

Focus

Meertens Instituut

Het Amsterdamse Meertens Instituut, een onderdeel van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW), onderzoekt en documenteert alledaagse verschijnselen in de Nederlandse taal en cultuur. *e-data* sprak met Sjef Barbiers, onderzoeksleider van de afdeling variatielinguïstiek, over een instituut in beweging.

Een gesprek over het Meertens Instituut kan niet anders dan beginnen met een vraag naar Voskuil, die met zijn zevendelige romancyclus *Het bureau* het Meertens Instituut bij velen bekend maakte. Barbiers herkent zich niet in de instelling zoals Voskuil die beschreef. 'Integendeel. Voskuil beschreef een in zichzelf gekeerde wereld, die niet deel nam aan het grotere wetenschappelijke debat. Onderzoekers van het huidige Meertens Instituut staan middenin de internationale academische wereld.'

Inhoudelijk is er natuurlijk wel sprake van continuïteit: het Instituut houdt zich nog steeds bezig met variatie in cultuur en taal. Maar waar vroeger papieren vragenlijsten naar alle hoeken van het land werden gestuurd, gingen voor de recente Syntactische Atlas van de Nederlandse Dialecten veldwerkers naar 267 plaatsen in het Nederlands taalgebied. Barbiers: 'Methodologisch was de grote vernieuwing dat we zoveel mogelijk met paren van lokale interviewers werkten. Daardoor treedt minder vervuiling op ten gevolge van het taalgebruik van de onderzoeker. Bovendien zijn de geluidsoptnamen on-line beschikbaar, opgelijnd met transcripties en verrijkt met woordsoortinformatie.' De gegevens zijn zo veel makkelijker doorzoekbaar dan voorheen. De resultaten van zoekvragen kunnen op kaarten worden gepresenteerd die inzicht geven in de verspreiding van dialectverschijnselen. Barbiers: 'Al onze bestanden komen op deze manier ter beschikking van onderzoekers en het bredere publiek.' Onder de databases van het Instituut zijn een paar zeer populaire bestanden: de database van Nederlandse familienamen bijvoorbeeld, maar ook de Volksverhalendatabank, de Liederenbank en Soundbites, een collectie van dialectopnamen. Onderzoek en collectievorming gaan voor het Meertens hand in hand. Barbiers verklaart: 'Het één kan niet zonder het ander. Dankzij grote digitale collecties ontstaan er nieuwe onderzoeksmogelijkheden. We verzamelen niet om het verzamelen.' Voor zover er nog vragenlijsten worden verzonden, loopt dat tegenwoordig via het web. Via het Meertens Panel worden regelmatig vragenlijsten voorgelegd aan meer dan 4500 respondenten. Ook is sinds vorig jaar het Meldpunt Taal online waar mensen interessante taalverschijnselen kunnen melden.



DIEDERIK VAN DER LAAN

Onderzoekleider Sjef Barbiers: 'Onze onderzoekers staan middenin de internationale academische wereld'

Het Meertens Instituut wil zich ontwikkelen tot centrum voor technologische innovatie binnen de geesteswetenschappen. Barbiers: 'Het is beleid van de KNAW, maar we waren zelf al hard op weg.' Het Meertens wordt één van de expertisecentra binnen de infrastructuur voor de geesteswetenschappen die het project CLARIN tot stand wil brengen. Er staat een groot aantal computationele taalkundige projecten op stapel, zoals bijvoorbeeld INPOLDER, voor het annoteren van historisch taalmateriaal – nu vaak onder andere vanwege afwijkende spelling nog een probleem in tekstgeoriënteerd onderzoek. Barbiers: 'De programmatuur die het resultaat is van dit soort projecten willen we via web services aan andere onderzoekers aanbieden. Het zijn trouwens niet alleen de taalkundigen die met computationele technieken werken: onderzoeker Theo Meder gaat bijvoorbeeld langs computationele weg de volksverhalen onderzoeken op motieven en thema's.'

Wenkend perspectief, voor zowel taalkundigen als etnologen, is de vraag naar het computationeel modelleren van menselijke cognitie. Barbiers: 'De musicologen willen weten hoe melodieën worden herkend. Taalkundigen zijn altijd al geïnteresseerd geweest in de representatie van taal in het brein. Digitale bestanden en onderzoeksmethoden kunnen een antwoord op zulke vragen dichterbij brengen.'

MESS migrantenpanel nu open voor wetenschappers

Een evaluatiecommissie van de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) oordeelde vorig jaar niet alleen gunstig over het verloop van het MESS-project (Measurement and Experimentation in the Social Sciences) van CentERdata. Op advies van dezelfde commissie heeft NWO nu 750.000 euro extra toegekend aan MESS.

Het eerste oordeel van de commissie werd gevraagd in het kader van een tussentijdse evaluatie van de vijf grootschalige onderzoeksfaciliteiten die vanuit BIG-programma van NWO zijn gefinancierd. MESS is een van die vijf, met als kern het LISS-panel (Longitudinal Internet Studies for the Social sciences).

Projectleider Annette Scherpenzeel van MESS is dubbel blij: 'De evaluatie van de commissie is natuurlijk een opsteker, maar we verheugen ons er vooral over dat we met de extra middelen weer nieuwe activiteiten kunnen opzetten die inspelen op recente wetenschappelijke en maatschappelijke ontwikkelingen.' Een deel van de extra subsidie zet CentERdata in om zogenoemde biomarkers in het LISS panel te ge-

bruiken: innovatieve apparaten en zelftests waarmee gezondheidsaspecten objectief gemeten kunnen worden. Op dit moment hebben duizend huishoudens in het LISS-panel bijvoorbeeld een geavanceerde weegschaal in gebruik die gewicht en vetpercentage meet en via internet aan CentERdata doorstuurt. Andere mogelijkheden zijn bloeddrukmeters en bewegingssensoren.

Met het tweede deel van de extra subsidie wordt het migrantenpanel, het tweede online panel binnen MESS, opengesteld voor wetenschappers. Dat panel omvat een extra steekproef van de verschillende migrantengroepen in Nederland die, net als de steekproef voor het LISS panel, getrokken is uit de Gemeentelijke Basis Administratie. Door de extra subsidie kunnen wetenschappers nu onderzoekvoorstellen indienen om gratis eigen data te verzamelen in het migrantenpanel. Ze zullen bovendien kosteloos over alle data kunnen beschikken die in het migrantenpanel verzameld zijn, net als bij het LISS panel. (EB)

 www.lissdata.nl

Op zoek naar gelijkennis in muziek

Mensen kunnen overeenkomsten tussen muziekfragmenten herkennen. Maar kunnen computers dat ook? Als het antwoord op die vraag bevestigend is levert dat nieuwe mogelijkheden op voor het elektronisch zoeken en vinden van muziek. Aan de universiteit Utrecht wordt naar dat antwoord gezocht, met een Vidi-subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek NWO.

Het herkennen van gelijkennis is cruciaal voor ons begrip van de wereld. In de cognitiewetenschap zijn formele modellen ontwikkeld die gelijkennis theoretisch beschrijven en empirisch meetbaar maken. Maar er bestaat nog geen algemeen model van gelijkennis in muziek. Dit is een cruciaal probleem in het jonge, snel groeiende vakgebied van *music information retrieval* (MIR). MIR onderzoekt methoden die het mogelijk maken om op basis van gelijkennis muziek te vinden in grote digitale muziekcollecties. Omdat er geen algemeen model voor gelijkennis in muziek bestaat, is het een fun-

damenteel probleem in MIR hoe de informatie die uit een digitaal muziekdocument wordt gehaald, gerelateerd kan worden aan de ervaring van gelijkennis door de gebruikers. In het project van de Universiteit Utrecht wordt een computationeel model ontwikkeld dat beschrijft hoe luisteraars gelijkennis in muziek ervaren. Het probleem wordt benaderd door een belangrijk principe in muziek, het variatieprincipe, als basis te nemen voor het modelleren van gelijkennis. De afbeelding toont een voorbeeld voor twee volksliederen uit de Nederlandse Liederenbank van het Meertens Instituut. De blauw gekleurde motiefjes tonen prominente patronen die door variatie met elkaar verbonden zijn (de toonhoogtes zijn dezelfde, maar de ritmische structuur verschilt). Omdat bij het luisteren naar één van de liedjes aan het gerelateerde motiefje van het andere liedje gedacht wordt, worden de twee liedjes als gelijkend beoordeeld. Op basis van dit fenomeen wordt de wisselwerking tussen lokale en globale kenmerken van

muziek (zoals melodie, ritme, harmonie) in een model ondergebracht om gelijkennis te bepalen. Het herkennen van de blauwe patronen tijdens het luisterproces is namelijk mede afhankelijk van de globale context (in de afbeelding in geel), waarin de patronen zijn ingebed.

Het project zal een model voortbrengen van gelijkennis voor drie muziekstijlen: klassiek, populair en volksmuziek. Hiervoor wordt gestreefd naar samenwerking tussen *music information retrieval*, muziekwetenschap en cognitiewetenschap. Niet alleen wordt een theoretisch kader ontwikkeld voor gelijkennis in muziek, ook worden computermethoden toegepast die het mogelijk maken gelijkende muziekdocumenten te vinden op basis van variatie. Daarmee zet het project een beslissende stap op weg naar een algemeen model van gelijkennis in muziek. Dat vormt dan weer de basis voor methoden om succesvol naar muziek te zoeken in online digitale collecties. (Anja Volk)

