

Nu alfabetisch!

RIK SCHUTZ IS WERKZAAM BIJ DE NEDERLANDSE TAALUNIE; VOORHEEN WAS HIJ UITGEVER BIJ VAN DALE LEXICOGRAFIE.

In 1864 was mijn overgrootvader een klein kind en de proto-Van Dale, het *Nieuw woordenboek der Nederlandsche taal* van Calisch en Calisch, was net nieuw. Pas vier generaties later, toen mijn zoon een peuter was, begon er iets – ten goede – te veranderen in de manier waarop Van Dale wordt bewerkt.

Generaties lang was de kaartenbak het voornaamste stuk gereedschap van de lexicograaf. Iemand, meestal een wat oudere meneer die moderne ontwikkelingen in de samenleving met een zekere bezorgdheid gadesloeg, noteerde woorden, woordbetekenissen en uitdrukkingen die hem opvielen omdat ze nieuw voor hem waren, zette ze op een kaartje en dat kaartje zette hij op de alfabetische plaats van het 'trefwoord' in een kaartenbak. Gemiddeld eens in de tien jaar kreeg hij van de uitgever het zetsel van de vorige druk van het woordenboek op grote vellen papier thuisgestuurd, waarop hij met de informatie op de kaartjes in zijn kaartenbak in de brede marge kon opschrijven wat er moest worden toegevoegd. Soms streepte hij ook wat tekst door, maar er kwam vooral veel bij. Die vellen werden in een drukkerij verwerkt in het loden zetsel van het boek en na een correctieronde werd een herziene, dikkere editie gedrukt. De uitgever bepaalde de oplage en de prijs, en soms besloot die uitgever tot een extra boekdeel. Door de decennia heen werd de dikke Van Dale steeds dikker, totdat in 1970 de 9^e editie in twee delen verscheen. In 1984 werden het er voor de 11^e editie zelfs drie.

In de dagen dat ik me uitgever van het *Groot woordenboek van de Nederlandse taal* mocht noemen, hebben we ooit een kostbaar marktonderzoek laten uitvoeren om het antwoord te vinden op de vraag hoe het publiek zou reageren op een nieuwe editie in vier boekdelen. Overigens vond onze directeur het toen niet bij de standing van de uitgeverij passen om van de Dikke van Dale te spreken. Dat was een volkse bijnaam en voor ons was het de Grote van Dale, ofwel de 'gvd', maar die afkorting was minder geschikt voor gebruik buiten de deur. Dat marktonderzoek was een leerzame kennismaking met onze doelgroep. Een van de consumentenpanels was geselecteerd op 'bekendheid met het merk' en 'bezit van het woordenboek'. Een klasje vol praatgrage woordenboekeigenaren vertelde gretig over hun trotse bezit: de een zou het ding aangeschaft hebben als deel van een encyclopedie (wat onmogelijk was), de ander dacht al een uitgave in vier delen te bezitten. Heel belangrijk vond iedereen ook de kleur van de boekband, maar niemand wist hoe de laatste editie eruit zag. Kortom, het deed er allemaal weinig toe. Wie de Van Dale dacht nodig te

hebben, zou de nieuwe editie toch wel kopen en wie hem als statusobject begeerde ook. Mooie tijden.

Maar al deed het er allemaal weinig toe, de kaartenbak van de woordenboekenmakers is niet meer. De overstap naar een moderne manier van werken werd ingezet door een nieuwe, in lexicografie gespecialiseerde uitgeverij. Zo'n vijfendertig jaar geleden, in de late jaren zeventig, werd door het voorname uitgevershuis Kluwer de uitgeverij Van Dale Lexicografie opgericht. Ruim een eeuw was Van Dale een boek geweest, en ineens werd het een merk en de naam van een bedrijf. De uitgeverij maakte een innovatieve reeks vertaalwoordenboeken met teams ervaren vertalers onder leiding van een groep jonge, ambitieuze taalkundigen. Net als de Dikke Van Dale, waren tot die tijd ook 'nieuwe' vertaalwoordenboeken gewoonlijk door gepensioneerde taaldocenten bewerkte versies van al bestaande woordenboeken. Maar de nieuwe Van Dales werden *from scratch* opgebouwd als gestructureerde gegevensbestanden en zouden dus 'in de computer' kunnen worden opgeslagen en bewerkt. Ambitie, visie en geld waren voorhanden, maar computers helaas nog niet. Toen het eerste deel in de nieuwe reeks zou verschijnen, was er in het hele uitgeversconcern nog maar één computer die de kopij van een heel boekdeel kon verwerken, het *IJsselbrein* in Deventer. Redacteuren hadden alle kopij met een pen op formulieren geschreven en datatypisten hadden die ingetikt op een enorme stapel *floppy disks*. Tegenwoordig past de inhoud van al die woordenboeken met gemak op een usb-stick en zijn er kant-en-klare programma's te koop voor het doelmatig beheren van complexe bestanden. Maar in de beginjaren moesten we alle 'databewerking' uitbesteden. Geleidelijk kwam er bij de uitgeverij een en ander in huis aan hard- en software en kwam er personeel dat daarmee kon omgaan. Toen werd het ook tijd voor het digitaliseren van ons kostbare erfstuk: de Grote van Dale. De nieuwe woordenboeken waren meteen al opgebouwd als gestructureerde gegevensbestanden: voordat ze werden gedrukt, werd voor elk bouwsteentje (trefwoord, woordsoort, definitie, synoniem, etc.) bepaald hoe het er in het boek zou uitzien: vet of cursief gedrukt, tussen rechte of spitse haakjes, met een komma ervoor of op een nieuwe regel. Maar om de tekst van de Grote van Dale in de computer te krijgen, moest de omgekeerde weg worden bewandeld. Van dat woordenboek wisten we hoe het eruit zag, maar niet wat de status was van de afgedrukte stukjes tekst. Dat moest worden afgeleid uit de – helaas niet consequent toegepaste – typografie. Allereerst werd het hele boek op floppy's overgetikt met codes voor vet en cursief gedrukte tekst. Vetgedrukte tekst was meestal een trefwoord. Om een voorbeeld te noemen: meer dan twintig keer kwam in het bestand de tekenreeks **...jam** of **-jam** voor. Uit het daaraan steeds voorafgaande voluit afgedrukte vetgedrukte woord (**aalbessendief, abrikozenboom, gemberbier, ...**), moest het gedeelte worden geïdentificeerd dat de plaats van de puntjes of het streepje kon innemen. Pas daarna wisten de computer en wij welke trefwoorden er in het woorden-

boek stonden (**aalbessenjam, abrikozenjam, gemberjam**, etc.) en konden we de redactionele bewerking systematisch aanpakken.

Pas toen, 120 jaar na verschijning van het woordenboek dat Johan Hendrik van Dale bewerkte tot tweede editie, werd een systematische manier van redigeren mogelijk. Het ligt immers voor de hand om soortgelijke woorden (bijvoorbeeld namen van schaakstukken of maanden van het jaar) op een soortgelijke manier te beschrijven. Door ze een code te geven, konden we ze bij elkaar zetten en in samenhang redigeren. Geen van de vroegere bewerkers had ooit de middelen en de mogelijkheden gehad om zulke clusters te bewerken en dat was altijd aan het woordenboek te zien geweest. Met de verkoop op cd-rom in het verschiep, zou het gebrek aan systematiek ook voor het publiek zichtbaar kunnen worden en dat maakte het des te noodzakelijker om te werken aan verbeteringen.

Het zichtbaar maken van de aanvankelijk ontbrekende systematiek mocht dan eenvoudig geworden zijn dankzij de computer, het uniformeren van definities en aanvullen van ontbrekende informatie bleef tijdrovend – en dus kostbaar – redactioneel handwerk. En dat leverde voor het commerciële bedrijf dat Van Dale nu eenmaal was en is, een dilemma op. Investeren in kwaliteitsverbetering loont als je in je marketing de resultaten ervan kunt gebruiken. Maar opscheppen over de verbeterde kwaliteit van een product dat al het aura van perfectie heeft, ligt gevoelig. Eén voorbeeldje: de 12^e editie uit 1992 was de eerste waarvoor gold 'nu in alfabetische volgorde'; meer dan 1.500 trefwoorden hadden voor die tijd niet precies op hun alfabetische plaats gestaan (overigens zonder dat iemand daar last van had). Hoe zou het overgekomen zijn als we die vernieuwing hadden gebruikt in de aanbiedingsprospectus?

Wat in dezelfde periode ook veranderde, was het aanbod van nieuwe informatie. Menig woordenboekartikel in oudere edities van Van Dale is gebaseerd op een enkele 'bewijsplaats', een krantenknipsel of een alinea uit een boek. De eruditie, de introspectie en de fantasie van de lexicograaf maakten er wel iets van; niet zelden een betekenis waarvan je je kunt afvragen of die buiten het citaat in kwestie in het Nederlands voorkomt. Een voorbeeldje? Bij de tweede betekenis van *potterig* ('geneigd tot het homoseksuele of het lesbische') stond lange tijd: '*omdat de directrice zo potterig was als een Aagje Deken* (van Swieten)'. Dankzij de systematische bewerking van alle literaire citaten door de huidige hoofdredacteur zijn er nu geen citaten meer van Ida van Swieten, auteur van *Penose en Méér over het rosse leven*, in Van Dale te vinden (maar vreemd genoeg is die obscure betekenis van *potterig* er nog wel).

Nog een voorbeeld, nu een dat illustreert welke onverwachte rol een dagblad kan spelen als bron van nieuwe woorden en citaten daarbij. Bij het werkwoord *kaperen* stond in de 13^e editie: 'al die willen te kaperen varen, moeten mannen met baarden zijn (NRC)'. Die bronvermelding is nu vervangen door 'bekend kinderliedje' (maar intussen is 'onbekend kinderliedje' misschien accurater).

Er is de laatste jaren tijdens systematische bewerkingen veel verbeterd en opgeruimd, bijvoorbeeld de voorbeeldzinnetjes waarin een lexicograaf de naam van zijn vrouw noemt ('Marijke zet een schitterende Badeloch neer'), de initiaal van zijn vriendin ('H. is gek op Y.'), of zijn adres ('baas, kun je ook zeggen waar de Rijnsburgerweg 22 is?').

Vanuit ons huidige digitale perspectief heeft de woordenboekenmaker zich generaties lang moeten behelpen met ontoereikende hulpmiddelen en gebrekkig zicht op hoe het woordgebruik in het Nederlands zich ontwikkelde. De laatste decennia heeft Van Dale kunnen profiteren van technische ontwikkelingen die redacteurs (m/v) in staat stellen van een woord dat hij of zij nooit eerder heeft gezien een reeks teksten te bekijken waarin het voorkomt. Er bestaat software die uit een grote verzameling digitale teksten niet alleen de nieuwe woorden kan zeven, maar ook kan tonen welke combinaties vaker voorkomen dan vroeger. Dat kan op nieuwe betekenissen duiden; tegenwoordig komt het woord *bereik* bijvoorbeeld vaak voor in combinatie met 'hebben' en dat rechtvaardigt een nieuwe definitie van dat woord.

Maar diezelfde overvloed aan digitale teksten staat ook de hedendaagse gebruiker van het woordenboek ter beschikking en dat heeft gevolgen voor de rol van het woordenboek. Vóór internet hadden we toegang tot de woorden die in onze directe omgeving werden gesproken en die voorkwamen in onze eigen lectuur. De verdienste van het woordenboek was dat het voor véél woorden véél informatie – die op een andere manier niet beschikbaar was – betrekkelijk gemakkelijk toegankelijk maakte. Waar nu behoefte aan is, is niet aan 'veel'. Integendeel. We willen dat een woordenboek de beschikbare overvloed kanaliseert en verwerkt tot overzichtelijke informatie. Redacteurs zullen nog betere technieken ter beschikking krijgen om ontwikkelingen in de woordenschat te volgen en zichtbaar te maken. Toekomstige woordenboeken zullen minder momentopnamen zijn en meer zicht geven op veranderingen in woordgebruik en -betekenissen in de tijd. Ook wie een woordenboek raadpleegt, profiteert natuurlijk van nieuwe techniek. Niet één boek, of drie of vier delen zullen de vorm zijn waarin de volgende generatie een woordenboek raadpleegt. We hebben de cd-rom al zien komen en gaan en tegenwoordig kunnen we ons abonneren op een internet-Van Dale en apps downloaden. Maar die vormen zijn allemaal maar manieren om de inhoud te kunnen raadplegen. Die inhoud, daar gaat het om.