

AUDREY ROUSSE-MALPAT, geboren Francaise, heeft aan de universiteit van Bordeaux een bachelor in Langues Etrangères Appliquées (Spaans en Engels) behaald. Daarna is ze naar Groningen gegaan om zich verder te verdiepen in de taalkunde. Inmiddels is ze in het bezit van een bachelor Romaanse Taal en Cultuur en een master Toegepaste Taalwetenschap. Ze is bijna klaar met de Research Master Linguistics en begint in september 2012 met haar PhD-project over de effecten van AIM. Daarnaast beschikt ze over onderwijskundige ervaring als lerares Frans op het Zernike College te Haren en is ze actief als docent aan de afdeling Frans van de Rijksuniversiteit Groningen. E-mail: <A.Rousse-Malpat@student.rug.nl>.

MARJOLIJN VERSPOOR is werkzaam als universitair hoofddocent toegepaste taalwetenschap aan de Rijksuniversiteit Groningen en research associate aan de University of the Free State (Bloemfontein, Zuid Afrika). Zij doet en begeleidt onderzoek naar tweedetaalverwerving. Naast onderzoek naar het effect van tweetalig onderwijs richt zij zich op de effecten van onderwijsmethoden, vooral de effecten van (extra) input in plaats van expliciete grammaticalessen. E-mail: <M.H.Verspoor@rug.nl>.

SASKIA VISSER is taalkundige en coördinator van de wetenschapswinkel Taal, Cultuur en Communicatie van de Rijksuniversiteit Groningen. E-mail: <s.j.visser@rug.nl>.

Het beoordelen van leerlingteksten met behulp van een schaalmodel

EVELINE POLLMANN, JOANNEKE PRENGER & KEES DE GLOPPER

Het onderzoeksproject 'Schrijven om te leren bij Nederlands, Geschiedenis en Science' (SotL-NGS) heeft tot doel experimentele evidentie te verzamelen voor de effectiviteit van het verbinden van het schrijfonderwijs bij Nederlands en het vakonderwijs bij geschiedenis en science in het eerste leerjaar van het vmbo, havo en vwo. In dit artikel zoomen we in op het beoordelen van de teksten uit dit onderzoeksproject, waarbij we vooral de ontwikkeling en het gebruik van een beoordelingsschaal zullen toelichten.

Schrijven is een belangrijke onderwijsdoelstelling, niet alleen bij Nederlands, maar ook bij andere vakken. Leerlingen moeten leren om informatieve, beschouwende en betogende teksten te schrijven. Dat stelt hoge eisen aan leerlingen, want schrijven is bij uitstek een complexe vaardigheid. Het is dan ook niet verwonderlijk dat het bereikte niveau in schrijfvaardigheid achterblijft bij de verwachtingen.

In de praktijk van het schrijfonderwijs in Nederland is er het probleem van de oefentijd en de transfer. Als het schrijven alleen bij Nederlands geïnstrueerd en geoefend wordt, worden de grenzen van de beschikbare

onderwijstijd snel bereikt. Bovendien wordt de vaardigheid niet wendbaar, omdat leerlingen alleen bij Nederlands en niet binnen andere domeinen en vakken en over andere onderwerpen leren schrijven. Voor zover er al sprake is van toepassingen van schrijven bij andere vakken, ontbreekt het meestal aan verbanden wat betreft instructie, oefening, feedback en beoordeling. Door een samenhangende didactiek voor het schrijfonderwijs bij Nederlands te verbinden met het schrijven bij andere vakken ontstaat er meer gelegenheid en tijd voor het leren van deze complexe vaardigheid. De vaardigheid van leerlingen wordt bovendien breder inzetbaar.

Schrijven om te leren

Leerlingen schrijven niet alleen om anderen te informeren, tot denken aan te zetten of te overtuigen. Schrijven heeft ook een conceptualiserende functie, op school en daarbuiten. Door te schrijven kunnen leerlingen hun kennis en begrip van zaken ontwikkelen, in allerlei verschillende vakken en leerdomeinen. De leereffecten van *writing-to-learn* zijn inmiddels redelijk goed gedocumenteerd, door verschil-

lende reviews (o.a. Klein, 1999) en door een meta-analyse (Bangert-Drowns, Hurley & Wilkinson, 2004). Deze inzichten over de rol van schrijven bij andere vakken sluiten nauw aan bij de in Nederland gangbare inzichten omtrent de waarde van taalgericht vakonderwijs en taalbeleid (Hajer & Meestringa, 2004). Daarin is eveneens het uitgangspunt dat de mondelinge en schriftelijke taalvaardigheid van leerlingen niet alleen bij Nederlands aan bod moet komen, maar ook bij andere vakken. Dat is niet alleen goed voor de taalontwikkeling; aandacht voor taal komt op zijn beurt het leren in de zaakvakken ten goede.

Vanuit die gedachte is in 2009 het Expertisecentrum taal, onderwijs en communicatie (Etoc) gestart met het project 'Schrijven om te leren bij Nederlands, Geschiedenis en Science' (verder te noemen SotL-NGS) op het Werkman College in Groningen. In dit project wordt naar experimentele evidentie gezocht voor een groei in schrijfvaardigheid.

Binnen dit project worden veel teksten geschreven door leerlingen. Om deze teksten te beoordelen is in het onderzoek gebruik gemaakt van een zogenoemde beoordelingschaal. In dit artikel zullen we uitleggen waarom er is gekozen voor een beoordelingschaal, en hoe deze schaal voor het beoordelen van teksten tot stand is gekomen. We doen dit vanuit de aanname dat het gebruiken van beoordelingschalen binnen het schrijfonderwijs Nederlands en voor het beoordelen van schrijfproducten bij andere vakken van grote waarde kan zijn.

Waarom een beoordelingschaal?

Om deze vraag te kunnen beantwoorden is het van belang verschillende beoordelingsmanieren die er bestaan kort toe te lichten. Het verzamelen van nauwkeurige oordelen over door leerlingen geschreven teksten is namelijk – zeker als het gaat om grote aan-

tallen teksten zoals in dit onderzoek – erg moeilijk (Blok & Hoeksma, 1984). Toch is het, vooral als het gaat om het vaststellen van schrijfontwikkeling in het kader van een experiment, erg belangrijk om een hoge mate van nauwkeurigheid in tekstoordelen te verkrijgen. Vanuit de praktijk van het onderzoek en het onderwijs zijn er verschillende manieren bekend waarop de kwaliteit van schrijfproducten beoordeeld kan worden (Wesdorp, 1981, Schoonen & De Gloppe, 1992).

Een eerste methode die gebruikt kan worden om teksten te beoordelen is het impressie-oordeel. Hierbij wordt de kwaliteit van de teksten vastgesteld door een globale impressie. De tekst wordt als geheel beschouwd: de verschillende aspecten van de tekst worden niet afzonderlijk onder de loep genomen. Deze methode is snel en eenvoudig te hanteren. Het blijft echter ondoorzichtig hoe de verschillende aspecten gewogen worden (Schoonen & De Gloppe, 1992).

Een tweede manier van tekstbeoordeling is de analytische methode. Hierbij geeft de beoordelaar een beoordeling aan het aspect dat hij wil waarderen: stijl, inhoud, genre, opbouw, enzovoort. Verschillende deelaspecten kunnen hierbij beoordeeld worden, waarbij het ene aspect zwaarder kan wegen dan het andere. De vraag die zich bij deze methode voordoet is welke aspecten je wilt beoordelen en of je voor de aspecten ook verschillende cijfers geeft (Schoonen & De Gloppe, 1992).

Een derde methode is de Primary Trait scoring. Hierbij richt de beoordelaar zich op het hoofdkenmerk van de tekst. In dit geval vindt er een soort globale beoordeling plaats op dit hoofdkenmerk. Wat dit hoofdkenmerk is, hangt af van het doel van de tekst. Deze methode kan snel toegepast worden, maar is door de beperking minder informatief (Schoonen & De Gloppe 1992, 190).

De beoordelingswijze die wij in dit onderzoek hebben gebruikt, is gebaseerd op een vierde methode waarbij gebruik gemaakt

wordt van voorbeeldopstellen. Een beoordelaar beoordeelt de teksten aan de hand van een voorbeeldopstel. Dit opstel dient als referentiepunt: het opstel kan bijvoorbeeld de grens van voldoende en onvoldoende illustreren. Er kan ook gekozen worden voor een aantal voorbeeldopstellen die ieder als voorbeeld dienen voor een bepaald niveau. In dat geval spreekt men van een beoordeling met behulp van een beoordelingsschaal. Een beoordelingsschaal bestaat uit een reeks in kwaliteit oplopende voorbeeldopstellen, die van een oordeel zijn voorzien. De beoordelaar vergelijkt de te beoordelen opstellen met de opstellen die in de schaal aanwezig zijn, om zo tot een oordeel te komen (Blok & Hoeksma, 1984).

Welke beoordelingsproblemen voorkom je als je een schaal gebruikt?

Verschillende bronnen beamen de effectiviteit van het werken met een beoordelingschaal (De Milliano e.a. 2012, Schoonen & De Gloppe 1992, Blok & Hoeksma 1984). Daarnaast voorkomt het schaalmodel als beoordelingsmethode verschillende beoordelingsproblemen. Ten eerste is er geen sprake van het sequentie-effect. Bij een sequentie-effect is de volgorde van de opstelbeoordeling van invloed op het oordeel. Er worden meerdere teksten tijdens één beoordelingssessie beoordeeld; zodra een docent een groot aantal slechte schrijfproducten heeft nagekeken zal hij een volgend schrijfproduct dat redelijk gemaakt is hoger waarderen. Dit effect kan zich ook andersom voordoen. Met een beoordelingsschaal kan dit voorkomen worden, doordat er voor elke categorie een voorbeeldtekst is.

Ten tweede wordt de normverschuiving voorkomen. Bij een normverschuiving past de beoordelaar zich – onbewust - aan het niveau van de teksten aan. Op die manier kan er

een scheve beoordeling ontstaan. Bovendien wordt het significief effect voorkomen. Met een significief effect worden de verschillen in de beoordeling tussen docenten aangeduid: zo kan een docent alleen op inhoud en opbouw letten, en een bepaald opstel daardoor hoger waarderen dan een andere docent (Erkens 2011, p.122). Wolowitsj-Schelvis (1975) (in Wesdorp 1981, p.27) liet in haar onderzoek zien dat docenten verschillende kenmerken van schrijven belangrijk vinden, waardoor er verschillende oordelen over dezelfde tekst ontstaan. Door een beoordelingsschaal gezamenlijk op te stellen kan dit probleem voorkomen worden.

Bovengenoemde punten waren voor ons doorslaggevend in onze keuze voor het gebruiken van een schaalmodel voor het beoordelen van de teksten binnen het project SotL-NGS. Het onderzoek kent een grote hoeveelheid teksten over hetzelfde onderwerp; dit rechtvaardigt de inspanning van het ontwikkelen van een beoordelingsschaal. Bovendien kan een beoordelingsschaal goed gebruikt worden wanneer er sprake is van meerdere beoordelaars: door de verschillende tekstoordelen van de beoordelaars na afloop te vergelijken, kunnen we ook een oordeel vellen over de betrouwbaarheid van de oordelen.

Onderzoeksopzet

In het onderzoek leren de leerlingen per schooljaar twee soorten teksten te schrijven. Er is gekozen voor tekstgenres die functioneel zouden kunnen zijn binnen zowel het schoolvak science als het schoolvak geschiedenis. Dit omdat leerlingen in beide vakken het geleerde bij Nederlands zouden moeten kunnen toepassen.

In fase I (van de herfstvakantie tot de kerstvakantie) leren leerlingen bij Nederlands een tekst te schrijven waarin ze een beschrijving

moeten geven van de overeenkomsten en verschillen tussen twee objecten of verschijnselen. In fase II (van de voorjaarsvakantie tot de meivakantie) leren ze een tekst te schrijven waarin ze de opeenvolgende gebeurtenissen bij een evenement of historische gebeurtenis moeten uitleggen. Daarbij is gekozen voor twee manieren waarop leerlingen een uitleg kunnen structureren: een uitleg waarbij ze kenmerken van een bepaald voorwerp of verschijnsel benoemen, of een uitleg waarbij ze stap voor stap beschrijven in welke volgorde een bepaald fenomeen of proces verloopt. Voor beide tekstsoorten is een module ontwikkeld van vijf of zes instructielessen Nederlands waarin de leerlingen de kenmerken van dit soort teksten leren. Hierbij ligt de focus vooral op de structuur van dit soort teksten, omdat uit het pilotjaar is gebleken dat de leerlingen op dit gebied beginnende schrijvers zijn.

Van de vier eerste klassen die aan het onderzoek meededen werd er per niveau (te weten vmbo-tl/havo en havo/vwo) steeds één aan de experimentele conditie en één aan de controleconditie toegewezen. Per onderzoeksfase (I en II) volgden alle vier de klassen vijf à zes lessen Nederlands waarbij het genre werd geïnstrueerd en geoefend. De experimentele interventie vond plaats bij de vakken. In de experimentele klassen werd het tekstgenre in verschillende schrijfp opdrachten toegepast: driemaal bij geschiedenis en driemaal bij science. De leerlingen in de controleklassen kregen bij Nederlands dezelfde lessen, maar bij geschiedenis en science verwerkingsopdrachten die niet om het schrijven van teksten vroegen, maar om het beantwoorden van vragen.

Van de onderzoekspopulatie werden achtergrondgegevens verzameld: de vaardigheid in tekstbegrip, de grootte van de woordenschat en de schrijffattitude/opvattingen over schrijven. In de voormeting van het experiment werd de schrijffvaardigheid Nederlands

gemeten door middel van een schrijftaak die leerlingen zonder enige instructie moesten uitvoeren. Ook werd hun algemene niveau van kennis en begrip in geschiedenis en science gemeten. In de nameting werd de schrijffvaardigheid Nederlands opnieuw gemeten. Dit gebeurde door de leerlingen dezelfde schrijftaak af te nemen als in de voormeting. De nameting bij geschiedenis en science had betrekking op kennis en begrip van de stof waarover door leerlingen geschreven is. Aan het eind van het schooljaar is de schrijffattitude opnieuw gemeten.

Ontwikkeling van de beoordelings-schaal

In het onderzoek schrijven de leerlingen in fase I bij de voor- en nameting een verslag en in fase II een verklaring. Voor beide tekstsoorten is een beoordelingsschaal met vijf voorbeeldteksten (ankerteksten) ontwikkeld. In dit artikel beschrijven we het ontwikkelproces van de schaal voor fase I. In totaal waren zo'n 119 leerlingen bij het onderzoek betrokken, 90 daarvan schreven zowel tijdens de voormeting als tijdens de nameting van fase I een tekst.

De schrijftaak in fase I hield in dat leerlingen een tekst moesten schrijven (aan de hand van steekwoorden) waarin ze de verschillen en overeenkomsten tussen de basisschool en middelbare school uitlegden aan leerlingen in groep 8. De 208 verzamelde teksten over de verschillen tussen de brugklas en de basisschool zijn nadat ze verzameld waren, geanonimiseerd en van een code voorzien. Zo was het voor de beoordelaars niet duidelijk of de teksten afkomstig zijn uit de voor- of nameting, of ze geschreven zijn door een vmbo-leerling of een havo-leerling, of door een jongen of een meisje.

Voor de selectie van de vijf voorbeeldteksten die uiteindelijk in de beoordelingsschaal

gebruikt worden, is eerst gebruik gemaakt van de impressie-analyse. De schaalontwikkelaar heeft 100 teksten op basis van een impressie-oordeel gecategoriseerd op een vijfpuntsschaal: zeer zwak, zwak, gemiddeld, sterk en zeer sterk. Vervolgens zijn uit elke categorie twee teksten geselecteerd die representatief waren voor de desbetreffende categorie. De tien op deze manier geselecteerde anonieme teksten zijn vervolgens door zeven deskundigen (taalwetenschappers) opnieuw geschaald op dezelfde vijfpuntsschaal aan de hand van een korte instructie, waarbij hen gevraagd werd vooral te letten op de structuurkenmerken. Dit om te voorkomen dat de beoordelaars zouden letten op spelling en formulering in plaats van op de vaardigheid van leerlingen om een goed gestructureerde tekst te schrijven. Tijdens de instructie bij Nederlands en tijdens de schrijftaken bij de vakken lag de focus immers op het aanleren en toepassen van de structuur die paste bij het gekozen tekstgenre, en om te zien of leerlingen daadwerkelijk gegroeid waren in hun vaardigheid van het structureren van hun teksten, wilden we dat de beoordelaars voornamelijk daar op zouden letten. Op basis van consensus over de indeling van de tien teksten door de deskundigen, zijn vervolgens vijf anker teksten geselecteerd die de basis vormden voor de beoordelingsschaal.

Alle schrijfproducten zijn vervolgens drie keer vermenigvuldigd, zodat elke schrijfproduct door drie beoordelaars beoordeeld kon worden. Deze beoordelaars waren derdejaarsstudenten taalbeheersing van de

Rijksuniversiteit van Groningen. De beoordelaars hebben een instructie gehad over het beoordelen van teksten aan de hand van een schaal. Ze kregen toelichting op de teksten die als anker tekst in de schaal waren opgenomen. In de volgende paragraaf worden de anker teksten van de schaal beschreven.

Beschrijving van de beoordelings-schaal

De voor dit project ontwikkelde beoordelingsschaal bestaat uit vijf voorbeeldteksten (ankerteksten). Deze vijf teksten vertegenwoordigen de categorieën 'zeer zwak', 'zwak', 'gemiddeld', 'sterk' en 'zeer sterk'. Om ervoor te zorgen dat de beoordelaars een genuanceerd oordeel kunnen geven, hebben we ervoor gekozen deze categorieën om te zetten in cijfers. De beoordelaars kregen de instructie om elke tekst op basis van vergelijking met de voorbeeldteksten van een score tussen de 50 (lager dan de zwakste anker tekst) en 150 (hoger dan de beste anker tekst) te voorzien (figuur 1).

De rangorde van de vijf teksten is bepaald door tekstopbouw (inleiding, kern en slot), alineaopbouw, en het gebruik van structuurkenmerken zoals signaalwoorden. We hebben hiervoor gekozen omdat deze aspecten ook als leerstof werden aangeboden in de les-serie die de leerlingen tussen de voor- en nameting gekregen hebben. Hieronder een korte toelichting op de keuze van de anker teksten van het schaalmodel.

score	50	70	85	100	115	130	150
categorie		Zeer zwak	Zwak	Gemiddeld	Sterk	Zeer sterk	

Figuur 1. Beoordelingsschaal

Tekst 1

De verschillen tussen de middelbare en de basisschool zijn: op de middelbare school heb je veel huiswerk en je gaat elk uur naar een andere klas en op de basisschool heb je haast geen huiswerk en blijf je de hele dag in een klas. Op de middelbare school heb je veel kortere pauzes dan op de basisschool. Op de middelbare heb je vakken Zoals Wiskunde, science, drama en beeldende vorming maar die vakken heb je op de basisschool niet. De overeenkomsten tussen de middelbare school en de basisschool zijn: je hebt een vaste klas en je hebt een ochtend pauze en een middagpauze. Sommige vakken tussen de middelbare school zijn wel hetzelfde zoals aardrijkskunde geschiedenis, en Nederlands (taal, spelling). Nog een overeenkomst is dat de basisschool en de middelbare school allebij gym hebben en dus ook allebij ook een gym zaal hebben.

Ankertekst 1: 70 punten (zeer zwak)

Wat deze tekst zeer zwak maakt, is dat de tekstopbouw en alineaverdeling ontbreken. Daarnaast ontbreken de titel, inleiding en het slot. Zo start de schrijver direct met het opnoemen van verschillen. In de tekst zijn weinig structurelementen gebruikt; signaalwoorden ontbreken op verschillende plaatsen in de tekst.

Tekst 2

Op de middelbare school en basisschool zijn overeenkomsten maar ook verschillen. De overeenkomsten zijn bijv:

- Je gaat elke ochtend op de fiets of lopent ofzo.
- Onder schooltijd heb je ook pauzes.
- Je kunt ook wel eens een toets hebben.
- Je hebt verschillende vakken.
- Je hebt natuurlijk een directeur en een congerge enz.

- Je kunt lange en korte dagen hebben. Maar je hebt natuurlijk ook verschillen. De verschillen zijn bijv.
- Op de middelbare school heb je huiswerk en op de basisschool niet.
- Op de middelbare school heb je een lesrooster en op de basisschool niet.
- En op de middelbare school heb je vaste docenten en op de basisschool heb je één of twee vaste leraar(leraren).
- Op de middelbare school heb je een kantine en op de basisschool niet.
- En op de middelbare school heb je huiswerk en dat heb je ook niet op de basisschool
- En op de middelbare school heb je iets meer vakken dan op de basisschool.

En er is nog een overeenkomst:

- Als school is afgelopen is toch iedereen wel weer blij om naar huis te gaan.

Ankertekst 2: 85 punten (zwak)

Op het eerste gezicht lijkt deze zwakke tekst een opsomming, maar wie goed kijkt, ziet dat deze tekst zich wel degelijk onderscheidt van de zeer zwakke tekst. Zo zijn er in deze tekst verschillende signaalwoorden als 'natuurlijk' en 'nog' te onderscheiden. De schrijver geeft de lezer houvast door deze woorden te gebruiken. Daarnaast heeft de schrijver een korte inleidende zin aan deze tekst toegevoegd. Hij vertelt de lezer dat deze tekst gaat over overeenkomsten en verschillen tussen de basisschool en middelbare school. Een slot en titel ontbreken bij deze tekst echter nog wel. Deze tekst scoort daarom lager dan de gemiddelde tekst.

Tekst 3

Nieuwe start!

Veel dingen zijn anders als je van de basisschool naar de middelbare school gaat. Het is allemaal nieuw!

Het raarste vond ik het 'huiswerk' op mijn vorige school had ik nooit huiswerk. Dus ik moest daar wel even aan wennen.

Op de middelbare school hebben we ook veel meer vakken zoals: Frans, Wiskunde en beeldende vorming op onze school hebben we nog drie andere vakken dans drama en muziek.

Op mijn vorige school werkten we niet echt met een lesrooster en er vielen ook geen lessen ofzo uit want we hadden maar één docent.

We hebben op deze school ook voor elk vak een andere leraar. Bij sommige leraren gaat het goed bij sommigen niet.

Pauze hebben we meestal 3 keer per dag die duren twintig minuten.

Mijn klas is super druk ze kunnen bijna nooit helemaal stil zijn alleen maar bij nederlands.

Dit was mijn verhaal over de middelbare school. Het is best wel anders dan de basisschool, maar je went heel snel.

Ankertekst 3: 100 punten (gemiddeld)

De gemiddelde tekst oogt meteen al heel anders. Zo staat er een titel boven de tekst en heeft de schrijver een alinea-indeling aangegeven. Echter, de suggestie van alinea's wordt wel gewekt, maar het zijn nog geen echte alinea's. Vaak bestaan ze namelijk maar uit één zin. De tekst bevat wel een inleiding en een slot. Dit maakt de tekst een gemiddeld schrijfsproduct.

Tekst 4

Naar de Middelbare school

Hallo kinderen van groep 8. Jullie denken vast dat het eerst heel eng is op de middelbare of zijn

er wel leuke docenten. Maar dat valt allemaal wel mee.

Op de basisschool krijg je altijd van 1 docent de hele dag les. Dat is niet zo op de middelbare school bij elk vak/uur krijg je een andere docent. In het begin is het wennen maar later is het wel leuk.

Alle vakken die je op de basisschool krijgt, krijg je ook wel op middelbare maar het wordt anders genoemd: Science is biologie, Nederlands: is eigenlijk gewoon taal maar rekenen is wel anders dan wiskunde. met rekenen tel je cijfers bij elkaar op en trek je af maar wiskunde is meer figuren tekenen.

Op de basisschool krijg je meestal een "speel" pauze en een "eet" pauze. Op de middelbare school krijg je om de 2 uur een korte pauze, je mag niet naar huis.

Wat je op de middelbare hebt en niet op de basisschool is een mentor, een mentor is een gewone docent die ook les geeft maar ook leiding heeft over zijn eigen klas. Met de mentor krijg je ook gesprekken over hoe het gaat met je op school.

Op de basisschool heb je wel huiswerk, maar niet zo veel. Op de middelbare school moet je elke dag huiswerk maken, in het begin is het even wennen maar het went snel.

Op de basisschool heb je meestal niet meer dan 3 toetsen in de maand. Op de middelbare school heb je soms 4 toetsen per week, het klinkt allemaal heel veel maar als je je huiswerk maakt valt het allemaal wel mee!

Ankertekst 4: 115 punten (sterk)

De sterke tekst onderscheidt zich direct met een duidelijke inleiding waarin ook de lezer wordt aangesproken. Daarnaast bevat deze tekst een duidelijke titel en kern. Het slot ontbreekt, al heeft de schrijver wel geprobeerd de laatste zin als een afsluitende evaluerende zin te laten functioneren. Wat opvallend is aan deze tekst is dat de schrijver ervan zich in tegenstelling tot de hiervoor gepresenteerde

teksten echt tot de lezers richt. Op meerdere momenten in de tekst legt de schrijver verschillen en overeenkomsten uit, waarbij uitgegaan wordt van de kennis van een basis-schoolleerling.

Tekst 5

Ga je binnenkort naar de middelbare? En vindt je het eigenlijk best wel eng? Maak je geen zorgen, het is niet eng. Maar er verandert wel veel, en sommige dingen blijven hetzelfde. Ik leg je nu uit wat de verschillen en overeenkomsten zijn!

Jij bent eraan gewend dat je één leraar hebt. Je leraar geeft je van elk vak les. Op de middelbare school is dit anders. Je hebt voor elke les een andere leraar.

Een overeenkomst is dat je gewoon een klas hebt. Je hebt dezelfde kinderen, elke les. Net als bij jullie! Alleen; de kinderen hebben jou niveau.

Op jou school ga je naar je klas lopen en wacht je tot je meester zegt wat je die dag gaat doen. Op de middelbare moet je op je lesrooster kijken. Daarop kun je zien wat je die hele week hebt en naar welke lokalen je moet.

Ook heb je gewoon twee pauze's waarin je kunt eten en drinken. Je hebt pauze met de hele onderbouw. Het is er altijd heel gezellig. Het enige verschil is dat je de tweede pauze niet naar huis mag.

Je weet het misschien al. Misschien ben je daar het meest bang voor...Huiswerk! Je hebt gehoord dat je er stapels van krijgt. Dit is natuurlijk niet altijd waar. Als je gewoon oplet in de klas komt alles goed!

Kortom het is een grote stap. Iedereen is wel een beetje bang... Maar toch is alles super leuk!

Ankertekst 5: 130 punten (zeer sterk)

In deze tekst is direct een duidelijke structuur te herkennen; er zit een echte alineaverdeling in. Per alinea komt niet meer dan één verschil of overeenkomst aan bod. Daarnaast is de structuur ook zichtbaar gemaakt door structurelementen. De tekst bevat een goede pakkende inleiding. Naast de inleiding is er ook een kort slot in de tekst terug te vinden. Bovendien is de tekst erg gericht op basis-schoolscholieren, de schrijver heeft rekening gehouden met de doelgroep.

Ondanks dat deze zeer sterke tekst geen titel bevat, is hij wel representatief voor de zeer sterke teksten van deze selectie. Omdat ook deze tekst voor verbetering vatbaar is, loopt de schaal tot 150 punten.

De resultaten na het gebruik van de beoordelingsschaal

Na een kennismaking met de beoordelingsschaal en de daarbij behorende toelichting en een korte instructie, zijn zes beoordelaars in twintig verschillende jury's aan de slag gegaan met het beoordelen van de 208 leerlingteksten.

Nadat alle teksten uit de voor- en nametingen waren geschaald, zijn er analyses uitgevoerd op de mate van overeenstemming en de jury-betrouwbaarheid bij de tekstoorden. De overeenstemming tussen de zes beoordelaars kan bepaald worden binnen vijftien verschillende paren. Deze overeenstemming wordt uitgedrukt in de productmoment-correlatie. De laagste correlatie bedraagt 0,49, en de hoogste 0,81. De gemiddelde correlatie is 0,65. Dit wijst op een behoorlijk goede overeenstemming tussen de zes beoordelaars. De zes beoordelaars vormen met elkaar 20 jury's van steeds drie beoordelaars. De jury-betrouwbaarheid, uitgedrukt in de α van Cronbach, varieert van 0,57 tot 0,92. De gemiddelde α is 0,83. Dit duidt op een goede betrouwbaarheid.

Vervolgens zijn de tekstoorden van de

		GEMIDDELDE	N	STANDAARDEVIATIE
Hele groep	voormeting	85,41	90	11,03
	nameting	102,30	90	13,66
Controlegroep	voormeting	85,82	90	11,15
	nameting	101,68	90	15,45
Experimentele groep	voormeting	84,78	90	11,03
	nameting	101,32	90	13,34

Tabel 1. Resultaten van de voor- en nameting

drie beoordelaars per jury geanalyseerd. Als criterium bij het vergelijken van de tekstoorden hebben we een verschil van minder dan 30 punten in tekstoordel als aanvaardbaar beschouwd. Bij een verschil van meer dan 30 punten (meer dan twee ankerteksten) moest de betreffende tekst opnieuw beoordeeld worden door een vierde (nieuwe) beoordelaar.

Uit de analyses bleek dat 80% van de tekstoorden minder dan 30 punten van elkaar verschilden. De beoordelaars verschilden bij de meeste teksten dus niet erg in hun tekstoordel. De overige 20% van de teksten moesten opnieuw beoordeeld worden. De verschillen tussen de beoordelaars bij deze teksten konden in de helft van de gevallen verklaard worden door praktische problemen zoals onleesbare of onvolledige kopieën van de leerlingteksten. Nadat de vierde beoordelaar de tekst had beoordeeld, werd de extreemste score van de vier verwijderd. Vervolgens werd het gemiddelde bepaald van de drie overgebleven scores. De uiteindelijke gemiddelde jurybetrouwbaarheid uitgedrukt in α is 0,89.

Nadat alle teksten op deze manier van een oordeel waren voorzien, zijn door de onderzoekers via de codering op de teksten de oordelen gesorteerd per leerling en gesplitst op teksten uit de voor- en uit de nameting. Op die manier konden de verschillen tussen de voor- en nameting voor alle leerlingen bere-

kend worden. In totaal bleken van 90 leerlingen zowel de voor- als de nameting aanwezig te zijn. Alle leerlingen hebben bij Nederlands instructie en oefening gehad in het schrijven van een uiteenzetting met daarin overeenkomsten en verschillen. In tabel 1 presenteren we de resultaten van de voor- en nameting.

Uit de tabel blijkt dat er zich tussen de voormeting en de nameting voor de groep als geheel zeer grote verschillen voordoen. De gemiddelde tekstscores ligt op de nameting 17 punten hoger dan op de voormeting. Het statistisch significante verschil bedraagt 1,5 standaarddeviatie (df 89, t 11,9, p .000), wat een zeer groot effect mag worden genoemd. Zoals uit de tabel overigens ook af te lezen is, zijn er nauwelijks verschillen te zien tussen de effecten voor de controleklassen en de experimentele klassen.

Implicaties voor de praktijk

De resultaten hebben laten zien dat de mate van overeenstemming en de betrouwbaarheid van de oordelen van de beoordelaars hoog waren. Dit geeft aan dat het werken met een schaalbeoordeling een nauwkeurige manier is om oordelen over leerlingteksten te kunnen geven. De beoordelaars gaven bovendien aan dat het werken met een beoordelingsschaal prettig is en relatief weinig tijd kost. Zodra ze de anker-

teksten goed kennen en de punten waarop de anker teksten van elkaar verschillen goed weten, is het geven van een tekstscore op basis van vergelijking vrij eenvoudig en doeltreffend.

Het ontwikkelen van een beoordelings-schaal is echter wel een arbeidsintensieve klus en kan pas gebeuren nadat de leerlingteksten verzameld zijn. Dit maakt dat het inzetten van een beoordelingsschaal voor het onderwijs wellicht alleen zinvol is als de schrijftaak waarvoor de beoordelingsschaal ontwikkeld wordt meerdere schooljaren bij verschillende leerlingpopulaties herhaald wordt. Een eenmaal ontwikkelde beoordelingsschaal is immers makkelijk opnieuw inzetbaar. De schrijftaak in het hier beschreven onderzoek zou zich goed voor een dergelijke herhaling lenen. Zo zouden leerlingen in de brugklas jaarlijks over de verschillen tussen de basisschool en de middelbare school kunnen schrijven, waarbij docenten dan een adequate en nauwkeurige manier van beoordelen voorhanden hebben. Daarnaast helpt het ontwikkelen van een beoordelingsschaal docenten om beter samen te werken en gezamenlijk na te denken over de beoordelingscriteria voor teksten.

LITERATUUR

- Bangert-Drowns, R. L., Hurley, M. M. & Wilkinson, B. (2004). The effects of school-based writing-to-learn. *Interventions on Academic Achievement: A Meta-Analysis. Review of Educational Research*, 74, 29-58.
- Blok, H. & Hoeksma, J.B. (1984). *Opstellen geschaald. De constructie van beoordelingsschalen voor vijf schrijfpoddrachten*. Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam.
- Erkens, T. (2011) *Toetsen op school*. Hoofdstuk 7: Het construeren van open vragen. Beschikbaar: <http://toetswijzer.kennisnet.nl/html/toetsenopschool/hoofdstuk7.pdf>
- Hajer, M. & Meestringa T. (2004). *Handboek taalgericht vakonderwijs*. Bussum: Coutinho.

Milliano, I. de e.a. (2012). Het belang van motivatie en lesparticipatie in taal- en zaakvakken voor de schrijfontwikkeling van vmbo-leerlingen. In: Jong, de N., Juffermans, K., Keijzer, M. & Rasier, L., *Papers of the Anéla 2012, Applied Linguistics Conference*. Delft: Eburon.

Wesdorp, H. (1981). *Evaluatietechnieken voor het moedertaalonderwijs. Een inventarisatie van beoordelingsmethoden voor de stelvaardigheid, het begrijpend lezen, de spreek-, luister- en discussievaardigheid*. 's-Gravenhage: Staatsuitgeverij.

Schoonen, R. & Glopper, C.M. de (1992). *Toetsing van schrijfvaardigheid: problemen en mogelijkheden. Levende Talen*, 470, 187-195.

EVELINE POLLMANN is afgestudeerd in de richting Taalbeheersing van de studie Nederlandse Taal en Cultuur aan de Rijksuniversiteit Groningen. In het kader van haar master was ze betrokken bij het onderzoek 'Schrijven om te leren' bij het Expertisecentrum taal, onderwijs en communicatie (Etoc) in Groningen. Sinds augustus 2012 is ze trainee bij Eerst de Klas, een programma voor excellente academici in het bedrijfsleven en het onderwijs. E-mail: <E.pollmann@umail.leidenuniv.nl>.

JOANNEKE PRENGER werkt sinds haar promotie-onderzoek naar de rol van tekst- en taalvaardigheden in het wiskundeonderwijs bij het Expertisecentrum taal, onderwijs en communicatie (Etoc) in Groningen. Daar doet zij onderzoek naar taalvaardigheidsontwikkeling en begeleidt ze scholen die willen werken aan het opzetten van taalbeleid of het implementeren van een taalgerichte didactiek. E-mail <J.Prenger@rug.nl>.

KEES DE GLOPPER is hoogleraar Taalbeheersing van het Nederlands aan de Rijksuniversiteit Groningen. Zijn onderzoek en onderwijs hebben betrekking op taalvaardigheidsontwikkeling en onderzoeksmethodologie. E-mail <c.m.de.glopper@rug.nl>.

Dialogisch leesonderwijs in de klas

AD STOOP, JOOP VAN DER KUIP & TANJA JANSSEN

Leesvaardigheidsonderwijs binnen het vak Nederlands bestaat voornamelijk uit het leren beantwoorden van metavragen, die specifiek op tekstanalyse gericht zijn en minder op tekstbegrip. Door tekstanalyse zou begrip moeten ontstaan. Die slag lijken leerlingen niet te slaan. Zij zijn gericht op het beantwoorden van de vragen en laten de tekst links liggen. In het hier beschreven quasi-experimentele onderzoek gebruikten we een dialogische benadering, waarbij de leerling zich juist op de inhoud van de tekst richt. Het bleek dat leerlingen bij deze benadering even goed presteren als leerlingen die reguliere leesvaardigheidslessen hebben gevolgd. Verder bleek dat leerlingen de dialogische aanpak waardeerden, maar wel tijdrovend vonden.

De Nederlandse jeugd slaat, afgaande op het PISA-rapport 2009 (Gille, Loijens, Noijons & Zwitser, 2010), internationaal gezien met leesvaardigheid geen slecht figuur, maar de scores zijn wel dalende ten opzichte van eerder onderzoek. PISA onderscheidt zes niveaus van leesvaardigheid. Leerlingen die op niveau 5 en 6 presteren worden toppresteerders genoemd. Slechts 9,8% van de

15-jarige Nederlandse leerlingen bereikt dat niveau.

Als we kijken naar de eisen die gesteld worden aan de leesvaardigheid, dan zien we dat het bij niveau 5 en 6 gaat om nauwkeurige tekstanalyse, om begrip van details, om kritische reflectie op teksten, het opstellen van een hypothese bij teksten, assimilatie van informatie en het gebruik van erudiete inzichten van buiten een tekst.

Traditionele leesvaardigheidsoefeningen lijken hier niet op aan te sluiten. Ze zijn gericht op een top-downanalyse (Drop, 1983) van de tekst, op het beantwoorden van metavragen, die vooral op tekstanalyse gericht zijn en minder op tekstbegrip. Zelfs de eindtermen van het vak Nederlands zijn gebaseerd op de traditionele aanpak. Geijkte vragen zijn bijvoorbeeld: Wat voor tekstsoort is dit? Wat voor soort verband vind je tussen alinea 4 en 5? Welke alinea's vormen de inleiding van de tekst?

PISA-2009 definieert leesvaardigheid als 'het begrijpen van, gebruiken van, reflecteren op en interesse tonen in geschreven teksten om je doelen te bereiken, je kennis en potentieel te verruimen, en deel te nemen aan de maatschappij' (Gille, Loijens, Noijons