

NOTEN

¹ In de doctoraalscriptie van Eva Wilders, getiteld 'Gij, je, jij, jou, jouw, u, *hoe is 't eigenlijk?*', blijkt *u* rond 1850 in toneelteksten *gij* te verdringen (Nederlands VU-Amsterdam).

² Ik volg blz. 127; blz. 84 vermeldt hier [+ status].

³ Bedoeld is 'kerkgenoodschap' of 'groep kerkgenoodschappen'.

Talig erfgoed : de zuidelijke Nederlanden in de 14de eeuw / red. Edward Vanhoutte. – Gent : Koninklijke Academie voor Nederlandse Taal- en Letterkunde, 2002. – 119 p. ; 24 cm. – (Ook verschenen in Verslagen en Mededelingen van de KANTL 111 (2001) ; afl. 3)
ISBN 90-72474-46-5 Prijs: € 9,-

Op 18 mei 2001 werd in Gent het colloquium 'Talig erfgoed' gehouden in verband met een nieuw project van de Vlaamse Koninklijke Academie voor Nederlandse Taal- en Letterkunde (KANTL) gewijd aan de veertiende-eeuwse ambtelijke taal. De auteurs bewerkten hun presentatie voor deze bundel. 'Het resultaat is een boeiende verzameling studies die het 14de-eeuwse talig erfgoed confronteert met haar verleden... en haar toekomst', zo introduceert redacteur Edward Vanhoutte op pagina 6 de acht bijdragen die volgen. Confronterend is de bundel zeker. Alhoewel het thema duidelijk lijkt – veertiende-eeuws oorkondenmateriaal uit Vlaanderen – wordt niet duidelijk aangegeven vanuit welke opdracht of welk uitgangspunt de auteurs hun bijdragen hebben geschreven en hoe het nieuwe project van de KANTL er nu eigenlijk uitziet. Dit maakt de samenhang tussen de bijdragen ondoorzichtiger dan ze bij een eerste waarneming lijkt. De ondertitel belooft bovendien een veel bredere blik dan daadwerkelijk gepresenteerd wordt: uitgangspunt is het schriftelijk overgeleverde taalmateriaal en niet 'de zuidelijke Nederlanden in de 14de eeuw'.

Georges De Schutter geeft zijn bijdrage 'Het talige erfgoed van Vlaanderen' de ondertitel 'een inleiding vol onvervulde wensen' (p. 7-20) en start met enkele opmerkingen over de waarde van taalmateriaal en 'corpusopbouw als wetenschapsgebied'. Hij stelt vast dat onderzoekers behoefte hebben aan teksten, veel teksten en dat het dertiende-eeuwse corpus dat door Maurits Gysseling is uitgegeven dient te worden opgevolgd door een veertiende-eeuws corpus. En daaraan – zo meent hij – is nog veel te weinig gedaan, terwijl er heel veel al bestaand digitaal materiaal beschikbaar gemaakt zou kunnen worden. De Schutter wil corpora bouwen die geschikt zijn voor zoveel mogelijk typen onderzoek en meldt dat de KANTL als ondersteuning daarvan het CTB, Centrum voor Teksteditie en Bronnenstudie in het leven heeft geroepen, vanwaaruit nieuwe projecten gecoördineerd zullen kunnen worden.

Marc Boone presenteert vervolgens in zijn bijdrage 'Bedenkingen van een historicus bij een corpus van laatmiddeleeuwse Nederlandse teksten' (p. 21-41) opmerkingen over hoe uit de enorme hoeveelheid veertiende-eeuws materiaal zodanig geselecteerd kan worden dat historici daar iets aan hebben. Liefst zou hij zich niet beperken tot de veertiende eeuw en slechts met tegenzin stemt hij in met het praktische van die keuze. En de voor hem meest bruikbare teksten stammen uit de rechtspraak: vonnissen, getuigenverklaringen, nota's van de griffler, en dergelijke. Verder zou Boone ook egodocumenten graag in het corpus zien. Omdat vanuit deze teksten de middeleeuwse maatschappij 'van onderen uit' zichtbaar wordt, zijn deze van belang voor historici; ook de mogelijkheid om het proces van globalisatie te onderzoeken maakt ze interessant. Amand Berteloot gaat in 'De bronnen en hun taal' (p. 43-52) eveneens in op teksten 'aan de onderkant', te weten goederenregisters. Die mogen niet ontbreken in een corpus van veertiende-eeuwse teksten. En uit eerdere edities is gebleken dat de transcripties beter alleen aangeleverd kunnen worden door taalkundig-geschoolden. Er is dus nog heel wat werk te doen, aangezien een aantal bestaande edities niet gebruikt kan worden. Oorkonden moeten wel in het corpus, maar zijn eigenlijk weinig nuttig voor taalkundig onderzoek. Het corpus zou dan ook aangevuld moeten worden met literaire teksten om de nadelen van de oorkonden op te vangen, en het liefst zelfs verschillende contemporaine afschriften van eenzelfde tekst, zo meent hij. Vervolgens gaat Berteloot helaas niet in op de wijze waarop de

problemen die opname van dergelijke teksten met zich meebrengen – bijvoorbeeld de diversiteit in taallagen en taalregisters die er zo zou ontstaan – het best opgelost kunnen worden.

De zojuist besproken drie eerste bijdragen laten de lezer in het duister over het referentiekader van de bundel. De inleiding sprak over een oorkondencorpus; in hoeverre ligt deze keuze al vast? Het is mogelijk dat Boone en Berteloot kritiek leveren op een al gemaakte keuze. Of doen ze suggesties voor de nadere invulling van het corpus?

Johan Taeldeman levert vervolgens 'Drie segmenteringsaspecten met betrekking tot de opbouw van een corpus van oude teksten (in casu 14de-eeuwse oorkonden)' (p. 53-58). Taeldeman geeft zijn wensen rond de ruimtelijke, de temporele, en de individuele (onderscheiding van handen) dimensie, waarbij in het eerste deel de al eerder gepresenteerde kritiek op de vlakkenverdeling zoals die door M. Mooijaart is gekozen voor haar *Atlas van Vroegmiddelnederlandse taalvarianten* uit 1992 nog eens breed wordt uitgemeten. Taeldeman kiest ook voor een vlakkenverdeling, maar dan met veel meer vlakken dan Mooijaart heeft gebruikt.

Met de bijdrage van Pieter van Reenen, 'Het 14de-eeuwse Middelnederlandse oorkonden-corpus als dynamisch-systematisch referentiekader voor taalkundig onderzoek' (p. 59-72) krijgen we vaste grond onder de voeten. Hij beschrijft wat er tot nu toe voor het Nederlandse gebied is gedaan en gaat in op het belang van oorkondencorpora voor taalkundig onderzoek: oorkonden kunnen functioneren als gidsdocumentatie en dynamisch-systematisch referentiekader. Aan de hand van een oorkonde uit Gent laat Van Reenen zien hoe deze tekst in het huidige Amsterdamse corpus is opgenomen. Ook meldt hij op welke punten het corpus nog niet goed en gelijkmatig gerepresenteerd is. Maar ook nu al helpt het corpus om teksten van onbekende herkomst te lokaliseren. Andere lokale bronnen dan oorkonden, bijvoorbeeld de ook door Berteloot genoemde goederenregisters e.d., zouden in principe wel wenselijk zijn in het corpus, maar ze zijn minder geschikt voor opname omdat ze niet gelijkmatig voorhanden zijn. Wel kunnen ze 'de betrouwbaarheid van oorkonden als vertegenwoordigers van plaatselijke taal confirmeren, maar zij zullen door hun beperkte vocabularium de netwerk-referentiefunctie van de oorkonden niet kunnen vervangen' (p. 69). Van Reenen legt verder uit voor welke carteringsmethode is gekozen: er is sprake van een binaire gegevenspresentatie – een bepaald gegeven is aanwezig of afwezig. Plaatsen en regio's zijn als punten voorgesteld en de vlakverdeling van de kaart wordt opgebouwd vanuit de punten, ofwel vanuit de gegevens zelf.¹ Van Reenen sluit af met de constatering dat veel delen van de zuidelijke Nederlanden nog onvoldoende in het oorkondencorpus vertegenwoordigd zijn en dat daar verandering in moet komen.

Hierna gaat Jozef van Loon 'Op zoek naar de ideale tekst' (p. 73-90). Oorkonden zijn belangrijk, zo meent hij, en het grootste probleem zal zijn om een verantwoorde keuze te maken uit de overvloed aan stukken. Het is volgens hem noodzakelijk om editoriale krachten te bundelen met historici. Ook Van Loon beveelt processtukken en egodocumenten aan als niet te versmaden voor een corpus dat ook de 'gewone' taal wil vertegenwoordigen; hij illustreert dit met een klein onderzoek naar persoonlijke voornaamwoorden in subject- en object-vorm in een aantal reisjournalen en het dagboek van de Kuringse pastoor Christiaan Munters.

De laatste twee bijdragen concentreren zich op enkele technische aspecten van het bouwen van een bruikbaar corpus. Tom De Herdt demonstreert 'Een hulpprogramma voor manuele lemmatisering van oudere teksten' (p. 91-103). Wil men een tekstcorpus kunnen bevragen, dan is het immers noodzakelijk om het materiaal erin beter bevragebaar te maken dan het in zijn gediversifieerde spelling en gelecteerde en geconjugeerde vormen zou zijn. Dat kan ondermeer door de teksten te lemmatiseren en op woordsoort te coderen. De Herdt gaat helaas niet in op andere opties, zoals technisch geavanceerde zoekmogelijkheden. Hij heeft een programma geschreven dat het de onderzoeker mogelijk maakt om teksten semi-automatisch te taggen. Het lijkt een goed doordacht programma – maar het is duidelijk dat alleen onderzoekers die al enigszins thuis zijn in technisch geavanceerdere benaderingen van tekstcorpora niet zullen worden afgeschrikt. De geïnteresseerde lezer zou willen weten of het programma ook door anderen aangeschaft dan wel gebruikt kan worden.

Edward Vanhoutte sluit de bundel af met een bijdrage over 'Literaire tekstedities, talig erfgoedonderzoek en humanities computing in Vlaanderen' (p. 105-119). Hij gaat in op de manier waarop een veertiende-eeuws corpus met het talig erfgoed elektronisch ontsloten zou

moeten worden, te weten in een diplomatische editie ‘waarbij de transcriptie is gebeurd m.b.v. *text encoding* technieken volgens een internationale standaard’ (p. 108). Op die wijze kan het corpus polyfunctioneel worden, zo stelt hij terugkerend naar een van de voorwaarden die De Schutter in zijn inleidende bijdrage had gesteld. Vervolgens legt Vanhoutte uit hoe zo’n corpus ‘projectgewijze’ uitgegeven kan worden (onder andere door gebruik te maken van de vele licentiaatsverhandelingen die ongepubliceerd blijven) en dat bij het coderen de TEI-standaard (Text Encoding Initiative) gehanteerd moet worden. Hij verhaalt uitvoerig hoe die tot stand gekomen is. Tot slot meldt de auteur dat het aan ‘de specialisten ter zake’ is om een prioriteitenlijst op te stellen, maar dat het Centrum voor Teksteditie en Bronnenstudie een belangrijke rol kan spelen in het van de grond krijgen van het project.

Het problematische van deze bundel is de ondoorzichtigheid van de ordening van de bijdragen en van de samenhang tussen de verschillende stukken. Het is onduidelijk wat de status is van het aangekondigde project van veertiende-eeuwse Zuidnederlandse taal. Gaat het nog opgestart worden, of zal het Corpus van Van Reenen en Mulder aangevuld en uitgebreid worden voor het Zuidnederlandse deel? Als dat laatste het geval is, dan zou de bundel het best met de bijdrage van Van Reenen zijn gestart. Of zijn de bijdragen bedoeld als aanzet tot een uitgewerkte projectaanvraag? Het zou handig zijn geweest als dit ergens vermeld was.

Een ander punt is dat het opvalt dat de auteurs die hun wensen (of hun kritiek?) melden wat de op te nemen teksten betreft, zich niet veel zorgen lijken te maken over de methoden en technieken die er nodig zijn om een tekstcorpus bestaand uit met name ouder taalmateriaal ook met profijt te kunnen gebruiken. Toch is het zo dat onderzoekers er niet zijn met digitale teksten alleen. Berteloot en Van Loon geven methodisch wel enkele interessante hints, met betrekking tot woordgeografie in handschriften van de *Legenda aurea* resp. het gebruik van persoonlijke voornaamwoorden, maar wat de technische kant betreft is het alleen De Herdt die met concrete hulpmiddelen komt. Het is te hopen dat binnenkort duidelijk wordt hoe het project veertiende-eeuwse oorkondentaal eruit ziet en welke zowel inhoudelijke als technische producten dit zal opleveren.²

Karina van Dalen-Oskam

NOTEN

¹ In Kaart 1 op p. 72 is er waarschijnlijk iets mis gegaan in de omschrijving; het betreft het voorkomen of afwezig zijn van het prefix *ghe-* in het voltooid deelwoord van *comen*. Niet de *aanwezigheid* daarvan maar de *afwezigheid* zal met donker zijn aangeduid.

² Inmiddels is er meer over het project ‘Het veertiende-eeuwse Talig Erfgoed van de Zuidelijke Nederlanden’ te vinden in: *Jaarboek van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde*, Gent, 2003, p. 48-51.

“Er waren een Marokkaan, een Turk en een Nederlander...” : volkskundige en taalkundige opstellen over het vertellen van moppen in de multiculturele wijk Lombok / onder red. van Theo Meder. - Amsterdam : Stichting beheer IISG, 2001. - VI, 210 p. : tab. ; 23 cm (Onderdeel van het onderzoeksprogramma TCULT) ISBN 90-6861-215-8 Prijs: € 22,65.

Kom op, vertel een mop: dat was het motto op een affiche waarmee TCULT-onderzoekers zich tijdens de tweede aflevering van het multiculturele festival *Salaam Lombok* in september 1998 aan het publiek presenteerden. TCULT is de afkorting van ‘Talen en Culturen in het Utrechtse Lombok en Transvaal’, een multidisciplinair samenwerkingsverband tussen de universiteiten van Utrecht, Amsterdam, Tilburg en Leiden en het Meertens Instituut. Behalve deze instellingen hebben ook de gemeente Utrecht en NWO dit inmiddels afgeronde project gefinancierd; het ressorteerde onder de NWO-paraplu ‘De multiculturele en pluriforme samenleving’. De bedoeling was om langs verschillende wegen antwoord te krijgen op de vraag: in hoeverre zijn talen en culturen onderhevig aan veranderingen door de specifieke contact-situaties in de wijk? In ‘*Er waren een Marokkaan, een Turk en een Nederlander...*’ worden