdomme* verschilt opmerkelijk. En dan hebben wij het niet eens over de generatie tot 25 jaar. Daarin lijken woorden als de hier genoemde geen enkele connotatieve waarde te hebben, waarschijnlijk omdat zij in die leeftijdscategorie tot betekenisloze stoplappen geworden zijn.

Het lijkt mij ondoenbaar om bij ieder niet-neutraal trefwoord of bij iedere niet-neutrale betekenis informatie te verstrekken van het type ‘door de meeste’ of ‘door grote groepen moedertaalsprekers als beledigend, schertsend, eufemistisch, informeel e.d. ervaren’. W(E)TEN is geen speeltuin van taalattitudinale experimenten. Zonder aarzeling betoog ik hier dat de taalgemeenschap de zorg voor een wetenschappelijk verantwoorde indicatie van de gebruiksmogelijkheden van woorden moet toevertrouwen aan de expertise

van de lexicografen. Zij zullen in de inleiding op hun woordenboek verantwoorden hoe zij deze lexicale klip genomen hebben. Ik weet dat dit een gezagsargument is, maar op basis van de vele citaten die hij moet interpreteren, ziet de lexicograaf nu eenmaal het beste wanneer een woord van kleur verschiet.

Wij hebben het al bij herhaling gehad over het streven van de lexicografen om bij voortdurend de wijze waarop zij de woordenschat beschrijven te verbeteren. Ik heb voor W(E)TEN in dat verband nog een grote wens. Dat woordenboek moet koppig kiezen voor een exhaustieve beschrijving van wat heet de ‘discoursewoorden’. Van partikels (Foolen 1993), van verraderlijke woorden als *gewoon*, *precies*, *want*, *ooit*, *pakweg*, *okay*, *maar*, *jeetje*, die vaak een emotioneel oordeel bevatten of de houding van de spreker jegens iets of iemand verklaren. Deze woorden, die onze communicatie smeren, verdienen een betere lexicale beschrijving.

Laten wij opnieuw een paar voorbeelden ter verduidelijking in slagotde opstellen. Het bijwoord *ooit* wordt tegenwoordig zeer vaak gebruikt ter versterking van een superlatief: *de rijkste Nederlander ooit*, *de grootste onderneming ooit*, *de jongste cricket-international ooit*, *onze duurste aankoop ooit*. In geen enkel woordenboek van het Nederlands wordt dit gebruik reeds beschreven.

Wie naar de schrale betekenisprofielen kijkt van een woord dat in gesproken taal te pas en te onpas gebruikt wordt, het Engelse leenwoord *okay*, vraagt zich af of dat woord in het Nederlands op dezelfde wijze gebruikt wordt als in het Engels. Dus om te zeggen dat je iemand wel mag of dat je een situatie, een gebeurtenis, iemands gedrag enz. aanvaardbaar vindt; of om te zeggen dat iemand in veiligheid is of dat een machine of apparaat goed werkt en gebruikt kan worden. Kunnen wij in het Nederlands *okay* ook gebruiken om te zeggen dat wij het met iets eens zijn? Kunnen wij het gebruiken om na te gaan of de persoon waartegen wij spreken, begrijpt wat wij juist gezegd

hebben? En wordt *okay* bij ons ook gebruikt om duidelijk te maken dat wij iets anders willen gaan doen of over iets anders willen gaan praten? (Sinclair 1987; vergelijk ook Hofland in *NRC/H*. 23-5-97). Vooralsnog weten wij het niet, want de bestaande woordenboeken zijn te oppervlakkig of laten anderszins verstek gaan.

Okay. Als de lexicografen hun beschrijvingen echt willen verbeteren, dan kunnen zij hier in het domein van de discoursewoorden in samenwerking met grammatici en pragmalinguïsten opmerkelijke resultaten boeken. Ik zou echt willen beschrijven of die Haarlemmerolie van onze gesprekken, de woorden die waardeoordelen van en over mensen bevatten, ook nauw met onze ethiek verbonden zijn (Vgl. Niemeier, Dirven 1997).

Zaken die met betekenis en gebruiksmogelijkheden te maken hebben, heb ik in het voorafgaande uitvergroot. Maar over de combinatoriek van woorden heb ik nog amper iets gezegd. Hoe zit het met woordgroepen als *chemische drugs*, *dubbelblind onderzoek*, *ecologische roes*, *witte rook*, *fiscale emigrant*, *koude sanering*, *psychedelische paddo*, *stille pijn*, *Zaanse verhoormethode*, *publieke omroep*, *georganiseerde verlapzwanzing*, *elektronische winkel*, *digitale ziel* en *zwarte kabel*?

In mijn oratie van 1986, die als titel droeg *Werk maken van woorden*, heb ik aandacht gevraagd voor de plaats waar vaste verbindingen in een lexicon opgenomen moeten worden. Afhankelijk van de aard van een woordenboek bleken daarover afspraken te maken. Geldt dat ook voor andere zaken? Bijvoorbeeld voor een beter onderscheid van de aard van de verschillende verbindingen?

In W(E)TEN zou ik groepen van verbindingen willen onderscheiden en identieke soorten bijeen willen zetten. In beginsel zou ik willen onderscheiden in: (a) grammaticale verbindingen en syntactische patronen, (b) lexicale verbindingen en (c) pragmatische routineformules.

Onder grammaticale verbindingen en syntactische patronen versta ik dan de combinaties die het behandelwoord aangaat met vaste voorzetsels, met infinitiefconstructies, met bijwoordelijke bepalingen, met afhankelijke bijzinnen, met *of*-constructies enz.

Lexicale verbindingen zijn dan die verbindingen waarin het behandelwoord een vaste eenheid vormt naar vorm en betekenis. De betekenis van die verbinding kan identiek zijn aan de betekenissen van de woorden die samen de verbinding vormen, die betekenis kan gedeeltelijk de betekenis zijn van de samenstellende leden, men spreekt dan van een endosemantische verbinding, en die betekenis kan volstrekt afwijken van de betekenissen van de samenstellende delen. In het laatste geval spreken wij van idioom en van exosemantische betekenis (Klimaszewska 1990). Het spreekt vanzelf dat in W(E)TEN bij de endo- en exosemantische verbindingen betekenisverklaringen gegeven moeten worden. Dat geldt voor hedendaagse verbindingen als *publiek gaan, in blessuretijd leven, rechthoekige ogen hebben of krijgen, zich de tandjes werken en in de kennisfile staan*. En wat te denken van de typisch Leidse creatie *voor mij mag je koperen hartkleppen krijgen, dan kan je je levenslang de tering poetsen!*.

Pragmatische toutineformules zijn verbindingen die in concrete communicatiesituaties hoogfrequent gebruikt worden: *even Apeldoorn bellen; gaan met die banaan; zo kannie wel weer even; mijn naam is haas; geef mijn portie maar aan fikkie; ik ben ook maar een mier met een rugnummer*. In W(E)TEN zal de pragmatica van deze clichés niet mogen ontbreken. Opgezocht moet kunnen worden waar, hoe en door wie deze verbindingen gebruikt worden. Het onderzoek in dezen door Van der Have kan nauwelijks afgewacht worden.

Waar het gaat om de beschrijving van betekenis en gebruiksmogelijkheden van verbindingen, geldt hetzelfde als bij die van de betekenis zelf. Ook hier is behoefte aan het verbeteren van die beschrijvingen, aan modellen voor een goede representatie. Maar ook hier wordt van de lexicograaf veel lenigheid

verwacht omdat de door de theoretische linguïstiek ontwikkelde modellen niet uitmunten door een superieure universele constructie en daardoor te veel ad hoc-oplossingen vergen. Voor de beschrijving van de pragmatische, semantische en syntactische eigenschappen van vaste verbindingen zijn wij niet gebaat bij de dictatuur van leerstelligheid, maar bij een descriptie die voor het toetsen van haar eigen consistentie eveneens gebruik kan maken van algebraïsche notaties.

Wat voor W(E)TEN in vergelijking met het WNT een enorme progressie kan betekenen, is dat de informatietechnologie de techniek en methode om verbindingen automatisch uit grote materiaalverzamelingen te halen, heeft verbeterd.

Twee statistische hulpmiddelen om te gebruiken bij een automatische selectie van die verbindingen wil ik hier niet onvermeld laten. Er is allereerst wat heet de *Mutual Information Index (MI)*. Deze index wordt gebruikt om te berekenen welke woorden in een corpus in elkaars onmiddellijke nabijheid voorkomen, a.h.w. een woordpaar vormen. Zij kan dus berekenen hoe vaak *zwart* met *kabel* voorkomt of *humaan* met *identiek* of *geest* met *fles*.

Als wij de MI meten, stellen wij niet alleen de waarschijnlijkheid van de afzonderlijke woordvormen vast, maar ook de waarschijnlijkheid van combinaties van twee woorden, van bigrammen. Dit leidt tot drie waarschijnlijkheidswaarden die vergeleken kunnen worden:

- (1) de waarschijnlijkheid (W) van de eerste woordvorm: $w1:W(w1)$;
- (2) de waarschijnlijkheid van de tweede woordvorm: $w2:W(w2)$;
- (3) de waarschijnlijkheid van een paar ($w1, w2$) dat bestaat uit de twee woordvormen: $W(w1, w2)$.

Deze drie waarden worden vergeleken: de waarschijnlijkheid dat $w1$ en $w2$ samen voorkomen (bijvoorbeeld naast elkaar) wordt gedeeld door het product van de afzonderlijke waarschijnlijkheid van $w1$ en $w2$ (Church 1991; Heid 1994).

Het kan voorkomen dat een zeldzaam woord, bijvoorbeeld *verlapzwanzing* in *georganiseerde verlapzwanzing*, zeer frequent voorkomt. Dat kan bepaald worden met behulp van de T-test. Die test werkt namelijk op woordparen. Zij vindt die additionele woorden die waarschijnlijk meer tegelijk voorkomen met een van de twee woorden van het paar dan met het andere. De resultaten van de T-test verschijnen als positieve en negatieve waarden. De hoogste en laagste waarden zijn significant: zij duiden op een sterke relatie van het ene of andere woord.

5. Opslag

Voor de samenstelling van W(E)TEN hebben de ministers die samen het Comité van de Nederlandse Taalunie vormen ons een termijn gegeven van 20 jaar. Dat is bijna 130 jaar minder dan voor het optrekken van het WNT, die andere lexicografische kathedraal, die startte in het voorwereldlijke jaar 1851. Die lange bewerkingstermijn is begrijpelijk. In de tweede helft van de negentiende en de eerste van de twintigste eeuw moest er nog heel wat afgeploeterd worden in dit lexicografische tranendal. Het selecteren van bronnen, het zoeken, uitschrijven en verifiëren van citaten, het alfabetiseren en het typen en meer dan eens corrigeren van de kopij, het vergde net zoveel tijd als in de landbouw het dorsen met de dorsvlegel.

De invloed van de technologie op de lexicografische arbeid bracht, zoals in het voorafgaande bleek, grote veranderingen. Het geestdodend routinewerk wordt gelukkig steeds meer overgenomen door computersystemen. Maar niet alleen computersystemen hebben de lexicografie veroverd. Ook de consumentenelektronica is binnen ons domein niet meer weg te denken. Vooral waar het de opslag van data betreft. Dat geldt voor de televisie (1951), de fax (1970), de beeldplaat (1972), de cd (1982), de digital audiotape (dat; 1987), de digital compact cassette (dcc; 1992), de dvd oftewel de digital versatile disk, een cd die tien maal zoveel informatie kan opslaan als een gewone

(1997-1998) en Webtv, een soort afstandsbediening met toetsenbord waarmee men via de televisie a.h.w. kan surfen op internet.

Mechanisering en technologie maken het productieproces nog meer beheersbaar en kunnen beter voorkomen dat de planningen voor woordenboeken altijd te optimistisch zijn. Maar het zou bijna heiligschennis zijn om de productietermijn die nodig was voor de standaard van onze lexicografische ambachtelijkheid, het WNT, te vergelijken met die van W(E)TEN. Die vergelijking is oneerlijk omdat de identiteit van beide producten anders is, omdat de periode die zij bestrijken indringend verschilt, omdat de investeringen die gedaan zijn en worden in mensen en machines onvergelykbaar zijn en omdat het WNT zich niet hoefde te bekommeren om een marktperspectief, maar uitsluitend om de wetenschap sui generis. Dat laatste is in onze dagen volstrekt ondenkbaar. Op de vleugels van de profeten schreef Bernard Al reeds in 1993: 'Zonder marktperspectief heeft de lexicografie geen toekomst' (Al 1993:83).

Ook de wetenschappelijke lexicografie moet nu de vraag beantwoorden voor wie zij producten maakt als het W(E)TEN.

De markt die W(E)TEN wil bedienen is die van degenen die professioneel met het Nederlands bezig zijn. Die professionals moeten kunnen putten uit de krachtigst denkbare descriptie van het hedendaags Nederlands. Voor hen is een selectie niet goed genoeg. Zij willen alle informatie en vele, vele toepassingen van woorden in taalgebruikssituatie. W(E)TEN is vooral nodig voor het nageslacht. Alleen een uitputtende beschrijving van het hedendaags Nederlands stelt onze taalgemeenschap in de toekomst in staat om het Nederlands dat wij nu produceren, inclusief de rariteiten en eendagsvliegen, te begrijpen. Wil ons taalgebied evenmin achterblijven op het gebied van taalanalyse en -generatie dan moeten er niet alleen grote digitale verzamelingen van het Nederlands aangelegd worden, maar ook elektronische woordenboeken die gebaseerd zijn op corpora. W(E)TEN en de materiaalverza-

meling waarop dat woordenboek berust, moeten de producten zijn die ook voor de taaltechnologie interessant zijn. Zij moeten een veelvoud aan expliciete morfologische, syntactische en semantische data bevatten, die voor tekstverwerking, dialoog- en expertsystemen e.d. begerenswaardig zijn (Al 1993). Tenslotte moet W(E)TEN een uitgangbestand zijn voor allerlei derivaten waaraan de Nederlandse taalgemeenschap behoefte heeft, hetzij om binnen Europa begrip en respect voor elkaars cultuur en vaardigheden te bewerkstelligen, hetzij om in de internationale Europese context de meertaligheid te handhaven. 'Woordenboeken zijn een stuk infrastructuur, net als wegen. Het zijn verbindingswegen tussen talen. Het is belangrijk dat die er zijn. Wie er van deze verbindingswegen gebruik maken, is een andere zaak, maar je maakt communicatie mogelijk', zegt Martin tegen Sanders (1997: 24).

Is dit een schot in de rug van al die uitgevers die al vele jaren hun producten slijten aan het onderwijs en de algemene publiekmarkt? Nee, natuurlijk niet. Die uitgevers produceren allereerst geen corpora waarop hun lexicale beschrijvingen getoetst worden en bovendien zoeken zij niet naar een uitputtende, maar naar een beredeneerde beschrijving van het hedendaags Nederlands. Vaak streven zij zelfs bewust naar een beperking van het aantal ingangen om voldoende ruimte over te houden voor een goede beschrijving van de geselecteerde woorden. Omdat hun woordenboeken voor het onderwijs gebruikt worden, zijn zij ook vaak pedagogischer van aard. Zij sluiten aan bij de norm. Natuurlijk bieden zij niet de veelheid aan zinnen met woorden in hun natuurlijke omgeving. En behoudens de Grote van Dale huldigen alle redacties het beginsel dat woordenboeken er uitsluitend voor het nut zijn en niet voor het genot. En redundantie aan materiaal is wetenschappelijk intrigerend, maar commercieel niet nuttig.

W(E)TEN is bestemd voor de professionals, heb ik zo-even gezegd. Voor de vakman die beroepshalve Nederlands produceert of data onderzoekt om er explicietere wetenschappelijke uitspraken over het Nederlands over te doen.

Hij zal daarbij niet tegen de maan blaffen, maar bij voorkeur zijn elektronische werkomgeving koesteren. Dat rechtvaardigt de vraag of W(E)TEN wel als folio, als boek op de markt gebracht moet worden. Het is moeilijk voorstelbaar dat een professional met een boek zittend in zijn luie stoel zal grijpen naar W(E)TEN als hij bij zijn lectuur een woord tegenkomt dat hij niet begrijpt. Ik durf dan ook de stelling aan dat W(E)TEN bijna uitsluitend voor wetenschappelijke doeleinden, zeer in het bijzonder voor het toetsen van hypothesen en het produceren van teksten gebruikt wordt en zelden voor receptieve doeleinden. Op grond daarvan houd ik staande dat W(E)TEN uitsluitend in elektronische vorm beschikbaar hoeft te zijn.

W(E)TEN beconcurrert als elektronisch woordenboek het debiet van de uitgevers niet. Waar wij natuurlijk van uitgaan, is dat ons elektronisch bestand zo aantrekkelijk is dat die uitgevers het als basisbestand voor afgeleide producten willen exploreren en exploiteren. Daar komt bij dat de bewindslieden willen dat wij onze verrijkte en gestructureerde data vermarkten en dat is wat anders dan in de tijd toen het WNT als staatsproduct vogelvrij was. Toen het WNT nog geen onderdeel vormde van de Stichting INL kon het naar willekeur gebruikt en geplagieerd worden door iedere uitgever. Op geen enkele wijze wierp het copyright barrières op. Mijn droomvoorstelling in 1997 is dat van W(E)TEN, zowel voor de publiekmarkt als de algemene markt zoveel producten gemaakt zullen worden, dat er op den duur geen onderscheid meer bestaat tussen commerciële en wetenschappelijke lexicografie.

6. Peroratie en dankbetuigingen

Met anderen weet ik dat woordenboeken geen gevoel van nationale identiteit en trots oproepen. Dat doet in Nederland alleen koninginnedag en deed in België alleen Eddie Merckx. Wanneer ik dan toch met een knipoog naar Jan Jacob Slauerhoff zeg: ‘Alleen in mijn moedertaal kan ik wonen’, dan wil

ik daarmee een gevoel tot uitdrukking brengen, waarvan ik weet dat u het allemaal herkent. In mijn moedertaal voel ik mij het meest op mijn gemak; in mijn moedertaal kan ik liefhebben, lachen, janken, juichen, denken, spotten en vloeken. Alleen in mijn moedertaal kan ik spontaan mijzelf zijn. Mijn moedertaal is mijn ‘genenpaspoort’.

Het Nederlands wordt nog op grote schaal door kinderen als eerste taal geleerd en het ziet er voorshands niet naar uit dat dit zal veranderen. Het Nederlands is de taal van ruim 20 miljoen sprekers. Het Nederlands is de officiële taal van Nederland, Vlaanderen, Suriname en de Nederlandse Antillen, het Nederlands wordt op grote schaal op verschillende niveaus onderwezen en dagelijks door de media als voertaal gebruikt. Er is een Nederlandse schrijftaal met een omvangrijke, levende letterkunde. Het Nederlands is bovendien productief en creatief. Die eigenschappen maken het mogelijk dat onze taal maatschappelijke veranderingen kan volgen en aan nieuwe communicatieve behoeften kan voldoen. Het Nederlands kan door genoemde eigenschappen veranderen (Uhlenbeck 1993: 29-30). Die veranderingen moeten altijd in woordenboeken beschreven worden. Want, mijn lexicografische credo blijft: het koloriet van de Nederlandse woordenschat is zo prikkelend, luisterrijk en adembenemend en toont zoveel levenskracht dat het loont blijvend te investeren in de beschrijving van onze eeuwenoude en zich steeds vernieuwende woordenschat (Van Sterkenburg 1993:9).

Aan het einde van mijn rede gekomen wil ik mijn dank uitspreken aan het College van Bestuur dat mij tot hoogleraar heeft willen benoemen. Het in mij uitgesproken vertrouwen hoop ik door mijn wijze van taakvervulling niet te beschamen.

Leden van het Bestuur van de Stichting Instituut voor Nederlandse Lexicologie,

Sedert 1977 mag ik als directeur en sedert 1978 ook als secretaris-niet-lid van het bestuur met u samenwerken aan het realiseren van twee zeer uitdagende opdrachten. Allereerst is er het voltooiën van het *Woordenboek der Nederlandsche Taal*, een lexicografisch symbool zo imposant als een Egyptische piramide. Maar er is meer. Ik voel mij bevoorrecht dat wij samen, en met succes, voor de continuïteit van het Instituut op de bres hebben gestaan. Wij hangen weliswaar aan een Nederlands-Vlaams subsidieinfuus, maar onze medewerkers hebben een lexicologische toekomst ook in de 21ste eeuw. Vooral door de ruimte die u mij geeft en gegeven hebt om de doelstellingen van onze Stichting te realiseren, zal ik uw gezelschap woordenschatplichtig blijven.

Dames en heren leden van het Algemeen Bestuur van de Stichting Leids Universiteits-Fonds,

Aan het begin van de jaren tachtig hebt u zich ingezet om bij de Kroon de bevoegdheidsverklaring te verwerven voor de stichting van de leerstoel lexicologie. In 1985 hebt u mij benoemd tot bijzonder hoogleraar om onderwijs in dit vak te verzorgen binnen de Faculteit der Letteren. Dat heeft geresulteerd in negen proefschriften, maar niet alleen daarin. De ontwikkeling van het vak in Leiden was bovendien van dien aard dat u met voldoening kon vaststellen hoe iets bijzonders tot gewoon kan worden. Uw vertrouwen was zeer stimulerend.

Mevrouw, mijne heren leden van het Curatorium van de leerstoel,

Op het moment dat ik mijn benoemingsbrief tot gewoon hoogleraar kreeg, viel er bij u een soort dankbrief voor bewezen diensten op de deurmat. Hoewel het Curatorium nu dus niet meer bestaat, toch een speciaal woord

van dank aan jullie adres. Marjolijn, Cor en Oscar, jullie waren meer dan betrokken bij mijn werk. De bijeenkomsten rondom het jaarverslag waren een feest, de maaltijden exquis en de humor relativerend. Maar vooral was er het oog voor de lexicologische proporties.

Waarde Weijnen,

De *Atlas Linguarum Europae*, beste Toon, verhinderde je in 1986 op deze plaats aanwezig te zijn. Ik heb toen je rol als stimulerend leermeester getekend. Ik zal dat vandaag niet herhalen, maar spreken over wat daarna kwam. Twee kloeke delen van jouw hand hebben wij voorbereid voor de reeks *Aan het Woord*. Je *Vergelijkende klankleer van de Nederlandse dialecten* en *De ziektenamen in de Nederlandse Dialecten* hebben onze vriendschap, en ik gedenk hier ook met piëteit de hartelijkheid van je vrouw Maria, alleen maar geïntensiveerd. Vol overtuiging kies ik hier dan ook voor dezelfde woorden die jij voor jouw leermeester Van Ginneken hebt gekozen. Toon, ik zou geen andere leermeester gehad willen hebben (Weijnen 1995).

Dames en heren van de Faculteit der Letteren,

Het *Studieprogramma Lexicologie* heeft ervoor gezorgd, dat ik mij de afgelopen elf jaar te midden van u in mijn element voelde.

Uw liberale en afstandelijke betrokkenheid heb ik als weldadig ervaren.

Dames en heren leden van de Vakgroep Nederlands,

Nooit heb ik mij in uw gezelschap als de *rara avis* gevoeld die een bijzonder hoogleraar krachtens zijn bescheiden opdracht misschien zou kunnen zijn.

U hebt mij alle ruimte gelaten om te ontsnappen aan vergaderingen en laat daarnaast niet na mij te betrekken bij zaken die goed zijn voor de Leidse Neerlandistiek. U moogt ook na de administratieve ombouw van mijn functie op mijn zelfde loyaliteit jegens de vakgroep blijven rekenen.

Dames en heren medewerkers van het Instituut voor Nederlandse Lexicologie,

Jarenlang is de historische lexicografie binnen ons Instituut gekoesterd. Terecht, want eens te meer bleek door het besluit van het Comité van Ministers van de Nederlandse Taalunie dat het verleden de sleutel was naar de toekomst. Wij mogen vanaf 1999 de fundamenten gaan leggen voor de lexicologie van de 21ste eeuw. Met deze rede heb ik iets willen laten zien van de lexicologische vergezichten die ik aan de horizon zie. Dat zijn beperkte inzichten. Samen met jullie hoop ik die beperkingen te overstijgen, zoals ik ook hoop met jullie de woordenschat die permanent in beweging is, te blijven dienen. Laat ons Instituut ook na 1999 wetenschappelijke producten maken van een indrukwekkende statuur.

Dames en heren studenten,

De Nederlandse lexicologie heeft internationaal veel crediet. Jullie zijn het wisselgeld dat in de toekomst de hardheid van dat crediet zal bepalen. Ik voel mij bevoorrecht in jullie te mogen investeren en ik verzeker een ieder dat aan het einde van de academische rit zal blijken dat de koersen van dit drogerende vak altijd minstens één fractie boven de Leidse kwaliteitsindex liggen.

Ik heb gezegd.

* Met dank aan Truus Kruyt, Ed Landman en Fons Moerdijk die een eerdere versie van deze tekst van kritisch commentaar voorzagen.

Geraadpleegde literatuur

Al 1993.

B.P.F. Al, 'Toekomst van de lexicografie, vanuit een commercieel perspectief'. In: A. van der Veen, *Op je woorden passen*. 75-84. INL. Leiden.

De Bies 1994.

Renata C.L. de Bies, *Op weg naar standaardisatie van het Surinaams-Nederlands. Een lexicale beschrijving van vijftien jaar administratietaal van de Republiek Suriname*. Proefschrift. U.I.A. Antwerpen.

Cavigneaux 1976.

A. Cavigneaux, *Die sumerisch-akkadischen Zeichenlisten. Ueberlieferungsprobleme*. Inaugural-Dissertation. Ludwig-Maximilians-Universität. München.

Cavigneaux 1989.

A. Cavigneaux, 'L'écriture et la réflexion linguistique en Mésopotamie. In: S. Auroux (ed), *Histoire des idées linguistique. Tome I: La naissance des métalangages en Orient et en Occident*. 99-118. Pierre Mardaga. Liège/Bruxelles.

Church 1991.

K.W. Church et al., 'Using Statistics in Lexical Analysis'. In: U. Zermik (ed.), *Lexical Acquisition: Using On-Line Resources to Build a Lexicon*. 115-163. Lawrence Erlbaum Associates. Hillsdale, New Jersey.

Claes 1985.

F. Claes, 'Van Dale Groot Woordenboek der Nederlandse Taal, elfde herziene en bijgewerkte druk'. In: *Nederlands van Nu* 33. 79-84.

Deygers 1996.

Katrien Deygers, *Het Belgisch Nederlands in de 'Grote van Dale'*.
Licentiaatsverhandeling. Gent.

Drewes 1976.

M.J. Koenen - J.B. Drewes, *Verklarend handwoordenboek der Nederlandse taal*. 27ste druk. Tjeenk Willink bv. Groningen.

DWB

J. und W. Grimm, *Deutsches Wörterbuch*. Leipzig 1854-1960.

Euralex 1994.

Willy Martin, Willem Meys, Margreet Moerland, Elsemiek ten Pas, Piet van Sterkenburg en Piek Vossen (Eds), *Euralex 1994. Proceedings*. Papers submitted to the 6th Euralex International Congress on Lexicography in Amsterdam, The Netherlands. Amsterdam.

Foolen 1993.

Ad Foolen, *De betekenis van partikels: een dokumentatie van de stand van het onderzoek met bijzondere aandacht voor 'maar'*. Proefschrift Nijmegen. Ponsen en Looijen. Wageningen.

Van Haeringen 1925.

C.B. van Haeringen, 'Eenheid en nuance in beschaafd-nederlandse uitspraak'. In: *NTg*. 18. 65 e.v.; ook in *Neerlandica*. 9-30. 's-Gravenhage 1949.

Heid 1994.

Ulrich Heid, 'On Ways Words Work Together. Topics in Lexical Combinatorics'. In: *Euralex 1994*. 226-258.

Klimaszewska 1990.

Zofia Klimaszewska, *Verbale Phraseologie des Niederländischen*. Warszawa.

Kruyskamp 1970.

K. Kruyskamp, *Van Dale Groot Woordenboek der Nederlandse Taal. Supplement op de achtste vermeerde druk*. Martinus Nijhoff. 's-Gravenhage.

Kruyt 1993.

J.G. Kruyt, *Design criteria for corpora construction in the framework of a European Corpora Network*. Technical Report. INL. Leiden.

Kruyt 1995.

J.G. Kruyt, 'Nationale tekstcorpora in internationaal perspectief'. In: *Forum der Letteren* 36. 47-58.

Kruyt 1996.

J.G. Kruyt en P. van Sterkenburg, 'Corpus Design Criteria'. In: Nicoletta Calzolari, Mona Baker, Johanna G. Kruyt, *Towards a Network of European Reference Corpora*. Linguistica Computazionale XI. Giardini. Pisa.

Landman 1997.

Ed Landman, *Van gasthuis tot ziekenhuis. Een lexicologische studie vanuit cognitief-semantisch perspectief naar een model voor een elektronisch historisch monolinguaal lexicon*. Proefschrift. Leiden.

Martin 1985.

W. Martin, 'Corpora voor woordenboeken'. In: A. Dees (red.), *Corpusgebaseerde woordanalyse. Jaarboek*. 91-99. VU Amsterdam.

Meys 1991.

Willem Meys, 'De empirische dimensie'. In: *Spektator* 20. 313-336.

Moerdijk 1989.

A. Moerdijk, 'Benaderingen van metonymie'. In: *Forum der Letteren* 28. 194-213.

Moerdijk 1990.

A. Moerdijk, 'Metonymie uit een ander vaatje'. In: *Traditie en Progressie. Handelingen van het 40ste Nederlands Filologencongres*. Gardata. Leersum. 111-122.

Moerdijk 1993.

'Metonymische polysemie en lexicografie'. In: A. van der Veen, *Op je woorden passen*. 25 Jaar Instituut voor Nederlandse Lexicologie 1967-1992. 23-35. INL. Leiden.

Moerdijk 1994.

A. Moerdijk, *Handleiding bij het Woordenboek der Nederlandsche Taal (WNT)*. Aan het woord. Sdu. 's-Gravenhage.

Niemeier & Dirven 1997.

Susanne Niemeier and René Dirven, *The Language of Emotions. Conceptualization, Expression and Theoretical Foundation*. John Benjamins. Amsterdam, Philadelphia.

OED

J.A.H. Murray et al., *The Oxford English Dictionary*. The Clarendon Press. Oxford 1882-1928.

Sanders 1997.

Ewoud Sanders, 'Vertaalwoordenboeken. Gaten in de taalmarkt. Grote behoefte maar een te kleine markt'. In: *Kwartaalschrift 1*. 24-25.

Sinclair 1987.

John Sinclair, *Collins Cobuild English Language Dictionary*. Collins. London, Glasgow.

Sinclair 1991.

J.M. Sinclair, *Corpus, Concordance, Collocation*. Oxford University Press. Oxford.

Van Sterkenburg 1986.

P.G.J. van Sterkenburg, *Werk maken van woorden*. Rede. Brill. Leiden.

Van Sterkenburg 1991(a).

P.G.J. van Sterkenburg, *Van Dale Groot woordenboek van hedendaags Nederlands*. Van Dale Lexicografie. Eerste druk 1984. Utrecht, Antwerpen.

Van Sterkenburg 1991 (b).

P.G.J. van Sterkenburg et al., *Van Dale Groot Woordenboek van Synoniemen en andere betekenisverwante woorden*. Van Dale Lexicografie. Utrecht, Antwerpen.

Van Sterkenburg 1995.

P.G.J. van Sterkenburg, *Van Dale Handwoordenboek hedendaags Nederlands*. Van Dale Lexicografie. Utrecht, Antwerpen.

Van Sterkenburg 1997a.

P.G.J. van Sterkenburg, *Neologismen*. Ter perse.

Van Sterkenburg 1997b.

P.G.J. van Sterkenburg, *Vloeken. Een cultuurbepaalde reactie op woede, irritatie en frustratie*. Sdu-Standaard. Den Haag, Antwerpen.

Stroop 1992.

J.P.A. Stroop, 'Welk Nederlands na 1992?' In: G.L.J. Leenders (eindred.), *Het Nederlands na 1992: teksten van de lezingen gehouden op het gelijknamige congres op 21 en 22 november 1991 georganiseerd door de Landelijke Vereniging van Neerlandistiek i.s.m. de Universiteit van Amsterdam*. 36-57.

Swanepoel 1994.

Piet Swanepoel, 'Problems, Theories and Methodologies in Current Lexicographic Semantic Research'. In: *Euralex 1994 Proceedings*. Amsterdam.

Taeldeman 1992.

Johan Taeldeman, 'Welk Nederlands voor Vlamingen?' In: *Nederlands van Nu* 40. 33-52.

Teubert 1995.

Wolfgang Teubert, 'Language Resources: The Foundation of a Pan-European Information Society'. In: Heike Rettig et al., *Telri. Trans-European Language Resources Infrastructure. Proceedings of the First European Seminar 'Language Resources for Language Technology'*. 105-129. Tihany, Hungary.

Thiering 1996.

Barbara Thiering, *Jesus the Man. A New Interpretation From the Dead Sea Scrolls*. Gorgi Books. London.

Van den Toorn 1997.

M.C. van den Toorn, 'Het Nederlands van 1920 tot nu'. Ter perse.

Uhlenbeck 1981.

E.M. Uhlenbeck, *Betekenis in linguïstisch perspectief. Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen. Afdeling Letterkunde, Nieuwe Reeks, Deel 44, No. 8*. B.V. Noord-Hollandsche Uitgevers Maatschappij. Amsterdam. Oxford, New York.

Uhlenbeck 1993.

E.M. Uhlenbeck, 'De dreiging van snelle taalsterfte: een sinds kort onderkend mondiaal probleem'. In: *Ons Erfdeel* 39. 23-32.

Verkuyl 1994.

Henk Verkuyl, *Knowledge Representation in Dictionaries. Keynote Lecture*.
Delivered at the 6th Euralex International Congress. 1-27. Amsterdam.

Verkuyl 1996.

Henk Verkuyl, 'Figuurlijk en stereotyperend taalgebruik'. In: *Trefwoord 11*.
109-117 en 125-128.

Verwijs en Verdam.

E. Verwijs en J. Verdam, *Middelnederlandsch Woordenboek*. 1885-1952.
's-Gravenhage.

VMNW

W.J.J. Pijnenburg et al., *Vroegmiddelnederlands Woordenboek*.
Conceptaflevering 1 (1989) - 8 (1996).

Van der Voort van der Kleij 1975.

J.J. van der Voort van der Kleij, 'Lexicologische verzamelingen en lexicografie'.
In: P.G.J. van Sterkenburg (red.), *De Nederlandse lexicologie tussen handwerk
en machine*. 58-68. H.D. Tjeenk Willink. Groningen.

Weijnen 1996.

A.A. Weijnen, 'Herinneringen aan Jacques'. In: Ad Foolen, Jan Noordegraaf
(Eds.), *De taal is kennis van de ziel. Opstellen over Jac. van Ginneken
(1877-1945)*. 35-51. Nodus Publikationen. Münster.

WNT

M. de Vries, L.A. re Winkel e.a., *Woordenboek der Nederlandsche Taal*.
Martinus Nijhoff etc. 's-Gravenhage etc. 1881-.