

Methoden voor tekstevaluatie onder de loep

Ter inleiding

De begrijpelijkheid van een bijsluiter, de geschiktheid van een voorlichtingsfolder voor de doelgroep, de toegankelijkheid van een website: in diverse studies is aangetoond dat formatief evaluatieonderzoek gericht op dergelijke punten een belangrijke stap is op weg naar effectieve (papieren en digitale) teksten. De tekstontwerper staat inmiddels een groot scala aan evaluatiemethoden ter beschikking: van expertonderzoek aan de hand van het CCC-model tot doelgroeponderzoek met hardopdenkprotocollen of de plus-en-minmethode.

Tijdens het VIOT-congres in december 2002 in Antwerpen vond een symposium plaats over 'Document design en methoden voor tekstevaluatie'. Daarin stonden diverse tekstevaluatiemethoden ter discussie. Een dergelijke discussie kan natuurlijk gevoerd worden op basis van praktijkervaringen. In vier bijdragen – nu in dit themanummer opgenomen – werd echter onderzoek gepresenteerd naar de waarde van een of meer methoden. Daarmee kreeg het symposium een soort Droste-effect, dat ook dit themanummer laat zien. Hier wordt onderzoek gepresenteerd naar de kwaliteit van een aantal teksten, maar het gaat de onderzoekers niet zozeer om die teksten maar om de methoden die gebruikt worden om tekstkwaliteit in een conceptfase vast te stellen – met het oog op revisie en verbetering.

Het onderzoek naar de waarde en beperkingen van tekstevaluatiemethoden is in de afgelopen 10 jaar een interessant en geschakeerd werkteerterrein geworden. In *Met het oog op de lezer* (De Jong & Schellens 1995) konden we nog niet veel meer doen dan een overzicht bieden over de verschillende beschikbare methoden. Die inventarisatie bracht ook aan het licht dat we van de meeste methoden nog heel weinig wisten: hoe makkelijk of lastig zijn ze in het gebruik in de praktijk, hoe betrouwbaar zijn ze, wat zeggen de resultaten ervan precies, hoe hanteerbaar zijn de resultaten als handvat voor revisie? Bij veel van de toen besproken methoden konden we dergelijke vragen alleen maar stellen en over de antwoorden speculeren. Zo werd de inventarisatie tegelijkertijd een onderzoeksagenda voor de toekomst. Uit dit themanummer blijkt dat inmiddels hard aan die agenda wordt gewerkt.

We signaleren hier enkele ontwikkelingen. Ten eerste is het terrein van onderzoek verbreed. Het ging ons destijds nog vooral om methoden die bij het ontwerp van voorlichtingsteksten konden worden gebruikt. In dit themanummer komen naast voorlichtingsteksten (brochures maar ook bijsluiters) nu ook websites aan de orde, hier vertegenwoordigd door een online catalogus en een wereldwijd gebruikt bibliografisch systeem.

Ten tweede is inmiddels een scala aan methoden in onderzoek: in de artikelen in dit themanummer gaat het niet alleen om vertrouwde instrumenten zoals de plus-en-minmethode, het daarvan afgeleide softwareprogramma Focus en het CCC-model van Renkema, maar ook

de hardopdenkmethode wordt nu in verschillende gedaanten als object van onderzoek onder de loep genomen.

Ten derde is in dit nummer vaak niet meer alleen de vraag aan de orde of een methode geschikt is voor zijn doel, maar treden achterliggende vragen naar voren: wat verklaart het verschil in resultaten die met twee uiteenlopende methoden worden geboekt? Wat is het effect van de ingreep in het 'normale' leesproces, die een methode met zich meebrengt? Leidt de methode tot dezelfde resultaten bij proefpersonen met uiteenlopende culturele achtergrond? Zijn de cognitieve vaardigheden en terminologische categorieën die de methode vooronderstelt, bij de lezer/proefpersoon wel beschikbaar? Met dergelijke vragen verdiept het onderzoeksterrein zich en worden de theoretische vooronderstellingen die elke methode noodzakelijkerwijs met zich meebrengt, duidelijk en onderzoekbaar.

Tenslotte stellen we vast – ook al blijkt dat in een Nederlandstalig tijdschrift voornamelijk uit de aangehaalde literatuur – dat het onderzoeksterrein internationaal aan gewicht wint. We gaven eerder een overzicht van het internationale onderzoek naar probleemopsporende tekst-evaluatiemethoden (De Jong & Schellens 2000, 2002). Dat onderzoek vindt plaats vanuit heel verschillende discipline achtergronden (document design, interface design, human-computer interaction, ergonomie, marketing research of onderwijskunde), maar bij elkaar genomen begint zich een gezamenlijke vraagstelling en methodologie af te tekenen. Het onderzoek uit Nederland (niet alleen van de onderzoekers die in dit nummer aan het woord komen) speelt in die ontwikkeling een vooraanstaande rol.

Wat mag de lezer hier verwachten?

Henk Pander Maat en Leo Lentz rapporteren over onderzoek naar de waarde van de hardop-leesmethode als methode om begripsproblemen op het spoor te komen. In vergelijking met de plus-en-minmethode en het computerprogramma Focus scoort het lezersprotocol zeer goed. Niet alleen is het aantal begripsproblemen dat met een lezersprotocol wordt opgespoord groter, ook de kans dat in een lezersprotocol reële begripsproblemen aan het licht komen blijkt groter. In een nader onderzoek met vier varianten van het lezersprotocol blijkt dat de productiviteit van de methode met name te danken is aan de onmiddellijkheid: proefpersonen geven meteen commentaar tijdens het lezen en stellen dat niet uit zoals in de plus-en-minmethode.

Sanne Elling en Leo Lentz deden onderzoek naar de waarde van het CCC-model van Renkema als alternatief voor een lezersgerichte tekstbeoordeling. De auteurs vergeleken de resultaten van experts die al dan niet het CCC-model gebruikten, met die van verschillende lezersgerichte methoden. De resultaten zijn niet bemoedigend voor experts: zij zijn het meestal niet eens over de lezersproblemen (ze voorspellen allemaal andere problemen), hun problemen komen slechts in geringe mate overeen met de problemen die doelgroepgerichte methoden aan het licht brengen en zij scoren mét CCC-model niet beter dan zonder. De 'experts' waren hier gevorderde studenten die een training van één uur met het CCC-model kregen. Mogelijk was dat te weinig, maar de resultaten bevestigen eerdere bevindingen: het blijkt niet eenvoudig om experts te brengen tot betrouwbare en goede voorspellingen van lezersproblemen.

Maaïke van den Haak, Menno de Jong en Peter Jan Schellens vergelijken synchrone en retrospectieve hardopdenkprotocollen in de evaluatie van een online catalogus. In de synchrone variant verrichtten de proefpersonen een aantal zoektaken hardopdenkend. In de retrospectieve variant verrichtten ze die taken stil, maar geven zij achteraf commentaar bij een opname van hun verrichtingen. De methoden blijken globaal tot dezelfde resultaten te leiden, maar ze

Methoden voor tekstevaluatie onder de loep

komen anders tot stand: synchroon hardopdenken leidt tot meer fouten en omwegen in de taakuitvoering; de gevonden problemen zijn daardoor merendeels observeerbaar en niet of niet alleen uit het commentaar van de proefpersoon af te leiden. Bij retrospectief hardopdenken presteren de proefpersonen beter. De problemen komen nu merendeels pas aan het licht door het commentaar achteraf, en zijn minder uit het gedrag af te leiden. Daarmee is de reactiviteit van hardopdenkonderzoek zeker bij complexe taken aan de orde.

Miranda Hall, Menno de Jong en Michaël Steehouder snijden, in een bijdrage die niet tijdens het VIOT-congres was te horen, een heel nieuw onderwerp aan. De resultaten van verschillende pretestmethoden zouden wel eens afhankelijk kunnen zijn van de cultuur waaruit proefpersonen afkomstig zijn. Zij vonden dat de plus-en-minmethode en retrospectieve hardopdenkprotocollen, ingezet om een website te evalueren, inderdaad anders functioneren. De plus-en-minmethode brengt bij proefpersonen uit een collectivistische cultuur minder problemen aan het licht dan de hardopdenkprotocollen. Daarnaast formuleren proefpersonen uit een collectivistische cultuur hun commentaar in de hardopdenkprotocollen indirecter dan proefpersonen uit een individualistische cultuur. Verder is interessant dat in het onderzoek de verwachte verschillen in gedrag wel werden gevonden en de verwachte verschillen in zelfrapportagematen niet. Dat wijst erop dat zelfrapportage (bijvoorbeeld in de beantwoording van vragenlijsten) veel gevoeliger is voor sociale wenselijkheid dan gedrag – een flinke complicatie in intercultureel onderzoek naar (o.a.) tekstkwaliteit.

Tenslotte buigen Leo Lentz en Menno de Jong zich over de vraag in welke termen lezers denken over tekstproblemen. In veel vragenlijsten die aan lezers worden voorgelegd, worden termen gebruikt die aan tekstonderzoek zijn ontleend; in sommige evaluatiemethoden (zoals in Focus) moeten lezers hun eigen commentaar benoemen in vooraf gegeven termen. Sluit die terminologie wel aan op categorieën en termen waarmee leken vertrouwd zijn? In hun onderzoek lieten Lentz en De Jong leken de lezersreacties uit eerder onderzoek op stapeltjes leggen en benoemen. Daaruit trekken zij de conclusie dat de termen die in het vakgebied worden gebruikt nog redelijk goed aansluiten op de lekenterminologie. De duidelijkst herkenbare klassen problemen die leken onderscheiden zijn: overbodige en ontbrekende informatie, geloofwaardigheid, formulering en structuur.

Ter afsluiting past dank en verontschuldiging. De redacteuren van de VIOT-congresbundel (Van Waes e.a., 2003) danken we hartelijk voor hun medewerking aan dit themanummer in de vorm van de bereidwillige afstand die zij van enkele zeer interessante papers voor hun bundel deden. De link met het congres en het daar gehouden symposium vormt tegelijkertijd de verontschuldiging voor het eenzijdig Utrechts-Twentse karakter van dit themanummer.

Bibliografie

- Jong, M. de & P.J. Schellens (1995).** *Met het oog op de lezer. Pretestmethoden voor schriftelijk voorlichtingsmateriaal.* Amsterdam: Thesis.
- Jong, M. de & P.J. Schellens (2000).** Toward a document evaluation methodology: What does research tell us about the validity and reliability of evaluation methods. *IEEE Transactions on professional communication*, 43, 242-260.
- Jong, M. de & P.J. Schellens (2002).** Tekstevaluatie. Onderzoek naar de validiteit van probleemopsporende methoden. *Tijdschrift voor Taalbeheersing*, 14, 146-166.
- Waes, L. van, P. Cuvelier, G. Jacobs, I. de Ridder (red.) (2003).** *Studies in Taalbeheersing* (vol.1). Van Gorcum: Assen