

De ontwikkeling van de Computergestuurde Toets voor Automatisering van het Lezen*

A. VAN DER LEIJ, H. SMEETS en
V.H.P. VAN DAAL
Sectie Speciale Pedagogiek,
Vrije Universiteit, Amsterdam

Samenvatting

In deze bijdrage aan de thema-serie 'Instructie van leerlingen met ernstige leerproblemen' wordt de ontwikkeling van de Computergestuurde Toets voor Automatisering van het Lezen (COTAL) beschreven als voorbeeld van onderzoek naar de diagnostiek van kinderen met ernstige leerproblemen volgens een taak- en procesgerichte benadering.

Er is een experimentele versie van een computergestuurde toets ontwikkeld die zeven subtoetsen bevat. Deze subtoetsen weerspiegelen de manieren waarop woorden als stimulus kunnen worden aangeboden (auditief; visueel) en diverse wijzen van antwoord geven (beoordelen op gelijkens; hardop benoemen; natypen). De moeilijkheidsgraad wordt opgevoerd door verkorting van de visuele stimulus dan wel verlenging van de tijd na de auditieve stimulus. Er worden woorden van zes verschillende orthografische structuren gebruikt.

In de gerapporteerde studie worden de toetsresultaten van 37 leerlingen met ernstige leesproblemen en 37 eerste klassers van overeenkomstig leesniveau vergeleken. De nadruk van deze vergelijking ligt op indicaties voor een verschillend verloop van het proces van leren lezen, toegespitst op directe woordherkenning, een van de pijlers onder het automatiseringsproces.

* Dit onderzoek is mede mogelijk gemaakt door een subsidie van het Ministerie van Onderwijs en Wetenschappen o.g.v. het Informatica Stimuleringsplan (INSP). De acteurs danken dr. E. J. Kappers voor het kritisch becommentariëren van een eerdere versie van dit artikel.

De resultaten met het instrument zijn beoedigend te noemen. Verschillen tussen zeer zwakke en normale lezers, o.a. gerelateerd aan het gebruik van de articulatorische code, blijken aantoonbaar. Implicaties voor verder onderzoek worden besproken.

1 Inleiding

Met dit artikel wordt de themaserie over instructie van leerlingen met ernstige leerproblemen afgesloten. Het past in de ontwikkelingslijn van het differentieel-diagnostisch onderzoek naar de kenmerken van het leerproces van leerlingen met ernstige leerproblemen (Van der Leij & Kappers, 1989). De bedoeling van het project is een instrument te construeren dat voor dergelijke leerlingen: a) de *instructiegevoeligheid* aangeeft: de mate waarin instructie effectief is (geweest), kwantitatief uitgedrukt in beheersingsniveaus van criteriumtoetsen, b) de *instructiebehoefte* indiceert, af te leiden uit de onderlinge verhouding van beheersingsniveaus (het 'leesprofiel') (o.a. de behoefte aan gebruik van verschillende modaliteiten van aanbieding (visueel/auditief) en van diverse mogelijkheden tot antwoord geven: hardop benoemen, naschrijven, beoordelen op gelijkens). Deze behoeften vormen de grondslag voor differentiatie van de instructie. Het instrument kan tevens gebruikt worden voor evaluatie van de effectiviteit van de instructie.

De belangrijkste aanleiding voor het onderzoek is geweest dat instrumenten die voor instructiegericht diagnostisch gebruik geschikt zijn, in het Nederlands taalgebied niet voor handen zijn, althans niet voor het stadium in het leesproces na het prille begin (i.c. het stadium van de automatisering). Kinderen met ernstige leesproblemen kunnen juist in dit stadium stagnaties vertonen in het leerproces (verg. o.a. Van der Leij, 1983; Van Aarle, 1988). Wat er aan instrumenten is, leent zich slechts voor algemene niveaubepaling (bijv. de EMT van Brus & Voeten, 1973) of is toegespitst op voorbereidend en het begin van het

aanvankelijk lezen (Sas & Wieringa, 1982; Struiksma, Van der Leij & Vieyra, 1987). Met de ontwikkeling van het instrument wordt getracht om het instrumentarium aan te vullen.

Dit artikel beschrijft het eerste gedeelte van het project, toegespitst op de vraag die ons primair relevant leek: in hoeverre indiceert het instrument verschillen in het ontwikkelingsverloop m.b.t. automatisering tussen leerlingen met ernstige leesproblemen en kinderen bij wie het leesproces zich op normale wijze ontwikkelt? Die verschillen vormen immers de grondslag voor realisering van het uiteindelijke doel: diagnostische informatie opleveren die aanpassing van de instructie aan de eigenaardigheden van kinderen met ernstige leesproblemen mogelijk maakt.

2 *Achterliggende theorie*

De resultaten van wetenschappelijk onderzoek (verg. o.a. Perfetti, 1985) indiceren dat er bij het leren lezen twee verwerkingswijzen van woorden zijn. In het beginstadium leert de leerling om woorden *om te zetten in de klankvorm* (ook wel fonologisch decoderen genoemd) en vervolgens verder te verwerken (koppelen aan de betekenis en/of het hardop uitspreken van de woorden). Daarnaast ontwikkelt zich in de loop van het leerproces een tweede verwerkingswijze. Steeds meer woorden worden via *directe woordherkenning*, dat wil zeggen zonder voorafgaande omzetting in de klankvorm, gekoppeld aan de betekenis op grond van ervaring in het lezen van die woorden. Beide processen worden in de loop van de eerste onderwijsjaren geautomatiseerd. Dat betekent dat de ervaren lezer in principe alle woorden snel en nauwkeurig kan verwerken. Bij bekende woorden geschiedt dat door directe herkenning, bij onbekende woorden door omzetting in de klankvorm, wat, vergeleken met bekende woorden, slechts een fractie van een seconde minder snel verloopt (Seymour, 1986). Kenmerkend voor dit automatische verloop is dat de processen geen of nauwelijks aanslag meer doen op de aandachtscapaciteit, die derhalve beschikbaar is voor andere activiteiten waarvan begrijpend lezen uiteraard de belangrijkste is.

Kinderen met ernstige leesproblemen blijven geen moeite te hebben met het waarnemen van visuele stimuli op zichzelf, zoals geschiedt

bij het beoordelen van visuele patronen (figuren; letters) op gelijkenis. Tekorten treden echter aan het licht wanneer visuele stimuli door middel van omzetting in een of andere klankcode verwerkt moeten worden, het eerste van de voornoemde twee processen (Bouma & Legein, 1980). Reitsma (1983) heeft aannemelijk gemaakt dat (zeer) zwakke lezers daarnaast grote moeite hebben met het zich eigen maken van 'woordspecifieke' kennis – de kennis van de unieke combinatie van lettertekens van een woord – die woordherkenning zonder letter-voor-letter omzetting in de klankvorm mogelijk maakt. (Normale lezers beginnen daarentegen al vroeg in het proces van leren lezen te beschikken over deze tweede verwerkingswijze.) Het gevolg is dat het automatiseren zowel van de omzetting in de klankvorm (het decoderen) als van de directe woordherkenning problematisch verloopt (Spear & Sternberg, 1987).

3 *Overwegingen bij de ontwikkeling van het instrument*

In algemene zin is voor de ontwikkeling van het instrument het idee als uitgangspunt genomen dat leerlingen met ernstige leesproblemen vooral stagnaties vertonen met het leren lezen van (losse)woorden. Stagnaties bij het lezen van zinnen en het begrijpend lezen zijn daarvan een afgeleide (Perfetti, 1985). Onze interesse ging primair uit naar het lezen van woorden die door veelvuldig gebruik in de kinderspreek- en schrijftaal qua betekenis in aanmerking komen om via directe woordherkenning te worden gelezen, het proces dat als tweede werd genoemd in het voorafgaande. De rationale achter de keuze voor dit proces was dat in de normale leesontwikkeling directe woordherkenning al vlak na het prille begin een steeds prominentere plaats inneemt en de belangrijkste pijler vormt onder het automatiseringsproces. De in de inleiding genoemde stagnatie in het stadium na het prille begin zou wel eens een gevolg kunnen zijn van het niet op gang komen van het automatiseringsproces door gebreken in de directe woordherkenning. Perfetti (in press) heeft aannemelijk gemaakt dat de opbouw van het zgn. *autonome lexicon* de belangrijkste voorwaarde voor automatisering in de zin van directe woordherkenning is. Dit lexicon bevat woorden

waarvan de verschillende aspecten in sterke mate geassocieerd zijn (orthografie, klankvorm, uitspraakvorm, betekenis). Een externe stimulus in de vorm van een van deze aspecten, faciliteert direct de toegang tot de andere aspecten die, afhankelijk van de bedoeling van het verwerkingsproces, in de vorm van opslag in het werkgeheugen of in de vorm van een antwoord kunnen worden omgezet. Woorden waarbij deze associaties nog niet (of niet allemaal) direct en onherroepelijk aangeleerd zijn, bevinden zich in het zgn. *functionele* lexicon.

Een manier om te weten te komen of een bepaald woord zich in het autonome lexicon bevindt, is de stimulus op verschillende wijzen aan te bieden en het antwoord op verschillende wijzen te laten geven. Zit het in het autonome lexicon, dan hebben dergelijke variaties geen invloed op de juistheid van het antwoord. Bovendien verloopt het proces snel en vrijwel zonder dat het aandacht vraagt. Voor de ontwikkeling van het instrument heeft dit idee tot gevolg gehad dat er verschillende *aanbiedings-* en *antwoordwijzen* werden ontworpen. Wat de aanbiedingswijzen betreft, een woord kan op twee manieren aangeboden worden: visueel-orthografisch en auditief. Voor het antwoord geven zijn er meer mogelijkheden. Bij de meest gebruikelijke antwoordvorm (hardop lezen) is dat de articulatie van het woord. Ter vergelijking werd spelling opgenomen omdat Perfetti (in press) voor het meten van het autonome lexicon spelling aanbeveelt: 'Reliable, confident, and facile spelling is an index of high quality representation. By hypothesis, it is also an index of autonomous representation' (p. 29). Echter, vanwege het betrekkelijk trage (want sterk sequentiële) antwoordproces bij spelling leek het ons raadzaam om ter vergelijking nog een derde antwoordwijze te introduceren. Vergelijken met hardop benoemen gaat spelling immers veel langzamer waardoor de invloed van bewuste antwoordstrategieën (bijv. analogie-regels toepassen) groter kan zijn. Bij autonome representatie wordt juist verondersteld dat dit soort strategieën niet meer toegepast hoeft te worden vanwege de automatisering van de verwerking. Besloten werd om ter vergelijking ook zogenaamde matchingstaken op te nemen die slechts ja/nee-beantwoording vergen en derhalve tijdens het antwoord geven geen aanslag hoeven te doen op bewuste ant-

woordstrategieën. Bij matchingstaken van twee elkaar opvolgende stimuli wordt beoordeeld of ze gelijk zijn of niet. Dit geschiedt zonder uitspraak en gaat sneller dan opschrijven. Gecombineerd met de twee aanbiedingswijzen zijn er in principe vier mogelijkheden: auditief-auditief; auditief-visueel, visueel-auditief en visueel-visueel matchen. Omdat er zich geen redenen opdrongen om a priori te selecteren, werd besloten om alle vier de mogelijkheden als subtoets in de experimentele toets op te nemen.

Een probleem als gevolg van de keuze voor onderzoek naar het autonome lexicon met verschillende aanbiedings- en antwoordwijzen was wel dat het niet mogelijk is *hetzelfde woord* op diverse manieren te laten verwerken, zonder dat er tijdens de testsessies een leereffect optreedt. Om die reden werd besloten om de analyse niet op individuele woorden uit te voeren, maar op woordklassen als geheel, waarbij de woordklasse bepaald werd door eenzelfde orthografische structuur. Kwamen leerlingen bijv. bij MKM-woorden (M = medeklinker, K = klinker) tot goede antwoorden, onafhankelijk van aanbiedings- en antwoordwijzen, dan werd die woordklasse beschouwd als opgenomen in het autonome lexicon. Daarnaast was het de vraag hoe er zeker van te zijn dat de stand van zaken in *directe woordherkenning* werd bepaald. Het is immers ook bij hoogfrequentie woorden mogelijk dat de leerling tot het goede antwoord komt door gebruik te maken van het eerste verwerkingsproces dat hij aanleert, omzetting in de klankvorm? Uit de theorie van perceptueel leren (verg. o.a. LaBerge & Samuels, 1976; Ehri & Wilce, 1983) is af te leiden dat een belangrijk aspect van het leren automatiseren de snelheid is waarmee de diverse informatieverwerkingsoperaties worden verricht. Zolang die nog traag verlopen, wordt dat als indicatie gezien van de hoeveelheid aandacht die de uitvoering kost, en daarmee van het niet-geautomatiseerd zijn. Een manier om na te gaan in hoeverre herkenningprocessen automatisch verlopen, is de tijd te meten die het geven van een accuraat antwoord kost. Echter, antwoord geven ontwikkelt zich continu sneller tot in de volwassenheid toe (Perfetti, in press) waardoor het uitgaan van antwoordtijden op jongere leeftijd minder aantrekkelijk is. Vanwege deze ontwikkeling zou het immers moeilijk zijn om een criterium

te bepalen ter indicatie dat de leerling zich 'op de goede weg' bevindt. Een alternatief is om de aanbiddingstijden van de visueel-orthografische stimuli (woorden) te verkorten. De rationale daarachter is, dat een leerling die woorden (nog) niet op grond van woordspecifieke kennis leest, tot het goede antwoord kan komen bij een relatief lange aanbiddingstijd. Hij voert dan de omzetting in de klankvorm uit terwijl de stimulus nog zichtbaar is. Echter, wanneer de stimulus slechts korte tijd beschikbaar blijft, is het noodzakelijk om snel – en dus tot op zekere hoogte automatisch – te werken. Besloten werd om de visuele stimulus met steeds kortere aanbiddingstijd te presenteren.

Een volgende overweging betrof het gebruik van het korte termijn geheugen, hier in navolging van Baddeley (1987) *werkgeheugen* genoemd. De sequentiële aanbidding van stimuli in de matchingstaken impliceert dat het werkgeheugen een rol in de verwerking speelt. Informatie van de eerste stimulus dient vastgehouden te worden tot de tweede stimulus vergelijking mogelijk maakt. Bij spelling is er, vanwege de relatief trage aard van het antwoord geven, ook een belasting van het werkgeheugen. Hoewel er volgens Spear en Sternberg (1987) nog geen definitieve bewijsvoering is op dit punt, wordt algemeen aangenomen dat het werkgeheugen dat gebruikt maakt van een fonetische code (een klankvorm dus), bij leerlingen met ernstige leesproblemen niet goed functioneert. Dat impliceert dat, naarmate het werkgeheugen meer belast wordt, deze leerlingen meer problemen met de taak zullen hebben. Het leek ons interessant om ook in dit tekort meer inzicht te krijgen door deze factor, die toch al opgenomen was door introductie van spellings- en matchingstaken, te variëren. Dit valt te doen door een langere periode tussen stimulus en response in te lassen zodat de stimulus langer in het werkgeheugen moet worden vastgehouden. Dat gebeurde in een tweetal subtoetsen.

Ten slotte zijn er keuzen gemaakt m.b.t. de inhoud en de vorm van het instrument. Vanwege de bedoeling om het aanleren van directe woordherkenning (i.c. het opbouwen van het autonome lexicon) in kaart te kunnen brengen, is *inhoudelijk* geconcentreerd op hetgeen door de normale leerling in de loop van de tweede helft van het eerste en de eerste helft van het tweede jaar van formeel leesonderwijs

wordt geleerd (resp. groep drie en vier jaar van het Basisonderwijs). Kennis van de letters is in de voorafgaande fase aangeleerd en wordt dus voorondersteld¹. Het materiaal bestond uit een- en tweelettergrepige woorden die in de bedoelde periode verreweg de grootste proportie van de aangeboden leesinhouden vormen.

Wat de *vorm* betreft is er om drie redenen voor gekozen om het instrument computer-gestuurd te maken. De methodologische voordelen kunnen in een onderzoek groot geacht worden in vergelijking met traditionele toetsen door de mate van controle op input en output (Torgesen, 1986). De eerder genoemde keuzen m.b.t. aanbiddings- en antwoordwijzen maakte een grote mate van controle gewenst. Bovendien hebben we plannen om op langere termijn te komen tot een combinatie van de toets met het Computergestuurde Orthodidactische Programma voor Aanvankelijk Lezen (COPAL: Van Daal, Van der Leij, Bakker & Reitsma, 1987). Aansluiting van diagnostische informatie en instructie kan dan via één medium geregeld worden. Ten slotte, de praktische voordelen van micro-computergebruik ten behoeve van gedifferentieerde instructie van kinderen met ernstige leesproblemen zijn evident, al behoeft de software-ontwikkeling de nodige aandacht (Reitsma, 1985). Het instrument heet de Computergestuurde Toets voor Automatisering van het Lezen (COTAL).

4 Onderzoeksopzet

4.1 Vraagstelling

Nagegaan is in hoeverre de COTAL-subtoetsen verschillen aan het licht brengen tussen leerlingen met ernstige leesproblemen en leerlingen met een normale leesontwikkeling. Door proefgroepen van (kwantitatief) vergelijkbaar algemeen leesniveau te selecteren, zijn eventueel optredende verschillen, te wegegebracht door de diverse aanbiddings- en antwoordwijzen, op te vatten als indicaties voor verschillen in de kwaliteit en het ontwikkelingsverloop van het proces van directe woordherkenning.

Tabel 1 *Proefpersonen*

groep	aant.	j	m	chron.leeft. ^a		EMT ^b		dle ^c	
				gem.	s.d.	gem.	s.d.	gem.	s.d.
ZZ	37	31	6	126.8	17.1	24.20	9.51	11	0.4
N	37	18	19	86.9	4.1	24.67	9.61	11	0.4
totaal	74	49	25						

a Leeftijd in maanden.

b Weergegeven is de ruwe score van de EMT

c Didactische leeftijdsequivalent (dle) is de didactische leeftijd waarop de betreffende EMT-score gemiddeld in het BaO ghaald wordt (verg. Struiksma, Van der Leij & Vieyra, 1987). Een dle van 11 wil zeggen: na een maand onderwijs in groep vier.

4.2 *Proefpersonen*

De proefpersonen met ernstige leesproblemen waren zeer zwakke lezers (ZZ) van vier LOM-scholen, de normale lezers (N) leerlingen van groep drie van vier basisscholen in de laatste maand van hun schooljaar (door de vakantiespreiding was dit eind juni/begin juli). Uit beide proefgroepen werden de leerlingen verwijderd die een EMT-score (Brus & Voeten, 1973) hadden < 10 teneinde problemen met letterkennis uit te sluiten, of > 40 om te voorkomen dat de leerlingen te goed waren. Aldus bleven 51 zeer zwakke en 66 normale lezers over. Proefpersonen in het hier gerapporteerde onderzoek zijn 37 paren van een normale en een zeer zwakke lezer, geselecteerd uit de laatstgenoemde groepen. Deze paarsgewijze matching op algemeen leesniveau was gebaseerd op vergelijkbaarheid op drie criteria: 1. EMT-ruwe score met een tolerantie voor verschil van plus of min twee punten, 2. algemene niveauscore volgens de AVI-toetsen van Van den Berg en Te Lintelo (1977) en 3. beheersing van het hardop lezen van MKM-woorden (een COTAL-subtoets die nog beschreven wordt; dit laatste criterium diende als extra controle op de onderdrempel: wanneer MKM-woorden beheerst worden, beheerst de leerling alle letters). Van matching op sekse is afgezien omdat er geen reden was om aan te nemen dat er tussen jongens en meisjes voor dit onderzoek relevante verschillen bestaan (verg. Van der Leij, 1983)². Het gevolg was dat de scheve verhouding tussen jongens en meisjes die gebruikelijk is voor LOM-scholen, weerspiegeld wordt door de samenstelling van de ZZ-proefgroep (Tabel 1).

De gemiddelde leeftijd van de ZZ-groep was die van leerlingen aan het einde van groep

zes van de basisschool, terwijl het leesniveau lag op begin groep vier: een gemiddelde achterstand van bijna drie jaar.

4.3 *Materiaal*

De COTAL bestaat uit subtoetsen die zijn ontwikkeld door variatie van drie dimensies: a) de aanbiedings- en antwoordwijze, b) de moeilijkheidsgraad van de aanbiedingswijze binnen de subtoets en c) de woordklasse.

a) *Aanbiedings- en antwoordwijze.* Er is een negental taken ontworpen waarvan er twee zijn afgevallen na de eerste toetsing³. In deze taken werd de stimulus (een woord) visueel of auditief aangeboden. De leerlingen antwoorden op een van de volgende drie manieren: door beoordeling van de gelijkens met een daarop volgende stimulus ('matchen'), door natypen, of door hardop benoemen. Omdat voor een computergestuurde toets is gekozen, kreeg natypen via het toetsenbord de voorkeur boven opschrijven. Eventuele invloed van problemen met de schrijfmotoriek werd daarmee tevens uitgesloten. Tabel 2 geeft een overzicht van de subtoetsen.

b) *Moeilijkheidsniveau.* Het verkorten van de aanbiedingstijd van de *visuele* stimulus werd als bemoelijkende factor opgenomen ter indicatie van de mate waarin woorden direct herkend worden. In vier subtoetsen (V1-V2, V-T, V-A, Fl.) liep het niveau van moeilijkheid op door een aanbiedingstijd van de visuele stimulus (V) van resp. 200 msec., 160 msec. en 100 msec..

Een andere bemoelijkende conditie was het verlengen van de tijd tussen de *auditieve* stimulus en wat erop volgde (een visuele sti-

Tabel 2 *Subtoetsen en moeilijkheidsniveaus*

subtoetsen ^a		moeil. niveaus			betekenis niveaus
		1	2	3	
V1-V2	visueel matchen ^b	200	160	100 ^c	aanb. tijd V1 en V2 neemt af
A-V	aud.-vis. match.	1''	3''	5''	tijd tussen A en V neemt toe
V-A	vis.-aud. match. ^b	200	160	100	aanb. tijd V neemt af
V-T	visueel, typen	200	160	100	aanb. tijd V neemt af
A-T	auditief, typen	1''	3''	5''	tijd tussen A en T neemt toe
Fl.	hardop lezen flitswoord	200	160	100	aanb. tijd V (Fl.) neemt af
L	hardop lezen woord	-	-	-	geen niveaus

a A = auditieve stimulus; V = visuele stimulus; T = natypen; Fl. = hardop lezen flitswoord; L = hardop lezen

b Bij V1-V2 en V-A is de tijd tussen eerste en tweede stimulus: 1.5''

c Bij de visuele stimulus is de aanbidingstijd in msec. gegeven.

mulus om mee te matchen of de mogelijkheid tot typen: resp. A-V en A-T), ter toetsing van de invloed van belasting van het werkgeheugen. Gekozen werd voor drie niveaus: resp. met een tussentijd van 1, 3, en 5 seconden⁴.

De zevende subtoets, L (woorden hardop lezen), was een gewone leestaak waarbij de stimulus op het scherm bleef staan totdat het antwoord was gegeven. Er waren bij deze subtoets dus geen moeilijkheidsniveaus.

c) *Woordklassen.* Alle stimuli waren hele, betekenisvolle woorden, zoveel mogelijk hoogfrequent. (Vanwege het grote aantal woorden dat voor de items nodig was, lukte het niet om uitsluitend met hoogfrequente woorden te werken; wat er aan laagfrequente woorden opgenomen werd, werd at random over de subtoetsen en moeilijkheidsniveaus verdeeld). Er werden drie soorten eenlettergrepige en drie soorten tweelettergrepige woorden opgenomen: a) MKM ('kop'), b) MKMM ('dorp'), c) MMKM ('stad'), d) twee lettergrepen: zonder regel ('keuken', 'deurmat'), e) gesloten ('lippen'), f) open ('bomen'). Voorondersteld werd dat de moeilijkheidsgraad van deze woordklassen van a) naar f) oploopt. Deze vooronderstelling was o.a. gebaseerd op het feit dat in leesmethoden deze volgorde gehanteerd wordt en op resultaten van eerder onderzoek (verg. Van Daal et al., 1987).

Items. De woorden werden als items per aanbodings- en antwoordwijze⁹, moeilijkheidsniveau en woordklasse per leerling at random geselecteerd uit een bestand dat opgebouwd was bij de ontwikkeling van COPAL (Van Daal et al., 1988). Voor de matchings-

taken zijn bij elk woord drie sterk gelijkende – zowel auditief als visueel – woorden gezocht. De COTAL bestond aldus uit 6 maal 6 maal 3 is 108 onderdelen (aantal aanbodings- en antwoordwijzen X aantal woordklassen X aantal moeilijkheidsniveaus). Daarbij kwamen nog de onderdelen van de taak zonder niveaus (L): 6 woordklassen. Totaal 114 onderdelen met 10 items elk. (Voor een vollediger beschrijving wordt verwezen naar Van Daal, Bakker, Van der Leij, Reitsma & Smeets, 1988).

4.4 *Werkwijze*

Aanbiding. De subtoetsen werden per microcomputer (en hoofdtelefoon) aangeboden (Macintosh Plus met harddisk). De visuele stimuli werden in het midden van het 17.5 X 11.5 cm grote scherm geprojecteerd met een grote, eenvoudige leesletter. Na aanbiding (200, 160 of 100 msec.) werden de visuele stimuli gemaskeerd m.b.v. een abstracte figuur die onmiddellijk, maar zeer kort, op het scherm verscheen. Dit had tot doel om de werking van het zogenaamde nabeeld op het netvlies te neutraliseren en zodoende de proefpersonen geen kans te geven om visuele kenmerken iets langer te analyseren.

De auditieve stimuli werden per hoofdtelefoon aangeboden. Het bestand aan klankvormen was verworven door de woorden op een audioband te laten inspreken door een logopediste. Door middel van een 'analog-to-digital converter' werd de natuurlijke spraak vervolgens gedigitaliseerd in de harde schijf opgeslagen.

Procedure. De subtoetsen werden stuk voor stuk afgewerkt onder leiding van een proefleider. De score was de accuratesse van de antwoorden (de instructie benadrukte dat element en niet de snelheid van antwoorden).

Antwoord gaven de leerlingen door:

- het pijltje met behulp van de 'muis' naar de juiste plaats op het scherm te brengen en de toets van de 'muis' in te drukken (matchen: ja = gelijk; nee = ongelijk),
- het toetsenbord te gebruiken waarop de relevante letters waren gemarkeerd met witte plakletters en de andere dysfunctioneel gemaakt (typen),
- hardop te antwoorden.

De proefleider toetste de juistheid van dit laatste soort antwoord in en regelde het goede verloop van de toetsafname, inclusief oefentrials.

Om het beheersingsniveau gerelateerd aan het oplopen van de moeilijkheidsgraad (woordklasse; moeilijkheidsniveau) te kunnen vaststellen, werd besloten om binnen elke subtoets de onderdelen aan te bieden in een volgorde van relatief gemakkelijk naar relatief moeilijk ('bloksgewijs'). Er zijn regels geconstrueerd voor instappen, afbreken en teruggaan naar een lager niveau teneinde het afnemen van te gemakkelijke of te moeilijke onderdelen zoveel mogelijk te voorkomen. Om daarvan een voorbeeld te geven, de V1-V2-subtoets werd als eerste afgenomen met MKM-woorden en een aanbiedingstijd van resp. 200 msec., 160 en 100 msec. Daarna volgden de MKMM-woorden, etc. Beheerste de proefpersoon deze subtoets op MKMM-niveau, dan werd bij de volgende subtoets ook begonnen met MKMM-woorden. Faalde hij toch, dan werd teruggegaan naar MKM-woorden.

Proefpersonen werden niet langer dan vijf tot tien minuten per dag getoetst. Het aantal sessies per proefpersoon beliep vier à vijf dagen.

Beheersingscriterium. Een van de gevolgen van 'bloksgewijze' aanbieding was dat er per onderdeel een criterium voor beheersing moest worden gehanteerd. Besloten werd dit op 90% te stellen om een toevallige fout niet te zwaar te laten wegen. Wanneer een leerling negen van de tien items goed had werd dat onderdeel als beheerst beschouwd en kreeg de score 1. Werd dit criterium niet gehaald dan was de score 0.

De scores zijn per subtoets en woordklasse getransformeerd naar scores van 0 tot en met 5: wanneer een leerling een subtoets bij een bepaalde woordklasse beheerste op alle niveaus van moeilijkheid, kreeg hij de score 5. Over zes woordklassen heen leverde dit per subtoets een maximale score-mogelijkheid van 30 op. Ook de scores van de subtoets 'hardop lezen' die geen niveaus kent (L), zijn op deze wijze getransformeerd.

5 Resultaten

5.1 Gemiddelde accuratesse per subtoets

In Tabel 3 is de gemiddelde accuratesse voor alle subtoetsen weergegeven op twee wijzen: gesommeerd over alle woordklassen en voor MKM-woorden apart. In beide gevallen is er gesommeerd over de drie moeilijkheidsniveaus (behalve 'hardop lezen' (L) waarvoor maar één niveau bestaat).

Alle woordklassen. Geconcludeerd kan worden dat de subtoetsen de nodige moeite opleverden. De maximale score van 30 werd bij geen van de subtoetsen benaderd: er zijn geen plafond-effecten. Dat valt ook af te leiden uit het feit dat geen enkele proefpersoon alle subtoetsen beheerste (bij alle woordklassen en op alle moeilijkheidsniveaus)⁵. Wel is er een tendens naar een bodemeffect, t.w. in het geval van V-T en A-T bij N-lezers.

Over de subtoetsen heen getoetst met een multivariate variantie-analyse is het verschil tussen de twee groepen significant ($F = 7.00$, $p < .001$). De ZZ-groep was op alle matchings- en type-taken beter dan N, de N-groep was beter in hardop lezen (L), de groepen verschillen niet in flitslezen (FL). De verschillen bij auditief-visueel matchen (A-V), en auditief en visueel typen (A-T en V-T) waren univariaat getoetst significant. Voor visueel matchen en hardop lezen (V1-V2 en L) was er een sterke tendens tot significantie op 5% niveau.

MKM-woorden. Interessant is het om de gegevens van de MKM-woorden, de eenvoudigste woordklasse, apart te analyseren. Gegeven dat een der drie criteria die als drempel fungeerden in de selectie, de beheersing betrof van het hardop lezen (L) van MKM-woorden, zou het immers kunnen zijn dat ook andere subtoetsen bij deze woordklasse door de leerlingen beheerst werden.

Tabel 3 Gemiddelden en standaarddeviaties van ZZ en N per subtoets, gesommeerd over drie moeilijkheidsniveaus, resp. voor alle woordklassen en MKM-woorden

subtoets	alle woordklassen ^a		F	p	MKM-woorden		F	p
	ZZ	N			ZZ	N		
V1-V2	7.2 (8.1)	3.4 (7.8)	3.45	n.s. ^b	2.6 (1.8)	0.8 (1.5)	20.59	<.001
A-V	17.2 (8.6)	10.8 (12.2)	7.23	<.01	3.7 (1.6)	0.8 (1.5)	11.82	<.001
V-A	8.5 (7.8)	5.5 (8.2)	1.98	n.s.	3.3 (1.7)	1.7 (2.0)	13.00	<.001
V-T	4.9 (6.4)	1.5 (3.3)	7.16	<.01	2.2 (1.9)	0.9 (1.7)	9.31	<.01
A-T	10.8 (9.6)	2.1 (5.0)	22.76	<.001	3.7 (1.6)	0.8 (1.5)	66.36	<.001
Fl.	6.0 (7.1)	5.9 (7.9)	0.00	n.s.	2.8 (1.8)	2.1 (2.2)	1.78	n.s.
L ^c	13.0 (8.9)	16.8 (7.8)	3.78	n.s. ^d	- ^e	-	-	-
Max. sc.	30	30			5	5		

a Standaarddeviaties tussen haakjes

b $p = .064$

c Deze subtoets kent geen moeilijkheidsniveaus.

d $p = .056$

e Het hardop lezen (L) van MKM-woorden werd als onderdrempel in de selectie gebruikt. Alle 74 leerlingen beheersten dit (score 5.0).

Tabel 3 toont echter dat dit bij geen van de andere subtoetsen voor de groepen als geheel het geval was. Nadere analyse leert, dat slechts vier van de 37 ZZ-lezers en niet één N-lezer alle subtoetsen bij MKM-woorden op alle niveaus van moeilijkheid beheersten.

De verschillen tussen ZZ en N waren bij MKM-woorden nog opvallender dan bij alle woordklassen tesamen (multivariaat getoetst: $F = 15.06$, $p < .001$). ZZ was significant beter in de matchings- en type-taken (voor de univariate toetsing, zie Tabel 3); het verschil in flitslezen was echter gering en niet significant.

5.2 Analyse van kwalitatieve verschillen

Teneinde na te gaan in hoeverre de verschillende aanbiedings- en antwoordwijzen, moeilijkheidsniveaus en woordklassen significante verschillen te zien geven tussen de groepen is een aantal multivariate interactie-analyses uitgevoerd (MANOVA's) met deze factoren als binnenproefpersoonvariabelen en ZZ/N als tussenproefpersoonvariabele. Voor de interpretatie is voorts gebruik gemaakt van univariate F's en de gemiddelde scores (zie o.a. Tabel 3). Omdat uit de gegevens van Tabel 3 is op te maken dat de beide typetaken voor de N-

groep te moeilijk waren (op dit punt wordt teruggelaten), zijn ze uit de analyse naar de invloed van antwoordwijze weggelaten.

Aanbiedingswijze. De interactie tussen ZZ/N en aanbiedingswijze (eerst auditief of visueel, twee niveaus) is onderzocht door A-V te contrasteren met V-A en bleek significant (F (df 2, 70) = 3.57; $p = .03$). Uit Tabel 3 valt op te maken dat dit komt omdat de ZZ-lezers relatief meer profiteren van de (eerst) auditieve aanbieding dan de N-lezers (het verschil was bij A-V significant, bij V-A niet).

Antwoordwijze. Over alle woordklassen gerekend, is de interactie tussen ZZ/N en matchen (V1-V2 en V-A gesommeerd) versus hardop benoemen (L en Fl. gesommeerd), significant (F (df 2, 70) = 8.87, $p < .0005$). Uit Tabel 3 is af te leiden dat dit enerzijds komt door een bijna significant verschil ten gunste van ZZ bij de V1-V2 subtoets, anderzijds door een bijna significant verschil ten gunste van N bij de L-subtoets.

Moeilijkheidsniveaus. Verkorting van visuele stimulus interacteert met ZZ/N wanneer de onverkorte aanbieding (L) en Flitslezen bij 200, 160 en 100 msec. (vier niveaus) in de analyse worden betrokken (F (df 4, 69) = 4.32;

$p < .005$). Wanneer de berekening wordt gereduceerd tot onverkort aanbieden/aanbieding van 200 msec. (twee niveaus), blijft de interactie significant ($F(df\ 2, 71) = 3.97; p = .02$). Op het niveau van 160 en 100 msec. is dat niet het geval. Gegeven het op 5% niveau bijna significante verschil ten gunste van N op de L-subtoets (Tabel 3) indiceert dit dat normale lezers relatief meer moeite hebben met verkorting van de stimulus tot 200 msec. dan de zeer zwakke lezers.

Verlenging van de tijd tussen de auditieve stimulus en de visuele stimulus (A-V, drie niveaus: 1'', 3'', 5'') interacteert ook met ZZ/N ($F(df\ 3, 70) = 11.76; p = .0001$), vooral ten gevolge van een relatief betere score van ZZ bij op het 1'' niveau ($F = 15.04; p = .0002$) en 3'' niveau ($F = 5.92; p < .02$). Bij 5'' zijn er geen significante verschillen tussen de groepen. De hogere score van ZZ bij A-V is dus niet toe te schrijven aan een surplus in de werking van het werkgeheugen – dan zou het verschil er ook zijn bij 5'' – maar aan het profijt dat ze trekken van auditieve aanbieding op het niveau dat het werkgeheugen minder belast wordt.

Woordklasse. Woordklasse (zes niveaus) interacteert met ZZ/N ($F(df\ 6, 67) = 9.65; p = .0001$), geheel toe te schrijven aan de relatief betere score van ZZ bij de eenvoudigste woorden (MKM) ten opzichte van gelijkwaardiger scores bij de moeilijker woordklassen. Deze uitspraak geldt echter wanneer alle subtoetsscores worden ingecalculleerd. Wanneer de twee subtoetsen (A-V, L) geselecteerd worden die, blijkens het voorafgaande, de duidelijkste verschillen te zien geven, is er een significante driewegsinteractie tussen ZZ/N, woordklasse en A-V/L. Daarbij zijn de woordklassen teruggebracht tot drie niveaus (MKM, MKMM + MMKM, tweelettergrepige woorden): $F(df\ 5, 67) = 3.87; p = .004$. Uit de univariate F 's en de gemiddelde scores valt op te maken dat zeer zwakke lezers met de A-V-subtoets significant beter zijn dan N wanneer het om eenlettergrepige woorden gaat (MKM en MKMM + MMKM; resp. $F = 11.82; p = .001; F = 4.12; p < .05$) maar niet bij tweelettergrepige woorden. Normale lezers zijn daarentegen relatief beter in het hardop lezen van MKMM + MMKM-woorden ($F = 5.95; p = .02$), maar niet van tweelettergrepige woorden. Bij het hardop lezen van MKM-woorden waren de groepen,

ten gevolge van de matchingsprocedure, gelijk. Waarschijnlijk zijn tweelettergrepige woorden voor beide groepen te moeilijk. Belangrijker is echter de constatering dat wanneer er eenlettergrepige woorden van MKMM- of MMKM-structuur worden aangeboden, zeer zwakke lezers significant beter zijn dan normale wanneer eerst een auditieve stimulus aangeboden wordt, te matchen met een visuele. Voor hardop lezen geldt juist dat normale lezers bij MKMM- en MMKM-woorden significant beter zijn. In de discussie wordt hierop teruggekomen.

5.3 *Mate van beheersing*

Tot nu toe betreffen de analyses de gemiddelde scores van de groepen. Hoewel de data interessante verschillen tonen, wordt daarmee geen inzicht verschaft in de mate waarin de leerlingen de subtoetsen op de diverse moeilijkheidsniveaus beheersen. Aangezien er criteriumtoetsen met afbreekregels gebruikt zijn, is ook dat te demonstreren, niet door de gemiddelde scores maar door de aantallen leerlingen die tot beheersing komen. Het is niet doenlijk om in het kader van een artikel een compleet overzicht hiervan te geven. Ter illustratie staan in Tabel 4 de aantallen leerlingen weergegeven die de subtoetsen beheersten waarbij hardop moest worden gelezen, apart per woordklasse (voor Fl. en L) en moeilijkheidsniveaus (voor Fl.).

Behalve bij het hardop lezen van MKM-woorden (L) werden de L- en Fl.-subtoetsen niet beheerst door alle leerlingen. Uit Tabel 4 blijkt dat er een groot verval was in aantal leerlingen dat de subtoetsen beheerste naar mate 1. de orthografische structuur van de woordklasse gecompliceerder is, en 2. het moeilijkheidsniveau van de visuele stimulus-aanbieding oploopt.

De eerste tendens, het *woordklasse-effect*, is geleidelijker bij de normale leerlingen, dan bij de zeer zwakke lezers die vooral met het hardop lezen van MKMM- en MMKM-woorden relatief meer moeite hadden. Dit grotere verval bij ZZ is in zekere mate ook terug te vinden in de scores op de flitsleestaak (Tabel 4) en de andere taken (niet weergegeven) en verklaart de bij woordklasse genoemde interacties. Dit is ook de reden dat het verschil tussen ZZ en N in Tabel 3 groter is bij MKM-woorden alleen dan bij alle woordklassen. Omgekeerd kan men ook zeggen dat ZZ-

Tabel 4 Aantallen leerlingen die de L- en Fl.subtoetsen beheersen, per woordklasse en moeilijkheidsniveau (max. = 37)

taak: woordklasse	L ^a		Fl. ²⁰⁰		Fl. ¹⁶⁰		Fl. ¹⁰⁰	
	ZZ	N	ZZ	N	ZZ	N	ZZ	N
MKM ^b	-	-	32	21	20	16	10	10
MKMM	22	31	17	15	9	12	5	9
MMKM	15	27	9	12	7	9	2	5
2 l.gr.	10	14	6	5	3	4	0	3
gesl. l.gr.	9	12	5	4	4	4	2	2
open l.gr.	3	3	2	1	1	1	1	0

a Deze subtoets kent geen moeilijkheidsniveaus.

b Het hardop lezen (L) van MKM-woorden werd als onderdrempel in de selectie gebruikt. Alle 74 leerlingen beheersten dit.

lezers beter waren dan N-lezers in het verwerken van MKM-woorden mits de response geen articulatie vereiste.

De tweede tendens, het *flitsaanbiedingseffect*, is juist geleidelijker bij ZZ wanneer L (moeilijkheidsgraad nul vanwege onbepaalde aanbidingstijd) wordt vergeleken met de reeks Fl.²⁰⁰ tot Fl.¹⁰⁰. Voor alle woordklassen geldt dat de verkorting van de aanbidingstijd tot 200 msec., vergeleken met een onbepaalde aanbidingstijd, meer N-leerlingen moeite gaf dan de ZZ-leerlingen, conform de bij moeilijkheidsniveaus eerder genoemde interactie. Omgekeerd kan men ook zeggen dat de N-lezers beter waren dan ZZ-lezers in het hardop lezen van woorden, die zichtbaar bleven.

Tabel 4 demonstreert behalve een ondersteuning van de interacties, nog twee punten. Ten eerste, de groepen van leerlingen verdelen zich keurig over de beheersingsniveaus van deze twee subtoetsen: het eenvoudigste niveau werd door allen beheerst en het moeilijkste door slechts 1 (uit de ZZ-groep). Het oplopen van het moeilijkheidsniveau door verkorting van de visuele aanbiding en aanbiding van moeilijker woorden doet de een na de ander afvallen. Daarmee is het diagnostisch differentiërende aspect van deze subtoetsen in kwantitatieve zin (van meer naar minder beheersing) evident. Ten tweede, inspectie van Tabel 4 leert ook dat de beheersingspatronen van ZZ en N over het geheel genomen, verschillen. Alle leerlingen beheersten het hardop lezen van MKM-woorden. Echter, van de 37 ZZ-leerlingen deden 32 dat ook bij Fl.²⁰⁰, maar slechts 22 bij het hardop lezen van MKMM-woorden. Van de 37 N-leerlingen beheersten er slechts 21 MKM-woorden bij

Fl.²⁰⁰, maar 31 het hardop lezen van MKMM. Deze verschillen zijn er resp. ten gevolge van het woordklasse- en het flitsaanbidingseffect. Combinatie met andere subtoetsen (A-V, V1-V2) ondersteunt het idee dat het instrument kwalitatieve verschillen in beheersingspatroon tussen ZZ en N indiceert, niet alleen als groep maar ook individueel.

6 Discussie en conclusies

6.1 Kwantitatieve analyse

Met betrekking tot de eerste bedoeling van het instrument – instructiegevoeligheid indiceren, uitgedrukt in beheersingsniveaus van criteriumtoetsen (zie de inleiding) – kan als voorlopige conclusie getrokken worden dat het instrument in *kwantitatief* opzicht bevredigende resultaten te zien geeft. De subtoetsen waren gemiddeld bepaald niet te gemakkelijk en er was een grote spreiding van scores van proefpersonen op een gemiddeld algemeen leesniveau van begin groep vier. Dat er geen plafondeffect is blijkt bijvoorbeeld uit de spreiding die Tabel 4 te zien geeft. Wordt de eenvoudigste taak (het hardop lezen van MKM-woorden) nog door 100% van de leerlingen goed gedaan – een gevolg van de selectieprocedure – het aanbieden van moeilijker woordklassen en kortere stimulustijden reduceert deze maximale proportie vrijwel tot nul. Sommige subtoetsen zijn echter aan de moeilijke kant: de beide type-taken leverden bij de N-groep lage scores op, ook bij MKM-woorden. Vanwege dit bodemeffect zijn deze subtoetsen niet in de interactie-analyses betrokken. Verondersteld kan worden dat de

ontwikkeling van de processen die betrokken zijn in deze taken in het normale geval plaatsvindt in de periode na de toetsperiode (eind groep drie) en dat de invloed van instructie daar groot op is. Dat betekent dat op den duur het geconstateerde verschil aan betekenis verliest, omdat het niet te herleiden is tot specifieke tekorten (bij N) of compensaties (bij ZZ) in onderliggende verwerkingsprocessen, maar tot een gebrek aan ervaring van de normale lezers. Ook op grond van de literatuur valt niet te verwachten dat zeer zwakke lezers superieur zouden zijn in spelling door toedoen van een surplus in capaciteiten (zie bijv. Seymour, 1986). Ze hebben in het algemeen met spelling (minstens) evenveel moeite als met hardop lezen.

Duidelijk is voorts dat de proefpersonen, op een paar uitzonderingen na, zelfs de eenvoudigste woorden (MKM) niet bij alle aanbiedings- en antwoordwijzen beheersten. Conform de ideeën van Perfetti (in press) indiceert dit dat ook deze woordklasse zich nog in het stadium van het functionele lexicon bevindt. Wanneer het ontwikkelingsproces normaal verloopt, valt te verwachten dat in de loop van de leerjaren die volgen op de toetsperiode (eind groep drie), de woordklassen een voor een in het autonome lexicon terecht komen. Dit is af te leiden uit het feit dat de woordklassen beheerst worden, onafhankelijk van aanbiedingsantwoordwijze en moeilijkheidsniveau. Het zal pas mogelijk zijn om een uitspraak te doen over de validiteit van de operationalisering van het begrip autonoom lexicon in de vorm van het instrument, wanneer deze verwachting uit blijkt te komen. Dit vereist vervolgonderzoek bij dezelfde leerlingen (zie de afsluitende paragraaf).

6.2 *Kwalitatieve analyse*

De tweede bedoeling van het instrument, indicatie van *kwalitatieve* verschillen tussen zeer zwakke en normale lezers van gelijk leesniveau, vormde de hoofdvraag van het gerapporteerde onderzoek. Uit de eerste analyse kan worden afgeleid dat 1. de ZZ-groep beter was in auditief-visueel matchen (A-V), 2. hardop lezen (L) door de normale lezers beter werd gedaan, 3. een grotere proportie van de ZZ-groep moeite had met het oplossen van de orthografische moeilijkheid van de (eenlettergeregde) woordklassen, en 4. een groter deel van de N-groep moeite had met het oplossen

van de moeilijkheid door verkorting van de aanbestedingstijd (de overgang van onbeperkte aanbesteding naar msec.). Uit deze opsomming kan geconcludeerd worden dat het variëren van aanbestedings- en antwoordcondities, moeilijkheidsniveaus en woordklassen verschillen in beheersingspatronen tussen zeer zwakke en normale lezers van hetzelfde leesniveau aan het licht brengt. In het navolgende zal geen poging gedaan worden deze gegevens te integreren tot een differentiële leestheorie voor normale en zeer zwakke lezers. Ontwikkeling van de beheersing van de leesprocessen, geïndiceerd door het vervolgonderzoek, zal daarover meer uitsluitsel kunnen geven. We beperken ons tot een specificatie en een voorlopige interpretatie van de gevonden verschillen.

Ad 1. Wat de A-V-taak betreft is het niet waarschijnlijk dat een verschil in ervaring het verschil bepaalt, zoals bij de type-taken het geval is: een dergelijke taak komt in het onderwijs niet of slechts sporadisch voor. Het lijkt er meer op dat de auditieve aanbesteding van de articulatorische vorm woordherkenning meer *faciliteert* bij zeer zwakke lezers dan bij normale lezers van gelijk leesniveau. De moeite die een behoorlijk groot deel van de ZZ-leerlingen blijkt te hebben met de overgang van MKM- naar MKMM- en MMKM-woorden (zie ad 3.) is in deze taak ook veel minder evident. Ter illustratie: op de A-Vsubtoets (niveau 1) beheersten 37 van de ZZ-leerlingen de MKM-woorden, 33 de MKMM-woorden en 32 de MMKM-woorden. Ter vergelijking: de aantallen die genoemd werden in Tabel 4 m.b.t. de hardop leestaak (L) zijn resp. 37, 22 en 15.

Het verschil bij A-V ten gunste van ZZ kan niet herleid worden uit een surplus in de werking van het *werkgeheugen* bij de zeer zwakke lezers, zoals de interactie-analyse liet zien. Wanneer er meer aanslag op dat geheugen wordt gedaan omdat de informatie 5'' moet worden vastgehouden, verdwijnen de verschillen immers. Het lijkt erop dat, zodra het werkgeheugen langduriger belast wordt er tussen de 3'' en 5'' aanbesteding een kritische grens wordt gepasseerd. De informatie in het werkgeheugen is niet meer zo duidelijk aanwezig en het verwerken van de visuele stimulus (dat veel aandacht kost) interfereert daarmee. In het vervolgonderzoek zal nagegaan worden in hoeverre dit een stabiel kenmerk

van zeer zwakke lezers is dat afwijkt van het normale ontwikkelingspatroon. Dan zal het mogelijk zijn om enig licht te werpen op de eerder genoemde meningen van Baddely (1987) en Spear & Sternberg (1987), die op dit punt contrasteren⁶.

Ad 2. Aangezien de leerlingen paarsgewijs gematched waren op het hardop lezen (op drie criteria: EMT, AVI, beheersing hardop lezen van MKM-woorden), is het resultaat ten gunste van de normale lezers bij het *hardop lezen* (L) opmerkelijk te noemen. Het moge echter duidelijk zijn dat de EMT en AVI niet identiek zijn aan L. Bij de EMT gaat het om het lezen van woorden in rijtjes, bij AVI om woordjes in zinnen. Snelheid speelt in beide toetsen een rol: bij de EMT is de score het aantal goed gelezen woorden in één minuut en bij AVI is er, per kaart, een beheersingscriterium voor de snelheid (en voor de accuratesse). De hardop lees-toets L betreft daarentegen het lezen van stuk voor stuk aangeboden woorden. Het snelheidsaspect speelt in de score geen rol, het gaat alleen om het behalen van het 90% goed criterium per woordklasse. In de EMT en AVI zijn bovendien allerlei processen van invloed die niet uitsluitend met het lezen van woorden te maken hebben: o.a. het verspringen van regel (vooral bij de EMT) en het gebruik maken van de context (AVI). Men kan L beschouwen als een zuiverder toets voor het lezen van losse woorden dan EMT en AVI.

Univariate toetsing over de woordklassen heen gaf bij L een marginaal significant verschil te zien (Tabel 3; $p = .056$). Uit nadere analyse per woordklasse bleek echter dat de verschillen bij MKMM- en MMKM-woorden wel significant waren (verg. ook de aantallen leerlingen die tot beheersing kwamen in Tabel 4.) Voor de interpretatie van het resultaat is het van belang te benadrukken dat, in tegenstelling tot de matchingssubtoetsen, in de hardop leestoets articulatie betrokken is. Gevozen kan worden op de parallel met de problemen die zeer zwakke lezers hebben met taken die het benoemen van overbekende stimuli vereisen (bijv. van cijfers, letters of plaatjes). Baddeley's (1987) conclusie dat het omzetten in de articulatorische code als onderdeel van het productieproces bij slechte lezers een zwak punt is, krijgt door ons resultaat ondersteuning (verg. ook Spring & Capps, 1974; Spring & Perry, 1983). Echter, Baddeley beschrijft de problemen als snelheidspro-

blemen. Aangezien snelheid van articulatie niet de score was waarmee we werkten, indiceert het resultaat met de L-subtoets o.i. dat zeer zwakke lezers ook met de *accurate productie* van de articulatorische code meer moeite hebben dan normale lezers van overeenkomstig leesniveau. Dat het een productie- en geen herkenningsprobleem is ontlenen we tevens aan vergelijking met de A-V-taak (zie ad 1). Aangezien in beide subtoetsen de visuele stimulus onverkort aangeboden werd, spitst het verschil in taakeisen zich toe op herkenning en opslag van de auditief aangeboden articulatievorm (A-V) versus produktie van die vorm (L). Zeer zwakke lezers waren met auditief-visueel matchen bij alle woordklassen beter dan met hardop lezen (resp. gem. 17.2 en 13.0), in contrast met N die voor A-V en L resp. gem. 10.8 en 16.8 behaalden (zie Tabel 3). Dit articulatorische productieprobleem is een belangrijk kenmerk van het leesproces van zeer zwakke lezers.

Ad 3. De sprong van MKM- naar MKMM- en MMKM-woorden vormde voor een groter deel van de ZZ-lezers een probleem dan voor N-lezers (het *woordklasse-effect*). Voor de hand ligt om dit toe te schrijven aan het feit dat ZZ-lezers de MKM-woorden significant accurater verwerkten in de matchingstaken, onafhankelijk van de aanbestedingswijze (V1-V2, A-V, V-A). Dit surplus valt – op A-V na – weg wanneer alle woordklassen ingecalculeerd worden (zie Tabel 3). (Bij L kan er geen verschil optreden bij MKM-woorden vanwege de selectieprocedure; het niet-significante verschil bij Fl. – ook een hardop leestaak – is daar waarschijnlijk een gevolg van). Echter, voor de A-V-subtoets geldt dat de zeer zwakke lezers ook MKMM- en MMKM-woorden significant beter verwerkten dan normale lezers, maar voor hardop lezen is het omgekeerde het geval. Dit verschil kan te wijten zijn aan het tekort in het produceren van de articulatorische code dat ad 2. als kenmerk van zeer zwakke lezers genoemd werd. Het productieprobleem is in de A-V-taak sterk gereduceerd: er wordt een uitspraakvorm gegeven (A) die moet worden onthouden en vergeleken met een orthografisch patroon (V). Anders gezegd: de auditief aangeboden uitspraakvorm compenseert het tekort door reductie van het aantal alternatieven in de 'inner speech store' waaruit gekozen moet worden tot één (een 'priming' effect).

Bij L is een dergelijke compensatie er niet: de leerling is volkomen afhankelijk van zijn vermogen om de juiste uitspraakvorm te genereren. Deze interpretatie impliceert echter dat ook in een matchingstaak als A-V de articulatorische code een rol speelt. Als theoretisch argument dat deze interpretatie ondersteunt kan gesteld worden dat het belang van de articulatorische code bij het leren lezen door vele auteurs groot wordt geacht (verg. Span, 1982, voor een overzicht). Het is echter op grond van onze gegevens niet te concluderen dat het verschil tussen zeer zwakke en normale lezers uitsluitend het gevolg is van het voornoemde tekort bij de zwakke lezers. Dat wordt geïllustreerd door het navolgende.

Ad 4. Verkorting van de visuele stimulus van onbeperkte aanbieding tot 200 msec. leidt er bij normale lezers toe dat ze hun relatieve voorsprong op zeer zwakke lezers bij hardop lezen verliezen (het *flitsaanbiedingseffect*). Dit geldt overigens alleen bij eenlettergrepige woorden – bij tweelettergrepige woorden zijn de verschillen bij onverkorte aanbieding gering omdat er in beide groepen veel leerlingen zijn die het hardop lezen niet op 90% niveau beheersen. Dit indiceert dat een behoorlijk proportie van de normale lezers voor de associatie van het visueel-orthografische patroon aan de uitspraakvorm nog vrij veel tijd nodig heeft. Meer dan de helft van degenen die het hardop lezen van MKMM- en MMKM-woorden nog beheersten, deden dat niet bij verkorting tot 200 msec. (zie Tabel 4). Het lijkt er op dat hun woordspecifieke kennis (nog) sterk gebonden is aan de mogelijkheid om het herkennen van de visueel-orthografische stimulus en het produceren van de uitspraakvorm simultaan uit te voeren. Voor zeer zwakke lezers geldt dat een groot aantal dit bij het hardop lezen van MKMM- en MMKM-woorden niet beheerste (resp. 15 en 22 van de 37; Tabel 4). Omdat het voor het overgrote deel om hoogfrequente woorden ging, kan aangenomen worden dat de woordspecifieke kennis van deze leerlingen geringer is, in overeenstemming met hetgeen Reitsma (1983) concludeert. Hoogfrequente woorden komen immers door veelvuldig gebruik in spreek- en schrijftaal het eerst in aanmerking om middels woordspecifieke kennis aangeleerd te worden.

De relatie tussen de beide genoemde verschillen tussen normale en zeer zwakke lezers,

in het aanleren van woordspecifieke kennis en het produceren van de articulatorische code, is een interessant theoretisch vraagstuk waar we in het kader van dit artikel niet op in kunnen gaan, op enkele suggesties na. Aangenomen dat de N-groep model staat voor een normale ontwikkeling, stellen we op grond van onze resultaten als hypothese dat de associatie van de orthografische structuur met de uitspraakvorm de kern vormt van woordspecifieke kennis en daarmee van de opbouw van het autonome lexicon. Dergelijke associaties leggen kost zeer zwakke lezers moeite, waarschijnlijk ten gevolge van een tekort in het produceren van de uitspraakvorm. Het eerst beschikken over de auditief aangeboden articulatievorm vergemakkelijkt woordherkenning bij zeer zwakke lezers: de noodzaak tot het zelf genereren van de uitspraakvorm wordt van buitenaf gereduceerd. Echter, dit is geen 'lezen' in de gebruikelijke zin van het woord: de visuele stimulus is er dan altijd eerst en er moet een antwoord geproduceerd worden. Wel kan er in therapeutisch-compenserende zin van dit kenmerk van zeer zwakke lezers gebruik gemaakt worden (zie de afsluitende paragraaf).

6.3 *Implicaties voor verder onderzoek*

Uit het voorafgaande moge duidelijk geworden zijn dat de analyse van de eerste gegevens van het onderzoeksproject een aantal interessante veronderstellingen heeft opgeleverd over de eigenaardigheden van ernstige leesproblemen die met COTAL kunnen worden geduid. Echter, het instrument (of een aangepaste versie daarvan) zal pas getoetst kunnen worden op differentiële validiteit in vervolgonderzoek. Zo is het noodzakelijk om veronderstelde verschillen in het ontwikkelingsverloop van het automatiseringsproces tussen normale en zeer zwakke lezers, aan te tonen. Inmiddels is de beschreven toetsronde gevolgd door nog twee toetsronden bij de N-groep. Aangenomen dat het automatiseringsproces in die periode in volle gang is, verwachten we dat in de loop van groep vier (om precies te zijn in november/december en in maart/april, ongeveer vijf en negen maanden na de eerste afname) de N-groep tot beheersing van steeds meer onderdelen komt en blijk geeft van een zich uitbreidend autonoom lexicon. Beheersing van een woordklasse (i.c. beheersing van de subtoetsen op alle moeilijk-

heidsniveaus binnen die klasse) zal, gegeven de opzet van het instrument, de indicatie zijn voor opname van die klasse in het autonome lexicon.

Ook de zeer zwakke lezers zijn opnieuw, zij het met een langere tussenpoos, getoetst. De vraag die o.i. van groot belang is voor onderscheid tussen normale en zeer zwakke lezers is in hoeverre zeer zwakke lezers in staat zijn om tot een autonoom lexicon te komen, dat wil zeggen welke grenzen daar in termen van woordklassen, moeilijkheidsniveaus en aanbiedings- en antwoordwijzen aan te ontdekken zijn. Naar verwachting zal de vooruitgang bij een groot aantal van de zeer zwakke lezers zo gering zijn in vergelijking met die van normale lezers, dat van een autonoom lexicon nauwelijks gesproken kan worden. Uiteraard zullen in de analyse van de longitudinale data de eerder geventileerde ideeën omtrent de verschillen in ervaring (typen), in woordspecifieke kennis en in produktie van de articulatoire code, aan een nadere toetsing onderworpen worden.

Een volgende stap is onderzoek doen naar de vraag of uit de gegevens de instructiebehoeften van zeer zwak lezende leerlingen zijn af te leiden: de condities waaronder het effect van instructie zo groot mogelijk zal zijn. Het lijkt voorlopig duidelijk dat auditieve aanbidding van de articulatievorm een essentieel compenserend instructieprincipe is. Onderzoek waarin dit op verschillende wijze is getoetst, ondersteunt deze uitspraak (verg. o.a. Van der Leij, 1983, 1986; Bouwhuis & Truin, 1986; Van der Leij & Van Daal, 1989; Van Daal & Reitsma, 1989). De belangrijkste vraag is uiteraard in hoeverre dergelijke instructie leidt tot 'autonomiseren' van het lexicon en daarmee tot automatiseren van het lezen.

Ten slotte wijzen we op de relatie tussen de toets en instructieprogramma's i.c. tussen COTAL en het Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen (COPAL) waarvan auditieve aanbidding een belangrijk onderdeel vormt (Van Daal et al., 1987). Een effectiviteitsonderzoek waarbij toets en programma gecombineerd worden staat als sluitstuk van het project op het programma.

Noten

1. Men kan zich afvragen in hoeverre deze vooronderstelling gerechtvaardigd is. Uit (nog ongepubliceerd) vervolgonderzoek blijkt dat leerlingen met een vergelijkbare leeszwakte als de proefpersonen die in de onderhavige studie zijn onderzocht, zich niet onderscheiden van normale lezers die gematched zijn op leesniveau, in accuratesse en snelheid van het benoemen van cijfers en letters. Ook herkenning of een stimulus een letter of een cijfer is gaf geen verschil te zien. Mocht de letterkennis nog niet geautomatiseerd zijn op dit leesniveau (eind groep drie/begin groep vier) dan zijn normale en zeer zwakke lezers in dit opzicht vergelijkbaar.
2. Een andere mogelijke matchingsvariabele, intelligentie, werd niet gebruikt omdat we aannamen dat er door het toelatingsbeleid van LOM-scholen – slechts normaal intelligente kinderen worden aangenomen – en door het verwijderen van BaO-leerlingen die te goed lezen (zeker zo'n 25% van de oorspronkelijke steekproef), weinig verschil in zou zijn. Het is echter mogelijk dat deze assumptie niet geheel gerechtvaardigd is. Om die reden zal in het follow-up gedeelte van dit longitudinale onderzoek alsnog een vorm van intelligentiemeting uitgevoerd worden. Mochten er significante verschillen optreden dat kan die meting hetzij tot correctie van de gematchte groepen, hetzij tot invoering van intelligentie als covariaat leiden.
3. In een eerdere versie van de toetsbatterij waren nog twee subtoetsen opgenomen: auditief matchen en nazeggen van woorden. De eerste subtoets (A-A) maakte het kwartet matchingsmogelijkheden tussen auditieve en visuele stimuli compleet. Analog aan de A-T- en A-V-subtoets werd de moeilijkheidsgraad verhoogd door tussen de stimuli het tijdsverloop te laten oplopen van 1 naar 3 naar 5 seconden. De tweede subtoets, nazeggen, had ten doel om na te gaan in hoeverre uitspraak op zichzelf problemen gaf. Bij 51 LOM-leerlingen die deze versie kregen bleek het percentage dat deze subtoetsen accuraat aflegde over alle woordklassen en moeilijkheidsniveaus heen, 95 te bedragen. Vanwege het niet-differentiërende karakter van deze subtoetsen waar het de groep van lezers betrof waarin we het meest geïnteresseerd waren, zijn ze in het verdere onderzoek weggelaten.
4. In een eerdere versie waren nog twee tussenliggende moeilijkheidsniveaus opgenomen: 180 en 130 msec. (V1-V2, V-A, V-T, Fl.) en 2'' en 4'' (A-V, AT). Omdat de scores bij het oplopen van het moeilijkheidsniveau bij alle woordklassen en subtoetsen lineair verliepen, zijn deze tussenliggende waarden weggelaten in het verdere onderzoek.

5. Vier ZZ-leerlingen beheersten één subtoets bij alle woordklassen, één ZZ-leerling beheerste er twee; vijf N-leerlingen beheersten er één. Van deze in totaal elf beheerste subtoetsen ging het in zes gevallen om L, de enige taak zonder moeilijkheidsniveaus. Vier keer betrof het A-V en één keer VI-V2, subtoetsen met drie moeilijkheidsniveaus.
6. Overigens is er, op basis van de gepresenteerde gegevens, geen uitspraak mogelijk over de kwestie in hoeverre de werking van het werkgeheugen bij zeer zwakke lezers afwijkt van die van leeftijdgenoten om dat het hier gaat om de vergelijking met jongere leerlingen van gelijk leesniveau. In het follow-up onderzoek zal deze vergelijking wel worden gemaakt.

Literatuur

- Aarle, E. J. M. van, *Behandeling van ernstige leesproblemen*. Een vergelijkend onderzoek naar de effectiviteit van zelfinstructie en training in decoderen en contextgebruik. Nijmegen: 1988 (diss.).
- Baddeley, A., *Working memory*. Oxford: Clarendon Press, 1987.
- Berg, R. M. van den & H. G. te Lintelo, *A.V.I.-pakket*. Den Bosch: K.P.C., 1977.
- Bouma, H. & C. P. Legein, Foveal and parafoveal recognition of letters and words by dyslectics and by average readers. *Neuropsychologica*, 1977, 15, 69-80.
- Bouma, H. & C. P. Legein, Dyslexia: a specific recoding deficiency. An analysis of response latencies for letters and words in dyslectics and average readers. *Neuropsychologica*, 1980, 18, 285-298.
- Bouwuis, D. G. & P. G. M. Truin, Woorden lezen met spraak naar keuze door dyslectische kinderen. In: S. Dijkstra & P. Span (Red.), *Leerprocessen en instructie*. Lisse: Swets & Zeitlinger, 1986.
- Brus, B. Th. & M. J. M. Voeten, *Een-minuuttest, vorm A en B*. Nijmegen: Berkhout, 1973.
- Daal, V. H. P., A. van der Leij, N. C. M. Bakker & P. Reitsma, Een Computergestuurd Programma voor Aanvankelijk Lezen. *Pedagogische Studiën*, 1987, 64, 364-376.
- Daal, V. H. P., N. C. M. Bakker, A. van der Leij, P. Reitsma & H. Smeets, *COPAL: Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen*. Verslag over de ontwikkeling en beproeving. Amsterdam: Vrije Universiteit & Paedologisch Instituut, 1988.
- Daal, V. van & P. Reitsma, Audio support in the remediation of reading problems. In: J. J. Dumont & H. Nakken (Eds.), *Learning disabilities. Volume 2. Cognitive, social and remedial aspects*. Lisse: Swets & Zeitlinger, 1989, 149-158.
- Ehri, L. C. & L. S. Wilce, Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology*, 1983, 75, 3-18.
- LaBerge, D. & S. J. Samuels, Toward a theory of automatic information processing in reading. In: H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark: IRA, 1976.
- Leij, A. van der, *Ernstige leesproblemen*. Een onderzoek naar mogelijkheden tot differentiatie en behandeling. Lisse: Swets en Zeitlinger, 1983 (diss.).
- Leij, A. van der, Inprentingsmethodiek, informatieverwerking en differentiële instructiegevoeligheid, een aanzet tot theorievorming. In: L. M. Stevens, & A. van der Leij, (Red.), *Dyslexie 86*. Verslag van het congres te Utrecht. Lisse: Swets & Zeitlinger, 1986.
- Leij, A. van der, & E. J. Kappers, Instructie van kinderen met ernstige leerproblemen: introductie op de thema-serie. *Pedagogische Studiën*, 1989, 66, 311.
- Leij, A. van der, & V. H. P. van Daal, Attacking dyslexia: the effect of preceding and simultaneous verbal practice. In: J. J. Dumont & H. Nakken (Eds.), *Learning disabilities. Vol. 2: cognitive, social and remedial aspects*. Lisse: Swets & Zeitlinger, 1989, 179-191.
- Perfetti, C. A., *Reading ability*. New York: Oxford University Press, 1985.
- Perfetti, C. A., The presentation problem in reading acquisition. To appear in: Ph. B. Gough (Ed.), *Reading acquisition*. New York: Erlbaum (in press).
- Reitsma, P., *Phonemic and graphemic codes in learning to read*. Amsterdam: Vrije Universiteit, 1983 (diss.).
- Reitsma, P., Automatische behandeling van leesen spellingsproblemen? In: A. van der Leij & L. M. Stevens (Red.), *Dyslexie*. Verslag van het congres te Den Haag. Lisse: Swets en Zeitlinger, 1985.
- Sas, J. F. & C. J. Wieringa, DVO. *Het deelvaardighedsonderzoek ten behoeve van het voorbereidend en aanvankelijk lees- en spellingsonderwijs*. Groningen: Wolters-Noordhoff, 1982.
- Seymour, P. H. K., *Cognitive analysis of dyslexia*. Londen: Routledge & Kegan Paul, 1986.
- Span, B., *Een informatieverwerkingsbenadering van cognitief functioneren. Deel 2: leren lezen*. Haren: RION, 1982.

- Spear, L. C., & R. J. Sternberg, An information-processing framework for understanding reading disability. In: S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1987.
- Spring, C. & C. Capps, Encoding speed, rehearsal, and probed recall of dyslexic boys. *Journal of Educational Psychology*, 1974, 66, 780-786.
- Spring, C. & L. Perry, Naming speed and serial recall in poor and adequate readers. *Contemporary Educational Psychology*, 1983, 8, 141-145.
- Struiksma, A. J. C., A. van der Leij & J. Vieyra, *Diagnostiek van technische lezen en aanvankelijk spellen*. Amsterdam: VU-boekhandel, 1987².
- Torgesen, J. K., Computer-assisted instruction with learning-disabled children. In: J. K. Torgesen & B. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities*. New York: Academic Press, 1986, 417-435.
- H. Smeets (1956) studeerde orthopedagogiek met bijvakken sociaal-wetenschappelijke informatica en wijsgerig-empirische pedagogiek aan de RU te Leiden. Hij participeerde tijdens de studie in het door NWO gesubsidieerde onderzoek 'De kwaliteit van de hechtingsrelatie aan moeders, vaders en crècheleidsters'. Vanaf 1988 werkt hij als wetenschappelijk onderzoeker bij de Vrije Universiteit, eerst met als opdracht de statistische dataverwerking van de Computergestuurde Toets voor Automatisering van het Lezen. In 1989 is hij verbonden aan het NWO-project 'Directe woordherkenning bij kinderen met leesproblemen'.
- V. van Daal (1953) studeerde psychologische functieer met bijvak onderwijskunde aan de Universiteit van Amsterdam. Verrichtte enige tijd onderzoek naar het bestuderen van studieteksten aan de Universiteit Twente. Vanaf 1985 werkt hij als onderzoeker aan de Vrije Universiteit (Sectie Speciale Pedagogiek) in projecten m.b.t. de ontwikkeling en beproefing van het Computergestuurd Orthodidactisch Programma voor Aanvankelijk Lezen en de Computergestuurde Toets voor Automatisering van het Lezen. Vanaf 1988 is hij verbonden aan het NWO-project 'De Ontwikkeling van het Orthografisch Geheugen'.

Curricula vitae

A. van der Leij (1946) studeerde psychologie aan de Universiteit van Amsterdam en werkte van 1971 tot 1977 bij het Gemeentelijk Schoolpsychologisch Bureau te Zwolle. Van 1977 tot heden is hij verbonden aan de Sectie Speciale Pedagogiek van de Vrije Universiteit, sinds 1984 als hoogleraar. Promoveerde in 1983 op een proefschrift onder de titel *Ernstige leesproblemen*.

Adres: Vrije Universiteit, Sectie Speciale Pedagogiek, Van der Boechorststraat 1, 1081 BT Amsterdam

Manuscript aanvaard 14-11-'89

Summary

Leij, A. van der, H. Smeets & V. H. P. van Daal. 'The development of the computer-assisted test for word recognition.' *Pedagogische Studiën*, 1990, 67, 163-178.

The development of the computer-assisted test for word recognition (COTAL) is an example of a research project on the diagnosis of children with severe learning disabilities within a conceptual framework of information processing.

To indicate the specific characteristics of the reading process of children with severe reading disabilities, a computer-assisted test has been developed. Seven subtests reflect the various ways to present words as stimulus (auditory; visual) and the ways to respond (matching; naming; typing). Six wordcategories (one, two syllables) were used. In the reported study, 37 children with severe reading disabilities were compared to 37 reading age controls.

The results indicate that the tests reveal significant differences between very poor and normal readers, related to the use of the articulatory code. Implications for further research are discussed.