

Het Antroponymisch Woordenboek van de historische Lage Landen Stand van zaken

In 2009 werd in dit tijdschrift (*Handelingen* 2009) het plan uiteengezet voor de realisatie van een Antroponymisch Woordenboek van de historische Lage Landen. In deze bijdrage wordt verslag gedaan van de werkzaamheden gedurende de afgelopen jaren en van het eerste nog zeer voorlopige resultaat, de website antroponymischwoordenboek.be.

Het Antroponymisch Woordenboek was voor Maurits Gysseling na zijn Toponymisch Woordenboek en zijn Corpus Middelnederlandse Teksten het derde grote desideratum voor de neerlandistiek. Gysseling had zelf alle materiaal daartoe al in de jaren 1944-1955 tijdens uitgebreide archiefzoektochten geëxcerpeerd. Op deze archiefexcerpten is niet enkel zijn onvolprezen Toponymisch Woordenboek gebaseerd, maar wilde hij ook een Antroponymisch Woordenboek realiseren. Gysseling moet zijn plannen daartoe echter nog in de jaren 1960 hebben opgeborgen. Zijn reusachtige verzameling transcripten, waarvan het bestaan slechts aan enkele bevriende vakgenoten bekend was, raakte in de loop der jaren verspreid en verdween langzaam in de collectieve vergetelheid. Ze kon in de jaren 2006-2009 voor vernietiging worden gered en opnieuw worden samengesteld. Een deel bestond uit 15.000 alfabetisch geordende naamsteekkaartjes, die duidelijk Gysselings aanzet waren geweest voor een Antroponymisch Woordenboek. Ze stelden echter te weinig voor voor een heus woordenboek vergeleken bij het Toponymisch Woordenboek, dat vijfmaal zoveel attestaties bevat. Het gros van de persoonsnamen (naar schatting zo'n 200.000) bleek nog opgeborgen te zitten in de duizenden velletjes met teksttranscripties die in de jaren 1944-1955 waren gemaakt.

In principe bevatten deze velletjes alle personen genoemd in historische documenten tussen 650 en 1225 uit een gebied zo groot als de Benelux, Noord-Frankrijk en de aangrenzende westrand van het Rijnland. Het materiaal was niet alleen belangrijk wegens zijn omvang, maar ook wegens zijn homogeniteit en accuraatheid. Het was immers gemaakt door één persoon, die bovendien bekend stond als een van onze beste paleografen.

Nadat eind 2008 Gysselings verzameling archiefexcerpten opnieuw volledig was samengesteld en als ongestructureerde tekst was gedigitaliseerd, kon met een FWO-krediet een onderzoeker worden aangetrokken die in zijn postdoctoraal project een databank en een zoekmachine met het voorhanden materiaal zou ontwikkelen. Het model was voorhanden met de digitale editie en databank van het Toponymisch Woordenboek, die in 2008 waren gerealiseerd (raadpleegbaar onder www.wulfila.be/tw).

Om de integriteit van de onderliggende gegevens te garanderen, moesten eerst door globale vervangopdrachten de talloze stereotiepe fouten worden verbeterd die bij de ocr-lezing waren ontstaan (zie echter infra). Er werden vervolgens programma's ontwikkeld om relevante tekstpatronen op de velletjes te herkennen en weg te schrijven naar genormaliseerde velden: persoonsnamen, plaatsnamen, dateringen, archiefsigels, paginnummer, sociale status (graaf, bisschop, mancipium ...), rollen (schenker, begunstigde, positie in getuigenlijst ...), persoonsrelaties (zoon, dochter van ...). Terloops mag hier worden vermeld dat bij die werkzaamheden honderden toponymische attestaties aan het licht kwamen die Gysseling niet in zijn Toponymisch Woordenboek heeft opgenomen. Vele liet hij kennelijk weg omdat de archiefdocumenten waarin ze voorkomen, niet precies te dateren waren. Sommige van deze attestaties bevatten nochtans opmerkelijke taalkundige varianten.

Bij de verwerking van het materiaal werd ook langzaam duidelijk waarom Gysseling er ondanks zijn werkkraft en scherp verstand niet in is geslaagd uit zijn transcriptenverzameling een Antroponymisch Woordenboek te distilleren. Het zal zeker wel te maken hebben met de omvang van het materiaal, die zo'n vijfmaal groter is dan die van het Toponymisch Woordenboek.

Het is echter vooral de chaotische heterogeniteit van naamvormen en het ontbreken van eenduidige referentievormen geweest die hem van de onderneming heeft doen afzien. Een uitermate weerbarstig probleem waarmee alle diachrone, en in de eerste plaats antroponymische thesaurussen geconfronteerd worden, is de lemmatisering. Bij het Toponymisch Woordenboek was dat probleem relatief gering omdat men voor plaatsnamen doorgaans beschikt over gestandaardiseerde spellingen en geografische coördinaten. Het antroponymisch woordenboek bevat echter namen uit zes eeuwen, uit diverse talen en taalfasen (Oudnederlands, Oudfrans, Medio-Latijn, Hebreeuws, Keltisch ...), in de meest uiteenlopende spellingen zonder dat men hier beschikt over een unieke en consistente referentievorm. Horen etymologisch identieke namen als *Chlothacharius*, *Hlotharius*, *Lotharius*, *Luther* thuis onder één lemma, en zo ja onder welke van de vier genoemde verschijningsvormen: de oudste, de frequentste of de bekendste? Bovendien duidt één persoonsnaam doorgaans verschillende personen aan (bv. de vele *Balduinus*, *Theodericus* enz.) en kan één persoon optreden onder verschillende namen, zoals de H. Bavo, die officieel *Allowinus* (of *Adlowinus*) heette. Hoe men zich daarover ook het hoofd blijft breken, een uniek en consistent lemmatiseringscriterium is met geen middelen te vinden. Alle antroponymische woordenboeken uit het verleden stonden daarbij voor een onmogelijke opgave zoals alleen al te zien is aan de ongewild rommelige typografische voorstelling van de lemma's in de antroponymische lexica van Schönfeld 1911, Morlet 1972, Reichert 1987-1990. Het is dan ook geen wonder dat Förstemann na meer dan honderd jaar nog steeds geen navolgers heeft gevonden en ... dat ook Gysseling zijn plannen voor een Antroponymisch Woordenboek heeft opgeborgen.

Met de moderne databanktechnologie is de langgezochte koningsweg echter gevonden. Het lemmatiseringsprobleem bij het Antroponymisch Woordenboek is opgelost door een lemmatisering op vier hiërarchisch geordende niveaus:

1. ‚Vindvormen‘ (of ‚Literals‘). Dit veld bevat de namen zoals ze letterlijk voorkomen in het archiefdocument. Het is in

de eerste plaats nuttig voor de historisch taalkundige, de filoloog of de paleograaf, die aldus kan zoeken op casusuitgangen, grafemen, afkortingen enz.

2. ‚Noemvormen‘ (nominatieven). Bevat de persoonsnamen in de grafemische standaardvoorstelling die de scriptor wellicht in gedachte had (‚competence‘) toen hij de naam neerschreef van een hem bekende persoon. Het veld maakt aldus een onderscheid tussen onder meer *UUillihelmus* en *Wilhelmus*, die kenmerkend zijn voor de oudste en de jongere Duitse scriptoria, *Willelmus*, de typisch West-Frankische vorm, *Guillelmus* en *Guillermus* als typisch Noord- en Zuid-Romaanse tegenhangers, het idiosyncratische *Willemmus* enz. Deze lemmatisering beantwoordt aan de klassieke indexen bij oorkondenedities.

3. ‚Etymologische vorm‘. Bevat de reconstructie van de Germaanse namen (die het merendeel van het woordenboek uitmaken) in hun Oudgermaanse vorm, met splitsing van de bestanddelen in oniemmen en suffixen. Oudtestamentische of Latijnse namen worden echter niet geëtymologiseerd. Waar de etymologie van een Germaanse naam onzeker is, wordt dat aangegeven door alternatieven. De vindvorm *Alwino* (ablatief) wordt aldus geëtymologiseerd als {al|alh|athal}/wini. Het veld is interessant voor naam- en taalkundig onderzoek, maar ook voor onderzoek van de vroegmiddeleeuwse mentaliteitsgeschiedenis.

4. Persoon. Op het hoogste niveau worden alle naamvarianten verenigd die op één persoon betrekking hebben. De diverse naamvarianten waaronder de H. Bavo bekend staat, *Allowinus*, *Adlowinus* en *Bavo*, worden op dit niveau met elkaar verbonden. De biografische lemmatisering is in de databank nog niet aanwezig, aangezien dat voor elke naam gedetailleerd historisch onderzoek vergt.

Antroponymie nieuwe stijl

De vraag naar een antroponymisch woordenboek als onderzoeksdesideratum gaat al terug tot 1846, toen Jakob Grimm het als onderzoeksproject door de Berlijnse Academie liet uitschrijven. Het project werd uitgevoerd door Ernst Förstemann

(1822–1906)⁽¹⁾ met een hele schaar medewerkers, onder wie de Duitse Antwerpenaar Heinrich Pottmeyer⁽²⁾, die de Nederlanden voor zijn rekening nam. Het resulteerde in het *Altdeutsches Namenbuch*, waarvan het deel *Personennamen* verscheen in 1856, met een herwerking in 1900. *Altdeutsch* in de titel betekent niet ‘Duits’, maar staat volgens de romantische opvattingen van die dagen voor Continentaal-Germaans. Förstemanns woordenboek bevat daarom niet alleen alle Germaanse persoonsnamen uit Duitsland en de Nederlanden, maar ook uit Frankrijk, Italië en Spanje, waar de persoonsnaamgeving in de vroege middeleeuwen bijna volledig van Germaanse oorsprong is. Förstemanns woordenboek is uiteraard verouderd, maar dwingt nog altijd bewondering af als organisatorische prestatie, nog meer wanneer men weet dat nadien er niemand nog in is geslaagd een werk van die allure neer te zetten. Een grootscheeps Duits project, de zogeheten ‘Neue Förstemann’, met satellieten in diverse andere landen, liep in de jaren 1980 uit op een jammerlijke mislukking, onder meer door een gebrek aan homogeniteit en coördinatie, maar ook wel aan nieuwe ideeën.

De behoefte aan een nieuw antroponymisch woordenboek bleef echter bestaan. In de jaren 1990 richtten enkele Duitse mediëvisten met die bedoeling de onderzoeksgroep *Nomen et Gens* op. De groep heeft sindsdien enkele boekenreeksen op haar actief, onder meer bij het prestigieuze *Reallexikon der germanischen Altertumskunde*, en bracht in 2008 een online databank uit met ca. 60.000 persoonsnamen van de vierde tot de achtste eeuw uit Continentaal Europa (www.neg.uni-tuebingen.de/?q=datenbank). Eveneens operationeel is sinds 2005 de imposante *Prosopography of Anglo-Saxon England* (pase.ac.uk). Het succes van beide ondernemingen is te danken aan de nauwe samenwerking van mediëvisten, historisch taalkundigen en specialisten in *humanities computing* en uiteraard ook aan de

⁽¹⁾ Förstemann was hoofd van de keurvorstelijke bibliotheek in Dresden. Hij wordt ook beschouwd als een van de grondleggers van de kwantitatieve linguïstiek en ontcijferde voor het eerst de Mayakalender.

⁽²⁾ Over deze verdienstelijke man zie R. Van Passen, Heinrich Pottmeyer (1862–1930), in HKCTD 74 (2002).

ruimhartige financiering vanwege resp. de *Deutsche Forschungsgemeinschaft* en *The Arts and Humanities Research Council*.

Opvallend bij die innoverende ondernemingen is dat het niet langer naam- of taalkundigen zijn die het voortouw hebben genomen, maar mediëvisten. Hun belangstelling voor persoonsnamen is niet verwonderlijk. Voor de hele periode tot ca. 1100, toch onze halve westerse geschiedenis, zijn de bronnen zo schaars dat namen vaak het enige concrete houvast zijn die de mediëvist heeft om verbanden te kunnen leggen. De tienduizenden personen die vóór ca. 1100 in onze bronnen staan vermeld, zijn echter tot overmaat van ramp voor het grootste deel alleen maar naam. Behalve wanneer het koningen of bisschoppen betreft, is meestal zelfs hun ambt, laat staan hun filiatie niet aangegeven. De vroegmediëvist beseft als geen ander de draagwijdte van de laatste woorden uit Umberto Eco's *Naam van de Roos: Nomina Nuda Tenemus*.

Vanuit onze moderne omgang met namen zijn we spontaan geneigd voornamen te beschouwen als louter arbitraire etiketten, die we volkomen autonoom naar eigen smaak en ingeving kunnen kiezen. Dat was echter allerminst het geval in de vroege middeleeuwen. Sinds de jaren 1960 is vooral onder impuls van Duitse mediëvisten het inzicht gegroeid dat persoonsnaamgeving in de vroege middeleeuwen allerminst vrijblijvend en allesbehalve een allerindividueelste aangelegenheid was. En precies dat houdt in dat een vroegmiddeleeuwse persoonsnaam onverwachte en heel verrassende informatie kan bevatten die ons door de teksten zelf niet expliciet wordt geleverd, over verwantschap, stand, leeftijd, politieke alliantie, itineraria, erfrechtelijke positie enz.

Het nieuwe inzicht heeft de laatste decennia tot een grondige koerswijziging geleid in de vroegmediëvistiek, waar prosopografie een vooraanstaande plaats is gaan innemen en waarbij persoonsnamen uiteraard een belangrijke rol spelen. Het grote belang van die prosopografische benadering blijkt uit het feit dat het niet de geringsten onder de historici zijn die hier de weg hebben gewezen. Om er slechts enkele te noemen: E. Ewig, K.-F. Werner, M. Heinzlmann, alle drie bij een ruimer publiek

beter bekend door de belangrijke rol die ze hebben gespeeld in de Frans-Duitse relaties, met fascinerende studies resp. over de merovingen en Gallo-Romaans Frankrijk; Karl Schmid, die met baanbrekende studies de libri memoriales van Fulda en Sankt Gallen met hun tienduizenden namen vanonder het stof heeft gehaald, recenter in Frankrijk onder meer M. Parisse en R. Le Jan over politiek en naamgeving bij de karolingische en Lotharingse adel in Frankrijk, in Engeland de mediëvisten verbonden aan het PASE-project enz.

De prosopografie laat zien hoe de antroponymie (en de toponymie) nieuw leven kan worden ingeblazen. De naamkunde heeft ooit betere tijden gekend, maar ze lijdt aan ernstige bloedarmoede en is als wetenschappelijke discipline eigenlijk op sterven na dood. Hopelijk brengt de realisatie van het antroponymisch woordenboek en de bruggen die het slaat naar andere disciplines, daar verandering in. Want zoals de Engelse mediëviste Cecily Clark ooit schreef: “*Names are too important to leave to linguists*”. Het Antroponymisch Woordenboek mag zich daarom niet louter tot *Nomina Nuda* beperken, maar moet de basis vormen voor een Facebook of *Liber Facierum* van de Lage Landen in de vroege middeleeuwen. Tot welke verrassende inzichten die nieuwe aanpak leidt voor onze vroegmiddeleeuwse geschiedenis, zal blijken uit een boek over de antroponymie van de Lage Landen waaraan op dit ogenblik de laatste hand wordt gelegd.

De (achter)stand van het project

Het Antroponymisch Woordenboek houdt dus grote beloften in, maar is jammer genoeg door technisch en menselijk falen nog ver van voltooid. De website antroponymischwoordenboek.be staat volgens afspraak sinds eind 2012 weliswaar online, maar achter de fraaie façade gaat voorlopig slechts een rommelige bouwwerf schuil. Diverse databanken zijn weliswaar voorbereid, maar de webapplicatie en de zoekmachine zijn door falende knowhow voorlopig nog onbruikbaar. Het enige veld dat voorlopig doorzocht kan worden, is dat van de Vindvormen (zie

hierboven 1), maar ook dat is zeer onvolledig en in feite onbruikbaar. Bij de eerste proeven met de databank in oktober 2012 bleek dat de tienduizenden leesfouten in de ocr-documenten in 2009 niet auto"magisch" waren verbeterd. De hoogdringende collationering kon op dat ogenblik echter niet meer ter hand worden genomen. De gegevens waren immers in 2012 in een voor filologisch editen inadequate auteurssoftware opgeslagen die het gebruik van speciale lettertekens en globale vervangingen van veel voorkomende ocr-fouten bemoeilijkte. De gegevens bleken bovendien in 2012 op een particuliere server te zijn ondergebracht en werden gedurende vele maanden geïmmobiliseerd. Dank zij de bemiddeling van enkele collega's werden de basisgegevens eind juli 2013 eindelijk aan de Universiteit Antwerpen terugbezorgd, zodat met een jaar achterstand eindelijk van wal worden gestoken met de collationering van transcript en typoscript. De collationering van de documenten zal naar schatting zo'n zes mensmaanden in beslag nemen. Door het ongelukkige formaat waarin de gegevens zijn opgeslagen, moeten namelijk de meer dan honderdduizend tekstfouten stuk voor stuk worden verbeterd (naast het verwijderen van de gebruikelijke ocr-vervuiling en de verbetering van de tekstopmaak). De verwachting is echter dat het Antroponymisch Woordenboek tegen het einde van 2013 jammer genoeg niet volledig, maar alleszins met integere gegevens operationeel zal zijn.

Jozef VAN LOON