

In welke termen denken lezers over tekstproblemen?

1. Inleiding

In welke termen denken leken over lezersproblemen? Wanneer zij moeite hebben met de interpretatie van een passage in een brochure, zien zij dat dan als een *begripsprobleem*? En beschouwen zij een betuttelende toon als een kwestie van *stijl*? Of zijn dat termen voor categorieën die wij als experts in tekstkwaliteit de afgelopen decennia hebben ontwikkeld om typen lezersproblemen aan te duiden, maar die niet aansluiten op de denkwereld van gewone lezers? In de literatuur zijn verschillende overzichten van categorieën gepresenteerd. De Jong en Schellens (2000) noemen in een recent overzicht bijvoorbeeld de volgende categorieën:

- begrip: hieronder vallen problemen op het niveau van een woord, zin of tekstfragment met het begrijpen en toepassen van de informatie;
- acceptatie: lezers betwisten als feiten gepresenteerde eenheden in de tekst, zijn het oneens met waardeoordelen of wijzen adviezen af;
- waardering: lezers geven de voorkeur aan een aantrekkelijker formulering zonder een probleem te ervaren met begrip of acceptatie;

Samenvatting

In diverse methoden voor tekstevaluatie leggen we aan proefpersonen categorieën voor van lezersproblemen, zoals *stijl*, *structuur* en *begrip*. Soms in de vorm van schaaltes waarop zij een score moeten aangeven of in open vragen zoals: "Wat vindt u van de stijl van deze tekst?" Soms ook in de vorm van aandachtspunten waarop lezers hun commentaar moeten richten. Onduidelijk is echter welke betekenis lezers die niet elke dag met tekstkwaliteit bezig zijn, aan deze categorieën toekennen. In welke termen formuleren zij zelf probleemttypen? En welke betekenis kennen zij toe aan die termen? In dit artikel rapporteren wij over een onderzoek waarin aan vijftig proefpersonen gevraagd is een ordening aan te brengen in een reeks problemen die andere lezers bij eerder onderzoek spontaan gerapporteerd hebben. De resultaten maken duidelijk dat men gemiddeld vijf tot negen probleemttypen onderscheidde. Deze verschillende probleemttypen worden besproken en met voorbeelden geïllustreerd.

- structuur: lezers hebben een probleem met de volgorde waarin eenheden gepresenteerd worden of met de markering van de structuur;
- relevantie: lezers vinden bepaalde informatie overbodig of een onderdeel te ver uitgewerkt;
- volledigheid: lezers vragen om meer informatie over een specifiek punt of een verdere uitwerking daarvan;
- vormgeving: lezers geven kritiek op de uiterlijke vorm van de tekst of op afbeeldingen in het document;
- correctheid: lezers geven kritiek op een passage wat betreft de spelling, interpunctie, zinsbouw of andere tekstconventies.

Dit overzicht is gepresenteerd met de bedoeling tot een categorisering te komen van allerlei mogelijke problemen die lezers in teksten kunnen ervaren. Andere overzichten, zoals in Maes, Ummelen en Hoeken (1996) zijn bedoeld om oordelen van lezers over teksten te genereren. Ook daar vinden we een categorisering van dimensies:

<i>begrijpelijkheid</i>	makkelijk	-	moeilijk
	eenvoudig	-	ingewikkeld
	duidelijk	-	onduidelijk
	overzichtelijk	-	onoverzichtelijk
	logisch opgebouwd	-	onlogisch opgebouwd
	bondig	-	omslachtig
<i>aantrekkelijkheid</i>	interessant	-	oninteressant
	afstandelijk	-	aansprekend
	afhoudend	-	uitnodigend
	saai	-	boeiend
	onpersoonlijk	-	persoonlijk
	eentonig	-	afwisselend

Blijkbaar vallen *opbouw* en *overzichtelijkheid* hier onder een categorie van een hogere orde zoals *begrijpelijkheid*, terwijl het verwante *structuur* bij Schellens en De Jong een zelfstandige categorie is. Wat zijn zinvolle rubrieken voor tekstevaluatie? Hoe ordenen en benoemen we die? En hoe werken we ze uit in afzonderlijke items voor vragenlijsten die getest zijn op betrouwbaarheid?

Een antwoord op deze vragen is niet goed mogelijk zonder een grondige bezinning op de vraag hoe leken die niet elke dag met tekstevaluatie bezig zijn, zelf denken over problemen met teksten. Welke categorieën onderscheiden zij? Hoe ordenen en benoemen zij die? Met iedere vragenlijst, met elke schaal en met elk lijstje met aandachtspunten doen we immers een beroep op hun interpretatie van onze termen. Ruim twintig jaar geleden vroeg Cicourel (1982) in een klassiek geworden artikel al aandacht voor het probleem van de ecologische validiteit in allerlei onderzoeksmethoden waarvoor geldt dat zij

In welke termen denken lezers over tekstproblemen?

“[...] presuppose something about the way people are able to analyze textual materials. For example, in responding to questions in a reading test, the respondent must utilize several sources of knowledge that the researcher used in interviews and surveys, therefore, presumes a theory of communication and comprehension that is seldom addressed [...].”

In het onderzoek¹ waarover we hier rapporteren, staat dat perspectief van validiteit centraal. Ons gaat het daarbij met name om de begripsvaliditeit van de instrumenten die in evaluatieonderzoek gangbaar zijn: vragenlijsten die met behulp van schaaltes de waardering van teksten meten of met open vragen een oordeel ontlokken over bijvoorbeeld de *stijl* of *geloofwaardigheid* van een document. Maar ook voor een evaluatiemethode als Focus is de vraag relevant, omdat proefpersonen met dat software-programma bij elk commentaar zelf een passende probleemcategorie moeten kiezen.

2. Opzet van het onderzoek

In essentie hebben de geformuleerde onderzoeksvragen betrekking op de categorisering van problemen die de lezer zelf ervaart bij het lezen van een tekst. Deze gedachte leidt tot een mogelijke opzet waarbij elke proefpersoon zijn of haar eigen commentaren op een tekst in probleemtijpen onderverdeelt. Een nadeel van die werkwijze is dat elke proefpersoon met specifieke eigen problemen komt. De ervaring leert ons dat in een pretest met bijvoorbeeld dertig proefpersonen hoogstens een vijftal problemen door meer dan tien proefpersonen gezamenlijk genoemd zou worden, en geen enkel probleem door meer dan twintig. Over de meeste problemen zou derhalve in zo'n opzet niet meer dan één uitspraak gedaan worden.

Wij kozen daarom voor een andere opzet, waarbij de lezersproblemen vooraf gegeven zijn en de proefpersonen dus een categorisering maken van *andermans* commentaren. We hebben twee groepen van 25 proefpersonen geconfronteerd met een deel van de feedback die in eerder evaluatieonderzoek verzameld was, verdeeld over telkens 40 kaartjes met op elk kaartje een nummer van de regel van de tekst en het probleem zoals lezers dat eerder verwoord hadden.

De proefpersonen lazen allereerst de tekst rustig door en namen vervolgens het eerste kaartje van de stapel en legden dat voor zich neer. Daarna pakten ze het tweede kaartje en stelden zichzelf de vraag of dat een vergelijkbaar soort probleem weergaf of dat het echt iets anders was. In het laatste geval legden ze het ernaast. Bij het derde kaartje gingen zij na of dat probleem beter bij het eerste of het tweede paste of dat het weer een nieuw soort probleem was. Zo vormden zij gaandeweg stapeltjes van lezersproblemen die naar hun idee bij elkaar hoorden. Tussentijds mochten ze zo vaak als ze wilden kaartjes veranderen van stapel. Ze mochten ook stapels samenvoegen en splitsen. Als alle kaartjes geordend waren, vroeg de proefleider of ze nog iets wilden veranderen. Zo nee, dan startte de tweede fase van de opdracht.

Er waren nu een aantal stapeltjes gevormd met lezersproblemen. Voor elk stapeltje werden de volgende vragen gesteld:

- Waarom horen deze kaartjes bij elkaar? Wat hebben ze gemeenschappelijk?
- Hoe zou u dit groepje problemen noemen, als u er een woord voor zou moeten kiezen?
- Is dat een term die u vaker gebruikt in het dagelijks leven?
- Welk kaartje vindt u het beste voorbeeld van die groep? En welke het slechtste?
- Kunt u een omschrijving geven van deze term? Hoe zou een definitie luiden?

Op deze manier werden de proefpersonen in feite gestuurd om in eigen termen een semantische analyse te maken van de categorieën die ze zojuist zelf gecreëerd hadden. Aldus ontstond voor elke proefpersoon een dataset met in groepjes geordende kaartjes, en voor elke categorie een term plus een karakterisering van de kenmerkende essentie.

2.1 Het materiaal. In twee sessies is gewerkt met verschillend materiaal. De eerste keer is gewerkt met een persuasieve tekst (*Kijk uit voor je huid*), waarop feedback was verkregen met behulp van het softwareprogramma Focus. De tweede keer is gewerkt met een meer informatieve en instructieve tekst (*Je eerste baan*), waarop feedback was verzameld met de plus-en-minmethode. De spreiding over de genres is nagestreefd teneinde te voorkomen dat we uitsluitend met categorieën te maken zouden krijgen die sterk aan een bepaald domein gebonden zouden zijn. De eerste tekst komt uit de gezondheidsvoorlichting en waarschuwt voor huidkanker als gevolg van zonnebrand. De tweede tekst is afkomstig van de Belastingdienst en geeft informatie aan jongeren die voor het eerst een baan krijgen en daardoor in fiscaal vaarwater terechtkomen.

Beide teksten zijn in eerder onderzoek met verschillende methoden geëvalueerd. De belastingtekst is onderworpen aan de plus-en-minmethode in het promotieonderzoek van De Jong (1998). De andere tekst is in eerder Utrechts onderzoek² geëvalueerd met behulp van het softwareprogramma Focus³. Proefpersonen die met Focus werkten, kregen de tekst op scherm aangeboden en konden in een aparte kolom hun commentaar noteren.

De pretest met Focus leverde 148 verschillende problemen op; de plus-en-minmethode leverde 157 problemen op. Omdat het niet realistisch is proefpersonen ongeveer 150 kaartjes voor te leggen, is voor beide teksten een selectie van veertig commentaren gemaakt. Daarbij zijn de volgende criteria gehanteerd.

- Het commentaar moest helder geformuleerd zijn, omdat een onbegrijpelijke reactie nu eenmaal niet goed te categoriseren valt.
- We hebben vermeden dat twee kaartjes naar dezelfde passage en een soortgelijk probleem verwijzen, omdat de toekenning van het ene kaartje aan een categorie dan van invloed zou kunnen zijn op de categorisering van het tweede kaartje; elk probleem moest dus duidelijk herkenbaar en uniek zijn.
- Problemen die door meer dan één proefpersoon genoemd zijn, hadden de voorkeur boven unieke probleemdetecties.
- Er is alleen gebruik gemaakt van commentaar dat betrekking had op tekst; reacties op illustraties en vormgeving zijn niet meegenomen.
- Het probleem moest duidelijk gelokaliseerd kunnen worden in de tekst, zodat de proefpersonen de betreffende passage (via de regelnummering) snel zouden kunnen vinden.

In welke termen denken lezers over tekstproblemen?

- De problemen moesten betrekking hebben op redelijk over de tekst verspreide passages en een grote variëteit aan problemen weerspiegelen.

De laatste twee criteria verdienen enige toelichting. Het vijfde criterium is gekozen om te vermijden dat proefpersonen langdurig op zoek zouden gaan naar de passage waar het commentaar betrekking op heeft. Dit leidt echter tot het gevaar dat op die manier de meer globale reacties buiten de categorisering blijven. En in het laatste criterium schuilt een groot gevaar van circulariteit. Wie alleen problemen met geloofwaardigheid en begrip selecteert, moet niet verbaasd opkijken als veel proefpersonen de kaartjes keurig in die beide categorieën sorteren, ofschoon menigeen een geheel eigen categorisering zal kiezen. Hier is onzes inziens echter sprake van een dilemma: wie geen aandacht schenkt aan spreiding van commentaren, loopt het gevaar van een toevallige eenzijdige selectie; en wie wel vooraf de commentaren wil spreiden over een brede waaier aan probleemttypen loopt het gevaar van een cirkelgang. Wij hebben in dit onderzoek voor beide opties gekozen: de feedback op de gezondheidstekst is mede geselecteerd op het criterium van spreiding volgens de categorieën van Focus; de feedback op de belastingtekst is vooral geselecteerd op grond van het derde criterium – reacties die door minimaal twee respondenten zijn geformuleerd, zonder rekening te houden met spreiding.

2.2 Proefpersonen en afname. Aan elk van beide sessies namen 25 proefpersonen deel. Alle proefpersonen waren hoger opgeleid (vwo-, hbo- of universitair niveau), maar hadden geen speciale expertise op het gebied van tekstschrijven. De leeftijd van de proefpersonen varieerde van 22 tot 66 jaar. In totaal namen 22 mannen en 28 vrouwen aan de sessies deel.

We hebben de procedure bij de afname hierboven al kort geschetst:

- eerst de tekst rustig lezen,
- dan de kaartjes een voor een op stapeltjes leggen,
- vervolgens elk stapeltje omschrijven met behulp van een term, een goed en slecht voorbeeld en een definitie.

Dit gebeurde in een rustige één-op-één situatie bij de proefleiders thuis, bij proefpersonen thuis en soms ook op een rustig moment op het werk van de proefpersoon. Elk kaartje was voorzien van een nummer. De proefleider noteerde de nummers waaruit de stapeltjes uiteindelijk bestonden op een scoreformulier en noteerde daar tevens de antwoorden op de vragen. Gemiddeld duurde de afname ongeveer een uur per proefpersoon.

2.3 De analyse. Van elk genummerd kaartje is in een tabel aangegeven hoe vaak het samengevoegd is met elk van de overige 39 kaartjes. Aldus ontstaat voor elk kaartje een rangorde met de meest frequente combinaties. In tabel 1 ziet u als voorbeeld vier rangordes met de hoogste frequenties voor de kaartjes 9, 24, 25 en 6.

Tabel 1. Rangordes van combinaties van vier kaartjes met tussen haakjes de frequentie waarmee die combinatie is gemaakt door proefpersonen ($N=25$)

9	24	25	6
25 (23)	25 (20)	9 (23)	24 (18)
24 (19)	9 (19)	24 (20)	9 (17)
6 (17)	6 (18)	6 (17)	25 (17)
14 (9)	33 (8)	14 (9)	33 (9)
36 (9)	14 (7)		2 (8)
33 (8)	36 (7)		4 (8)
26 (7)			14 (7)

Kaartje 9 is dus 23 keer gecombineerd met kaartje 25 en is 19 keer met kaartje 24 gecombineerd. In de volgende kolommen zien we dat die beide kaartjes ook weer vaak met elkaar zijn gecombineerd. Aldus ontstaat in deze tabel een beeld waarin de kaartjes 9, 24, 25 en 6 een kerngroep vormen. Het kaartje met nummer 14 wordt daar niet meer toe gerekend, omdat het verval in de frequentie tussen kaartje 6 (minimaal 17 keer gecombineerd in deze groep) en kaartje 14 (maximaal 9 keer) precies op deze grens intreedt. Een kerngroep van combinaties is vastgesteld telkens wanneer zo'n verval zichtbaar werd. Soms was echter geen plotseling verval zichtbaar omdat de frequenties geleidelijk terugliepen; in die gevallen is afhankelijk van de hoogste frequentie een streep getrokken ergens tussen een frequentie van 6 tot 10.

Nadat zulke kerngroepen waren gevormd, is vervolgens nagegaan welke proefpersonen die combinatie in een stapeltje gemaakt hadden. In het bovenstaande voorbeeld zijn dus de gegevens verzameld van alle proefpersonen die de combinatie 9, 24, 25 en 6 gemaakt hebben. Van die proefpersonen zijn alle namen genoteerd die aan de combinatie zijn gegeven en de omschrijvingen die daarvoor zijn geformuleerd. Bovendien is genoteerd welk kaartje als beste vertegenwoordiger van die categorie genoemd is. Aldus ontstond voor elke kerngroep een reeks benamingen en omschrijvingen. De analyse van die gegevens leidde tot de resultaten die in het vervolg van dit artikel gerapporteerd worden.

3. Resultaten

Hoeveel groepjes creëerden de proefpersonen? De gemiddelden verschilden in de beide settings. De kaartjes met reacties op de belastingtekst werden gemiddeld in 5.4 categorieën verdeeld, terwijl de feedback op de gezondheidstekst in gemiddeld 8 stapeltjes werd verdeeld. Het kleinste aantal categorieën was 3 (door vier proefpersonen) en het grootste aantal was 11 (door twee proefpersonen). De meeste proefpersonen maakten tussen de vijf en negen groepjes.

Wat zijn dat voor groepjes? Hieronder vatten we voor de beide settings de resultaten samen per categorie. Niet elke categorie was echter even duidelijk. Sommige probleemcategorieën vertoonden duidelijk minder samenhang dan andere. We beginnen met de vijf categorieën waarover grote duidelijkheid bestond.

In welke termen denken lezers over tekstproblemen?

3.1 Overbodige Informatie. In de teksten over huidkanker en belasting zaten de volgende kaartjes die tot de kerngroep gerekend werden.

Kijk uit voor je huid	Je eerste baan
<i>Leuk om te weten, maar je hebt er niks aan.</i>	<i>De mededeling dat de belasting wordt afgedragen staat er twee keer in.</i>
<i>Het lijkt me dat iedereen dit wel weet.</i>	<i>De informatie over terugvragen staat er twee keer in.</i>
<i>Dit is overbodig als het niet wordt toegelicht.</i>	<i>In dit tekstblok staat twee keer dezelfde zin over aanvragen.</i>
<i>Dit is al eerder gezegd, kan dus weg.</i>	

Er waren in totaal 26 proefpersonen die de kern van deze categorie in een groep bij elkaar brachten. Zij gaven er de volgende namen aan:

- overbodige informatie/overbodigheid
- onnodige informatie
- dubbelop/doublures/dubbelingen
- overlap

De eerste naam krijgt de voorkeur van een meerderheid van achttien proefpersonen. In de definities van de proefpersonen komen twee elementen vaak terug: het herhalen van informatie en het geven van informatie die niet ter zake doet, niet relevant is, niet van toepassing is, of onnodig is. Wij komen tot de volgende omschrijving:

Er is sprake van overbodige informatie als een passage voor de lezer niet ter zake doet of onnodig wordt herhaald.

Overigens waren er negen proefpersonen die bovenstaande kerngroep gevormd hadden, maar toch niet bij de nadere analyse betrokken zijn. Bij nader inzien bleek dat zij de bovenstaande groep hadden samengevoegd met de onderstaande groep, en aldus een ruimere overkoepelende categorie hadden gevormd.

3.2 Ontbrekende informatie. Ook de tweede categorie kwam als een duidelijke groep naar voren. Hiervoor werden kaartjes samengevoegd met de volgende commentaren van lezers.

Kijk uit voor je huid	Je eerste baan
<i>Er staat niet bij dat je je hiertegen ook moet beschermen.</i>	<i>Waar vind je het dichtstbijzijnde belastingkantoor?</i>
<i>Ik mis voorbeelden van huidziektes waar het hier om gaat.</i>	<i>Er moet direct gezegd worden wat je moet doen als je je soft-nummer kwijt bent.</i>
<i>Er staat "hoofdzakelijk" maar waar UV nog meer van komt, staat er niet bij.</i>	<i>Wat is de sanctie als je geen formulier aanvraagt?</i>

Er waren in totaal 31 proefpersonen die de kerngroep van deze categorie gevormd hadden. Zij gaven daarvoor de volgende benamingen:

- ontbrekende informatie/ontbreking
- onvolledigheid
- gemiste informatie/informatieleemte
- vragen, vragen n.a.v. de tekst, onbeantwoorde vragen.

In de omschrijvingen kwamen de woorden “ontbreken” en “onvolledig” erg vaak terug. Bovendien gaven proefpersonen vaak het gevolg van die onvolledigheid aan: lezers begrijpen de tekst niet goed meer, omdat er iets ontbreekt. Daarom luidt voor deze categorie de omschrijving:

Er is sprake van ontbrekende informatie als de lezer meer informatie nodig heeft om de tekst goed te begrijpen.

3.3 Geloofwaardigheid. De derde categorie kwam als een iets minder duidelijke groep uit de data naar voren. We vinden in deze categorie de volgende commentaren als kerngroep.

Kijk uit voor je huid	Je eerste baan
<i>Ik vraag me af of het waar is dat de meeste mensen verstandig zonnen.</i>	<i>Het is onterecht dat je geen studiefinanciering meer krijgt.</i>
<i>Er staat dat afwijken van de instructies de gezondheid nauwelijks kan schaden.</i>	<i>Soms moet je premies betalen, maar kun je er niet van profiteren.</i>
<i>Volgens mij klopt dat niet.</i>	<i>Dit feit klopt niet: het kost nooit geld,</i>
<i>“Die mensen” verwijst naar een groep voor wie de folder toch juist niet bedoeld is?</i>	<i>je verdient altijd wel iets.</i>
	<i>De opmerking dat je ook over vakantiegeld moet betalen is onterecht.</i>

Er waren 19 proefpersonen die de kern van deze categorie in één stapeltje vormden. De terminologie voor deze groep lezerreacties was niet erg eenduidig. We kwamen de volgende benamingen tegen:

- geloofwaardigheid
- oordeel/persoonlijke mening/waardeoordeel/opinie
- waarheid/betrouwbaarheid/correctheid/ervaringen/incorrecte details
- inhoudelijke discussie.

De term “geloofwaardigheid” is slechts vier keer genoemd door alle vijftig proefpersonen, hetgeen betekent dat we hier allermindst met een voor de handliggende benaming te maken hebben. De alternatieve namen werden echter nog minder vaak genoemd. Sommige proefpersonen kozen namen waarmee ook voorgaande categorieën meegenomen werden, zoals bijvoorbeeld “inhoud”. In de omschrijvingen vonden we eveneens een grote verscheidenheid. Maar er waren wel gemeenschappelijke elementen, zoals: een inhoudelijk oordeel over de tekst, de lezer gaat in discussie met standpunten in de tekst, betwijfelt of de tekst inhoudelijk juist is, de lezer gelooft niet wat er in de tekst beweerd wordt, vraagt zich af of de tekst wel klopt. Wij maken een onderscheid tussen twee aspecten in deze omschrijvingen:

In welke termen denken lezers over tekstproblemen?

de twijfel over de waarheid van een feitelijke mededeling en het betwisten van een standpunt in de tekst over een kwestie. Daarmee komen we tot de volgende omschrijving:

Er is sprake van een probleem met de geloofwaardigheid van de tekst als de lezer betwijfelt of de tekst inhoudelijk klopt of als de lezer het niet eens is met standpunten in de tekst.

3.4 Formulering. De voorgaande categorieën hebben alle met de inhoud van de tekst te maken. In de volgende stapeltjes die proefpersonen maakten, gaat het meer om tekstkenmerken. De volgende commentaren hadden met name betrekking op formuleringskwesties.

Kijk uit voor je huid	Je eerste baan
<i>Ik zou eerder kiezen voor “Kijk uit met je huid” dan voor “Kijk uit voor je huid”.</i>	<i>De zin is kinderachtig: natuurlijk krijg je loon als je gaat werken.</i>
<i>“Er zijn natuurlijk zanaanbidders” lijkt mij mooier.</i>	<i>De eerste zin is te ingewikkeld.</i>
<i>Als je kiest voor “slecht”, zou ik “gunstig” door “goed” vervangen.</i>	<i>De mededeling bruto-netto is kinderachtig.</i>

Deze groepering van lezersfeedback is door 36 van de vijftig lezers aangebracht, hetgeen een hoge score is. Zij gaven zeer uiteenlopende benamingen aan deze groep:

- formulering/woordformulering
- stijl/taalgebruik
- schrijfwijze
- woordkeuze
- redactioneel
- tekstoordeel
- doelgroep

De term “doelgroep” wekt misschien enige verbazing, maar de omschrijving hierbij luidde: “de wijze waarop in de richting van de doelgroep wordt geschreven”. De term “formulering” werd het meest gekozen, gevolgd door “stijl”. In de omschrijvingen komt duidelijk naar voren dat het niet om de inhoud van de tekst gaat, maar om nuances in de woordkeus of in de zinsbouw, waarbij overigens woordkeus veel vaker genoemd wordt. Een proefpersoon gaf als omschrijving: “de manier waarop zinnen en woorden zijn geschreven”. Veel proefpersonen gaven in hun formulering een stilistisch verfraaiingsperspectief aan en wezen tegelijkertijd op een aspect van duidelijkheid: “Door de keuze van een ander woord wordt de tekst duidelijker of mooier.” Dat aspect van verduidelijking verklaart waarom sommige kaartjes niet alleen bij deze categorie terecht kwamen, maar ook in een ander groepje functioneerden dat meer betrekking had op het begrip van de tekst. Dit roept de vraag op of verduidelijking een onderdeel moet zijn van de omschrijving. Wij kiezen ervoor die beide aspecten mee te nemen in de omschrijving:

Er is sprake van een formuleringsprobleem als de lezer moeite heeft met bepaalde woorden in de tekst of met de manier waarop de zin is geschreven, bijvoorbeeld omdat iets te ingewikkeld of juist te kinderlijk gezegd wordt of omdat de stijl lelijk gevonden wordt.

3.5. Structuur. In de vijfde categorie brachten de proefpersonen de volgende kaartjes met lezerreacties in één groep bij elkaar.

Kijk uit voor je huid	Je eerste baan
<i>Deze toelichting komt te laat. Ik zou de voor- en nadelen eerder noemen.</i>	<i>De functie van de twee subparagrafen is niet helder.</i>
<i>Ik vind deze informatie over ziektes niet goed geordend. Ik zou een duidelijker onderscheid aanbrengen.</i>	<i>De structuur van het onderdeel is niet helder. Vreemd dat er halverwege opeens subparagrafen zijn.</i>
<i>Het is logischer om eerst UV-A en dan UV-B te noemen</i>	<i>De structuur van het onderdeel is vreemd: begonnen moet worden met de loonbelastingverklaring.</i>

Er zijn 31 proefpersonen die deze groep gevormd hebben. Zij geven daarvoor redelijk eenduidige termen:

- (tekst)structuur
- (tekst)opbouw/ordening/(tekst)indeling
- informatievolgorde

Opvallend is dat in de feedback op de belastingtekst het woord “structuur” zelf al twee keer voorkomt in de probleemomschrijving. Deze term is door onze proefpersonen het meest genoemd en lijkt daarmee de beste kandidaat om de categorie mee aan te duiden. In de omschrijvingen komen telkens de woorden indeling, opbouw, volgorde en structuur weer terug, soms ook vergezeld van vormgeving en lay-out. Wij kiezen op grond van de gepresenteerde formuleringen voor de volgende omschrijving:

Er is sprake van een structuurprobleem wanneer de volgorde van de informatie in de tekst niet goed is, waardoor de lezer geen helder beeld krijgt van de opbouw van de tekst en het overzicht verliest.

Tot zover de vijf categorieën waarvoor de proefpersonen duidelijk herkenbare groepen gevormd hadden. Een zesde categorie staat in evaluatieonderzoek vaak centraal, maar is in onze data minder duidelijk zichtbaar naar voren gekomen. Het gaat hier om een type problemen dat wij doorgaans met de term “begripsproblemen” aanduiden.

3.6 Onduidelijkheid. In de analyse van de data die verkregen zijn met de reacties op de gezondheidstekst is wel een herkenbare categorie zichtbaar geworden die op begripsproblemen betrekking lijkt te hebben, maar in het onderzoek met de belastingtekst is een dergelijke groep niet duidelijk naar voren gekomen. Het gaat om de volgende reacties.

In welke termen denken lezers over tekstproblemen?

Kijk uit voor je huid

Wordt hier met “zonnebrand” zonnebrandcrème bedoeld of verbranden door de zon?

Wat is “buiten bescherming”? Is dat altijd wanneer je buiten bent, ook in de regen?

Betekent dit dat je dus ook onder de zonnebank moet beschermen?

Gaat het hier nu over de zon of over zonnebanken?

Ik snap niet wat hiermee bedoeld wordt.

Deze groepering van reacties is dus slechts door één van beide groepen gemaakt, en wel door 17 van de 25 proefpersonen. Zij gaven aan deze categorie de volgende namen:

- onduidelijkheid
- begrip
- betekenisprobleem
- onduidelijke formulering
- boodschapprobleem

Op grond van deze resultaten lijkt duidelijk sprake te zijn van een categorie die betrekking heeft op problemen met het vormen van een mentale representatie van een deel van de tekst. Maar hoe komt het dan dat bij de feedback op de belastingtekst deze categorie niet naar voren is gekomen? Een eerste verklaring zou kunnen zijn dat in de kaartjes onvoldoende reacties verwoord werden die op deze categorie duiden. Daar is echter geen sprake van. Die kaartjes waren er in voldoende mate, maar ze zijn op een andere manier gegroepeerd. Een deel van die kaartjes kwam terecht in de categorie “ontbrekende informatie”. Sommige proefpersonen gaven die categorie de naam “vragen” of “vraagstellingen”. Het ging om bijvoorbeeld onderstaande reacties.

Je eerste baan

Waar vind je het dichtstbijzijnde belastingkantoor?

Er moet direct gezegd worden wat je moet doen als je je sofí kwijt bent.

Wat is de sanctie als je geen formulier aanvraagt?

Anderen brachten die kaartjes onder in de categorie “formulering”. Een gebrekkige formulering kan immers een begripsprobleem tot gevolg hebben, zoals misschien het geval is bij de reactie *De eerste zin is te ingewikkeld*. Deze tweede verklaring lijkt onaannemelijk omdat de kern van de groep “formulering” in de belastinggroep eerder bestaat uit kritiek op een te kinderachtige stijl. Een nadere analyse van de feedback (op de belastingtekst) die door de proefpersonen is ondergebracht bij “ontbrekende informatie” leidde tot twee sterk verwante clusters van kaartjes die beide betrekking lijken te hebben op dezelfde categorie. Hieronder geven we de kerngroepen van die beide stapeltjes weer.

Je eerste baan

<i>Waar vind je het dichtstbijzijnde belastingkantoor?</i>	<i>De beschrijving van de tariefgroep roept vragen en onzekerheid op.</i>
<i>Hier staat dat je je paspoort mee moet nemen, maar is er geen ander identiteitsbewijs mogelijk?</i>	<i>De concrete gevolgen voor partner/ouder zijn onduidelijk.</i>
<i>Wat is de sanctie als je geen formulier aanvraagt?</i>	<i>De relatie tussen loonheffing en andere premies is niet duidelijk.</i>

Een mogelijke verklaring voor de ordening van deze reacties in twee groepen is dat in de linkerkolom vragen geformuleerd zijn, die vooral klinken als een roep van de lezer om meer informatie, terwijl in de rechterkolom meer de stem van een beschouwende criticus doorklinkt die een diagnose formuleert van een mankement in de tekst. Uiteraard waren er ook proefpersonen (6) die dit onderscheid niet maakten. In de keuzes voor benamingen voor beide subgroepen of voor de overkoepelende groep komt “begrip” niet één keer voor. Wel komen we de term “onduidelijk(heid)” negen keer tegen. Wanneer we de kaartjes op een rij zetten die door deze negen proefpersonen als beste vertegenwoordiger van de groep “onduidelijkheid” worden genoemd, ontstaat het volgende overzicht.

Je eerste baan

<i>In de folder staat niet wat er wordt gedaan met de belasting.</i>
<i>De beschrijving van de tariefgroep roept vragen en onzekerheid op. De zin over reiskosten is niet duidelijk.</i>
<i>Er moet direct gezegd worden wat te doen als je je soft kwijt bent.</i>
<i>De relatie tussen premies en loonheffing is niet helder.</i>
<i>De structuur van het onderdeel is vreemd: begonnen moet worden met de loonbelastingverklaring.</i>
<i>Vreemd dat er halverwege opeens subparagrafen zijn.</i>
<i>De functie van de twee subparagrafen is niet helder.</i>

Dit overzicht wijst niet onmiddellijk in de richting van een verborgen groep begripsproblemen, alhoewel die wel vertegenwoordigd zijn. De laatste drie lezerreacties zijn echter van een andere orde, en de eerste reactie heeft toch meer betrekking op een wens om meer informatie dan op een echt begripsprobleem.

In de omschrijvingen die proefpersonen geven voor de term “onduidelijkheid” komen we de volgende definities tegen:

- je weet niet wat wordt bedoeld met iets dat in de tekst wordt gezegd;
- de lezer snapt niet wat met de tekst wordt bedoeld;
- wat met de tekst bedoeld wordt, komt niet duidelijk over op de lezer;
- als je een tekst leest, dan begrijp je die niet zomaar, want de tekst roept vragen op over de bedoeling of betekenis van woorden en zinnen.

En voor “begrip” komen we (in de data van de gezondheidstekst) de volgende omschrijvingen tegen:

In welke termen denken lezers over tekstproblemen?

- onbegrip voor tekstinhoud ontstaat, de tekst roept vragen op;
- de lezer begrijpt niet wat er in de tekst staat;
- het is niet duidelijk wat ergens mee bedoeld wordt, de lezer snapt het niet;
- een term, zin of uitdrukking is niet duidelijk, de lezer begrijpt niet wat de schrijver bedoelt.

Uit deze omschrijvingen komt een duidelijk onderscheid naar voren met de categorie “ontbrekende informatie”. Het gaat hier immers om informatie die wel in de tekst te vinden is, maar die niet goed begrepen wordt. In de laatst geciteerde omschrijving wordt een relatie zichtbaar met “formulering”, omdat uiteraard een gebrekkige formulering daar de oorzaak van kan zijn.

Wij concluderen dat er toch wel reden is om een categorie begripsproblemen te onderscheiden, die we echter op grond van de data liever met de term “onduidelijkheid” aanduiden. Deze term is immers door respectievelijk zeven en negen proefpersonen naar voren gebracht en in de eerste groep is er ook daadwerkelijk een als zodanig herkenbare categorie gevormd van lezerreacties. De resultaten van de tweede groep leiden echter tot voorzichtigheid, omdat de term mogelijk door proefpersonen als overlappend wordt ervaren met de categorie “ontbrekende informatie”. In de omschrijving van de term zou dit onderscheid dan ook duidelijk geformuleerd moeten worden, bijvoorbeeld op de volgende manier.

Er is sprake van onduidelijkheid wanneer de tekst ergens wel informatie over geeft, maar als die passage vragen oproept, omdat de lezer de inhoud ervan niet goed begrijpt.

3.7 Leestekens en spelling. In de gezondheidstekst waren in het eerdere evaluatieonderzoek bewust een aantal fouten aangebracht in de spelling en interpunctie. Dit was gedaan met het doel om na te gaan of hoger en lager opgeleide proefpersonen hier verschillend op zouden reageren (zie Lentz en De Jong, 2000). Op een aantal kaartjes waren lezerreacties zichtbaar die hierop betrekking hadden. Deze kaartjes zijn door sommige proefpersonen in twee groepen ondergebracht en door andere in één gezamenlijke groep. Het gaat om de volgende reacties.

Interpunctie	Spelling / Correctheid
<i>Hier hoort ook een puntkomma. Deze komma is overbodig.</i>	<i>Volgens mij moet er spelbreker staan in plaats van speldbreker.</i>
<i>De dubbele aanhalingstekens maken het erg onduidelijk.</i>	<i>Er moet verhoogd staan in plaats van verhoogt. Daarmee moet Daarvan zijn.</i>
<i>Ik zou dat kan niet gezond zijn tussen aanhalingstekens zetten.</i>	

Het is de vraag of dit een zinvolle categorie is, omdat we in evaluatieonderzoek immers geen lezers inschakelen teneinde na te gaan of de tekst goed gespeld is. In principe leggen we een incorrect gespelde tekst niet aan lezers voor. Toch weten we uit ervaring dat lezers vaak de behoefte hebben commentaar te leveren op komma's en punten, en op de spelling

van woorden. Ook onder leken bevindt zich menige meester Pennewip, die het niet kan laten zijn deskundigheid op dit punt tentoon te stellen.

Hoe noemen de proefpersonen deze categorie? Wij kregen de volgende benamingen:

- leestekens / interpunctie
- spelfout / spelling
- taalfout / taal- en spelfout.

In de omschrijvingen werden deze termen vaak herhaald, maar er was ook vaak sprake van “punten en komma’s” en één creatieve proefpersoon schonk ons het prachtige “taalelastiekjes”. Wie in evaluatieonderzoek deze categorie aan proeflezers aan wil bieden, zou volgens ons het best interpunctie en spelling samen kunnen nemen onder de naam *Leestekens en Spelling*. Daarbij past de volgende omschrijving.

Er is sprake van een probleem met de leestekens als die op een verkeerde plaats in de tekst staan of als ze ontbreken.

En er is sprake van een probleem met de spelling als de woorden niet gespeld zijn volgens de regels van het Nederlands.

4. Conclusie en discussie

Leken die niet beroepsmatig met taal en tekst bezig zijn, kunnen in evaluatieonderzoek geconfronteerd worden met de volgende categorieën en de daarbij passende omschrijvingen. Er is sprake van:

<i>overbodige informatie</i>	als informatie voor u niet ter zake doet of onnodig wordt herhaald.
<i>ontbrekende informatie</i>	als u meer informatie nodig heeft om de tekst goed te begrijpen.
<i>onduidelijkheid</i>	wanneer de tekst ergens wel informatie over geeft, maar als die passage vragen oproept, omdat u de inhoud ervan niet goed begrijpt.
<i>formuleringsprobleem</i>	als u moeite heeft met bepaalde woorden in de tekst of met de manier waarop de zin is geschreven, bijvoorbeeld omdat iets te ingewikkeld of juist te kinderlijk gezegd wordt of omdat u de stijl lelijk vindt.
<i>structuurprobleem</i>	wanneer de volgorde van de informatie in de tekst niet goed is, waardoor u geen helder beeld krijgt van de opbouw van de tekst en het overzicht verliest.
<i>geloofwaardigheidsprobleem</i>	als u betwijfelt of de tekst inhoudelijk klopt of als u het niet eens bent met bepaalde standpunten in de tekst.
<i>foutieve leestekens spellingsfouten</i>	als die op een verkeerde plaats in de tekst staan of ontbreken. als de woorden niet gespeld zijn volgens de regels van het Nederlands.

In welke termen denken lezers over tekstproblemen?

Twijfel is mogelijk over de groep “onduidelijkheid”, omdat die in één van de beide sessies niet als zodanig duidelijk naar voren kwam, hetgeen verklaard kan worden door de overlap met “ontbrekende informatie” en in mindere mate met de categorie “formuleringsprobleem”. Bovendien is twijfel mogelijk over de zinvolheid van een categorie die op spelling en leestekens is gericht, omdat dat een type reacties is waar we in evaluatieonderzoek niet direct behoefte aan hebben.

Uit bovenstaand overzicht moet niet geconcludeerd worden dat elke proefpersoon deze categorieën spontaan creëerde tijdens het sorteren van kaartjes met lezerreacties. Het overzicht is tot stand gekomen door een analyse van de meest sterke verbanden tussen de verschillende kaartjes. Dat er ook volkomen andersoortige categorieën zijn genoemd, blijkt uit onderstaand overzicht.

• wijzigingen	• persoonlijk	• opmaak
• algemene informatie	• arbeidsproces	• kader
• verdieping	• belastingontduiking	• omschrijvingsfouten
• eerlijkheid	• aanspreekniveau	• tekstueel
• redactioneel	• persoonlijke gevolgen	• inhoudelijk
• belangrijke opmerking	• flauwekul	• inhoudelijke details
• onbelangrijke opmerking		

Sommige proefpersonen sorteerden volgens een dimensie van relevantie van het commentaar met als polen belangrijk – onbelangrijk. Anderen sorteerden op een dimensie als inhoud – tekst – vorm. En nog weer anderen kozen inhoudelijke rubrieken die uit het thema van de tekst voortkwamen; zij bedachten groepen met titels als *arbeidsproces* en *persoonlijke gevolgen*. Misschien had voor die proefpersonen de instructie duidelijker geformuleerd moeten worden.

Het onderzoek levert een grootste gemene deler op voor de categorisering van problemen. Toch is het duidelijk dat ook met deze categorieën lezers soms zullen twijfelen in welke rubriek ze een probleem moeten indelen. Bijvoorbeeld omdat zij een ander indelingssysteem hanteren, of omdat zij voor andere labels zouden kiezen. Het onderzoek bevestigt echter in grote lijnen de bruikbaarheid van de lijst met probleemcategorieën, zoals die aan het begin van dit artikel gepresenteerd is, en waar we in Focus al jaren mee werken. Het helpt ons daarnaast om de probleemomschrijvingen en -afbakening meer op de kennis van leken af te stemmen en daarmee de ecologische validiteit van ons instrumentarium te vergroten.

Meer onderzoek is nodig: (1) naar de mate waarin mensen met de hier gepresenteerde categorieën uit de voeten kunnen en de problemen die ze daarbij ervaren; (2) naar de precieze afbakening tussen probleemcategorieën; (3) naar de mate waarin de resultaten afhankelijk zijn van het gekozen teksttype (brochure), omdat denkbaar is dat bij een handleiding of een website een gedetailleerder arsenaal aan probleemttypen nodig is, dat specifiek gericht is op problemen met de toepassing van die informatie; (4) naar de mate waarin de resultaten afhankelijk zijn van het opleidingsniveau van de lezers, dat in dit onderzoek relatief hoog was.

Noten

- 1 Het betreft onderzoek dat aan de Universiteit Utrecht is uitgevoerd door Viola Teepe en Aart van Kapel in het kader van een scriptieproject van Lentz en De Jong.
- 2 Meer over dit onderzoek is te vinden in Lentz en De Jong (2000), waarin de verschillen worden besproken tussen de feedback op deze tekst van hoger en lager opgeleide proefpersonen.
- 3 Over Focus als instrument voor tekstevaluatie is meer te vinden in De Jong en Lentz (2001), waar o.a. een vergelijking wordt gemaakt met feedback die verkregen is met de plus-en-minmethode.

Bibliografie

- Cicourel, A.V. (1982).** Interviews, surveys, and the problem of ecological validity. *The American Sociologist*, 17 (2), 11-20.
- Jong, M. de (1998).** *Reader feedback in text design. Validity of the plus-minus method for the pretesting of public information brochures.* Amsterdam: Rodopi.
- Jong, M. de & L. Lentz (2001).** Focus: Design and evaluation of a software tool for collecting reader feedback. *Technical Communication Quarterly*, 10, 387-401.
- 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.
- Kapel, A. van (2000).** *Taalgebruikers, teksten en probleemcategorieën. Een onderzoek naar het lexicon van taalgebruikers voor categorieën van lezersproblemen aan de hand van bestaande feedback op een tekst.* Doctoraalscriptie Universiteit Utrecht, afdeling Taalbeheersing.
- Lentz, L. & M. de Jong (2000).** Hoger en lager opgeleide proefpersonen en hun feedback op een voorlichtingstekst. In: R. Neutelings, N. Ummelen & A. Maes (red.) *Over de grenzen van de taalbeheersing* (pp. 317-326). Den Haag: SDU.
- Maes, A., N. Ummelen & H. Hoeken (1996).** *Instructieve teksten. Analyse, ontwerp en evaluatie.* Bussum: Coutinho.
- Teepe, V. (1999).** *Wat is dat voor probleem? Een onderzoek naar de categorieën waarin tekstproblemen ingedeeld worden en de termen die aan die categorieën worden gegeven.* Doctoraalscriptie Universiteit Utrecht, afdeling Taalbeheersing.