



# Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

## Toekomstig woordenboek

van der Sijs, N.

### **published in**

NRC Handelsblad  
2013

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

### **citation for published version (APA)**

van der Sijs, N. (2013). Toekomstig woordenboek. *NRC Handelsblad, Wetenschapsbijlage, 2-2*.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[pure@knaw.nl](mailto:pure@knaw.nl)



# Toekomstig woordenboek

Op 5 november vorig jaar maakte de Engelse woordenboekuitgeverij Macmillan het einde van de papieren uitgaven bekend onder de kop ‘*Stop the presses – the end of the printed dictionary*’. Bij dit bericht veerde ik op, want lexicografie is een oude liefde sinds ik als 20-jarige student redacteur werd van een Nederlands-Russisch woordenboek.

‘De toekomst van het woordenboek is digitaal’, aldus de uitgeverij. Een open deur, zou je zeggen, maar dan heeft u buiten de woordenboekschrijvers gerekend: het nieuws verspreidde zich razendsnel over lexicografische internetgroepen, waar heftige discussies werden gevoerd, en sommige professionele lexicografen zwolgen in herinneringen aan juweeltjes waarop zij vroeger al bladerend in papieren woordenboeken waren gestuit.

Eind november kwamen Nederlandse lexicografen bijeen om te praten over de toekomst van hun vak. Vertegenwoordigers van Van Dale Lexicografie en Prisma zagen die vooral zonnig in: voorlopig bleven ze, zo beweerden ze, nog veel boeken verkopen, en afgeleide producten daarvan leidden een eigen online leven. Van vernieuwende ideeën gaven zij geen blijk.

Misschien moeten we die eerder verwachten uit de wetenschappelijke hoek. Op het Leidse Instituut

voor Nederlandse Lexicologie wordt sinds 2000 gewerkt aan het Algemeen Nederlands Woordenboek, een online woordenboek dat in 2019 voltooid moet zijn. Redacteuren beschrijven de ‘kernwoordschat’ van het Nederlands uit de periode 1970 tot 2019. Bij de start was het project vernieuwend, maar inmiddels vind je op internet meer en betere informatie dan wat de redacteurs moeizaam handmatig bijeengezocht hebben. Bovendien bevat het woordenboek veel te veel eendagsvliegen (ooit gehoord van *curlingouder of shampoogeneratie?*) en samenstellingen als *januariavond*, *-cijfers*, *-dag*, *-maand*, *-middag*, *-nacht*, *-nummer*, *-ochtend*, *-temperatuur*, *-week*, *-wind*, *-zon* – voor alle twaalf maanden. Hoezo kernwoordschat?

Het is de vraag welke gebruikers dergelijke informatie opzoeken. Want voor de gebruikers, zowel de algemene als de wetenschappelijke, doen de lexicografen het toch. De meeste gebruikers reageerden begripvol op de beslissing van Macmillan om te stoppen met het uitgeven van gedrukte woordenboeken. En dat ligt voor de hand. Op internet bestaan voor Engelstaligen prachtige informatiebronnen die de papieren woordenboeken in kwaliteit en kwantiteit verre overstijgen. Daar vindt de innovatie plaats.

Ik noem slechts enkele voorbeel-

den. Om te beginnen is er het interactieve Engelse onlinewoordenboek Wordnik ([www.wordnik.com](http://www.wordnik.com)). Hier wordt het lexicografische handwerk op slimme wijze gecombineerd met computertechnieken die automatisch gegevens op internet verzamelen. Als men op Wordnik een woord opzoekt, krijgt men definities uit een groot aantal gangbare woordenboeken voorgeschoteld, geïllustreerd met voorbeeldzinnen uit kranten, boeken, films en tweets, afbeeldingen van Flickr en audiobestanden, zodat men uitspraakverschillen in verschillende variëteiten van het Engels kan beluisteren. Ook krijgt men lijsten met gerelateerde woorden te zien, bijvoorbeeld met dezelfde of tegengestelde betekenis.

Even innovatief is The Visual Thesaurus ([www.visualthesaurus.com](http://www.visualthesaurus.com)), die zich met name van Wordnik onderscheidt door de manier waarop de gegevens worden getoond: op een zogenaamde ‘woordkaart’ worden de verschillende betekenissen van een woord verbonden aan gerelateerde woorden. Een Duits vergelijkbaar concept tot slot is het prachtige Digitale *Wörterbuch der deutschen Sprache* (DWDS: <http://retro.dwds.de>).

Tijdens de bijeenkomst van de Nederlandse woordenboekmakers werd de wens uitgesproken dat er meer samenwerking zou komen tussen de commerciële en de academi-

## In een Digitaal Woordenboek van het Nederlands moet alles staan uit alle moderne woordenboeken

sche wereld. Zo’n wens blijft meestal steken in goede voornemens. Maar stel nu eens dat die samenwerking van de grond komt, hoe zou het woordenboek van de toekomst er dan uit kunnen zien?

Het gedroomde Digitale Woordenboek van het Nederlands bevat de inhoud van alle moderne Nederlandse woordenboeken – die van hoge kwaliteit en heel compleet zijn – en van belangrijke digitale woordbestanden. Behalve woorden zijn ook persoons- en plaatsnamen opgenomen: het onderscheid tussen woorden en namen is ooit vanwege ruimtegebrek gemaakt. Er zijn links naar beeld, geluid, biografische bronnen en encyclopedische gegevens uit bijvoorbeeld Wikipedia.

De trefwoorden zijn automatisch gekoppeld aan tekstbestanden in kranten, boeken, tweets. Plaatjes laten zien met welke woorden een trefwoord wordt gecombineerd, zodat men direct antwoord krijgt op vragen als: *zeggen de meeste mensen de commerciële en de academi-*

computer zien welke woordvormen en woordcombinaties in de bronnen van de laatste 10-50 jaar een stijgende dan wel dalende frequentie vertonen: op die manier spoort de computer automatisch veranderingen in betekenis en woordgebruik op. Alle trefwoorden bevatten gedetailleerde informatie over woord- en betekeniskenmerken. Zo kunnen gebruikers antwoorden vinden op vragen als: *Wat is de Latijnse naam voor blindedarmontsteking? Met welke voorzetsels worden werkwoorden van beweging gecombineerd?*

Veel van deze informatie is wel ergens beschikbaar, maar computerspecialisten zouden, samen met lexicografen, de gegevens op inhoudelijk zinnige wijze met elkaar moeten verbinden. Daarmee wordt lexicografie alleen maar spannender. En woordenboekuitgevers hoeven niet te vrezen voor een faillissement: zij kunnen het immense Digitale Woordenboek gebruiken voor de ontwikkeling van gespecialiseerde digitale producten als woordenlijsten voor spellingcontrole, spraakherkenning en vertaalapplicaties in telefoons, computers en tablets.

Nicoline van der Sijs is per 1 januari 2013 benoemd als (deeltijd)hoogleeraar historische taalkunde van het Nederlands in de digitale wereld aan de Radboud Universiteit, Nijmegen.