

Selecteren van kerngedeelten in teksten*

Bruikbaarheid van tekstanalyse- en beoordelingsmethoden voor onderzoek naar de verbetering van tekstbestudering

B. VAN HOUT WOLTERS
EN B. WELTENS

*Facultair Instituut Algemene Onderwijskunde
voor de Lerarenopleiding (A.O.L.O.),
Katholieke Universiteit Nijmegen*

Samenvatting

Bij veel onderzoek naar de verbetering van tekstbestudering is het nodig vooraf vast te stellen welke de kerngedeelten zijn in de onderzoeksteksten. Dit gebeurt vaak op intuïtieve, subjectieve wijze; er bestaat echter behoefte aan een meer objectieve methode. In dit artikel worden een aantal methoden beschreven die daarvoor in aanmerking komen: tekstanalyse-methoden (propositionele methoden en schemamethoden) en beoordelingsmethoden. Deze methoden worden getoetst op hun bruikbaarheid voor het vaststellen van kerngedeelten in langere studieteksten, het type teksten dat meestal in het onderzoek naar de verbetering van tekstbestudering wordt gebruikt. Geconcludeerd wordt dat de beoordelingsmethoden, onder enkele voorwaarden, vooralsnog de meeste perspectieven lijken te bieden.

1 Probleemstelling

Er wordt al jaren onderzoek verricht naar maatregelen om het bestuderen van studieteksten te verbeteren. Het meeste onderzoek op dit terrein heeft betrekking op wijzigingen en aanvullingen die door de auteur of docent in het studiemateriaal worden aangebracht om de leerresultaten van de studenten of leerlingen te verbeteren: docentbewerkingen of 'experimenter-provided activities' (EPA's). Er vindt echter ook steeds meer onderzoek plaats naar

de effecten van bepaalde tekstbewerkingen en bestuderingsactiviteiten door de lerende zélf: studentbewerkingen of 'subject-generated activities' (SGA's). De onderzoeken in deze twee groepen zijn verder in te delen in vier categorieën; er kan sprake zijn van:

- het wijzigen van de tekst zelf, met name van de structurering en het taalgebruik;
- het wijzigen van de vormgeving van de tekst, bijv. het vet of cursief drukken, onderstrepen, kleuren of omlijnen van gedeelten (markeren);
- het toevoegen van verduidelijkende gedeelten aan de tekst, bijv. overzichten, schema's, samenvattingen of figuren;
- het toevoegen van sturende of instruerende gedeelten aan de tekst, bijv. doelen, voortoetsen of vragen.

De onderzoeken op deze gebieden hebben in het algemeen een vergelijkbare basisopzet. Men zou deze als volgt kunnen omschrijven. Enkele groepen proefpersonen bestuderen een studietekst. Voor de experimentele groep(en) zijn in de tekst bepaalde wijzigingen of aanvullingen aangebracht door de proefleider (EPA), of er wordt van de proefpersonen gevraagd om de tekst op een bepaalde manier door te nemen of te bewerken (SGA). Na bestudering wordt aan de hand van een bepaalde vorm van toetsing gemeten wat de proefpersonen hebben onthouden van de inhoud van de tekst. Ook wordt soms vastgelegd wat de invloed is van andere variabelen op de resultaten (bijv. persoons- en tekstvariabelen).

Hoewel het onderzoek op de genoemde terreinen heel wat gegevens heeft opgeleverd, is er ook kritiek op te leveren. In Van Hout Wolters (1980) worden met name kanttekeningen geplaatst bij de onderzoeksupzet en de beperkte theoretische achtergronden van de onderzoeken. Ook wordt signaleerd dat het onderzoek in het algemeen nog sterk produktgericht is; er wordt nog weinig experimenteel nagegaan op welke cognitieve processen de onderzochte tekstbewerkingen precies van invloed zijn en door welke variabelen deze invloed beperkt of versterkt kan worden. Naast

* Veel dank aan prof. dr. L. F. W. de Klerk, prof. dr. A. M. P. Knoers, drs. G. Seegers en drs. H. Tillema voor hun opmerkingen bij een eerdere versie van dit artikel.

deze algemene problemen zijn er ook specifieke problemen bij dit type onderzoek, zoals bijv. de geringe representativiteit van het gebruikte tekstmateriaal in het onderzoek voor het normale studiemateriaal van de lerende, of het vaak achterwege laten van een voortoets.

Dit artikel handelt echter over een ander specifiek probleem. De vraag is namelijk op welke manier de onderzoeker kan bepalen welke de kerngedeelten zijn in de studieteksten die in het onderzoek worden gebruikt. Het is om verschillende redenen nodig deze kerngedeelten of kernpunten vast te stellen:

- In het docentbewerkingsonderzoek worden de bewerkingen van de studietekst (de vragen, doelen, onderstrepingen, schema's e.d.) op deze kerngedeelten afgestemd. Hierbij wordt de speciale aandacht van de lerende gevestigd op belangrijke gedeelten in de tekst.
- In het docent- en studentbewerkingsonderzoek worden de toetsresultaten van de lerenden veelal geanalyseerd op het onthouden van de kerngedeelten t.o.v. de overige inhoud van de tekst. Voor het samenstellen en analyseren van de (gesloten en open) toetsen over de tekst is het onderkennen van kerngedeelten dus eveneens van belang.

In het meeste onderzoek naar de verbetering van tekstbestudering zijn er rond dit punt een aantal problemen te signaleren. Meestal is er in de artikelen nauwelijks iets vermeld over de methode die de onderzoekers hanteerden om vast te stellen welke tekstgedeelten speciale nadruk zouden krijgen door bepaalde docentbewerkingen en in de toetsen. Zo wordt bijvoorbeeld in onderzoeksverslagen met betrekking tot het onderstrepen in studieteksten vaak alleen gezegd dat de onderzoekers in de docentbewerkingsconditie zelf de 'main points' in de gebruikte tekst hadden onderstreept en niet welke methode ze daarvoor gebruikten, noch wat ze onder 'main points' verstonden. Onduidelijk is bijvoorbeeld of de 'main points' als vaststaande grondgegevens van de tekst werden opgevat, die gemakkelijk en eenduidig door iedere individuele onderzoeker uit de tekst te halen waren, of dat er bij de definiëring van dit begrip ook lezersvariabelen een rol speelden, zoals de doelstelling van de lezer. Deze punten worden eveneens signaleerd door Foster (1979).

Men kan zich afvragen of er voor het onderzoek naar de verbetering van tekstbestudering geen explicietere en objectievere manieren zijn om de kerngedeelten in een tekst vast te stellen. Twee typen methoden komen daarvoor in aanmerking. In de eerste plaats zou gewerkt kunnen worden met één van de *tekstanalyse-methoden* die de laatste jaren ontwikkeld zijn (bijv. Meyer, 1975 a en b; Thorndyke, 1977). Deze worden, voor zover bekend, in het onderzoek naar docentbewerkingen in studieteksten nog niet gebruikt. Foster (1979) en Meyer en McConkie (1973) verwachten hiervan echter veel voor het identificeren van de kerngedeelten in een tekst. In de tweede plaats zou gedacht kunnen worden aan het werken met *externe beoordelaars*. Met name in het onderzoek naar het markeren in studieteksten wordt door een aantal onderzoekers van deze methode gebruik gemaakt (bijv. Coles en Foster, 1975 en Foster, 1979). Bij onderzoek naar andere tekstbewerkingen zoals vragen bij teksten, gebeurt dat echter vrijwel niet.

Het doel van dit artikel is om deze twee methoden, de tekstanalyse-methode en de beoordelingsmethode, te beschrijven en te beoordelen op geschiktheid voor het selecteren van kerngedeelten in teksten die in het onderzoek naar de verbetering van tekstbestudering worden gebruikt.

2 Tekstanalyse-methoden

Grofweg kunnen er op theoretische gronden twee typen tekstanalyse-methoden worden onderscheiden: de propositionele methoden en de schema-methoden. De propositionele methoden komen voort uit de propositiebenadering in de tekstverwerkingsliteratuur, die benadrukt dat de informatieverwerking bij het lezen primair gestuurd wordt vanuit de tekst. De schema-methoden sluiten aan bij de schema- of netwerk-benadering, waarin benadrukt wordt dat vooral de aanwezige netwerken van kennis bij de lezer de informatieverwerking sturen.

Door verschillende onderzoekers werden de achtergronden van deze methoden de laatste jaren beschreven (in het Nederlandse taalgebied o.a. door Van Dam, 1979; Kok, Boonman en Beukhof, 1980; Tillema en Boonman, 1979). Gezien onze probleemstelling zullen wij deze methoden vanuit een ander gezichtspunt

benaderen en beschrijven dan deze onderzoekers. Voor ons is het vooral van belang te weten hoe de methoden precies 'werken', welke doelen de auteurs zelf hadden bij het construeren van hun methoden en of ze ook voor ons doel (het selecteren van kernpunten in langere studieteksten) te gebruiken zijn.

2.1 De propositionele methode

De propositionele methode vindt haar oorsprong bij Meyer en McConkie (1973), die de structuur van een tekst weergeven in een grafische boomstructuur, daarmee aansluitend op de boomdiagrammen die in de grammatische wetenschap voor de representatie van zinnen gebrûkt worden. Deze boomstructuren werden verder ontwikkeld door o.a. Meyer (1975 a en b), Frederiksen (1975) en Kintsch (1974), en ze heten dan respectievelijk 'content structure', 'graph structure' en 'coherence graph'.

De methode gaat uit van de veronderstelling dat sommige gedeelten in een tekst een meer centrale plaats innemen dan andere en dat er meerdere niveaus van belangrijkheid te onderscheiden zijn. Hoe belangrijker de tekstgedeelten (kerngedeelten), des te hoger zullen ze in de boomstructuur worden weergegeven.

De knooppunten in de boomstructuren worden gevormd door 'content words' of 'inhoudswoorden' van de tekst. Dit zijn woorden met een beschrijvende en naar andere begrippen verwijzende functie, zoals zelfstandige en bijvoeglijke naamwoorden. De lijnen tussen deze knooppunten geven ruimtelijk weer hoe de inhoud is georganiseerd.

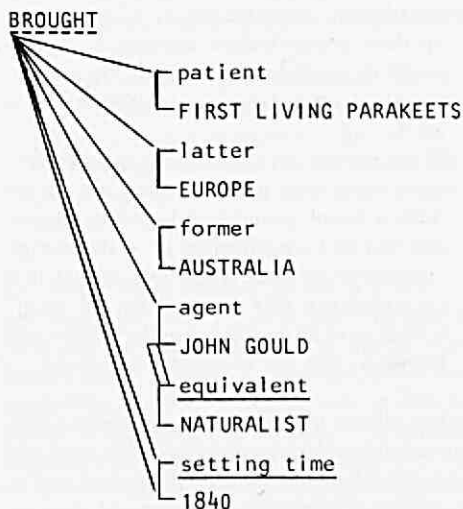
De methode Meyer (1975 a en b) kort beschreven

Omdat de methode van Meyer wat verder uitgewerkt is dan de methoden van de andere genoemde auteurs, wordt deze hier ter illustratie kort beschreven.

Meyer begint met het herschrijven van alle proposities van de tekst met behulp van de grammatica-regels van Grimes (1972, 1974) tot predicaten en argumenten. De predicaten geven de relaties weer tussen de inhoudswoorden; de argumenten zijn de inhoudswoorden zelf. Ze onderscheidt lexicale ('lexical'), functie-aanduidende ('role') en stilistische ('rhetorical') predicaten, die elk hun eigen hiërarchisch gewicht hebben. Lexicale predicaten zijn inhoudswoorden die een relatie tussen andere inhoudswoorden aangeven. Ze staan een ni-

veau hoger dan deze inhoudswoorden, die als hun argumenten fungeren. Functie-aanduidende en stilistische predicaten geven ook relaties tussen inhoudswoorden weer, maar zijn zelf geen inhoudswoorden.

In haar boek van 1975 analyseert Meyer een alinea van 7 zinnen (de zgn. 'Parakeet Paragraph'; zie Meyer, 1975b, 44), hetgeen resulteert in een boomstructuur van anderhalve pagina. Als voorbeeld geven we hier de analyse van de derde zin van deze alinea: 'The first living parakeets were brought to Europe from Australia by John Gould, a naturalist, in 1840'. De structuur is weergegeven in Figuur 1.



Figuur 1 Analyse van de derde zin van de 'Parakeet Paragraph' van Meyer (1975b, 47)

Ter verduidelijking:

- inhoudswoorden: in hoofdletters
- predicaten: in kleine letters
 - lexicale predicaten: met stippellijn
 - stilistische predicaten: onderstreept
 - functie-aanduidende predicaten: overige

Doelen van Meyer (1975 a en b) en van de andere auteurs

Meyer ziet deze analyse als een mogelijkheid om, enerzijds, leerresultaten bij het bestuderen van teksten te kunnen evalueren (dit zowel in het kader van tekstverwerkingsonderzoek, als in de onderwijspraktijk); anderzijds om overeenkomsten en verschillen in de structuur van teksten te kunnen bepalen. Ook dit kan binnen onderzoek én onderwijs van belang zijn, bijv.

voor het vergelijken van experimentele teksten en het samenstellen van equivalente leestests (parallelvormen van leestests) voor gebruik op school (Meyer, 1975a, 7-10). Zij wijst er op dat de tot dan toe gebruikelijke methoden voor deze doelen ongeschikt of ontoereikend zijn: 'verbatim' gescoorde protocollen en multiple-choice vragen geven geen goed inzicht in het leerresultaat en leesbaarheidsformules 'deal with surface factors of prose and are not concerned with the meaning of prose and the manner in which the content is organized to convey this meaning to the reader' (Meyer, 1975a, 8).

Ook Frederiksen (1972, 211-212; 1975, 158) ziet als doel van dit type tekstanalyse het beter kunnen evalueren van leerresultaten (verbale protocollen) na het bestuderen van een tekst. Hij heeft echter daarnaast ook het doel te onderzoeken of de 'responses which do not simply reproduce text content' te verklaren zijn vanuit verwerkings- of recall-processen. Kintsch (1974) gaat veel verder: hij wil een aanzet geven tot het ontwikkelen van een tekstverwerkingsmodel dat cognitieve processen verklaart die plaatsvinden bij comprehension, retrieval en recall van teksten.

Gebruik van de propositionele methode voor het onderzoek naar het bewerken van studieteksten
Op basis van een analyse van de tekst in grafische boomstructuren is het in principe mogelijk om kerngedeelten in een tekst aan te wijzen die gemarkeerd moeten worden of waarop bijv. vragen en doelen kunnen worden gericht. Men kan dan uitgaan van de predicaten en argumenten die het hoogst in de hiërarchie staan.

Ook zal het met behulp van deze boomhiërarchieën mogelijk zijn om toetsvragen op te stellen en recall-protocollen te analyseren.

Er zijn echter praktische problemen bij het gebruik van deze methode; ze is namelijk vooral bij heel korte teksten te gebruiken. Een alinea van 7 zinnen levert bij Meyer (1975, a en b) al een boomstructuur van 1½ pagina op. Het is bijna ondoenlijk om teksten van 10 of 20 pagina's om te zetten in één of meer boomstructuren. In feite is deze propositionele methode voor ons doel veel te gedetailleerd.

Bovendien is er nog een ander probleem. Deze methode pretendeert dat teksten eenduidig geanalyseerd kunnen worden zodat ieder die deze methode toepast, tot eenzelfde hiërarchie komt. Het is echter te verwachten dat ook bij deze wijze van analyseren van teksten en

aanwijzen van kernpunten allerlei lezersvariabelen een ongecontroleerde rol spelen, zoals bijv. het leesdoel en de voorkennis van de lezer. Hierdoor zullen er toch verschillen in resultaten kunnen ontstaan. Dit wordt ook door Kintsch (1974, 20) gesignaleerd.

2.2 De schema-methode

Bij de schema- of netwerk-benadering speelt de notie 'schema' een grote rol. Het is een begrip dat oorspronkelijk afkomstig is uit de 'mentalistiche' psychologie van Bartlett (Van Dam, 1979, 67), waarin het staat voor een complex netwerk van kennis waarin nieuwe informatie wordt geïntegreerd, maar dat ook kan dienen als een soort 'mal' die op de informatie gelegd wordt om er structuur in aan te brengen (zie bijv. Anderson, 1978 en Rumelhart en Ortony, 1977). Met name deze laatste eigenschap wordt in de tekstwetenschap benadrukt.

In de schema-benadering is men, voortbordurend op de zogenaamde herschrijfgeregels uit de (zins-)grammatica, herschrijfgeregels gaan formuleren voor bepaalde typen teksten. De meeste aandacht is daarbij besteed aan de structuur van een eenvoudig verhaal (narrative) met één hoofdpersoon, dat een aantal gebeurtenissen beschrijft zonder van dialogen gebruik te maken. De geformuleerde 'Story Structure Models' genereren dus door middel van een systeem van herschrijfgeregels de globale structuur van een verhaal. De kernpunten van het verhaal zouden in de hogere niveaus van deze structuur gezocht kunnen worden. Structuur-modellen van verhalen verschijnen in min of meer dezelfde vorm bij Thorndyke (1977; 'story grammar'), Mandler & Johnson (1977; 'a grammar of simple stories') en Van Dijk (1977; 'narrative super structures').

De methode Mandler & Johnson kort beschreven

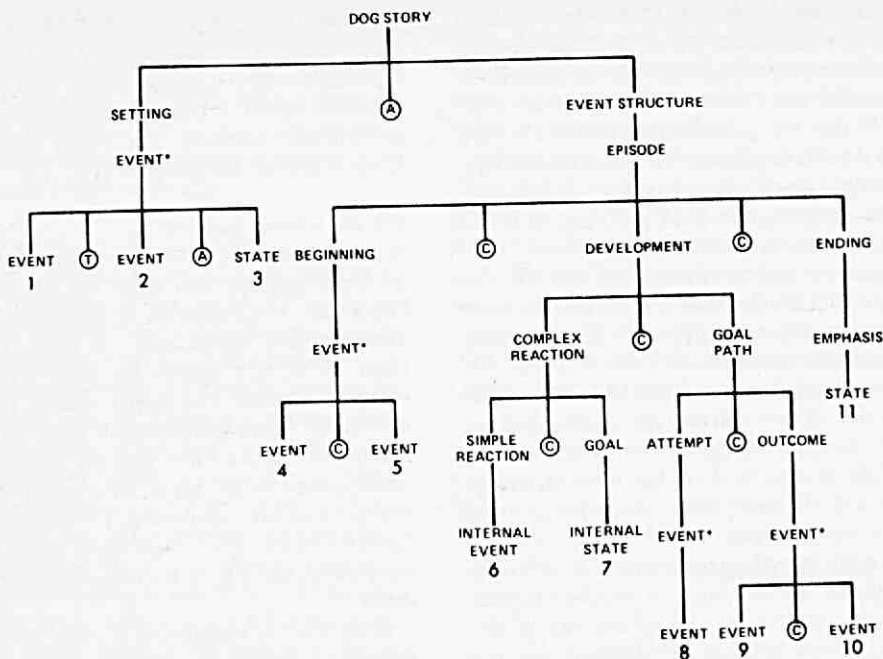
Ter illustratie wordt hier een korte impressie gegeven van enkele herschrijfgeregels van Mandler & Johnson (1977, 117) voor de globale structuur van een verhaal of vertelling:

FABLE → STORY AND MORAL

STORY → SETTING AND EVENT STRUCTURE
(.....)

EVENT STRUCTURE → EPISODE ((THEN EPISODE)^a)

Ter verduidelijking: de buitenste haakjes geven



Figuur 2 Analyse van de 'dog story' van Mandler & Johnson (1977, 119 e.v.) Ter verduidelijking: De nummers 1-11 verwijzen naar de proposities; de letters A, T en C verwijzen naar de zgn. AND, THEN en CAUSE-relaties tussen onderdelen van de tekststructuur; de sterretjes geven aan dat het repetitieve elementen betreft.

aan dat deze elementen niet noodzakelijk aanwezig hoeven te zijn; met ()ⁿ wordt aangegeven dat er één of meer van deze elementen kunnen voorkomen.

Met behulp van deze regels wordt de verhaalstructuur van een tekst gevormd. Deze krijgt de vorm van een boomdiagram en is opgebouwd uit zgn. 'slots' of invulplaatsen, waar de proposities van de tekst kunnen worden ingevuld. Een boomdiagram de zgn. 'dog story', een verhaaltje dat bestaat uit 11 proposities (zie Mandler & Johnson, 1977, 119 e.v.), is weergegeven in Figuur 2.

Doelen van Mandler & Johnson en van de andere auteurs

Mandler & Johnson stellen dat deze analyse van de onderliggende structuur van eenvoudige verhalen weergeeft welke schemata een lezer gebruikt bij 'encoding' en 'retrieval' van teksten (Mandler & Johnson, 1977, 111). Hun uiteindelijke doel is dan ook precies te voorspellen wat mensen zich wel en niet herinneren van teksten.

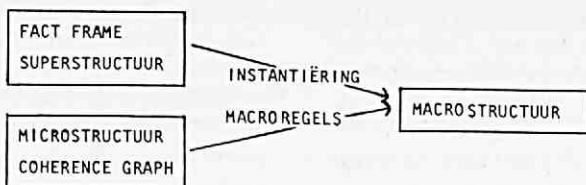
De boomstructuren die Thorndyke (1977)

en Van Dijk (1977, 16-18) gebruiken zijn vergelijkbaar met die van Mandler & Johnson; zij benoemen echter geen relaties tussen tekstdelen (de AND, THEN en CAUSE-relaties van Mandler & Johnson). Ook zij willen nagaan of deze boomstructuren psychologische (cognitieve) realiteit hebben. Zij voorspellen o.a. dat bij de verwerking van een verhaal de gelezen informatie door de lezer wordt 'ingevuld' in een verhaalschema ('instantiëring'; zie bijv. Van Oostendorp, 1980, 75), en dat bij het zich herinneren van het verhaal het bestaande schema wordt geactiveerd en aldus als 'retrieval-plan' wordt gebruikt.

Gebruik van de schema-methode voor het onderzoek naar het bewerken van studieteksten

Als methode om de kerngedeelten van een tekst vast te stellen is deze methode minder geschikt, en wel om de volgende redenen:

- de methode is voornamelijk beperkt tot korte eenvoudige verhalen (narratives); voor lange informatieve teksten is zij moeilijk te gebruiken;
- er worden stereotype structuren opgelegd



Figuur 3 Schematische weergave van het model van Kintsch en Van Dijk (1978)

aan de tekst, die lang niet altijd aanwezig hoeven te zijn;

- de boomdiagrammen suggereren weliswaar dat er sprake is van een hiërarchische structuur, waarin de kernpunten 'bovenaan' staan, maar de hypothesen van Mandler & Johnson (1977, 133) over de herinnering van tekstgedeelten geven aan dat de kerngedeelten slechts zeer indirect uit het verhaal-schema zijn af te leiden.

2.3 Beide tekstanalyse-methoden + macroregels bij Kintsch en Van Dijk (1978)

In het model van Kintsch en Van Dijk (1978) wordt geprobeerd de bovenbeschreven benaderingswijzen te integreren. Ook Van Dijk wijst in diverse publikaties (o.a. 1977, 1978, 1979) op een mogelijke combinatie van beide benaderingswijzen. Enerzijds wordt uitgegaan van het formele schema van het type tekst, bijv. van een verhaal of wetenschappelijk artikel (superstructuur of 'fact frame'), anderzijds van de proposities van de tekst zelf (microstructuur of 'coherence graph'). In hun model geven zij weer hoe volgens hen de tekstverwerking verloopt. Ze stellen zich tot doel: 'to describe the system of mental operations that underlie the processes occurring in text comprehension and the production of recall and summarization protocols' (Kintsch en Van Dijk, 1978, 363).

Aan de ene kant gebruiken lezers bij het lezen hun kennis van de formele superstructuur van teksten. De informatie uit de tekst wordt door hen ingepast (geïncantierd) in een reeds aanwezig tekstschema. Aan de andere kant gaat de lezer uit van de volledige propositionele inhoud, de microstructuur. Deze wordt bewerkt met zogenaamde macroregels, die tot doel hebben de informatie op efficiënte wijze te selecteren en hergroeperen. In combinatie leiden beide processen tot een globale betekenisstructuur van de tekst: de macrostructuur. In schema zou men dit beknopt kunnen weerge-

ven (zie Figuur 3).

Voor het selecteren van kernpunten in een tekst t.b.v. het onderzoek naar de verbetering van tekstbestudering, kunnen vooral de macroregels van dit model van belang zijn. Bij alle andere tekstanalyse-methoden rangschikt men de letterlijke tekst in een boomstructuur en hanteert men geen macroregels. Kintsch en Van Dijk gaan echter verder dan de letterlijke tekst; ze zijn van mening dat het irreal is om te veronderstellen dat lezers woord voor woord proberen de tekst te ordenen en op te slaan. Ze zullen de tekst ook bewerken: ze selecteren informatie, brengen informatie onder een gemeenschappelijke noemer of vervangen de informatie door nieuwe, meer algemene, informatie. De kernpunten hoeven namelijk niet altijd expliciet in de tekst te staan. Ze kunnen van een hogere orde zijn dan de letterlijke proposities ('higher order propositions' of 'macro propositions'), aldus Van Dijk (1979) in zijn kritiek op Meyer (1975 a en b). Voor het vormen van de macrostructuur geven ze een drietal macroregels: selectie, generalisatie en constructie (zie bijv. Van Dijk, 1978, 197). Dat het hier geen strikt wetenschappelijke generatieve regels betreft, onderkennen ze zelf. Van Dijk (1977, 11) formuleert ze 'provisionally'; hij spreekt over basisoperaties wier toepassing beslist niet expliciet en eenduidig is (1978, 58). Het aantal macroregels, de gebruikte terminologie en de definiëring verschilt ook in de publikaties van Kintsch en Van Dijk.

Om deze redenen zijn ze toch nog moeilijk te gebruiken om op eenduidige wijze kernpunten op te sporen in teksten die in tekstbestuderingsonderzoek gebruikt worden.

2.4 Schematisch overzicht van een aantal tekstanalyse-methoden

Samenvattend kan men de bovenbeschreven methoden, met het type produkt en de achterliggende doelen, als volgt indelen (Tabel 1).

Tabel 1 *Schematisch overzicht van een aantal tekstanalyse-methoden*

METHODE	AUTEURS	TYPE PRODUKT	DOELEN
Propositie- methode	Frederiksen (1975)	linguïstische structuur met relatie-labels	- evalueren tekst- reproducties - inzicht in comprehen- sion en recall
	Kintsch (1974)	linguïstische structuur met relatie-labels	- inzicht in de cog- nitieve processen bij tekstverwerking en reproductie
	Meyer (1975 a en b)	linguïstische structuur met relatie-labels	- evalueren tekstre- producties - overeenkomsten/ver- schillen tussen teksten vaststellen
Schema- methode	Van Dijk (1977, 16-18)	verhaalstructuur	- inzicht in comprehen- sion en recall
	Mandler & Johnson (1977)	verhaalstructuur met relatie- labels (AND, THEN en CAUSE)	- inzicht in comprehen- sion en recall
	Thorndyke (1977)	verhaalstructuur	- inzicht in comprehen- sion en recall
Beide methoden	Kintsch & Van Dijk (1978)	linguïstische structuur, macro- regels en verhaalstructuur	- inzicht in de cog- nitieve processen bij tekstverwerking en reproductie

2.5 Het gebruik van tekstanalyse-methoden voor onderzoek naar het bewerken van studieteksten

Uit het voorgaande blijkt al dat er een aantal bezwaren zijn om de huidige tekstanalyse-methoden te gebruiken voor het aanwijzen van kerngedeelten in teksten die in het onderzoek naar markeren, vragen, doelen, samenvattingen e.d. gebruikt worden. Een belangrijk bezwaar bij alle methoden is, dat ze pretenderen dat teksten eenduidig geanalyseerd kunnen

worden, terwijl daarbij toch allerlei variabelen een rol kunnen spelen, zoals de doelstelling en de voorkennis van de lezer.

Bovendien geldt voor alle bovenstaande tekstanalyse-methoden dat ze zeer arbeidsintensief en zeer gecompliceerd zijn. Het betekent dat het analyseren van één alinea al een moeilijke opgave is; het analyseren van langere teksten is welhaast ondoenlijk. Ook Reder & Anderson (1980) komen tot deze conclusie bij het zoeken van een dergelijke methode om op

een verantwoorde wijze samenvattingen van teksten te maken. Voor hun doel blijken deze methoden ongeschikt: 'We could not find a published procedure for generating summaries that was sufficiently general in application and clearly specified to allow one to represent complex texts such as these (nl. wetenschappelijke teksten van ca. 5000 woorden). All systems for representing text seem to require involved, tedious techniques that are difficult to use for one paragraph. Trying to adapt them to chapters over 13 pages long would border on the impossible' (Reeder & Anderson, 1980, 123).

Ook Frederiksen (1975) signaleert dat tekstanalyse-methoden voor bepaalde onderzoeksopzetten te gedetailleerd kunnen zijn. Zelf beperkt hij zich in zo'n geval tot een globale analyse: 'In the present study, it was decided not to work with a detailed model of text structure, but rather to attempt to represent a major portion of the relational structure of a text . . .' (Frederiksen, 1975, 161).

Sanford & Garrod (1981, 73-84) bespreken de propositionele en de schema-methoden ook beknopt, en wel in het licht van geheugenopslag van verbale informatie. Zij toetsen de methoden echter ook als middel om de kern van een tekst te bepalen ('to establish gist'). Ze komen tot de slotsom dat deze methoden van een voorlopige aard zijn, en nog onvoldoende empirisch getoetst zijn om ze op een verantwoorde manier te kunnen beoordelen (Sanford & Garrod, 1981, 78).

Concluderend willen we stellen dat de huidige tekstanalyse-modellen wél gebruikt kunnen worden als uitgangspunt voor onderzoek naar de cognitieve processen bij tekstverwerking en reproductie, maar dat ze minder geschikt zijn voor het vaststellen van kerngedeelten in langere studieteksten. Onderzoekers als Foster (1979) en Meyer en McConkie (1973) moeten in dit opzicht teleurgesteld worden (zie pag. 335).

3 Beoordelingsmethoden

In plaats van tekstanalyse-methoden zou men ook beoordelingsmethoden kunnen hanteren bij het selecteren van kernpunten in teksten die in het onderzoek naar de verbetering van tekstbestudering worden gebruikt.

Een belangrijke reden voor onderzoekers om te kiezen voor beoordelaars in hun onder-

zoek is vaak, dat de taak nog zó complex en ondoorzichtig is, dat men er niet in kan slagen om een objectieve maatstaf, formule of expliciete strategie te geven om deze taak te verrichten. Van de beoordelaar wordt dan aangenomen dat deze bewust of intuïtief, een systeem volgt van redelijk vaste, maar onuitgewerkte normen (De Groot, 1972, 226-227). Empirisch kan dit blijken uit de betrouwbaarheid van de beoordelingen bij herhaling van de procedure.

Er zijn echter ook garanties nodig dat het systeem zelf niet te subjectief is. Het belangrijkste criterium hiervoor is de overeenstemming met de oordelen van andere beoordelaars: de interbeoordelaars-betrouwbaarheid.

In deze paragraaf zullen we eerst ingaan op de verschillende beoordelingsmethoden die in het onderzoek naar de verbetering van tekstbestudering (met name in het markerings-onderzoek) gebruikt worden. We zullen hierbij enige kritische kanttekeningen plaatsen en vervolgens ingaan op de geschiktheid van de beoordelingsmethode voor het eerder beschreven doel.

3.1 *Welke beoordelingsmethoden worden er in het onderzoek naar de verbetering van tekstbestudering gehanteerd voor het selecteren van kerngedeelten*

Hoewel in het meeste onderzoek naar de invloed van bepaalde bewerkingen van studieteksten alleen door de onderzoeker zelf wordt bepaald welke de kerngedeelten van de gebruikte tekst zijn en op welke gedeelten de bewerkingen (vragen e.d.) gericht zullen zijn (zie pag. 335), worden er in een aantal onderzoeken hiervoor ook beoordelingsprocedures gehanteerd. Dit laatste geldt met name voor het onderzoek naar de invloed van markeringen in teksten.

In Van Hout Wolters (1981) werden 20 onderzoeken geanalyseerd, waarin de tekst door docent of proffeider was gemarkeerd (onderstreept, gekleurd e.d.). In 6 daarvan, waarvan 4 door dezelfde auteurs, werd met behulp van externe beoordelaars vastgesteld welke gedeelten in de tekst gemarkeerd zouden worden (in de overige onderzoeken bepaalde de proffeider dat, op verschillende manieren, zelf).

In de onderzoeken van Coles & Foster (1975, 3 onderzoeken) en Foster & Coles (1977), waarin dezelfde tekst gebruikt werd,

werd deze tekst (1050 woorden) door de onderzoekers getransformeerd in een lijst van 61 beweringen ('statements'). De tekst en de statements werden aan 10 onafhankelijke beoordelaars (docenten) voorgelegd, die het belang van ieder statement voor de hoofdgedachte van de tekst op een driepuntsschaal moesten weergeven. De statements werden in volgorde van hun totaalscore gezet. De equivalenten van de eerste 28 statements werden in de tekst onderstreept (ongeveer 30% van de tekst).

Foster (1979) gebruikte een andere methode bij een tekst van 172 zinnen. Hij vroeg 20 psychologie-studenten en 6 psychologie-docenten de tekst te lezen en bij herlezing dié zinnen te onderstrepen die naar hun gevoel het meest belangrijk waren voor de gehele passage, m.a.w. die de kerngedachten bevatten die de auteur trachtte over te brengen. De helft van de groep mocht niet meer dan 16 zinnen onderstrepen, de andere helft niet meer dan 8. Er bleek hierbij weinig overeenstemming tussen de beoordelaars. Voor de keuze van te markeren zinnen (10%) werd het aantal keren dat iedere zin geselecteerd was in beide condities gesommeerd, en werden 16 zinnen gekozen die door 6 of meer beoordelaars geselecteerd waren.

Rickards & August (1975) gebruikten ook een beoordelingsmethode voor hun tekst, die uit 80 zinnen bestond. De methode was afkomstig van Johnson (1970). Hierbij werden 43 studenten in drie subgroepen verdeeld die resp. de taak hadden $\frac{1}{4}$, $\frac{1}{2}$ of $\frac{3}{4}$ deel van dié zinnen weg te strepen, die het minst belangrijk waren voor de inhoud van de passage. Het aantal keren dat een zin in alle condities was overgebleven (dus wel belangrijk geacht werd), verschafte een maat voor het structurele belang van de zin.

Ook Reder & Anderson (1980), die de tekstanalyse-methoden voor hun doel onge-schikt vonden (zie pag. 341), besloten te werken met beoordelaars. Ze lieten 'several people' samenvattingen maken, waarna over verschillpunten werd gediscussieerd totdat er een consensus ontstond. Men zou dit als een variant kunnen beschouwen van bovenbeschreven beoordelingsmethoden.

Sanford & Garrod (1980, 68) bevelen deze samenvattingsmethode aan om op deze manier de 'gist' van de passage vast te stellen. In een 'goede' samenvatting komen immers die ge-

deelten uit de tekst voor die het hoogste structurele belang hebben. Opgemerkt moet worden dat dit echter meestal géén letterlijke tekstdelen zullen zijn, omdat er bij het maken van samenvattingen meestal niet alleen selectie, maar ook generalisatie en constructie zal plaatsvinden (in termen van Van Dijk, 1978). Voor het vaststellen van (letterlijke) kerngedeelten voor het onderzoek naar de verbetering van tekstbestudering is deze methode daarom minder te prefereren.

3.2 *Het onderscheiden van teksteenheden*

In de vorige paragraaf werden een aantal beoordelingsmethoden beschreven die in het onderzoek naar het markeren van studieteksten worden gebruikt om vast te stellen welke tekstdelen in de experimentele conditie gemarkeerd moeten worden en waarover in de toets vragen kunnen worden gesteld (zie ook Tabel 2).

Bij al deze methoden ligt de omvang en inhoud van de te beoordelen teksteenheden vast: Foster & Coles (1977) en Coles & Foster (1975) gaan bij de beoordeling uit van beweringen of statements, die ze zelf opgesteld hebben, terwijl Foster (1979) en Rickards & August (1975) zinnen als teksteenheden gebruiken. In geen van deze onderzoeken worden beoordelingsmethoden gebruikt om, voorafgaand aan het oordeel over de kerngedeelten, vast te stellen welke tekst- of kenniseenheden ('linguistic units') de tekst precies omvat. Als dit echter wel zou gebeuren zouden deze teksteenheden ook gebruikt kunnen worden voor het analyseren van free-recall protocollen van de proefpersonen over de tekst.

Johnson (1970) probeerde de kenniseenheden in een tekst wél met behulp van beoordelaars vast te leggen. Hij liet daarbij beoordelaars een tekst segmenteren in zgn. 'pause acceptability units', m.a.w. hij liet hen plaatsen aanwijzen waar men zou kunnen pauzeren. In de instructie werd uitgelegd dat pauzeren om drie redenen kan plaatsvinden:

- om adem te halen,
- om nadruk te geven aan het verhaal,
- om de betekenis te ondersteunen.

Deze methode resulteerde in een betrouwbare segmentatie: er was een correlatie van .98 tussen de beoordelingen van twee verschillende groepen (resp. 23 en 50 'student raters').

Naast het onderscheiden van teksteenheden, zoals ook Levitt (1956) dat probeerde, stelde Johnson (1970) ook een rangorde op van het

Tabel 2 Een aantal gebruikte beoordelingsmethoden in tekstbestuderingsonderzoek

Onderzoekers	Selecteren van kerngedeelten		Onderscheiden van teksteenheden	
	Aantal en type beoordelaars	Methode	Type eenheden	Methode
Coles & Foster (1975)	10 docenten	3-puntsschaal	statements	door proefleiders onderscheiden
Foster (1979)	20 studenten en 6 docenten	8, 16 zinnen onderstrepen	originele zinnen	gegeven in tekst
Foster & Coles (1977)	10 docenten	3-puntsschaal	statements	door proefleiders onderscheiden
Johnson (1970) (2 onderzoeken)	96 resp. 99 studenten	1/4, 1/2, 3/4 deel wegstrepen	'pause acceptability units'	door 73 resp. 38 beoordelaars vastgesteld
Reder & Anderson (1980)	'several people'	maken van samenvattingen	n.v.t.	n.v.t.
Rickards & August (1975)	43 studenten	1/4, 1/2, 3/4 deel wegstrepen	originele zinnen	gegeven in tekst

structurele belang van de gevonden eenheden. Zijn methode hiervoor (wegstrepen van tekstdelen) werd in de vorige paragraaf bij het onderzoek van Rickards & August (1975) al beschreven.

In een aantal experimenten toonde Johnson vervolgens aan dat het vastgestelde structurele niveau van de teksteenheden samenhangt met de reproductie ervan (Johnson, 1970, 18-19).

3.3 Schematisch overzicht van een aantal gebruikte beoordelingsmethoden in tekstbestuderingsonderzoek

In Tabel 2 wordt een overzicht gegeven van de beoordelingsmethoden die de onderzoekers in de vorige paragrafen hanteerden.

3.4 Kanttekeningen bij de beschreven beoordelingsmethoden

Bij een aantal aspecten van de hier beschreven beoordelingsmethoden zijn kritische kanttekeningen te plaatsen. Deze zijn gedeeltelijk ontleend aan De Groot (1972, 244-250).

Representativiteit van de beoordelaars

In onderzoek naar het bewerken van studie-

teksten dienen de beoordelaars representatief te zijn voor degenen die normaal aanwijzingen geven in de studietekst over de kerngedeelten en de gedeelten waarover toetsvragen worden gesteld: de auteur en/of docent. Toch gebruiken Foster (1979) en Rickards & August (1975) geen docenten maar studenten als beoordelaars.

Aantal beoordelaars, interbeoordelaars-betrouwbaarheid en herhaalde metingen

Omdat de beoordelaars kunnen verschillen in hun oordelen, dienen meerdere beoordelaars in het onderzoek betrokken te worden. Het gewenste aantal hangt af van de verwachte grootte van deze verschillen. Door geen van de auteurs wordt echter een verantwoording van de aantallen gegeven. Ook wordt niet altijd de interbeoordelaars-betrouwbaarheid vermeld. Om de betrouwbaarheid te vergroten kunnen ook herhaalde beoordelingen plaatsvinden met een zeker tijdsverloop daartussen. Dit gebeurt in geen van de beschreven onderzoeken.

Onafhankelijkheid van beoordelaars

Om betrouwbare gegevens te verkrijgen is het van belang dat de beoordelaars onafhankelijk

werken. Ze mogen dus niet, zoals bij Reder & Anderson (1980), hun resultaten vergelijken en erover discussiëren.

Instructie van de beoordelaars

De instructie voor de taak dient zo expliciet mogelijk te zijn, liefst voorzien van voorbeelden en een trainingsperiode. Wat ons inziens hierbij vaak vergeten wordt, is het geven van een duidelijk doel van bestudering. Toch zullen de geselecteerde kerngedeelten bij het lezen met één bepaald doel kunnen verschillen van de delen die geselecteerd worden als met een ander doel gelezen wordt. Dit kan verklaren waarom er soms grote verschillen zijn tussen beoordelaars, zoals bij Foster (1979).

3.5 Het gebruik van beoordelingsmethoden voor onderzoek naar de verbetering van tekstbestudering

Er zijn allerlei redenen om van beoordelingsprocedures gebruik te maken in het onderzoek naar verbetering van tekstbestudering, als er tenminste met bovenstaande kanttekeningen rekening wordt gehouden.

- De beoordelingsmethoden zijn tamelijk eenvoudig toe te passen, ook bij langere teksten.
- De methoden zijn niet gekoppeld aan één bepaald type tekst (bijv. verhalen) met een bepaalde stereotype structuur; allerlei typen teksten kunnen op die manier geanalyseerd worden.
- Ook al zijn alle variabelen en strategieën die een rol spelen bij het selecteren van kerngedeelten in teksten nog niet geëxpliciteerd, ze zijn wel allemaal ingecalculeerd als er gewerkt wordt met beoordelingsmethoden. De beïnvloedende lezersvariabelen kunnen, in ieder geval gedeeltelijk, gecontroleerd worden door een precieze doel-instructie te geven en door te werken met meerdere beoordelaars (waarbij de interbeoordelaarsbetrouwbaarheid wordt vastgelegd en eventueel herhaalde metingen plaatsvinden).
- Uit Johnson's (1970) onderzoeken blijkt dat het door beoordelaars vastgelegde structurele belang van de teksteenheden ook psychologische realiteit heeft. Zowel onmiddellijk na het lezen als na enige tijd (7, 21 en 63 dagen) worden de eenheden met het hoogste structurele belang (de kerngedeelten) beter onthouden dan de andere eenheden.

Dit was het geval bij verschillende teksten.

Een interessante vraag is welke strategieën de beoordelaars hanteren bij het selecteren van de kerngedeelten in de teksten. Van de voornoemde onderzoekers gaan alleen Reder & Anderson (1980) daarop in. Zij trachten na te gaan volgens welke principes hun samenvattingen afgeleid zijn van de oorspronkelijke tekst en ze constateren dat die principes 'look like a composite of Van Dijk's (1977) summarization rules for shortening sentences and a rule for deleting "lower level" paragraphs and propositions' (Reder & Anderson, 1980, 124). De beoordelaars zouden volgens hen dus impliciet een soort macroregels gebruiken. Het lijkt ons van groot belang dat met procesgericht onderzoek meer inzicht verkregen wordt in deze gehanteerde strategieën.

4 Conclusie

Zo lang de taak van het vaststellen van het relatieve belang van tekstgedeelten nog zo ondoorzichtig is en de huidige tekstanalysemethoden nog zo beperkt en arbeidsintensief, zal het aan te bevelen zijn om te werken met externe beoordelaars.

Met deze conclusie worden de tekstanalysemethoden niet helemaal overboord gegooid: voor het verkrijgen van meer inzicht in het verwerken en reproduceren van tekstgehelen hebben ze zeker waarde, met name het model van Kintsch & Van Dijk (1978). Als we echter willen weten welke de kerngedeelten in een studietekst zijn,

- waarop we door markeringen, vragen, doelen, samenvattingen, schema's, e.d. extra aandacht willen richten,
 - waarop we de toetsvragen willen afstemmen en de resultaten willen analyseren,
- zijn beoordelingsprocedures met externe beoordelaars meer geschikt. Bij de keuze en het gebruik van een beoordelingsmethode zal echter wel rekening gehouden moeten worden met de gemaakte kanttekeningen.

Literatuur

Anderson, R. C., Schema directed processes in language comprehension. In: A. M. Lesgold e.a. (eds), *Cognitive psychology and instruction*. New York: 1978, 67-83.

- Coles, P. en J. Foster, Typographic cueing as an aid to learning from typewritten text. *Programmed Learning and Educational Technology*, 1975, 12, 102-108.
- Dam, G. van, De invloed van theoretische ontwikkelingen in de psychologie op de methoden van het onderzoek naar het onthouden van teksten. *Tijdschrift voor Taalbeheersing*, 1979, 1, 61-74.
- Dijk, T. A. van, Semantic macro-structures and knowledge frames in discourse comprehension. In: M. A. Just en P. Carpenter (eds.), *Cognitive processes in comprehension*. Hillsdale, N.J.: 1977, 3-32.
- Dijk, T. A. van, *Tekstwetenschap; een interdisciplinaire inleiding*. Utrecht/Antwerpen: 1978.
- Dijk, T. A. van, Recalling and summarizing complex discourse. In: W. Burghardt en K. Hölker (eds.), *Text processing; papers in text analysis and text description*. Berlin/New York: 1979, 49-118.
- Foster, J. J., The use of visual cues in text. In: P. A. Kolers, M. E. Wrolstad en H. Bouma (eds.), *Processing of visible language*. New York en Londen: 1979, 189-203.
- Foster, J. en P. Coles, An experimental study of typographic cueing in printed text. *Ergonomics*, 1977, 20, 57-66.
- Frederiksen, C. H., Effect of task-induced cognitive operations on comprehension and memory processes. In: R. Freedle en J. B. Carroll (eds.), *Language Comprehension and the Acquisition of Knowledge*. Washington: 1972, 211-246.
- Frederiksen, C. H., Acquisition of semantic information from discourse: effects of repeated exposures. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 158-169.
- Grimes, J. E., *The thread of discourse*. Ithaca: 1972.
- Grimes, J. E., *The thread of discourse*. Den Haag: 1974.
- Groot, A. D. de, *Methodologie. Grondslagen van onderzoek en denken in de gedragswetenschappen*. Den Haag: 1972 (7e druk).
- Hout Wolters, B. van, Verbetering van tekstbestudering; een analyse van onderzoek. In: C. Boonman en J. Zwarts, *Tekstbestudering*. Den Haag: SVO-reeks 1980, 7-52. (Interne uitgave: AOLO-rapport nr. 2).
- Hout Wolters, B. van, *Markeren van kernpunten in studieteksten door docent en lerende; een analyse van verricht onderzoek en een onderzoeksplan*. AOLO-rapport nr. 4, Universiteit Nijmegen, 1981.
- Johnson, R. E., Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 12-20.
- Kintsch, W., *The representation of meaning in memory*. Hillsdale, N. J.: 1974.
- Kintsch, W. en T. A. van Dijk, Toward a model of text comprehension and production. *Psychological Review*, 1978, 85, 363-394.
- Kok, W., C. Boonman en G. Beukhof, Studievaardigheden voor de basisschool: leren omgaan met teksten (1). *Pedagogische Studiën*, 1980, 57, 417-432.
- Mandler, J. M. en N. S. Johnson, Remembrance of things parsed: story structure and recall. *Cognitive Psychology*, 1977, 9, 111-151.
- Meyer, B. J. F., Identification of the structure of prose and its implications for the study of reading and memory. *Journal of Reading Behavior*, 1975a, 7, 7-47.
- Meyer, B. J. F., *The organization of prose and its effects on memory*. Vol. 1. New York: 1975b.
- Meyer, B. J. F. en G. W. McConkie, What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, 65, 109-117.
- Oostendorp, H. van, Tekststructuur bekeken vanuit de cognitieve psychologie. *Tijdschrift voor Taalbeheersing*, 1980, 2, 75-78.
- Reder, L. M. en J. R. Anderson, A comparison of texts and their summaries: memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 121-134.
- Rickards, P. en J. August, Generative underlining strategies in prose recall. *Journal of Educational Psychology*, 1975, 67, 860-865.
- Rumelhart, D. E. en A. Ortony, The representation of knowledge in memory. In: R. C. Anderson e.a. (eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: 1977, 99-137.
- Sanford, A. J. en S. C. Garrod, *Understanding written language; explorations of comprehension beyond the sentence*. Chichester: 1981.
- Thorndyke, P. W., Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, 9, 77-110.
- Tillema, H. en C. Boonman, *Het leren van teksten; literatuuronderzoek naar de structurering van verbale informatie voor gedifferentieerd onderwijs*. Project Procesgerichte differentiatie, deelrapport 1, R.U.U., 1979.

Curricula vitae

B. van Hout Wolters behaalde in 1969 het doctoraal examen psychologie aan de K.U. Nijmegen. Ze werkte tot 1978 bij het Centrum voor Didactiek en Onderzoek van Onderwijs (C.D.O.) van de T.H. Twente. Sindsdien is ze verbonden aan het Facultair Instituut Algemene Onderwijskunde voor de Lerarenopleiding (A.O.L.O.) van de K.U. Nijmegen. Ze publiceerde met name over de verbetering van lees- en studiemethoden en de samenstelling van schriftelijk studiemateriaal. Momenteel werkt ze aan een promotie-onderzoek m.b.t. de beïnvloeding van leerprocessen bij het bestuderen van studieteksten.

Adres: Facultair Instituut A.O.L.O., Katholieke Universiteit Nijmegen, Erasmusplein 1, 6500 HD Nijmegen.

B. Weltens behaalde in 1981 zijn doctoraal examen in de Algemene Taalwetenschap aan de K.U. Nijmegen. Gedurende zijn laatste studiejaar werkte hij als student-assistent bij het Facultair Instituut A.O.L.O. van de K.U. Nijmegen. Van september 1981 tot april 1982 verbleef hij als 'research student' bij het Department of Linguistic Science van de University of Reading (Engeland). Momenteel werkt hij bij het Instituut voor Fonetiek van de K.U. Nijmegen aan het gebruik van visuele feedback bij het aanleren van intonatie.

Adres: Instituut voor Fonetiek, Katholieke Universiteit Nijmegen, Erasmusplein 1, 6500 HD Nijmegen

Manuscript aanvaard 12-1-'82