

MIDDELNEDERLANDSCH WOORDENBOEK

Algemeen

Het *Middelnederlandsch Woordenboek* van Verwijs en Verdam (hierna: *MNW*) beschrijft de middeleeuwse Nederlandse woordenschat uit de periode ca. 1250 tot ca. 1550. Het woordenboek bevat 60.000 trefwoorden (exclusief 14.000 verwijslemmata) en telt negen delen, waarvan het eerste in 1885 verscheen, het laatste in 1929. Ondanks zijn respectabele leeftijd, is het woordenboek nog steeds onmisbaar als werkinstrument voor iedereen die zich bezighoudt met de taal- en letterkunde van onze middeleeuwen.

In 1941 verscheen nog een deel met *Aanvullingen en verbeteringen op het gebied van dijk- en waterschapsrecht, bodem en water, aardrijkskunde enz.*, samengesteld door A. Beekman. Vanwege zijn inhoudelijk en formeel zeer afwijkende karakter zal dit deel pas in een volgende release in de applicatie opgenomen worden.

Het *MNW* is een alfabetisch geordend woordenboek. Een probleem voor de samenstellers daarbij was de grote spellingvariatie binnen het Middelnederlands. Om dit probleem op te lossen zijn alle trefwoorden gespeld in een gereconstrueerd, “genormaliseerd” Middelnederlands. De gebruikers moeten er o.a. rekening mee houden dat woorden die met *c-* beginnen onder de *K* opgenomen zijn wanneer de *c-* als *k-* wordt uitgesproken (b.v. *confessie, costume*); woorden met een *z-* als beginletter zijn te vinden onder de *S* (*sac, so, seggen, sonde*). O.m. om deze en andere spellingproblemen van het Middelnederlands zoveel mogelijk uit te schakelen is in de internetapplicatie aan elk lemma ook een trefwoord in de moderne spelling toegevoegd.

Net zoals in het *ONW*, *VMNW* en *WNT* worden bij elke betekenis citaten gegeven ter illustratie en controle van de redactionele tekst. Deze aanhalingen zijn in het grafische *MNW* doorgaans niet voorzien van lokalisering- en dateringsinformatie. Deze informatie is, via de bronaanduiding bij een citaat, met enig zoekwerk wel terug te vinden in de *Bouwstoffen* of bronnenlijst (*MNW*, dl. 10). Ook dit aspect is in de applicatie gebruikersvriendelijker gemaakt.

Van *MNW* naar *MNW online*

Het digitale leven van het *Middelnederlandsch Woordenboek* begon in 1998. Samen met het *Corpus Gysseling* (het bronmateriaal voor het *VMNW*) en een collectie van nog eens bijna 300 Middelnederlandse rijm- en prozateksten, verscheen het dat jaar op de *Cd-rom Middelnederlands*. Het digitale bestand van het woordenboek op deze door het INL uitgebrachte cd-rom vormt de basis voor het *MNW* binnen de internetapplicatie. Er komt echter nog altijd behoorlijk wat bij kijken vooraleer een dergelijk woordenboekbestand optimaal binnen een systeem als de woordenboekapplicatie kan functioneren: 1) elk woordenboek moet op dezelfde manier bevraagd kunnen worden, 2) men moet identieke onderdelen en informatiecategorieën (grammaticale en etymologische gegevens, de woordbetekenissen, het citatenmateriaal) op identieke wijze kunnen onderzoeken, en 3) een bepaalde zoekvraag moet desgewenst ook tegelijkertijd aan verschillende woordenboeken binnen de applicatie gesteld kunnen worden. Een extra probleem hierbij is dat de applicatie zowel woordenboeken bevat die volledig digitaal tot stand gekomen zijn (het *ONW* en het *VMNW*), als gedigitaliseerde grafische producten (het *WNT* en het *MNW*) die ook nog eens geschoeid blijken op een 19^{de}-eeuwse lexicografische leest. Werken aan het *MNW online* biedt echter ook, net zoals het werk aan het *WNT online*, extra kansen. Het is de voor de hand liggende gelegenheid tot verrijking en correctie van het bestand, waarbij ook op een uiterst vruchtbare manier gebruik gemaakt kan worden zowel van computerlinguïstische toepassingen (het “draaien” van diverse scripts op het bestand) als van de combinatie van semi-automatische en handmatige bestandsbewerkingen. Dit heeft, in een ruimere zin, mede geleid tot het niet onbelangrijk inzicht, dat een van de INL-troeven bij het werken met grote digitale bestanden juist ook de efficiënte combinatie van computer en handwerk kan zijn. In wat volgt worden enkele voorbeelden gegeven van bewerkingen van dit combinatietype die er de afgelopen jaren op het *MNW*-bestand uitgevoerd werden ten behoeve van opname ervan in de GTB-applicatie. Of: hoe er gewerkt werd aan de digitale

toekomst van het verleden, onontbeerlijk om een 19^{de}-eeuws lexicografisch monument als het *MNW* overeind en bruikbaar te houden.

Modern trefwoord

De 60.000 lemmata van het *MNW* zijn alle, net zoals gebeurd is bij het *ONW*, *VMNW* en *WNT*, voorzien van een trefwoord in de hedendaagse spelling. Dit moderne trefwoord vergemakkelijkt het zoeken in het *MNW* en functioneert tegelijkertijd – wanneer het woord in kwestie ook in de andere woordenboeken van de applicatie voorkomt - als een *passe-partout* waarmee de andere woordenboeken apart of gecombineerd geraadpleegd kunnen worden. De toekenning van dit moderne trefwoord gebeurde automatisch, o.m. door er de corresponderende moderne trefwoorden van het *WNT* en het *VMNW* bij te betrekken of door gebruik te maken van een aantal substitutieregels om Middelnederlandse spellingen om te zetten in hedendaagse. De resultaten van deze automatische omzettingen vereisten echter wel een grondige (semi-handmatige) controle en correctie.

Streven naar een consistente microstructuur

Een *MNW*-artikel bestaat standaard uit het zogenaamde *kopje* en de *betekenisafdeling*. Het kopje opent bij het trefwoord en eindigt waar de eerste betekenisomschrijving begint. Het biedt naast het trefwoord en de eventuele vormvarianten, grammaticale informatie (m.n. de woordsoort) en etymologische gegevens; de betekenisafdeling bevat de definities en de illustrerende citaten daarbij. Al deze onderdelen en informatiecategorieën van het *MNW* moeten zowel afzonderlijk als in combinatie met de andere historische woordenboeken binnen de applicatie geraadpleegd en opgevraagd kunnen worden. Ze moeten dan ook voor een computer in dezelfde mate herkenbaar zijn als nu al het geval is bij bv. *VMNW* en *WNT*.

Een consequent aangehouden microstructuur is uiteraard de eerste voorwaarde bij het omzetten van het *MNW* naar een *MNW online*, waarin de codering van de diverse elementen o.m. gebaseerd is op consistente typografische eigenschappen van het gedrukte boek. Voor een menselijke gebruiker oogt de microstructuur van de lemma's in het grafisch *MNW* behoorlijk consistent. Zo staat het kopje b.v. doorgaans binnen een aparte alinea vóór de betekenisafdeling, zijn binnen de betekenisafdeling de definities cursief gedrukt en wordt het citatenblok in de regel van de definitietekst gescheiden door een zg. *hek* (||).

Bij automatiseringstoepassingen blijken de *MNW*-structurering en de typografische weergave daarvan echter een stuk minder consistent te zijn. Zeker wanneer men dieper gaat dan het niveau van de gewone genummerde betekenissen is het verre van eenvoudig om de structurering voor de computer herkenbaar te maken. Ook het typografisch onderscheid tussen redactionele tekst en de Middelnederlandse citaten is op dat niveau vaak niet helder en dus 'computeronvriendelijk'. Maar ook wanneer men zich toespitst op het eerste betekenisniveau is de situatie dikwijls complexer dan gedacht. Bij het automatisch aanbrengen in de woordenboekartikelen van de scheidslijn tussen het kop- en het betekenisgedeelte en bij het automatisch afbakenen van de definities, bleken er in elk geval 15 verschillende types onderscheiden te moeten worden. Bij tien daarvan kon de begrenzing van kop en betekenisafdeling en de herkenning van definities automatisch gebeuren, maar semi-handmatige controle en correctie bleven daarbij noodzakelijk; dit betrof ca. 45.000 lemmata. Bij de overige vijf types, ca. 15.000 lemma's en zo goed als steeds artikelen waarin kop en betekenisafdeling samen binnen één alinea staan, moesten de afbakening van die twee onderdelen en de definitieherkenning ook weer semi-handmatig uitgevoerd worden, o.m. ook omdat er zich bij die artikelen vaak nog na de definitietekst of na het citatenblok extra etymologische gegevens bevonden, die een gebruiker ook als kopinformatie aangeboden moesten kunnen worden.

Verrijking van de citaten

Het *MNW* telt bijna 400.000 citaten. Het is bij het *MNW online* de bedoeling om deze aanhalingen ook zo optimaal mogelijk te presenteren als een corpus met historisch taal materiaal. Daarom wordt bepaalde informatie uit de

Bouwstoffen (vnl. daterings- en lokaliseringsgegevens, maar ook – indien bekend - auteursnamen) aan de citaten toegevoegd. Deze gegevens staan echter ongemarkeerd in de tekst van de bronbeschrijvingen, zodat ze niet automatisch te herkennen of met de *MNW*-citaten te verbinden zijn. Om dit probleem op te lossen werd in de *Bouwstoffen* bij elke bronbeschrijving de betreffende informatie geformaliseerd door ze binnen een aantal specifieke velden te herhalen. Alle citaten die uit die bron afkomstig zijn kunnen vervolgens automatisch met de gegevens uit die velden verrijkt worden. De *Bouwstoffen* tellen 1417 artikelen. Binnen deze groep zijn er echter een 400 nummers die zelf ook weer uit verschillende onderdelen bestaan die elk afzonderlijk als bron kunnen fungeren. Deze 400 zg. complexe of meerdelige *Bouwstoffen*-nummers (vooral tijdschriften en verzamelwerken) tellen samen een 10.000 onderdelen, elk weer met een eigen bronbeschrijving. Theoretisch lag hier dus een enorme klus: het invullen van de bovengenoemde velden bij ruim 11.000 items. Onderzoek via de *Cd-rom Middelnederlands* wees echter uit dat niet elk beschreven item in de *Bouwstoffen* in de praktijk als bron was gebruikt: ‘slechts’ bij 4.100 *Bouwstoffen*-nummers of onderdelen daarvan moest de veldenset ingevuld worden, waardoor dit werk een stuk haalbaarder bleek dan aanvankelijk gedacht. Voor de eerste release van het *MNW online* is de citaatverrijking al geïmplementeerd bij de aanhalingen uit de zg. enkelvoudige bronnen (bronnen die niet verder op te delen zijn); in een volgende release zal ook voor de citaten uit de zg. complexe *Bouwstoffen*-nummers deze informatie voor de gebruiker volledig beschikbaar zijn.

Correctie van digitaliserings- en parsingsfouten

Ten behoeve van het *MNW online* is ook een aantal digitaliserings- en parsingsfouten verbeterd in het bestand dat de basis was voor de cd-romversie van het *MNW*. Zo werden o.m. een 640 gevallen van ontbrekende tekst aangevuld; het ging daarbij om lacunes variërend van enkele letters tot een bijna complete kolom in *MNW*-deel 8, met de lemmata *verporrenen* t.e.m. *verposen*. Ook in de *Bouwstoffen* bleken in het totaal een 600 regels tekst te ontbreken, vaak met informatie die onmisbaar was om citaten correct te kunnen dateren.

De Rotterdamse softwarefirma AND, die bij de ontwikkeling van de *Cd-rom Middelnederlands* de technische kant van het project voor haar rekening nam, heeft destijds het gedigitaliseerde *MNW* voorzien van een specifieke XML-achtige codering om het bestand op allerlei wijzen te kunnen ontsluiten met de door hen ontwikkelde zoekprogrammatuur. Tijdens dit parsen zijn er onbedoeld ook weer nieuwe fouten in het bestand terechtgekomen. Het betreft daarbij vooral het correct herkennen + afbakenen van de citaten en het linken van die citaten aan de correcte bronbeschrijving in de *Bouwstoffen*; ook werd tekst soms onterecht als citaat gemarkeerd. Al deze onvolkomenheden werden in de mate van het mogelijke gecorrigeerd zodat het citatenmateriaal van het *MNW*, zeker na de verrijking met de lokaliserings- en dateringsgegevens, nog beter als onderzoekscorpus met historisch taalmateriaal te gebruiken is dan dat bij het *MNW* op cd-rom al het geval was.

Aanpassing van de tekst van het grafische MNW

Zonder ons meteen op het gladde ijs van inhoudelijke correctie te begeven, hebben we het werken aan het *MNW online* ook gebruikt als de gelegenheid om een aantal onvolkomenheden van het grafisch woordenboek aan te pakken. Zo werden de honderden aanvullingen en verbeteringen verwerkt die in het *MNW* zelf (in dl. 1, 7 en 9) en in de *Bouwstoffen* al genoemd werden; ook de acht pagina's aanvullingen op de bronnenlijst die G.I. Lieftinck in de jaargangen 72 (1954) en 75 (1957) van *TNTL* publiceerde werden aan het bestand toegevoegd. Het betreft hier voornamelijk kleinere tekstcorrecties, maar ook extra citaten, de signalering van lemmagedeeltes die weggevallen of op de verkeerde plaats terechtgekomen waren en een extra lemma (het lemma *Stuut*, ‘stuit’) dat nog in dl. 7 ingevoegd moest worden. In de *Bouwstoffen* wordt verder ook twintig keer gewezen op lemma's die geschrapt moeten worden, aangezien ze gebaseerd waren op een verkeerde lezing van het als bron gebruikte handschrift. Deze zg. ‘spookwoorden’, waarbij Verdam vaak scherpzinnige maar nu dus geheel overbodig gebleken hypothesen ontwikkelde, werden niet uit het bestand verwijderd, maar voorzien van een eigen codering, waardoor ze bij bepaalde zoekvragen uitgesloten worden en dus niet zorgen voor vervuiling van de zoekresultaten.

Nog enkele andere onrechtmatigheden die rechtgezet konden worden: bij ca. 100 artikelen werden zetfouten in de lemmaingang gecorrigeerd, woordsoortinformatie werd toegevoegd waar die in het grafische woordenboek ontbrak en

een kleine 200 “blinde” citaten (citatoren zonder bronvermelding) werd geïdentificeerd en gelinkt aan het correcte artikel in de *Bouwstoffen*.

Geraadpleegde literatuur:

D. Geeraerts en G. Janssens, *Wegwijs in woordenboeken*, Van Gorcum, Assen, 1982, p. 93-95.

D. Geirnaert, ‘Werken aan de toekomst van het verleden: van MNW naar MNW_online’. In: E. Beijk, L. Colman e.a. (red.), *“Fons Verborum”*. Feestbundel voor prof. dr. A.M.F.J. (Fons) Moerdijk, Leiden – Amsterdam, 2009, p. 201-213.

Geschiedenis van het MNW

Het idee voor een woordenboek van het Middelnederlands was oorspronkelijk van Matthias de Vries, die eveneens de grondlegger was van het *Woordenboek der Nederlandsche Taal*. In 1864 en 1865 verschenen twee door hem vervaardigde afleveringen, met daarin artikelen van A tot *Anxene*. Omwille van o.a. zijn werkzaamheden aan het *Woordenboek der Nederlandsche Taal* kon De Vries zijn plannen voor een Middelnederlands woordenboek echter niet verder ten uitvoer brengen.

Jacob Verdam en Eelco Verwijs namen in 1873 de fakkel van De Vries over. Het werk dat door hem reeds was gedaan, incorporeerden zij in de eerste afleveringen van hun *Middelnederlandsch Woordenboek*. De aanpak van De Vries, zo vermeldt Verdam in zijn inleiding, namen zij vervolgens ook over als voorbeeld voor hun verdere bewerking. Verwijs, die samen met Verdam de bouwstoffen (het materiaal voor het woordenboek) verzamelde, stierf echter twee jaar vóór het verschijnen van de eerste aflevering, in 1880. Hij had toen het grootste deel van de letter A en de artikelen beginnend met BA bewerkt. Verdam ging vervolgens alleen door. Toen ook hij in 1919 stierf was het woordenboek bijna af: tot *weigerlike* lag de kopij klaar. F.A. Stoett, die al vanaf 1888 de drukproeven kritisch meelas, maakte het negende deel van het woordenboek verder af.

Na de afronding door Stoett (in 1929) zouden er nog twee delen worden toegevoegd. In 1941 verscheen van de hand van A. Beekman deel elf, waarin *Aanvullingen en verbeteringen op het gebied van dijk- en waterschapsrecht, bodem en water, aardrijkskunde enz.* opgenomen waren. Willem de Vreese verzamelde uit het MNW de ca. 23.000 door Verdam in zijn woordenboek aangebrachte verbeteringen van opgenomen citaten en publiceerde die in 1941 als de *Tekstcritiek van J. Verdam*; samen met het eerste stuk van de bronnenlijst (de bronnen beginnend met de letter A, tot en met de bronnen beginnend met F) vormde deze *Tekstcritiek* De Vreeses aandeel in het tiende MNW-deel (de *Bouwstoffen*); het tweede stuk ervan, de bronnen met beginletters F-Z, werd samengesteld door G.I. Liefstinck en verscheen in 1952. Hiermee was het grafische MNW zoals wij dat nu nog kennen, voltooid.

Handwoordenboek MNW en het Supplement daarop

Negenendertig jaar lang heeft Jacob Verdam in zijn eentje aan het MNW gewerkt. Hij was verantwoordelijk voor ca. 19.800 kolommen. Dat alleen al maakt het tot een indrukwekkende prestatie. Maar niet minder indrukwekkend zijn de hoge kwaliteitseisen die hij aan zijn werk stelde. Hij verzekerde zich bijvoorbeeld van advies van specialisten als J.A. Fruin en S.J. Fockema bij de verklaring van rechtstermen of G.J. Boekenoogen als het ging om wapen- en zegelkunde, eigennamen of volkskunde. Hij keek ook heel kritisch naar de door hem gebruikte edities van Middelnederlandse teksten, hetgeen resulteerde in het reeds genoemde totaal van ca. 23.000 verbeteringen. Tijdens de bewerking voegde hij ook regelmatig nieuwe bronnen toe, getuige de aanvullende bronnenlijsten. Nieuwe bronnen betekenden automatisch nieuw materiaal voor reeds gepubliceerde letters. Verdam wilde dergelijke aanvullingen en ook materiaal dat hij gaandeweg van anderen kreeg niet laten liggen. Daarom kondigde hij al in het eerste verschenen deel een supplement aan: ‘Nieuwe woorden, nieuwe betekenissen van opgenomen woorden, en nieuwe bewijplaatsen voor reeds vermelde betekenissen zullen worden bijeengevoegd in het supplement achter het laatste deel’ (MNW I, p. XXXVII). Hij bleef materiaal verzamelen, ook voor de reeds gepubliceerde letters. Dat resulteerde in een verzameling van uiteindelijk ca. 9100 strookjes papier met aantekeningen (uit ca. 700 bronnen), die nu nog steeds bij het INL bewaard worden in Verdams fraaie eikenhouten loketkast.

Toen de uitgever er in 1907 bij Verdam op aandrang om een *Middelnederlandsch Handwoordenboek* te maken, woog Verdam bezwaar dat het grote woordenboek nog niet verder dan de letter *S* was gekomen, en dat het handwoordenboek dus ongelijkmatig van omvang zou zijn, uiteindelijk niet op tegen de daarmee ontstane mogelijkheid om reeds aanvullingen en verbeteringen op basis van het supplementmateriaal, in het handwoordenboek op te nemen. Zo zou het kleinere woordenboek alvast als voorlopig supplement of vermeerderde, verbeterde herdruk van het grote woordenboek kunnen functioneren, en dat doet het in feite nu nog steeds. De eerste uitgave van het *Handwoordenboek* verscheen in 1911. De tweede editie verscheen in 1932, waarvoor het gedeelte van *sterne* tot *wuwe* door C.H. Ebbing Wubben opnieuw bewerkt werd.

Een eigenlijk supplement op het *MNW* is nooit verschenen. Aan J.J. Mak, die zelf in het bezit was van een persoonlijke verzameling van aantekeningen op het *MNW*, werd in 1965 de opdracht gegeven om aan dat supplement te werken. Hij begon zijn werkzaamheden met het exciperen van bronnen uit de aanvullende bronnenlijsten uit het *MNW*, van het supplementmateriaal van Verdam zelf, van De Vreeses handexemplaar van het *MNW*, waarin allerlei aanvullingen en verbeteringen stonden, en van een aantal sinds 1929 verschenen tekstedities. In 1972 werd het project ondergebracht bij het INL te Leiden. Toen was Mak nog volop bezig met exciperen en was er nog niet één artikel verschenen. Uiteindelijk liet hij bij zijn pensionering in 1973 een verzameling na van ca. 25.400 fiches uit ca. 900 bronnen. Er werd toen vastgesteld dat het nog tien jaar zou duren voor het exciperen van de overige aanvullende bronnenlijsten en de nieuw verschenen tekstedities voltooid zou zijn. Daarna moest het woordenboek nog geschreven worden. Vanwege de toenmalige ontwikkelingen in de lexicologie zag men meer heil in nieuwe woordenboeken die de woordenschat van een afgebakende periode behandelen (synchrone woordenboeken of periodewoordenboeken) en die uiteindelijk het oude *MNW* zouden moeten vervangen. Er werd daarom besloten om het verzamelde materiaal zoals het er lag te verwerken in een supplement op het *Handwoordenboek*. Dit *Supplement op het Middelnederlandsch Handwoordenboek* verscheen in 1983 en was van de hand van INL-medewerker J.J. van der Voort van der Kleij. Het bevat ca. 16.500 trefwoorden, vergezeld van een cijfer dat aangeeft welk type aanvulling of verbetering er ten opzichte van het *MNW* aangeboden wordt; geen cijfer (10.703 keer) betekent dat het woord in het *MNW* ontbreekt.

Geraadpleegde literatuur:

K. Depuydt, 'Het staat in het woordenboek, dus het klopt'. In: R. Jansen-Sieben, J. Janssens, F. Willaert (red.), *Medioneerlandistiek. Een inleiding tot de Middelnederlandse letterkunde*. Hilversum, Verloren, 2000, p. 317-325
J.J. van der Voort van der Kleij, *Verdam Middelnederlandsch Handwoordenboek, Supplement*, Leiden – Antwerpen, 1983, p. xiv-xvi.