

Kunnen we de moeilijkheid van teksten meten?

Naar concrete maten voor de referentieniveaus

HILDE HACQUEBORD & KLARIEN LENTING-HAAN

In dit artikel wordt verslag gedaan van een onderzoek waarbij paneloordelen worden vergeleken met objectief meetbare tekstkenmerken als het gaat om een ordening van teksten volgens de referentieniveaus uit het Referentiekader Taal. Nederlandstalige teksten zijn geordend op basis van expertbeoordelingen in termen van de referentieniveaus en geanalyseerd op verschillende kwantitatieve linguïstische maten. Uit het onderzoek blijkt dat er een aantal kwantitatieve maten zijn die een belangrijke rol spelen bij het bepalen van het tekstniveau, zoals dat door een expertpanel is beoordeeld. Een kwantitatieve analyse van teksten is daarom een goed hulpmiddel bij het selecteren van geschikt tekstmateriaal voor leerlingen.

Tekstniveaus in het Referentiekader Taal

Met het Referentiekader Taal (Expertgroep Taal en Rekenen, 2009) krijgen de referentieniveaus een belangrijke plek in het onderwijs. Docenten en leerlingen hebben daarbij behoefte aan een hanteerbare indeling in tekstniveaus in relatie tot deze referentieniveaus. In het Referentiekader Taal is beschre-

ven wat voor teksten leerlingen op de verschillende drempels van 1F tot 4F moeten kunnen lezen en op welke manier ze daarmee moeten kunnen omgaan. Zo is duidelijk wat leerlingen op bepaalde momenten in hun schoolloopbaan moeten kunnen, en weten docenten van vervolgopleidingen wat zij minimaal van hun leerlingen mogen verwachten. Het Referentiekader is echter niet altijd even duidelijk over wat een leerling concreet moet kunnen lezen op een bepaald niveau. Ook zijn de verschillen tussen de niveaus niet altijd helder geformuleerd. Bovendien blijkt dat het kader te algemeen is voor het vmbo omdat de prestatieverschillen daar te groot zijn (Van den Bergh, Evers-Vermeul & De Haas, 2009).

Voor het vmbo is er het Raamwerk Nederlands (Bohnen et al., 2007). Dit referentiekader, dat opgesteld is aan de hand van het Europees Referentiekader (Raad van Europa, 2001) en al bestond voor het Referentiekader Taal, geeft beschrijvingen voor het vak Nederlands op het vmbo. Het Raamwerk geeft voor de lagere referentieniveaus ook meetbare linguïstische maten als tekstkenmerken, zoals zinslengte en het aantal woorden. Zo wordt bijvoorbeeld aangegeven dat teksten op niveau A2 een

gemiddelde zinslengte hebben van rond de 10 woorden per zin, en dat teksten op dit niveau veel enkelvoudige zinnen bevatten. Het is echter niet duidelijk in hoeverre deze meetbare criteria empirisch zijn onderbouwd, en ook is niet precies duidelijk hoe de indeling van het ERK gekoppeld kan worden aan het Referentiekader Taal.

Omdat de beschrijvingen van de referentieniveaus in het Referentiekader Taal op sommige punten vrij abstract zijn, zou het mooi zijn als er eenduidige en objectieve maten gevonden kunnen worden die aan de beschrijvingen kunnen worden toegevoegd. Daarmee is een objectieve basis voorhanden waarmee teksten (automatisch) kunnen worden gescreend. Dit maakt het voor docenten eenvoudiger om geschikt tekstmateriaal te zoeken zodat hun leerlingen gericht naar de referentieniveaus toe kunnen werken. In dit artikel doen wij verslag van een onderzoek naar dergelijke concrete kwantitatieve maten als grondslag voor de beoordeling van teksten in relatie tot het Referentiekader Taal.

Leesbaarheidsonderzoek

In het zogenaamde leesbaarheidsonderzoek (Staphorsius, 1994; Staphorsius en Krom, 2008, Kraf & Pander Maat, 2009) wordt veel gebruik gemaakt van kwantitatieve linguïstische maten voor het bepalen van tekstmoeilijkheid. In het kader van het NWO-project Begrijpelijke Taal is de kennisbank Begrijpelijke Taal ontwikkeld met een overzicht van leesbaarheidstoetsen (www.kennisbank-begrijpelijketaal.nl). Ook voor het verzamelen van tekstmateriaal voor tekstbegripstoetsen wordt gebruik gemaakt van leesbaarheidsformules. Zo maakt het CITO gebruik van de CLIB, een index die aangeeft hoe moeilijk een tekst te begrijpen is (Staphorsius, 1994), en wordt voor het taaltoetspakket Diataal ook een door de onderzoekers ontwikkelde leesbaarheids-

formule gebruikt (Andringa & Hacquebord, 2000). Beide formules hanteren de maten gemiddelde zinslengte (GZL), gemiddelde woordlengte (GWL) en het dekkingspercentage frequente woorden (% dekking) in een tekst.

Leesbaarheidsonderzoeken stuiten op veel scepsis (Jansen & Lentz, 2008). Het bepalen van de leesbaarheid van een tekst op basis van uitsluitend kwantitatieve maten is te beperkt. Deze studie is dan ook niet bedoeld om de leesbaarheid van een tekst in linguïstische maten te vangen, maar om maten te vinden die kunnen helpen bij het toekennen van een tekstniveau naast het oordeel van experts.

Vraagstelling

In dit onderzoek is eerst door een panel van experts een ordening van teksten gemaakt op basis van het Referentiekader Taal. In het Referentiekader is een aantal tekstkenmerken beschreven. Deze kenmerken waren van doorslaggevend belang bij het bepalen van het niveau van de tekst en zijn weergegeven in tabel 1.

Een beoordeling aan de hand van deze kenmerken is vrij globaal en intuïtief. Een beoordeling aan de hand van meer objectieve en meetbare maten heeft grote praktische voordelen. Door docenten geselecteerde teksten zouden dan eerst onderworpen kunnen worden aan een computer-screeningsprogramma om het (referentie)niveau ervan te kunnen bepalen.

De vraag in dit artikel is in hoeverre de ordening van teksten door experts op basis van het Referentiekader Taal voorspeld wordt door de genoemde linguïstische maten en of er daarnaast nog andere objectief meetbare tekstkenmerken zijn die iets kunnen bijdragen aan de voorspelling van het niveau van de tekst. Bij deze vraag worden de volgende subvragen gesteld:

ONDERWERP	1F	Het onderwerp sluit aan bij de leefwereld van de leerling.
	2F	Het onderwerp sluit aan bij de leefwereld van de leerling of staat daar iets verder van af.
	3F	Het onderwerp is van maatschappelijke aard of uit de (beroeps)opleiding.
	4F	Het onderwerp van de tekst maakt niet uit.
TEKSTSTRUCTUUR	1F	De tekst is eenvoudig van structuur.
	2F	De tekst heeft een heldere structuur.
	3F	De tekst heeft een vrij complexe structuur, maar een duidelijke opbouw.
	4F	De structuur van de tekst is niet altijd even duidelijk.
INFORMATIE-DICHTHEID	1F	De tekst heeft een lage informatiedichtheid en belangrijke informatie is gemarkeerd of wordt herhaald. Er wordt niet teveel nieuwe informatie tegelijkertijd gegeven.
	2F	De tekst heeft een lage informatiedichtheid en is niet te lang. Verbanden in de tekst worden duidelijk aangegeven.
	3F	De informatiedichtheid van de tekst is vrij hoog.
	4F	De informatiedichtheid van de tekst is hoog. De tekst is complex.
TEKSTSOORTEN	1F	De tekst is eenvoudig.
	2F	a. In de informatieve tekst worden verschillende dimensies gecombineerd. b. De instructieve tekst is veel voorkomend c. De betogende tekst is (meestal) redundant en vaak licht opiniërend
	3F	a. In de informatieve tekst wordt formeel taalgebruik gebruikt. Schema's en rapporten over het eigen werkterrein kunnen ingewikkeld zijn. b. De instructieve tekst is ingewikkeld. c. De betogende tekst is bijvoorbeeld een tekst uit een schoolboek of een opiniërend artikel.
	4F	a. De informatieve tekst heeft een hoge informatiedichtheid, is lang en ingewikkeld. b. De betogende tekst heeft een ingewikkelde argumentatie, de schrijver neemt (impliciet) een standpunt in of geeft een beschouwing.
TECHNIEK EN WOORDENSCHAT	1F	In de tekst staan voornamelijk alledaagse frequente woorden. De betekenis van een onbekend woord kan uit de context afgeleid worden.
	2F	Wanneer nodig kan de betekenis van onbekende of moeilijke woorden uit de context worden afgeleid.

Tabel 1. Tekstkenmerken die beschreven zijn in de referentieniveaus.

- Hoe sterk is het onderlinge verband tussen de verschillende tekstkenmerken?
- In hoeverre wordt de expertbeoordeling van teksten in termen van de niveaus uit het Referentiekader Taal voorspeld door de objectief meetbare tekstkenmerken GZL, GWL en %dekking?
- Zijn er andere objectief meetbare voorspelers die bijdragen aan een betere voorspelling van de expertbeoordeling van teksten?

Methode

Tekstcorpus

Het tekstcorpus bestaat uit 19 teksten die oorspronkelijk geselecteerd zijn om leraren een idee te geven hoe gemakkelijk of moeilijk verschillende teksten kunnen zijn. Criteria voor de selectie van deze teksten waren dat de teksten informatief waren, dat het onderwerp verband hield met een schoolvak en dat de teksten een goed voorbeeld vormden voor één van de niveaus uit het Referentiekader Taal. De teksten zijn afkomstig uit examens Nederlands, uit kranten en van informatieve websites.

Deze teksten zijn beoordeeld door een panel van ervaren tekstbeoordelaars bestaande uit Masterstudenten Toegepaste Taalwetenschap die stage liepen bij CPS en onderwijsadviseurs met ervaring als taaldocent in het voortgezet onderwijs. Het panel heeft de teksten gezamenlijk gelezen en vervolgens in een gesprek beoordeeld aan de hand van de tekstkenmerken uit het Referentiekader Taal.

Aan alle teksten is een referentieniveau toegekend. Bij de teksten die volgens het panel tussen twee niveaus in vielen, is dat in het referentieniveau terug te zien. Zo heeft een tekst die tussen 2F en 3F inzit niveau 2,5 gekregen. Als in het gesprek bleek dat panelleden twee verschillende niveaus aan de tekst wilden toekennen, is eveneens gekozen voor een tussenniveau. De beoor-

delingen van verschillende panelleden lagen in geen van de gevallen meer dan één niveau uit elkaar. Een voorbeeld van een tekst met een beoordeling is te vinden in de bijlage.

Tekstkenmerken

Uit onderzoek is bekend dat GWL, GZL en %dekking goede maten zijn om tekstmoeilijkheid te meten (Andringa & Hacquebord, 2000). De kwantitatieve maten die het Raamwerk Nederlands naast de genoemde maten hanteert, zijn Tekstlengte en de mate waarin enkelvoudige zinnen en verbindingswoorden voorkomen.

Naast de bekende maten op het niveau van zinnen en woorden hebben wij in dit onderzoek ook een aantal maten meegenomen op het niveau van de tekst. De maten die we hebben gehanteerd zijn ingedeeld in drie groepen: lexicale complexiteit (micro-niveau), zinscomplexiteit (mesoniveau) of tekstcomplexiteit (macroniveau). Voor elke groep zijn minimaal twee maten geselecteerd. De groep waarbij het kenmerk ingedeeld kon worden was een belangrijk selectiecriteria. Daarnaast was de meetbaarheid van de kenmerken een belangrijk criterium voor opname in het onderzoek. De kenmerken die in het onderzoek zijn gebruikt, zijn weergegeven in tabel 2.

In het onderzoek worden drie maten voor *lexicale complexiteit* gebruikt om de teksten te analyseren. De eerste is gemiddelde woordlengte (GWL). Dat is de som van het aantal tekens in de tekst gedeeld door het aantal woorden in de tekst. De tweede maat is het dekkingspercentage basiswoorden (%dekking). Dit is het percentage woorden in de tekst dat voorkomt in de woordenlijst van De Kleijn & Nieuwborg (1991). GWL en het dekkingspercentage zijn beide maten die ook in het onderzoek van Andringa & Hacquebord (2000) gebruikt zijn. De laatste maat voor *lexicale complexiteit* is Type Token Ratio (TTR). Dit is het aantal verschillende

GROEP	TEKSTKENMERK	INDEX
LEXICALE COMPLEXITEIT	Gemiddelde woordlengte	Het aantal lettertekens in de tekst gedeeld door het aantal woorden
	Dekkingspercentage basiswoorden	Het percentage woorden in de tekst dat gedekt wordt door de woordenlijst van De Kleijn & Nieuwborg (1991)
	Type token ratio (TTR)	Het aantal verschillende lemma's (types) gedeeld door het totaal aantal woorden (tokens)
ZINS- COMPLEXITEIT	Gemiddelde zinslengte	Het aantal woorden in de tekst gedeeld door het aantal zinnen.
	Gemiddeld aantal deelzinnen per tekst	Het aantal deelzinnen gedeeld door het aantal woorden
	Afhankelijkheidslengte subject-werkwoord	De afstand tussen het finiete werkwoord en het subject
	Afhankelijkheidslengte object-werkwoord	De afstand tussen het finiete werkwoord en het object
TEKST- COMPLEXITEIT	Percentage verbindingswoorden	Het aantal verbindingswoorden gedeeld door het totaal aantal woorden.
	Tekstlengte	Het totaal aantal woorden in een tekst

Tabel 2. De tekstkenmerken die gebruikt zijn voor dit onderzoek

woorden (types) gedeeld door het totaal aantal woorden (tokens). De types zijn de verschillende lemma's.

Voor het meten van zinscomplexiteit worden vier verschillende maten gebruikt. De eerste is gemiddelde zinslengte (GZL). Dat is de som van het aantal woorden in de tekst gedeeld door het aantal zinnen in de tekst. Dit is een redelijk sterke voorspeller van leesbaarheid (Staphorsius, 1994). De tweede voorspeller is een maat voor deelzinnen waarmee de complexiteit van zinnen gemeten kan worden: het aantal deelzinnen gedeeld door het totaal aantal woorden. De

derde en vierde tekstkenmerken om zinscomplexiteit te meten zijn maten voor tangbreedte. Afhankelijkheidslengtes maken tangconstructies moeilijk. Als er meer woorden tussen het subject of object en het werkwoord staan, betekent dat dat de zin complexer is. De afhankelijkheidslengtes zijn gemeten door het aantal woorden te tellen tussen het subject en het finiete werkwoord (Tang1), en de afstand tussen het object en het finiete werkwoord (Tang2).

Twee maten voor tekstcomplexiteit zijn van belang voor dit onderzoek. De eerste is het percentage verbindingswoorden (verbindw)

TEKSTKENMERK	N	GEMIDDELDE	STANDAARDDEVIATIE
Tekstlengte	19	621,06	339,14
GZL	19	14,71	4,28
GWL	19	5,32	0,32
Dekkingspercentage	19	81,53	3,54
Verbindingswoorden	19	2,79	1,17
Deelzinnen	19	1,43	0,30
Tang₁	19	2,66	1,21
Tang₂	19	2,25	1,15
TTR	19	0,60	0,06

Tabel 3. Gemiddelden en standaarddeviaties voor alle tekstkenmerken die geoperationaliseerd zijn zoals beschreven in tabel 2

in de tekst; dat is het aantal verbindingswoorden in de tekst gedeeld door het totaal aantal woorden. Eerder onderzoek heeft uitgewezen dat teksten met expliciete verbindingen beter begrepen worden dan teksten zonder connectieven (Land, Sanders, Lentz & van den Bergh, 2002; Land, 2009). De tweede maat voor tekstcomplexiteit is de totale tekstlengte; dat is het totaal aantal woorden in een tekst. Deze maat is opgenomen in de studie omdat lange teksten wellicht moeilijker te lezen zijn dan kortere teksten.

Voor de analyse is gebruik gemaakt van twee computerprogramma's: het programma Textscreen dat geschreven is door studenten van de Rijksuniversiteit Groningen en dat gebruikt is door de auteurs van Diataal (Andringa & Hacquebord, 2000) en een programma uit Utrecht: T-Scan (Kraf & Pander Maat, 2009). Het eerste programma meet gemiddelde woordlengte, gemiddelde zinslengte en het percentage basiswoorden. Met T-scan zijn TTR, afhankelijkheidslengtes en het aantal deelzinnen geanalyseerd.

Procedure

Aan elk van de teksten is een niveau uit het Referentiekader Taal toegekend door een panel van experts. Op basis daarvan zijn de teksten geordend en vervolgens vergeleken met de ordening op basis van de verschillende tekstkenmerken uit tabel 2. Vervolgens is het verband tussen elk van de linguïstische tekstkenmerken en de paneloordelen onderzocht. Om het verband te meten zijn rangcorrelaties gebruikt. Deze geven aan in hoeverre de rangordeningen van teksten op basis van verschillende maten overeenkomen, bijvoorbeeld of een rangordening op zinslengte overeenkomt met een ordening op basis van percentage onbekende woorden. Zo kon het onderlinge verband tussen de gemeten tekstkenmerken worden vastgesteld, en de relatie van elk van die kenmerken met de paneloordelen.

Vervolgens is onderzocht of de tekstkenmerken van voorspellende waarde waren voor de tekstordening door het panel. Dat is gedaan met behulp van een meervoudige regressieanalyse. Deze statistische techniek

analyseert de onderlinge samenhang van meerdere variabelen, in dit geval tekstkenmerken, en geeft aan in welke mate die ieder afzonderlijk en samen bijdragen aan de voorspelling van (in dit geval) tekstmoeilijkheid. In deze analyse zijn eerst alle tekstkenmerken meegenomen om te onderzoeken hoeveel van de variantie in de paneloordelen maximaal voorspeld kon worden door alle tekstkenmerken samen. Vervolgens is onderzocht wat het beste model is voor de voorspelling.

Resultaten

De tekstkenmerken zijn eerst met elkaar vergeleken om duidelijk te krijgen hoe ze met elkaar samenhangen en om te onderzoeken

of er een verband is tussen de verschillende kenmerken en de paneloordelen. Vervolgens is onderzocht of de tekstordening voorspeld kon worden door de linguïstische maten en als dat zo was, welke voorspellers het sterkst waren. De gemiddelden en standaarddeviaties van de tekstkenmerken zijn weergegeven in tabel 3.

Om de verschillende kenmerken met elkaar te vergelijken, zijn de rangcorrelaties berekend. Deze zijn weergegeven in tabel 4. Omdat de correlatie tussen TTR en TTR logischerwijs 1 is, is deze maat alleen in de horizontale reeks weergegeven.

Er zijn significante correlaties tussen tekstlengte en GZL, tekstlengte en TTR, GZL en TANG₁ en GWL en dekkingpercentage. Deze correlaties impliceren dat deze maten op een-

	GZL	GWL	%DEKKING	VERBINDW	DEELZIN	TANG ₁	TANG ₂	TTR
TEKSTLENGTE	0,487*	-0,167	0,196	-0,020	-0,090	0,176	0,400	0,744**
GZL		0,408	-0,157	-0,344	-0,058	0,515*	0,143	0,419
GWL			-0,530*	0,194	0,031	0,234	0,052	0,016
%DEKKING				-0,104	0,016	0,248	-0,170	0,055
VERBINDW					-0,164	-0,260	-0,249	0,198
DEELZIN						-0,043	0,347	0,493*
TANG ₁							0,256	0,090
TANG ₂								0,490*

Tabel 4. Rangcorrelaties tussen alle linguïstische kenmerken voor alle teksten; * $p < 0,05$, ** $p < 0,01$

TEKSTKENMERK	PANEOORDEEL
Tekstlengte	0,425
GZL	0,713**
GWL	0,444
Dekkingspercentage	-0,226
Verbindingswoorden	-0,014
Deelzinnen	-0,248
Tang ₁	0,493*
Tang ₂	0,090
TTR	0,490*

Tabel 5. Rangcorrelaties tussen de tekstkenmerken en de paneeloordeelen in termen van de referentieniveaus; * $p < 0,05$, ** $p < 0,01$

zelfde manier tekstmoeilijkheid meten. Om te achterhalen welke maten mogelijk van belang zijn voor de referentieniveaus, zijn ook de correlaties met de paneeloordeelen berekend. Deze zijn weergegeven in tabel 5.

Er zijn significante correlaties gevonden tussen enerzijds de paneeloordeelen en anderzijds GZL, Tang₁ en TTR. Volgens deze analyse zijn deze drie maten dus de belangrijkste voorspellers van de tekstniveaus zoals beschreven in het Referentiekader Taal.

De voorspellende waarde van de tekstkenmerken

De volgende vraag is hoeveel variantie in de tekstordening op basis van paneeloordeelen verklaard kan worden door de verschillende tekstkenmerken. GZL, Tang₁ en TTR hangen samen met de paneeloordeelen (zie tabel 5), maar ook met elkaar (zie tabel 4). De verwachting is daarom dat in ieder geval één van deze maten een goede voorspeller is voor de paneeloordeelen. Met een regressieanalyse is

eerst onderzocht hoeveel variantie alle maten samen kunnen verklaren. Hieruit bleek dat 80,6% van de variantie wordt verklaard door alle tekstkenmerken samen.

Vervolgens is gezocht naar het beste regressiemodel voor de voorspelling van de paneeloordeelen, dat wil zeggen dat onderzocht is welke maten samen de paneeloordeelen het best voorspellen. Als eerste is een model gebouwd waarin de drie bekende maten GZL, GWL en %dekking zijn opgenomen. Door deze drie maten wordt 52,6% van de variantie verklaard. Omdat deze drie maten echter slechts een deel van de variantie verklaren, is gezocht naar een beter model. %dekking is een bekende maat uit de literatuur, maar heeft vermoedelijk teveel overlap met GWL om genoeg bij te kunnen dragen aan de voorspelling. Het lijkt daarom beter om een andere maat in het model op te nemen. Een goed meetbare maat is tekstlengte. Daarom is een tweede regressiemodel gebouwd waarin de maten GZL, GWL en Tekstlengte zijn opgenomen. Hieruit bleek dat GZL, GWL en TL samen 71,1% van de variantie in de paneeloordeelen verklaren. Met een uitbreiding van het model met TTR en Tang₁, de andere veelbelovende maten uit tabel 5, bleek de voorspellende waarde vrijwel niet toe te nemen. We kunnen dus volstaan met GZL, GWL en TL als voorspellers van de tekstordening in termen van het Referentiekader Taal volgens het expertpanel.

Discussie en conclusie

De resultaten van dit onderzoek laten zien dat er een intern verband is tussen verschillende tekstkenmerken, en dat er een verband is tussen enkele tekstkenmerken en de tekstbeoordelingen in termen van de niveaus uit het Referentiekader Taal. De regressieanalyse wijst uit dat een model met GZL (gemiddelde zinslengte), GWL (gemiddelde woordlengte)

en TL (tekstlengte) een groot deel van de variatie in de paneloordelen verklaart. Het feit dat er niet meer voorspellers zijn gevonden kan twee oorzaken hebben: 1. GZL, GWL en TL zijn sterk genoeg om de paneloordelen te voorspellen of 2. de overige tekstkenmerken uit dit onderzoek zijn niet goed genoeg om iets bij te kunnen dragen aan de voorspelling. Type token ratio, %dekking, het aantal deelzinnen, afhankelijkheidslengtes, en het aantal verbindingswoorden konden allemaal niet (genoeg) bijdragen aan de voorspelling van tekstmoeilijkheid. Deze maten hebben voor een deel teveel overlap met andere maten om de paneloordelen in termen van het Referentiekader Taal te kunnen voorspellen naast de al bekende maten.

Voor *lexicale complexiteit* blijkt alleen GWL een goede voorspeller te zijn, zoals uit eerder onderzoek ook is gebleken. De correlatie die gevonden is tussen TTR en de paneloordelen was niet van dien aard dat deze in het regressiemodel paste. Verder is het opvallend dat %dekking niet als sterke voorspeller uit dit onderzoek komt, terwijl dit in eerdere onderzoek (Andringa & Hacquebord, 2000) wel een goede maat bleek te zijn. Een mogelijke verklaring is dat in dit onderzoek teksten van verschillende niveaus (van 1F tot voorbij 3F) geanalyseerd zijn, terwijl in het eerdere onderzoek alleen teksten van de lagere niveaus (1F tot 2F) zijn onderzocht, en dat het criterium dekkingpercentage van een basiswoordenschat geen rol meer speelt voor leerlingen die de hogere tekstniveaus (vanaf 2F) beheersen. Voor wat betreft zinscomplexiteit spelen het aantal deelzinnen en de afhankelijkheidsrelaties geen significante rol in de voorspelling van de referentieniveaus. Dit onderzoek geeft geen antwoord op de vraag wat moeilijk is in lange zinnen, maar de relatie die gevonden is tussen GZL en TangI geeft aan dat een grotere afstand tussen het subject en het werkwoord gepaard gaat met een grotere zinslengte. Ook de andere maten op zinsniveau worden blijk-

baar afgedekt door GZL.

Op het niveau van de tekst blijkt alleen de tekstlengte een belangrijke voorspeller te zijn voor het referentieniveau dat experts aan een tekst toekennen. De andere maat op tekstniveau die we hebben onderzocht, het percentage verbindingswoorden, blijkt geen goede voorspeller te zijn. Deze maat heeft echter geen enkele correlatie met Tekstlengte. Kraf en Pander Maat (2009) wijzen erop dat deze maat bij moeilijkere teksten een rol speelt. Het zou dus kunnen zijn dat in dit onderzoek het gemiddelde niveau van de teksten niet hoog genoeg was om deze maat als voorspeller te identificeren.

De resultaten van dit onderzoek laten zien dat er een onderling verband is tussen de onderzochte tekstmaten, en dat de variatie in de tekstoordelen van een expertpanel voor 80,6% wordt voorspeld door alle maten samen. De maten gemiddelde zinslengte, gemiddelde woordlengte en tekstlengte blijken samen 71,1% van de variatie in de tekstoordelen te verklaren. Onderzoek wijst uit dat er voor het bepalen van het tekstniveau meer factoren van invloed zijn dan alleen een tekst met zijn (eenvoudige) taalkundige kenmerken. Toch kan leesbaarheidsonderzoek helpen bij het bepalen van wat lezers moeilijk vinden in een tekst. Dit onderzoek laat zien dat de eenvoudige en objectieve kwantitatieve maten gemiddelde zinslengte, gemiddelde woordlengte en tekstlengte belangrijke voorspellers zijn voor tekstoordelen in termen van de niveaus uit het Referentiekader Taal. Voor een aantal andere maten is wel een verband met de beoordelingen in termen van referentieniveaus gevonden, maar dit verband bleek niet significant voorspellend te zijn. Onderzoek met meer teksten en meer referentiebeoordelingen zal misschien meer meetbare tekstkenmerken opleveren, die eventueel kunnen worden toegevoegd aan de al bekende maten. Deze kunnen dan ook worden opgenomen in een referentiekader.

LITERATUUR

Andringa, S. J. & Hacquebord, H. I. (2000).

De moeilijkheidsgraad van schoolboekteksten als grondslag voor het vaststellen van tekstbegripvaardigheid. *Toegepaste Taalwetenschap in Artikelen*, 64, 83–94.

Bergh, H. van den, Evers-Vermeul, J. & Haas, T. de (2009). Doorlopende leerlijnen voor het taalonderwijs. In A. Backus, M. Keijzer, I. Vedder & B. Weltens (Eds.). *Artikelen van de zesde Anéla-conferentie* (pp. 408–415). Delft: Eburon.

Bohnen, E., Jansen, F., Kuijpers, C., Thijssen, R., Schot, I. & Stockmann, W. (2007). *Raamwerk Nederlands, Nederlands in (v)mbo-opleiding, beroep en maatschappij*. 's-Hertogenbosch: CINOP.

Expertgroep Taal en Rekenen (2009). *Referentiekader taal en rekenen*. Enschede: SLO.

Jansen, C. & Lentz, L. (2008). Hoe begrijpelijk is mijn tekst? *Onze taal*, 77, 4–7

Kleijn, P. de, & Nieuwborg, E (1991). *Basiswoordenboek Nederlands*. Groningen: Wolters-Noordhoff.

Kraf, R. L. & Pander Maat, H. L. W. (2009). Leesbaarheidsonderzoek: oude problemen, nieuwe kansen. *Tijdschrift voor taalbeheersing*, 31, 97–123

Land, J. (2009). Zwakke lezers, sterke teksten? *Effecten van tekst- en lezerskenmerken op het tekstbegrip en de tekstwaardering van vmbo-leerlingen*. Delft: Eburon

Land, J., Sanders, T., Lentz, L. & Bergh, H. van den (2002). Coherentie en identificatie in studieboeken. *Tijdschrift voor taalbeheersing*, 24, 281–302.

Pander Maat, H. L. W. (2002). *Tekstanalyse: wat teksten tot teksten maakt*. Bussum: Uitgeverij Coutinho.

Raad van Europa (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Staphorsius, G. (1994). *Leesbaarheid en leesvaar-*

digheid. De ontwikkeling van een domeingericht meetinstrument. Arnhem: Cito.

Staphorsius, G. & Krom, R. S. H. (2008). *CLIB, CILT en AVI: leesbaarheidsindexen*. Arnhem: Cito.

HILDE HACQUEBORD is universitair hoofd-docent Toegepaste Taalwetenschap aan de Rijkuniversiteit Groningen. Daar heeft zij o.a. Diataal ontwikkeld. Vanaf 2011 is Diataal verzelfstandigd in een universitaire BV onder haar leiding.

KLARIEN LENTING-HAAN schreef haar masterscriptie voor Toegepaste Taalwetenschap aan de Rijksuniversiteit Groningen over de rol van tekstkenmerken bij het bepalen van tekstmoeilijkheid. Sinds 2011 is ze medewerker onderzoek en ontwikkeling bij Diataal. E-mail: <klarienlenting@diataal.nl>.

Bijlage: Voorbeeld van een tekst en de beoordeling

Smelten van het ijs

Door het opwarmen van de aarde smelten de ijsvoorraden. Dit wereldwijde smelten is een duidelijk teken dat ons klimaat drastisch verandert. De gevolgen gaan veel verder dan alleen het verlies van sneeuw en ijs. Miljarden mensen en hele ecosystemen zijn afhankelijk van deze ijslagen. Gletsjerijs zorgt voor zoetwatervoorraden en zee-ijs vormt de leefomgeving voor flora en fauna. Het wereldwijde smelten draagt bij aan de stijging van de zeespiegel: 90 procent komt door het uitzetten door verwarming van het zeewater en 10 procent door smeltwater.

Stijging zeespiegel

Scenario's laten zien dat de zeespiegel in de toekomst met enkele meters zal stijgen. Alleen deze eeuw tussen de 9 en 88 cm. Dit zal wereldwijd rampzalige gevolgen hebben. Wanneer de zeespiegel met 1 meter stijgt moeten in Bangladesh alleen al meer dan tien miljoen mensen verhuizen. Wanneer al het ijs smelt zal het onze oceanen verdunnen, en het zoutgehalte zo veranderen dat het water schadelijk wordt voor de visstanden. Ook zullen oceaanstromen ontregelen.

Minder ijs, hogere temperaturen

Opwarming van de aarde is een zichzelf versterkend proces. Het ijs smelt, en met minder ijs wordt de aarde ook weer warmer. Sneeuw en ijs reflecteren namelijk het zonlicht beter dan kale grond en water. Dus minder ijs betekent weer warmere temperaturen op aarde, die weer zorgen voor meer smeltend ijs enzovoort.

Bron: <http://www.opwarmingvandeearde.nl/jongeren/opwarming.aspx>

KENMERK	WAARDE
Tekstlengte	219
GZL	13,1
GWL	5,3
Dekkingspercentage	78,1
Verbindingswoorden	1,92
Deelzinnen	1,6
Tang1	2
Tang2	0,33
TTR	0,58

Paneloordeel: 2,5

ONDERDEEL REFERENTIEKADER	NIVEAU	UITLEG
Onderwerp	2-3	Het onderwerp staat verder van de leerling af en is maatschappelijk van aard
Tekststructuur	2	De structuur is helder en wordt aangegeven door middel van kopjes
Informatiedichtheid	3	De informatiedichtheid van de tekst is vrij hoog
Tekstsoorten	Informatief: 2-3	De tekst zou in een schoolboek kunnen staan, kan ook tot voorlichtingsmateriaal gerekend worden. Het taalgebruik is niet echt formeel
Techniek en woordenschat	>2	Vrij veel lastige formuleringen en moeilijke woorden (bijvoorbeeld: Ook zullen oceaanstromen ontregelen)