

Eindwerkstuk Communicatiestudies

Effect van woordmoeilijkheid op tekstbegrip
bij sterke en zwakke lezers

Naam: Emmar Sinnige

Studentnr.: 3349160

Instelling: Universiteit Utrecht

Opleiding: Communicatie & Informatiewetenschappen

Opdracht: Bachelorthesis

Onderwerp: Leesbaarheid van schoolteksten

Begeleider: Rogier Kraf

Datum: November 2010



Abstract

Dit onderzoek test of de woordmoeilijkheid van een tekst invloed heeft op tekstbegrip bij sterke en zwakke lezers. Om tekstbegrip te meten is een cloze-test, ofwel gatentekst ontworpen die gemanipuleerd is op woordmoeilijkheid. Deze test is samen met een leesvaardigheidstoets en een woordenschattoets afgenomen bij 90 leerlingen, waarvan de helft 2^e klas Gymnasiumleerlingen waren en de andere helft 2^e klas Vmbo/Havo-leerlingen. Uit de woordenschattoets en leesvaardigheidstoets bleek dat de Gymnasiumleerlingen zoals voorspeld sterke lezers zijn en de Vmbo/Havo-leerlingen zwakkere lezers. De resultaten wijzen uit dat woordmoeilijkheid van een tekst inderdaad effect heeft op tekstbegrip. De makkelijke versie van de cloze-test werd significant beter gemaakt dan de moeilijke versie. Daarnaast blijkt dat de Gymnasiumleerlingen beter op zowel de moeilijke als de makkelijke tekstversie scoren dan de Vmbo/Havo-leerlingen. Ook het verschil tussen de scores op de moeilijke en makkelijke versie was een stuk groter bij de zwakkere lezers dan bij de sterke lezers. De invloed van woordmoeilijkheid lijkt uiteindelijk een groter effect te hebben op het tekstbegrip van zwakkere lezers dan op tekstbegrip van sterke lezers.



Inhoudsopgave

Inleiding.....	4
1. Theoretisch kader	5
1.1 Woordmoeilijkheid	5
1.2 Voorkennis.....	6
1.3 Cloze test.....	6
1.3.1 Validiteit.....	7
1.3.2 Cloze-item moeilijkheid.....	8
1.3.3 Mogelijke cloze-testen.....	8
1.3.4 Antwoordopties.....	10
1.4 Hypotheses.....	11
1.4.1 Woordmoeilijkheid	11
1.4.2 Lezersniveau.....	11
1.4.3 Interactie tussen woordmoeilijkheid en lezersniveau	11
1.5 Belang.....	12
2. Methode	13
2.1 Materialen.....	13
2.1.1 Manipulaties.....	13
2.1.2 Bepalen van cloze-items	13
2.1.3 Andere testen.....	14
2.2 Proefpersonen en procedure	15
2.3 Design.....	15
2.4 Scores	16
3. Resultaten	17
3.1 Betrouwbaarheden.....	17
3.1.1 Leesvaardigheidstoets.....	17
3.1.2 Cloze-test.....	18
3.2 Effect van schoolniveau en tekstversie op de verschillende toetsen	18
3.2.1 Woordenschat	18
3.2.2 Leesvaardigheid.....	19
3.2.3. Cloze.....	20
3.3 Schooltype-effect nader bekeken.....	22
3.3.1 Correlaties.....	23
3.3.2 Co-varianten.....	23
3.4 Cloze-items nader bekeken	25
3.4.1 Moeilijke cloze-items	25
3.4.2 Makkelijke cloze-items.....	27
3.4.3 Verschil in resultaten van beide versies.....	27
4. Conclusie	29
5. Discussie	30
Literatuurlijst	33
Bijlagen	35
Bijlage 1: Originele tekst 'Tweede Wereldoorlog'	36
Bijlage 2: Regels voor Cloze-ontwerp	37
Bijlage 3: Gemanipuleerde teksten inclusief uitleg gekozen cloze-items.....	42
Bijlage 4: Moeilijke en makkelijke tekstversie inclusief gaten	45
Bijlage 5: Verdeling teksten en versies per schooltype	46
Bijlage 6: Goed gerekende niet-letterlijke antwoorden bij cloze-test.....	47
Bijlage 7: SPSS uitdraai.....	49



Inleiding

'Het is in het jaar 1535. Op het plein in de stad is een tribune gebouwd. Ook is een schuur opgericht. Enkele palen steken omhoog. Wat is er aan de hand? Er zullen pannen rechtgezet worden. Ze zijn in optocht door de stad naar het plein gebracht.'

Dit stukje tekst klinkt raar en onsamenhangend. Toch is dit wat een Vietnamese leerling begreep van een stukje tekst uit een geschiedenisboek omdat hij enkele woorden niet kende. De originele tekst luidde namelijk:

"Het is in het jaar 1535. Op het plein in de stad is een tribune gebouwd. Ook is een schavot opgericht. Enkele palen steken omhoog. Wat is er aan de hand? Er zullen ketters terechtgesteld worden. Ze zijn in optocht door de stad naar het plein gebracht."

Het komt niet zelden voor dat leerlingen een tekst niet begrijpen omdat ze enkele woorden niet kennen. Docenten Nederlands stuiten vaak op slecht tekstbegrip en slechte woordenschat bij hun leerlingen. De woordkennis van bijna 15% van de leerlingen op middelbare scholen, ongeacht het niveau, blijkt te laag te zijn voor de moeilijkheidsgraad van de schoolteksten (Hacquebord, Linthorst en Stellingwerf 2004), wat betekent dat een deel van de leerlingen een te kleine woordenschat heeft of dat het niveau van de schoolteksten te hoog ligt. Woordenschat en begrijpend lezen zijn volgens mevrouw dr. H.I. Hacquebord, universitair hoogdocent Toegepaste Taalwetenschap aan de Rijksuniversiteit Groningen, echter van cruciaal belang voor schoolsucces.¹ Met begrijpend lezen bedoelt zij het begrijpend lezen van schoolboekteksten. Aangezien schoolboekteksten tot doel hebben informatie over te dragen aan leerlingen, is het van groot belang dat leerlingen deze teksten snappen.

Dit onderzoek richt zich dan ook op de leesbaarheid van schoolteksten, waarbij gekeken wordt naar de relatie tussen moeilijke woorden en tekstbegrip. Uit onderzoek van Stahl (2003) blijkt namelijk dat moeilijke woorden in een tekst grote invloed hebben op de begrijpelijkheid van de tekst. Daarnaast wordt in dit onderzoek bekeken of deze tekstenmerken verschillende invloed hebben op tekstbegrip bij sterke en zwakke lezers. De hoofdvraag van dit onderzoek is dan ook:

Heeft de woordmoeilijkheid van een tekst invloed op tekstbegrip bij sterke en zwakke leerlingen?

In hoofdstuk 1 worden eerdere bevindingen over dit onderwerp in kaart gebracht. In hoofdstuk 2 wordt de methode, het gebruik van cloze-testen, uitgelegd. De resultaten worden in hoofdstuk 3 weergegeven, waarop vervolgens in hoofdstuk 4 de conclusie volgt. Het onderzoek wordt afgesloten met een discussie en aanbevelingen voor vervolgonderzoek.

¹ H.I. Hacquebord, lezing 'Maakt taal verschil?' 24 sept 2003
<http://www.rug.nl/let/voorzieningen/etoc/lezinghacquebord>



1. Theoretisch kader

Om leerlingen iets te laten leren is het, zoals in de inleiding is vermeld, belangrijk dat de leerlingen de schoolteksten die ze te lezen krijgen snappen.

Er zijn verschillende manieren om tekstbegrip te voorspellen en te meten. Een van de manieren om tekstbegrip te voorspellen is een 'readability formula'. Een leesbaarheidsformule toetst de leesbaarheid van teksten aan de hand van tekstkenmerken (Chall 1995 en Zakaluk & Samuels 1988). Leesbaarheid is volgens Staphorsius (1994) de door de tekst veronderstelde leesvaardigheid. Door middel van woordfactoren en zinslengte probeert de leesbaarheidsformule woordmoeilijkheid en grammaticale complexiteit te meten (Chall 1958).

De leesbaarheidsformule lijkt op het eerste gezicht een bruikbare formule, er kunnen echter een aantal kanttekeningen geplaatst worden bij de betrouwbaarheid van deze formule. Zo wordt bij de leesbaarheidformule bijvoorbeeld alleen gekeken naar tekstkenmerken, terwijl juist de interactie tussen tekst en lezer belangrijk is bij de leesbaarheid van een tekst. De specifieke kenmerken van de lezer, zoals voorkennis en vaardigheden, worden in de leesbaarheidformule geheel buiten beschouwing gelaten (Jansen en Woudstra 1979). Daarnaast is de leesbaarheidsformule gebaseerd op het gemiddelde kind. Volgens de leesbaarheidsformule zijn teksten over bekend veronderstelde onderwerpen makkelijker te lezen dan teksten over minder bekende onderwerpen. Een gemiddeld kind bestaat echter niet. Wat bekend en dus eenvoudig te lezen is voor één kind, kan voor een ander kind onbekend en moeilijk zijn (Stahl 2003). Uit de leesbaarheidformule blijkt echter wel één bruikbaar gegeven voor dit onderzoek: de leesbaarheidformule heeft aangetoond dat moeilijkheid van woorden de belangrijkste factor is bij het bepalen van de moeilijkheid van een tekst. Woordfactoren hebben hierop veel meer invloed dan zinsfactoren (ibidem). Tevens blijkt uit onderzoek van Stahl et al. (1989) dat woordmoeilijkheid begripelijkheid beïnvloedt. Dit is de reden waarom ik in dit onderzoek ga kijken naar de moeilijkheid van woorden in relatie tot tekstbegrip.

1.1 Woordmoeilijkheid

Het is lastig te bepalen wat precies moeilijke woorden zijn en wat niet. Eén manier om woordmoeilijk te bepalen is te kijken naar het aantal lettergrepen per woord. Hoe meer lettergrepen een woord bevat, hoe moeilijker het woord zou zijn. Dit stuit echter op problemen omdat langere woorden niet altijd moeilijker hoeven te zijn dan korte woordjes (Stahl 2003).

Een andere manier om woordmoeilijkheid te bepalen is gebruik te maken van 'makkelijke woordenlijsten' zoals de Dale-Chall (Chall 1995) of de Harris-Jacobson lijsten (Harrison 1980). Op deze lijsten staan frequent gebruikte woorden, die beschouwd worden als makkelijke woorden. Als een woord niet op de lijst staat wordt deze als moeilijk gezien. Dit lijkt een snelle manier om moeilijke van makkelijke woorden te onderscheiden, maar er kleven een aantal nadelen aan deze lijsten. Ten eerste rijst de vraag: wat zijn frequente woorden en wat niet? Als een lijst 4000 woorden bevat, is het 4000^e woord voor een beginner dan even makkelijk als het eerste woord op een lijst (Stahl 2003)? Ten tweede, woorden kunnen meerdere (moeilijkere) betekenissen hebben, waar in de woordenlijst geen rekening mee gehouden wordt (ibidem). Ten derde hebben deze lijsten problemen met sommige zelfstandige naamwoorden zoals namen van personen. Als veel verschillende namen in een stukje



tekst voorkomen zal deze tekst waarschijnlijk als moeilijker worden ervaren dan een stukje tekst waarin maar weinig namen voorkomen. Dit is niet opgenomen in alle woordenlijsten (ibidem).

Tot slot kunnen woorden ingedeeld worden in functiewoorden en inhoudswwoorden. Functiewoorden zijn woorden van een gesloten klasse, die een syntactische functie hebben. Voorbeelden van functiewoorden zijn voegwoorden, voornaamwoorden en lidwoorden. Inhoudswwoorden construeren, zoals het woord *al* zegt, de inhoud van de zin. Ze bevatten semantische informatie. Voorbeelden van inhoudswwoorden zijn bijvoeglijke naamwoorden, werkwoorden, bijwoorden en zelfstandige naamwoorden (Stahl et al. 1989). Volgens Taylor, Aborn et al., Tannenbaum et al., Klein-Braley, Perkins en German en Brown (in: Abraham en Chapelle 1992) zijn functiewoorden over het algemeen makkelijker dan inhoudswwoorden.

Onder het kopje '1.3.2. Cloze-item moeilijkheid' worden een aantal tekstfactoren genoemd die invloed kunnen hebben op de moeilijkheid van een in te vullen woord in een gatentekst.

1.2 Voorkennis

Naast woordmoeilijkheid bepalen ook voorkennis en interesse de mate van tekstbegrip. Stahl et al. (1989) hebben onderzoek gedaan naar de interactie tussen voorkennis en moeilijke vocabulaire bij tekstbegrip. Uit de resultaten blijkt dat er geen interactie bestaat tussen voorkennis en woordmoeilijkheid, maar dat deze onafhankelijk van elkaar effect hebben op tekstbegrip. Volgens hen beïnvloedt moeilijke vocabulaire microprocessen. Microprocessen gaan over de ontwikkeling van een microstructuur, of een tekstfundament, wat een samenhangend en geordend stel gegevens is die ideeën uit de tekst weergeven (Kintsch en van Dijk, 1978). Woordmoeilijkheid heeft volgens Omanson (1985) invloed op het verwerken van individuele gegevens, wat daardoor de creatie van een samenhangende tekstrepresentatie kan beïnvloeden. Dit valt te goed begrijpen: als kinderen een tekst lezen die ze moeilijk kunnen begrijpen, zullen ze zich minder goed een voorstelling kunnen maken van de tekst dan kinderen die de een makkelijkere tekst lezen en deze wel goed begrijpen. Onder microprocessen wordt letterlijk begrip van de tekst verstaan, zoals het herhalen van informatie uit de tekst, het beantwoorden van expliciete vragen over de tekst of het invullen van de exacte functiewoorden in een gatentekst. Op deze aspecten heeft woordmoeilijkheid invloed. Uit dit onderzoek blijkt dus dat woordmoeilijkheid een belangrijke indicator is voor letterlijk tekstbegrip.

Uit onderzoek van Stahl et al. (1989) komt ook naar voren dat voorkennis juist macroprocessen beïnvloedt. Macroprocessen hebben te maken met het ontwikkelen van een macrostructuur. Dit is het vermogen de belangrijkste dingen uit de tekst te kunnen selecteren, wat overeenkomt met het maken van een samenvatting. Iemand die bekend is met een onderwerp, door voorkennis bijvoorbeeld, heeft minder moeite een goede samenvatting van een stuk te schrijven dan iemand waarvoor een tekstonderwerp onbekend is (ibidem). Macroprocessen hebben dus niet te maken met letterlijk naar de zinnen en woorden in een tekst kijken, maar de tekst in een groter geheel te zien en de grote lijnen van een tekst te herkennen.

1.3 Cloze test

Een cloze-test is een veel gebruikte testmethode om tekstbegrip te meten (Taylor 1953). Bij een cloze-test worden woorden uit een tekst weggehaald, waardoor een 'gatentekst' ontstaat. De lezer wordt gevraagd de tekst te reconstrueren door op elk gat het weggelaten woord in te vullen. Het idee achter



de cloze-test is dat bij het invullen van de gaten dezelfde denkprocessen meespelen als bij het lezen van een tekst. Het percentage goede reconstructies wordt gezien als relatieve aanwijzing van de leesbaarheid van de tekst. Naarmate een tekst redundanter is, zou deze volgens Taylor gemakkelijker te lezen zijn en daardoor ook gemakkelijker te reconstrueren (Staphorsius en Sanders 2008). Klare (1984) heeft aangetoond dat cloze-testen het effect van woordmoeilijkheid op tekstbegrip van een tekst meten.

1.3.1 Validiteit

Er kleven echter wel een aantal nadelen aan de resultaten van de cloze-testen. Ten eerste meten cloze-testen vooral lokale begripelijkheidsprocessen en geen globale begripelijkheidsprocessen ofwel macroniveau. Dit blijkt uit onderzoeken van onder andere Kintsch en Yarbrough (1982), Shanahan, Kamil en Robin (1982, beide in: Stahl et al. 1989). Zij vonden namelijk dat cloze-testen alleen lokale zinsbegripelijkheid testen en niet de integratie tussen de zinnen. Uit onderzoek van Alderson (1979b) blijkt eveneens dat cloze-testen taalkundige ('lower-order') in plaats van conceptuele ('higher-order') taalvaardigheden meten en ook Kamalski (2007) toont aan dat 70% van de resultaten van een cloze-test toegeschreven kunnen worden aan op tekst gebaseerde (linguïstieke) representaties. Ten tweede zijn tekstbegrip en het in kunnen vullen van de gaten niet hetzelfde, het laatste vereist namelijk grotere productieve vaardigheden (Kobayashi 2002). Ten derde blijken ook andere factoren naast taalvaardigheid bij te dragen aan de testperformance (Abraham & Chapelle 1992).

Toch blijken cloze-testen wel valide meetinstrumenten op gebied van tekstbegrip te zijn. Kamalski (2007) heeft onderzoek gedaan naar welke testmethodes het beste een situatiemodelrepresentatie meten. Een situatiemodel is vergelijkbaar met een mentale representatie van iets. Door een mentale representatie te vormen van, in dit geval, een tekst, is het makkelijker zinnen aan elkaar te relateren en op deze manier een gehele tekst te begrijpen. Hierbij wordt eigen kennis gekoppeld aan wat gelezen wordt. Volgens het situatiemodel wordt de informatie uit een tekst dus geïntegreerd met eigen kennis (Endsley en Georgia, 2000). Als een tekst niet begrepen wordt, kan iemand een heel ander of onsamenhangend beeld van die tekst vormen dan werd bedoeld en daardoor een ander beeld over die tekst ontwikkelen. Het situatiemodel kent daarom een nauwe relatie met tekstbegrip.

Kamalski (2007) concludeert dat van de vier geteste methodes de cloze-test op de derde plaats komt te staan. Cloze-testen meten volgens haar gedeeltelijk (voor 30%) situatiemodel representaties, alleen is dat niet het voornaamste dat een cloze test meet, het meet namelijk voor 70% linguïstieke kennis. Wat voor dit onderzoek van belang is, is dat naar voren komt dat de internal reliability van de cloze-test zeer hoog is: zowel de scores op de Cronbach's alpha en op de split-half betrouwbaarheidsmeting lagen bij vier geteste methodes tussen de .79 en de .91. Dit houdt in dat de verschillende items (in dit geval de gaten in de tekst) allemaal hetzelfde construct meten. De cloze-test blijkt intern dus zeer betrouwbaar te zijn en eveneens veel betrouwbaarder dan een question asking test, sorting task test of een mental model test. Als een toets een hele lage interne betrouwbaarheid heeft, is deze eigenlijk niet bruikbaar omdat de verschillende items niet hetzelfde toetsen. Je kan dan geen betrouwbare resultaten uit de toets halen. Bij de cloze-test is dit dus gelukkig wel het geval.

Daarnaast blijkt de cloze-test een acceptabele known-group validity te hebben: de cloze-test kan effectief onderscheid maken tussen groepen. Zo blijkt uit de cloze-test van Kamalski (2007) dat 12^e



klassers gemiddeld significant hoger scoren dan 10^e klassers, wat logisch is aangezien 12^e klassers een hoger niveau zouden moeten hebben. Dat de cloze-test onderscheid kan maken tussen groepen is van belang aangezien ik in mijn onderzoek het verschil tussen Gymnasium en Vmbo/Havo leerlingen ga onderzoeken. De cloze-toets blijkt dus een betrouwbaar instrument te zijn om deze twee niveaus van elkaar te onderscheiden.

Tot slot blijkt dat voorkennis geen grote invloed heeft op de scores van een cloze-test.

1.3.2 Cloze-item moeilijkheid

Uit bovenstaande wordt duidelijk dat cloze-testen, ondanks dat enige kanttekeningen geplaatst kunnen worden, wel degelijk valide meetinstrumenten kunnen zijn om lokaal tekstbegrip te meten. Abraham en Chapelle (1992) vermelden dat de resultaten op de cloze-testen voorspeld kunnen worden door (contextuele) item kenmerken. De belangrijkste factoren die cloze-item moeilijkheid voorspellen zijn (1) de frequentie waarin het weggelaten woord elders in de tekst voorkomt, (2) het aantal lettergrepen in de zin waarin het weggelaten woord staat, (3) aantal letters van het woord, (4) de categorie waarin het woord ingedeeld kan worden: functiewoorden of inhoudswoorden en (5) de hoeveelheid vereiste context dat nodig is om het weggelaten woord te kunnen bepalen.

Uit hun onderzoek komt naar voren dat bij een fill-in cloze-test² een woord makkelijker te herplaatsen is als deze vaker voorkomt in de tekst (#1). De lengte van de zin waarin het weggelaten woord staat blijkt echter geen indicator te zijn voor itemmoeilijkheid (#2). Bij een rational cloze-test³ zijn weggelaten woorden makkelijker te herplaatsen als de lengte van het te voorspellen woord kort is (#3) en weinig context nodig is (#5). Daarnaast blijkt ook, eveneens als uit onderzoek van Stahl et al. (1989), dat woordmoeilijkheid een significant verschil heeft op het herplaatsen van weggelaten functiewoorden. Woordmoeilijkheid heeft echter geen invloed op het herplaatsen van inhoudswoorden (#4).

Tot slot maken Campbell (1961) en Carrol (1989) onderscheid tussen intrinsieke en extrinsieke tekstkenmerken om de moeilijkheid van cloze-items te bepalen. Intrinsieke tekstkenmerken zijn objectieve kenmerken zoals inhoudselementen, functie/inhoudswoorden en aantal lettergrepen. Intrinsieke kenmerken vallen onder te verdelen in inhoudelijke stimulus, cognitieve processen en type vereiste antwoorden. Extrinsieke tekstkenmerken zijn subjectieve kenmerken zoals voorkennis en oplosstrategieën van de lezer (Abraham en Chapelle 1992).

Om puur en alleen woordmoeilijkheid te meten is het bij het ontwerpen van de cloze-test en de bijbehorende tekstmanipulaties van belang dat met deze aspecten rekening gehouden wordt.

1.3.3 Mogelijke cloze-testen

Cloze-testen kunnen op verschillende manieren uitgevoerd worden. Abraham en Chapelle (1992) hebben in hun onderzoek drie cloze-testen met elkaar vergeleken om te onderzoeken wat de

² Nadere uitleg over een fill-in clozetest volgt in het volgende deelpopje

³ Nadere uitleg over een rational clozetest volgt ook in het volgende deelpopje



verschillende cloze-testen werkelijk meten. Hiervoor gebruikten zij een fixed-ratio cloze-test, een rational cloze-test en een multiple-choice cloze-test.

Fixed-ratio cloze-test

Een fixed-ratio is een objectieve test waarbij random, maar met een vast interval, woorden weggelaten worden. De eerste twee zinnen blijven in tact en vervolgens wordt bijvoorbeeld elk 8^e woord geëlimineerd. De lezer moet zelf bedenken welk woord het meest toepasselijk is op de open plek in de tekst.

Uit het onderzoek kwam naar voren dat de scores op de fixed-ratio test opgevat kunnen worden als de vaardigheid inhoudelijke woorden vanuit het lange termijngeheugen op te vragen, of antwoorden letterlijk uit de tekst te kunnen halen. Kenmerken die effect hebben op de scores van de fixed-ratio test zijn namelijk: frequentie van voorkomen van letterlijke weggelaten woord in de tekst, inhoud-/functiewoorden en aantal vormen dat het juiste antwoord aan kan nemen.

Rational cloze-test

Een rational cloze-test is een subjectievere test omdat zorgvuldig uitgezocht wordt welke woorden weggelaten worden. De twee criteria hiervoor waren in het onderzoek van Abraham en Chapelle: in dezelfde passage als waarin het woord verwijderd wordt, moeten twee duidelijk identificeerbare aanwijzingen aanwezig zijn en er moet genoeg bruikbare informatie te vinden zijn buiten de betreffende zin waarin het verwijderde woord staat.

De scores op de rational test tonen aan in hoeverre de lezer contextaanwijzingen kan gebruiken, oftewel informatie uit de context kan halen, zowel binnen de zin waarin het weggelaten woord staat als in de rest van de tekst. Kenmerken die effect hebben op de scores van een rational cloze-test zijn: de hoeveelheid contextuele informatie, lengte, afweging van inhoudswoorden en functiewoorden, aantal mogelijke antwoorden en aantal vormen dat het juiste antwoord aan kan nemen.

Uit onderzoek van Sachs, Tung en Lam (1997) blijkt de rational cloze-test betrouwbaardere resultaten op te leveren dan een fixed-ratio. Omdat de rational cloze-test een subjectieve test is, biedt dit de mogelijkheid preciezer te testen: in het onderzoek is gekeken naar welk proces gemeten wordt als een bepaalde woordsoort weggelaten wordt. Hierbij is onderscheid gemaakt tussen analytische processen en perceptuele processen. Een analytisch proces houdt in dat de lezer het probleem expliciet kan definiëren, de informatie in de tekst analyseert en daarna tot een antwoord voor het weggelaten woord komt. Een perceptueel proces is een automatisch proces waarbij de lezer de tekst leest en zonder expliciete definitie van het probleem het antwoord invult. Hierbij gaat het vooral om aangeleerde uitdrukkingen en standaard constructies. Uit het onderzoek bleek dat het weglaten van inhoudswoorden en het weglaten van random woorden met een vast interval (bijvoorbeeld elk 7^e woord) het analytische proces testen. Hierbij is het weglaten van inhoudswoorden (rational cloze-test) een betrouwbaardere methode om het analytische proces te meten dan de 'every seventh word deletion' methode (fixed-ratio). Random woorden weglaten (zonder vast interval) en het weglaten van voorzetsels test het perceptuele proces. Hierbij is het weglaten van voorzetsels (rational cloze) een betrouwbaardere methode dan het random weglaten van woorden.



Multiple choice cloze-test

Tot slot hebben Abraham en Chapelle een multiple-choice cloze-test vergeleken. Hierbij werden dezelfde woorden weggelaten als bij de rational test, alleen hoefden de lezers de in te vullen woorden niet zelf te bedenken, maar konden ze kiezen uit 4 invulmogelijkheden. Uit het onderzoek blijkt dat de scores van de multiple-choice test niet hetzelfde meten als de twee andere invultesten. Wat de multiple-choice test daadwerkelijk precies meet wordt niet duidelijk.

Aangezien ik me in dit onderzoek richt op het daadwerkelijke tekstbegrip en niet direct op het lange termijngeheugen, en de rational-cloze betrouwbaardere resultaten oplevert dan een fixed-ratio, lijkt de rational cloze-test het meest geschikt te zijn. Dit kan nog onderbouwd worden door het onderzoek van Kamalski (2007) waarin wordt vermeld dat cloze-testen representaties van situatiemodellen preciezer kunnen meten als woorden niet random uit de tekst worden gehaald, maar zorgvuldig worden uitgekozen op pragmatische en tekstuele componenten. Voorbeelden hiervan zijn gaten op woorden waarvoor eigen wereldkennis gecombineerd moet worden met tekstuele aanwijzingen en gaten op woorden waarvoor lezers zinnen moeten linken door de tekststructuur te volgen. Hierbij is het nodig dat informatie uit verschillende delen van de tekst geïntegreerd wordt. Bij een fixed-ratio tekst vallen gaten namelijk dikwijls op plekken waarvoor alleen grammaticale en lexicale kennis nodig is om de woorden in te vullen, zonder dat de tekst ook daadwerkelijk begrepen hoeft te worden.

1.3.4 Antwoordopties

Bij een rational cloze-test is het de bedoeling dat de lezer zelf de gaten in de tekst invult. Het is mogelijk dat de lezer niet het precieze woord invult dat bedoeld was, maar bijvoorbeeld een synoniem. Er zijn verschillende scoringsmethoden om de antwoorden op rational en fill-in cloze-testen te beoordelen op juistheid. Kobayashi (2002) heeft de interactie tussen drie verschillende scoringsmethoden en itemkenmerken met elkaar vergeleken. De scoringsmethoden die hij onderzocht heeft zijn (1) exacte scoringsmethode, waarbij alleen het exact bedoelde antwoord goed gerekend wordt, (2) semantisch en syntactisch acceptabele woord scoringsmethode en (3) semantisch acceptabel, maar syntactisch onacceptabele woord scoringsmethode. Uit de resultaten blijkt dat op de 'acceptabele' scoringsmethodes (2 en 3) de hoogste gemiddelden werden behaald, waarbij de semantisch acceptabele woord scoringsmethode (3) lichtelijk boven de andere (2) uitstak. Het belangrijkste is dat de twee laatste methodes ook een consistent hogere betrouwbaarheid kenden en daardoor geschiktere methodes zijn dan de exacte scoringsmethode.

Kobayashi (2002) heeft daarnaast getest welke invloed de verschillende scoringsmethodes hebben op de resultaten van zeer goede en minder goede leerlingen. Daaruit blijkt dat de zeer goede leerlingen bij de laatste twee scoringsmethodes veel hoger scoren dan bij de exacte scoringsmethode. De minder goede leerlingen scoorden gemiddeld lager dan de goede leerlingen en bij de minder goede leerlingen was het verschil in scores tussen de drie methodes minder groot. Hieruit zou afgeleid kunnen worden dat zeer goede leerlingen het vermogen hebben om meerdere geschikte woorden te verzinnen die ook passend zijn in de context, maar niet het exact bedoelde woord zijn. Bij gebruik van de exacte scoringsmethode worden deze leerlingen dus onterecht 'gestraft' door hun bredere woordenschat. Omdat in dit huidige onderzoek ook naar het verschil tussen goede en slechte leerlingen wordt gekeken, zou het op basis van het onderzoek van Kobayashi niet eerlijk zijn de exacte



scoringmethode te gebruiken. Een van de twee 'acceptabele' woord scoringmethodes is dus een bruikbaarere methode in dit geval.

1.4 Hypotheses

Het hoofddoel van dit onderzoek is bekijken of het tekstkenmerk 'woordmoeilijkheid' invloed heeft op tekstbegrip en of hierbij verschil bestaat tussen goede en slechte lezers.

1.4.1 Woordmoeilijkheid

Uit onderzoek van Zakaluk en Samuels (1988) blijkt dat makkelijkere woorden in een tekst de begrijpbaarheid van de tekst niet meteen verhogen. Dit valt te wijten aan de relatie tussen vocabulaire en kennis van de wereld. De wereldwijze manier van denken van de lezer, in plaats van specifieke woorden, bepaalt volgens Stahl (2003) of de lezer een tekst goed begrijpt. Uit ander onderzoek van Stahl et al. (1989) blijkt echter dat woordmoeilijkheid wel degelijk invloed heeft op tekstbegrip. Woordmoeilijkheid en voorkennis hebben volgens hen namelijk onafhankelijk invloed op tekstbegrip. Daarnaast komt uit onderzoek van Chall (1958) naar voren dat teksten makkelijker te lezen zijn als woorden vervangen worden door makkelijkere synoniemen. Aangezien Chall (1958) ook heeft aangetoond dat de vocabulaire factor in de leesbaarheidformule een van de sterkste voorspellers is van tekstmoeilijkheid, verwacht ik dat de makkelijke tekstversie beter begrepen wordt dan de moeilijke tekstversie.

H1: De makkelijke tekstversie wordt beter begrepen dan de moeilijke tekstversie.

1.4.2 Lezersniveau

Uit onderzoek van Kobayashi (2002) blijkt dat de hoge proficiency groep beter scoort op de testen dan de lage proficiency groep. Ook uit onderzoek van Kamalski (2007) blijkt dat 12^e klassers significant hoger scoren dan 10^e klassers. Dit zijn voor de hand liggende resultaten omdat een hoger leesniveau vrij waarschijnlijk leidt tot beter tekstbegrip. Aangezien deze resultaten zeer aannemelijkheid zijn, verwacht ik ook dat goede leerlingen, dus de Gymnasiumleerlingen, hoger scoren op zowel makkelijke als moeilijke teksten dan de Vmbo/Havo-leerlingen.

H2: Gymnasiumleerlingen scoren gemiddeld beter op beide tekstversies dan Vmbo/Havo-leerlingen.

1.4.3 Interactie tussen woordmoeilijkheid en lezersniveau

Ik verwacht dus dat een hoger lezersniveau leidt tot beter tekstbegrip. Ik heb in paragraaf '1.4.1 Woordmoeilijkheid' beschreven dat ik verwacht dat woordmoeilijkheid een rol speelt, dus een tekst met een hoge woordmoeilijkheidsgraad slechter begrepen wordt dan een tekst met makkelijke woorden. Woordmoeilijkheid en lezersniveau spelen natuurlijk samen een rol in tekstbegrip. Dit zou betekenen dat ook de Gymnasiumleerlingen meer moeite zouden moeten hebben met een moeilijke tekstversie, daarom verwacht ik dat bij zowel de Gymnasiumleerlingen als de Vmbo/Havo-leerlingen een tekst met een hoge woordmoeilijkheid leidt tot slechter tekstbegrip.



H3: Zowel de Gymnasiumleerlingen als de Vmbo/Havo-leerlingen scoren gemiddeld lager op de moeilijke versie.

Daarbij verwacht ik dat bij de Vmbo/Havo-leerlingen het verschil tussen de scores op de makkelijke en moeilijke versie groter zullen zijn dan bij de Gymnasium leerlingen. Het niveau van de Vmbo/havo-leerlingen is namelijk een stuk lager dan bij de Gymnasiumleerlingen. De Vmbo/Havo-leerlingen zullen de makkelijke tekstversie waarschijnlijk wel begrijpen, maar meer moeite hebben met de moeilijke tekstversie. Ik verwacht daarom dat deze leerlingen veel lager scoren op de moeilijke tekstversie. Het niveau van de Gymnasiumleerlingen is een stukje hoger, vandaar dat ik verwacht dat zij minder moeite zullen hebben met de moeilijke versie dan de Vmbo/Havo-leerlingen en daar dus hoger op scoren, terwijl zij ook hoog zullen scoren op de makkelijke versie. De resultaten van de moeilijke en de makkelijke versie zullen daarom dichterbij elkaar liggen.

H4: Het verschil tussen de scores op de moeilijke en makkelijke versie zal bij de Vmbo/Havo-leerlingen groter zijn dan bij de Gymnasiumleerlingen.

1.5 Belang

Verschillende onderzoeken naar de relatie tussen tekstbegrip en tekstkenmerken leiden dikwijls tot verschillende resultaten. Dit zou kunnen worden verklaard door de verschillende doelen of nadruk, manipulaties, testmethodes, antwoordopties en te vergelijken variabelen die gebruikt zijn in de onderzoeken. Zakaluk & Samuels (1988) hebben ondervonden dat ondanks dat woordmoeilijkheid leesbaarheid voorspelt, een tekst met makkelijkere woorden die tekst niet meteen begrijpbaarder maakt. In het artikel van Stahl et al. uit 1989 worden drie aan elkaar gerelateerde onderzoeken beschreven. Uit het eerste onderzoek blijkt echter dat een tekst minder begrijpelijk is als woorden vervangen worden door moeilijke woorden. Uit het tweede onderzoek van Stahl et al. (1989) komt, tegengesteld aan de resultaten van het eerste onderzoek, naar voren dat woordmoeilijkheid geen invloed heeft op het herplaatsen van inhoudswoorden in een cloze-test. Dit terwijl van Sachs, Tung en Lam (1997) juist aantonen dat het analytische leesproces gemeten kan worden in een cloze-test waarin inhoudswoorden uit een tekst weggelaten worden. Aangezien deze resultaten niet eenduidig zijn, is verder onderzoek op dit gebied noodzakelijk.

Daarnaast incorporeren vele onderzoeken naar tekstbegrip meerdere componenten als zinsfactoren, voorkennis, contextuele informatie, verschillende soorten testen en methodes. Er is weinig onderzoek verricht dat zich alleen op woordmoeilijkheid zelf richt. Dit is de reden waarom in dit onderzoek alleen naar woordmoeilijkheid wordt gekeken.

Tot slot is nog geen onderzoek gedaan naar woordmoeilijkheid en tekstbegrip waarbij Gymnasiumleerlingen en Vmbo/Havo leerlingen met elkaar vergeleken worden. Dit zijn juist de leerlingen die de meeste baat hebben bij goed op hen afgestemde schoolboekteksten.



2. Methode

2.1 Materialen

Om tekstbegrip te meten heb ik voor dit onderzoek een cloze-test ontworpen. Hiervoor is één tekst gebruikt die zo is gemanipuleerd dat er twee verschillende versies ontstonden; een moeilijke en een makkelijke tekst. Vervolgens zijn hier zorgvuldig woorden uitgekozen die weggelaten gaan worden en die door de lezers ingevuld moeten worden, de zogenaamde 'gaten' in de tekst.

2.1.1 Manipulaties

Een tekst over de Tweede Wereldoorlog uit het Historisch Canon is als brontekst gebruikt, waarvan het niveau is vastgesteld op eind basisschool/begin middelbare school. De originele tekstversie bevat 298 woorden en is te vinden in bijlage 1. Deze tekst is gemanipuleerd op woordmoeilijkheid door woorden in de moeilijke versie te vervangen door moeilijkere woorden en woorden in de makkelijke versie te vervangen door makkelijkere woorden. Vooral werkwoorden, zelfstandige naamwoorden, bijvoeglijke naamwoorden en een enkel bijwoord zijn gemanipuleerd. Dit is gedaan aan de hand van een synoniemenlijst. Bij de prestaties op de cloze scores spelen naast woordmoeilijkheid nog vele andere factoren een rol zoals zinslengte, hoeveelheid verbindingswoorden, de plek van de clozes, de frequentie waarin het weggelaten woord voorkomt en de hoeveelheid context nodig is om het gat in te vullen (zie hiervoor paragraaf 1.3.2 'Cloze-item moeilijkheid'). Natuurlijk moet alleen de invloed van woordmoeilijkheid op tekstbegrip gemeten worden, vandaar dat erop gelet is dat de zinsstructuur in beide versies zoveel mogelijk hetzelfde bleef en er alleen enkele woorden of woordgroepen werden vervangen door moeilijkere of makkelijkere synoniemen. In elke zin bleef tenminste één zelfstandig naamwoord hetzelfde. Op deze manier kunnen de cloze-gaten later in beide versies op dezelfde niet-gemanipuleerde woorden vallen. Als de zinnen compleet veranderd worden is het namelijk lastig om in beide versies dezelfde woorden weg te laten vallen. Tot slot is ervoor gezorgd dat minimaal één op de tien woorden is gemanipuleerd, wat er op neerkomt dat één op de vijf inhoudswoorden is gemanipuleerd. Er is bewust gekozen voor dit aantal vanwege de publicatie van Freebody en Anderson (1983a), waarin staat dat één manipulatie op de drie inhoudswoorden te veel is. De woordmoeilijkheid is in dat geval zo hoog dat leerlingen niet in staat zijn dat de compenseren met hun eigen kennis.

2.1.2 Bepalen van cloze-items

Na het manipuleren van de tekst moest vastgesteld worden welke woorden weggelaten zouden worden. De rational cloze-test heeft een aantal voordelen ten opzichte van de fixed-ratio cloze-test: De rational cloze-test test het vermogen om contextuele informatie uit een tekst te gebruiken in plaats van vooral grammaticale en lexicale kennis te testen. Daarnaast is de rational cloze-test een veel betrouwbaardere test dan de fixed-ratio cloze-test om het analytische vermogen te testen. Door een rational cloze-test te gebruiken en de weg te laten woorden zorgvuldig uit te zoeken, kan het lokale karakter van de cloze-test, wat als groot kritiekpunt gebruikt wordt tegen de cloze-test, voor een deel weggenomen worden waardoor het meetgebied van de cloze-test uitgebreid wordt. Door bijvoorbeeld vooral inhoudswoorden weg te laten kan het analytische vermogen van de lezer getest worden. Daarnaast kan hierbij rekening gehouden worden met de pragmatische en tekstuele componenten. Op



deze manier test de cloze-test meer dan alleen linguïstieke componenten. Voor dit onderzoek heb ik dan ook een rational-cloze test gemaakt. Dit is gedaan aan de hand van het document 'Regels voor cloze-ontwerp', opgesteld door H. Pander Maat en R. Kraf. Zij hebben een procedure opgesteld om een rationele cloze-test te ontwikkelen die op een intersubjectief betrouwbare manier tekstbegrip meet. Dit document is te vinden in bijlage 2.

De eerste zin van het document is in tact gelaten om de lezer de mogelijkheid te bieden in de tekst te komen. Verder zijn alleen niet-gemanipuleerde woorden weggelaten omdat de manipulatie anders niet meer bijdraagt aan de moeilijkheid van de tekst. Om het analytische vermogen van de lezer te testen zijn vooral inhoudswoorden weggelaten. Als woorden uit vaste uitdrukkingen of phrasal verbs weggelaten worden en ingevuld moeten worden, worden daarmee alleen linguïstische vaardigheden en lokale zinsbetekenis kennis gemeten. Een leerling hoeft namelijk alleen een zinsdeel of zin te begrijpen om dat woord in te vullen, zonder de context te snappen of verbanden tussen zinnen te leggen. Daarom zijn die woordgroepen niet weggelaten als 'in te vullen gat'. Zelfstandige naamwoorden en werkwoorden bevatten wel vaker centrale informatie. Als deze woorden uit de tekst geëlimineerd worden en de leerling weet het 'gat' van het juiste woord te voorzien, kan gezegd worden dat de leerling de tekst heeft begrepen. Om het juiste woord in te vullen heeft hij/zij waarschijnlijk verbanden moeten leggen met andere delen uit de tekst. Daarom zijn vooral zelfstandige naamwoorden en werkwoorden weggelaten, in plaats van veel bijwoorden en bijvoeglijke naamwoorden. Als meerdere zelfstandige naamwoorden in één zin in aanmerking kwamen zijn de woorden veelal gekozen volgens de syntactische ranglijst: subject, object, indirect object, bepalingen. Tot slot is ernaar gestreefd één woord op elke twaalf woorden weg te laten. In bijlage 3 is zowel de moeilijke als makkelijke tekstversie in een schema geplaatst, inclusief de gekozen weg te laten woorden (gaten), voorzien van uitleg over waarom voor die woorden is gekozen. Bijlage 4 bevat de moeilijke en makkelijke versie zoals de leerlingen deze hebben gekregen.

2.1.3 Andere testen

Niet alleen de gatentekst over de Tweede Wereldoorlog wordt gebruikt, maar ook nog een tweede gatentekst: 'De eerste spoorlijn'. Deze gatentest is van een collega en neem ik alleen af. Hier wordt in dit onderzoek verder niets mee gedaan.

Daarnaast wordt een leesvaardigheidstoets afgenomen. Deze toets is ontwikkeld door het CITO en is afgestemd op het niveau van 2^e klassers. Het CITO is het Nederlandse nationale examinatie bureau en staat in de top drie van internationale organisaties in testontwikkeling en het meten van leerprestaties. De leesvaardigheidstest bestaat uit 25 korte tekstjes waarover per tekst één meerkeuze vraag wordt gesteld. De toets wordt afgenomen om het leesniveau van de lezer te meten om aan de hand daarvan te kunnen bepalen wat sterke en wat zwakke lezers zijn. Deze resultaten van de leesvaardigheidstoets zijn een voorspellende factor voor de cloze-test. Hoe beter de leesvaardigheidstoets gemaakt wordt, hoe beter de antwoorden op de cloze-test waarschijnlijk zijn. Door deze twee testen te correleren kan gemeten worden of leesvaardigere lezers tot een beter tekstbegrip komen. Leesvaardigheid kan echter ook een onbedoelde beïnvloedende factor zijn op de cloze-test. Om ook de resultaten van de cloze-test te bekijken zonder dat deze beïnvloed wordt door de leesvaardigheid, kan het effect van de leesvaardigheid op tekstbegrip uiteindelijk gecorrigeerd



worden door de resultaten van de leesvaardigheidstoets als co-variaat te gebruiken. Daardoor heeft leesvaardigheid geen onbedoeld effect op de manipulaties van tekstbegrip.

Tot slot wordt een woordenschattoets afgenomen. Deze woordenschattoets is ontwikkeld door de Rijksuniversiteit Groningen. De woordenschattoets bevat 50 zinnestukjes waarin één woord is onderstreept. Middels een meerkeuze-antwoord kunnen de leerlingen aangeven wat de betekenis is van het onderstreepte woord. Ook voor de resultaten van de woordenschattoets geldt dat deze een voorspelling geven voor de resultaten op de cloze-testen. Hoe beter de woordenschattoets gemaakt wordt, dus hoe hoger de woordenschat van een lezer is, hoe beter deze de moeilijke cloze-test versie zal begrijpen. Ook hier geldt dat de woordenschat een onbedoeld effect kan hebben op tekstbegrip en als co-variaat meegenomen kan worden, zodat de scores op de woordenschattoets gelijk getrokken worden naar het gemiddelde en aan de hand daarvan de scores op de cloze-items worden berekend. Zowel de woordenschattoets en leesvaardigheidstoets zijn voorzien van aparte antwoordbladen.

2.2 Proefpersonen en procedure

Voor dit onderzoek zijn 90 proefpersonen getest. Alle 90 proefpersonen vallen in de leeftijdscategorie 13/14 jaar en zitten in de tweede (brug)klas van de middelbare school 'het OSG Erasmus' in Almelo. Van de proefpersonen hebben 46 leerlingen het niveau Vmbo/Havo en zijn verdeeld over twee klassen. De andere 44 leerlingen zijn leerlingen uit twee Gymnasiumklassen. Er is gekozen voor deze twee niveaus aangezien ik in dit onderzoek ook het verschil tussen goede en minder goede lezers wil testen. Deze niveaus liggen vrij ver uit elkaar waardoor hopelijk duidelijke verschillen naar voren kunnen komen.

Bij elke leerling zijn vier toetsen afgenomen waarvoor zij één blokketijd (100 minuten) de tijd hadden. De Gymnasiumleerlingen hebben alle toetsen binnen 70 minuten kunnen afronden, de Vmbo/Havo-leerlingen deden daar wat langer over. Bijna alle Vmbo/Havo-leerlingen hadden wel genoeg aan 100 minuten om alle toetsen af te ronden. Elke klas heeft dezelfde uitleg gekregen waarin nadrukkelijk werd gevraagd eerst één toets af te ronden voordat ze aan de volgende toets begonnen, zodat ze niet in de war raakten of elke toets maar half afronden. Daarbij is een indicatie van benodigde tijd voor elke toets gegeven waar ze zich aan moesten proberen te houden. Daarnaast is verteld dat het geen proefwerk is voor een cijfer, maar dat de docenten wel op de hoogte worden gesteld van de resultaten, dus dat serieus invullen belangrijk is. Bij elke klas bleef de docent het gehele blokketijd surveilleren. Aan het einde van het blokketijd zijn alle testen en antwoordformulieren tegelijkertijd ingenomen. Om de leerlingen extra te motiveren om serieuze antwoorden te geven werden ze naderhand voor hun bijdrage beloond met iets lekkers.

2.3 Design

Zoals is gezegd werd de leerlingen gevraagd 4 toetsen te maken. Elke leerling kreeg een identieke woordenschattoets en leesvaardigheidstoets en twee gatenteksten. Van elke gatentekst is een moeilijke en een makkelijke versie gemanipuleerd. Het is natuurlijk niet mogelijk om de leerlingen zowel de moeilijke als de makkelijke versie van dezelfde gatenteksten in te laten vullen. In dat geval hebben de manipulaties geen zin meer omdat in beide teksten de gaten op dezelfde woorden vallen. Elke leerling heeft dus één versie van beide gatenteksten gekregen, waarbij ervoor werd gezorgd dat



ze één makkelijke versie en één moeilijke versie kregen. In tabel 1 zijn het aantal afgenomen moeilijke en makkelijke versies te zien per tekst.

Tabel 1: Aantal leerlingen per tekst per tekstversie

	Makkelijk	Moeilijk	Totaal
Tekst 1	41	49	90
Tekst 2	49	41	90
Totaal	90	90	

Eén Vmbo/Havo klas en één gymnasiumklas hebben de makkelijke versie van 'De Tweede Wereldoorlog' (tekst 1) gekregen en de moeilijke versie van 'De eerste spoorlijn' (tekst 2). De andere Vmbo/Havo klas en Gymnasiumklas hebben de moeilijke versie van 'De Tweede Wereldoorlog' en de makkelijke versie van 'De eerste spoorlijn' gekregen. Omdat in elke klas zowel betere als minder goede leerlingen zijn, zal de spreiding in de versies ongeveer gelijk zijn. In bijlage 5 is deze verdeling van teksten en tekstversies over de verschillende klassen weergegeven. In bijlage 5 valt ook te zien hoeveel makkelijke en moeilijke versies per tekst bij beide niveaus zijn ingevuld. De tekst 'De eerste spoorlijn' zal ik verder niet meenemen in dit onderzoek. In tabel 2 valt de verdeling van tekstversie en schooltype te zien voor de tekst 'De Tweede Wereldoorlog'.

Tabel 2: Verdeling schooltype en versietekst voor de cloze-test

Schooltype	Versietekst		
	Makkelijk	Moeilijk	Totaal
Vmbo/Havo	20	26	46
Gymnasium	21	23	44
Totaal	41	49	90

2.4 Scores

Omdat uit de resultaten van Kobayashi (2002) blijkt dat de acceptabele woord scoringsmethode een bruikbaarere methode is dan de exacte woord scoringsmethode omdat goede leerlingen daardoor niet onterecht gestraft worden, kijk ik de cloze-test na aan de hand van de acceptabele woord scoringsmethode. Dit houdt in dat niet alleen de exacte weggelaten woorden als antwoord goed zijn gerekend, maar ook synoniemen van de weggelaten woorden, woorden met een semantische overeenkomst of woorden die logischerwijs passend en correct zouden zijn in de context van de tekst. In bijlage 6 zijn deze goed gerekende antwoorden per cloze-item te vinden. Bij de leesvaardigheidstoets en woordenschattoets is uiteraard maar één dezelfde antwoordoptie per vraag goed gerekend.



3. Resultaten

Nadat alle gegevens uit de testen zijn ingevoerd in SPSS heb ik eerst de betrouwbaarheden uitgerekend van de leesvaardigheidstoets en de cloze-test (paragraaf 3.1). Als deze betrouwbaar blijken te zijn kan gekeken worden of het schoolniveau en de tekstversie van invloed is op de uitslagen van de testen (paragraaf 3.2). Vervolgens wordt het effect van de leesvaardigheid en woordenschat op tekstbegrip bekeken en in hoeverre de moeilijkheidsgraad van de tekst invloed heeft op tekstbegrip als gecorrigeerd is voor leesvaardigheid en woordenschat (paragraaf 3.3). Tot slot ga ik bekijken aan welke (op leerling toegespitste) aspecten het zou kunnen liggen dat het ene cloze-item beter is gemaakt dan het andere cloze-item (paragraaf 3.4).

3.1 Betrouwbaarheden

De resultaten van de leesvaardigheidstoets en de cloze-test zijn per antwoord ingevuld in het databestand. Dit is gedaan omdat het bij de cloze-test belangrijk is te bekijken of de verschillende items hetzelfde construct meten. De leesvaardigheidstoets is de laatst gemaakte toets, hierdoor is het voorgekomen dat enkele leerlingen de toets niet af hebben gekregen of hun concentratie aan het eind is verminderd. Daarom is het belangrijk om te kijken of ook de laatste vragen serieus zijn ingevuld. Tevens bevatte deze toets verschillende vragen over teksten, dus is het zinvol te meten of alle antwoorden wel correleren. De resultaten van de woordenschattoets zijn bij elkaar opgeteld omdat elke vraag hetzelfde was geconstrueerd, deze toets in het midden heeft plaatsgevonden en daardoor waarschijnlijk door elke leerling wel serieus en geconcentreerd is ingevuld. Alleen de totaalscore van het aantal goede antwoorden op de woordenschattoets is in het databestand gezet, vandaar dat daar geen betrouwbaarheidsmeting per item op uitgevoerd kan worden.

3.1.1 Leesvaardigheidstoets

De leesvaardigheidstoets kent een gemiddelde correlatie tussen alle items van $\alpha=0.714$. Aangezien $\alpha>0.650$, betekent dit dat de items van de leesvaardigheidstoets samenhangen, en dit daardoor een betrouwbare toets is. Vraag 18 uit de leestest is slecht gemaakt. Door het weglaten van deze vraag verhoogt de betrouwbaarheid naar 0.738. Omdat ik de antwoorden op de cloze-test later ga correleren met de leesvaardigheidantwoorden en ik hieraan het niveau van de leerlingen ga bepalen vind ik het belangrijk dat de betrouwbaarheid zo hoog mogelijk is. Omdat het verwijderen van vraag 18 de betrouwbaarheid met 0.024 verhoogt heb ik ervoor gekozen deze vraag te verwijderen zodat de uiteindelijke betrouwbaarheid van de leesvaardigheidstoets uitkomt op 0.738.

Uit de gemiddeldes blijkt dat de laatste vragen slechter zijn gemaakt dan de eerste vragen. Het gemiddelde resultaat op de eerste helft (vraag 1 t/m 14) is 9.7 en op de tweede helft 5.3. Een T-test wijst uit dat er sprake is van een significant verschil: $p<0.001$. Dit zou kunnen komen door vermindering van concentratie. De betrouwbaarheid blijkt echter niet te vergroten door het weglaten van de laatste vragen ($\alpha=0.650$ bij items 1 t/m 14). De laatste vragen laat ik daarom gewoon mee tellen.



3.1.2 Cloze-test

De gemiddelde correlatie van de antwoorden op de cloze-test is $\alpha=0.675$. Aangezien $\alpha>0.650$, maakt dit de cloze-toets betrouwbaar. Alle items correleren met elkaar. De betrouwbaarheidsgegevens zijn te vinden in bijlage 6. Hierbij is geen onderscheid gemaakt tussen de makkelijke en de moeilijke versie. Voor de makkelijke versie geldt $\alpha=0.702$, voor de moeilijke tekstversie is $\alpha=0.634$. Item 19 is gemiddeld zeer goed gemaakt in de moeilijke versie (mean= 0.94). Als vraag 19 daarin weggelaten wordt stijgt de α naar 0.652, wat precies voor betrouwbaar door zou gaan. De betrouwbaarheid voor de makkelijke en de moeilijke versie samen zou nauwelijks verschillen als vraag 19 geëlimineerd wordt ($\alpha=0.673$). De betrouwbaarheid voor de makkelijke versie zou echter sterk dalen naar 0.672. Ik heb er daarom ook niet voor gekozen vraag 19 te elimineren.

3.2 Effect van schoolniveau en tekstversie op de verschillende toetsen

Nu de testen betrouwbaar blijken te zijn, kan gekeken worden of het niveau van de leerlingen invloed heeft op de leesvaardigheid, woordenschat en tekstbegrip van de leerlingen. Tevens kan worden bekeken of schoolniveau en tekstversie invloed hebben op de resultaten van de cloze-testen en of de ene tekstversie door leesvaardigere lezers of lezers met grotere woordenschat is gelezen dan de andere tekstversie. Als dit het geval is zal de leesvaardigheidstoets en/of woordenschattoets als co-variant gebruikt moeten worden om het effect van de leesvaardigheid en woordenschat te corrigeren.

3.2.1 Woordenschat

Hiervoor is een meervoudige ANOVA uitgevoerd waarin het effect van schooltype en tekstversie op woordenschat (afhankelijke variabele) getoetst wordt. In tabel 3 worden de gemiddeldes en standaarddeviatie weergegeven van de woordenschattoets per schoolniveau en tekstversie. De woordenschattoets kent geen moeilijke of makkelijke versie, dus het zou raar zijn als er grote verschillen zouden bestaan tussen beide versies. Nul procent van de variantie in de woordenschattoets wordt dan ook verklaard door de tekstversie ($F(1): 0.001$; $p=0.978$ en $\eta^2= 0.000$). Dit betekent dat het niveau van woordenschat en goede testindicator zou kunnen zijn als er verschillen optreden tussen beide versies in de cloze-test.

Tabel 3: Gemiddeldes en standaarddeviaties van de woordenschattoets per schoolniveau per tekstversie

Schooltype	Versietekst	Gem.	Std. Deviatie	Aantal leerlingen
Vmbo/Havo	Makkelijk	37,40	3,95	20
	Moeilijk	37,46	4,85	26
	Totaal	37,43	4,44	46
Gymnasium	Makkelijk	42,67	3,51	21
	Moeilijk	42,65	3,10	23
	Totaal	42,66	3,26	44
Totaal	Makkelijk	40,10	4,55	41
	Moeilijk	39,90	4,85	49
	Totaal	39,99	4,69	90



Schooltype blijkt een groot en significant effect te hebben op de woordenschat ($F(1): 39.00; p < 0.001$). Gymnasium scoort ongeveer 5 punten hoger dan het Vmbo/Havo. 31,2% van de variantie in de woordenschattoets wordt verklaard door het schooltype ($\eta^2=0.312$). Dit betekent dat de Gymnasiumleerlingen een grotere woordenschat hebben dan de Vmbo/Havo leerlingen en is daardoor een indicatie voor sterkere lezers. Omdat tekstversie geen enkele invloed heeft op woordenschat bestaat er geen interactie-effect tussen schooltype en versietekst.

3.2.2 Leesvaardigheid

Ook voor de leesvaardigheidtoets is een meervoudige ANOVA uitgevoerd om te bekijken of schoolniveau invloed heeft op leesvaardigheid. De Leesvaardigheidtoets is de afhankelijke variabele, schoolniveau en tekstversie zijn de onafhankelijke variabelen. In tabel 5 staan de gemiddeldes en standaarddeviaties van de antwoorden op de leesvaardigheidtest per schoolniveau en per tekstversie.

Tabel 4: Gemiddeldes en standaarddeviaties van de leesvaardigheidtoets per schoolniveau per tekstversie

Schooltype	Versietekst	Gem.	Std. Deviatie	Aantal leerlingen
Vmbo/Havo	Makkelijk	13,40	2,48	20
	Moeilijk	12,85	3,75	26
	Totaal	13,09	3,24	46
Gymnasium	Makkelijk	14,38	4,38	21
	Moeilijk	19,39	2,81	23
	Totaal	17,00	4,40	44
Totaal	Makkelijk	13,90	3,57	41
	Moeilijk	15,92	4,67	49
	Totaal	15,00	4,30	90

Het gymnasium scoort gemiddeld bijna 5 punten hoger op de leesvaardigheidtoets dan de Vmbo/Havo-leerlingen. Het schoolniveau heeft dan ook een significant effect op de leesvaardigheid ($F(1): 26.48; p < 0.001$). 23,5 procent van de variantie wordt verklaard door het schoolniveau ($\eta^2=0.231$), wat een redelijk effect is. Dit bevestigt samen met de resultaten van de woordenschattoets dat de Gymnasiumleerlingen in dit onderzoek de sterke lezers zijn en de Vmbo/Havo-leerlingen de zwakkere lezers.

Ook hier blijken de gemiddeldes bij de makkelijke en moeilijke versie bij het vmbo/havo niveau dicht bij elkaar te liggen. De gemiddeldes bij het gymnasium liggen verder uit elkaar: 14,4 tegenover 19,4. Aangezien $p > 0.05$, is dit significant effect ($F(1): 9.28; p=0.003$). Maar een zeer klein percentage (9,7 procent) van de variantie kan verklaard worden door de tekstversie ($\eta^2=0.097$).

Toch is het vreemd dat er een significant verschil bestaat tussen de tekstversies, aangezien de leesvaardigheidtoets geen moeilijke en makkelijke versies kent, maar één dezelfde versie. Uit een one-way ANOVA test blijkt dat de mannen significant slechter scoren op de leesvaardigheidtoets dan vrouwen ($F(1): 10.89; p=0.001$). De gemiddeldes en standaarddeviaties per toets per geslacht staan



uitgewerkt in tabel 5. Vrouwen scoren gemiddeld 3 punten hoger dan mannen bij de leestest, wat betekent dat zij leesvaardiger zijn dan mannen.

Tabel 5: De gemiddeldes en standaarddeviaties per toets per geslacht

Toets	Geslacht	Aantal leerlingen	Gem.	Std. Deviatie
Cloze_totaal	man	47	18,06	2,56
	vrouw	35	18,34	2,80
	Totaal	82	18,18	2,65
Leestest_totaal	man	47	13,83	3,95
	vrouw	35	16,83	4,23
	Totaal	82	15,11	4,31
Woordenschat_totaal	man	47	40,28	4,16
	vrouw	35	40,37	4,90
	Totaal	82	40,32	4,46

De verdeling van geslacht is niet gelijk over de verschillende versies. In de makkelijke versie zitten 25 mannen en 14 vrouwen, de moeilijke versie kent 22 mannen en 21 vrouwen. Aangezien vrouwen significant hoger scoren dan mannen en de moeilijke versie veel meer vrouwen bevat dan de makkelijke versie, kan deze verdeling over de versies zorgen voor een significant versieverschil in de resultaten van de leesvaardigheidstoets.

Er bestaat tevens een significant interactie-effect tussen schooltype en tekstversie bij de leesvaardigheidstoets ($F(1,1): 14,47; p < 0.001$). In Vmbo/Havo klassen zaten samen 13 vrouwen en 27 mannen, Gymnasium kent 20 mannen en 22 vrouwen. Dit interactie-effect valt dus waarschijnlijk ook te verklaren door de scheve geslachtsverdeling in de beide tekstversies én in beide schoolniveaus. Dit effect kan echter maar voor 14,1% procent verklaard worden door de interactie, wat zeer weinig is.

Bij de cloze-test en woordenschattoets liggen de gemiddeldes erg dicht bij elkaar en valt geen significant verschil op te merken tussen geslacht (Cloze: $F(1): 0.22; p = 0.640$; Woordenschat: $F(1): 0.01; p = 0.925$). Dit zou de reden kunnen zijn waarom bij de woordenschattoets geen significant interactie effect tussen schooltype en versietekst waargenomen is. Omdat de cloze-test ook geen significant verschil tussen man en vrouw kent, hoeft bij de tekstbegripresultaten in het volgende kopje (3.2.3 Cloze) geen rekening gehouden te worden met de schuine verdeling van het geslacht.

3.2.3. Cloze

De gemiddeldes en standaarddeviaties van de antwoorden op de cloze-test staan in tabel 6 per schooltype en per tekstversie uitgewerkt. Figuur 1 geeft een visuele weergave. De volledige variantie-analyse op de cloze-test staat in de SPSS uitdraai in bijlage 6. Ook hier zien we dat schooltype een significant effect heeft op de cloze-totalen ($F(1): 28.21; p < 0.001$). 24,7% van de variantie in de cloze-



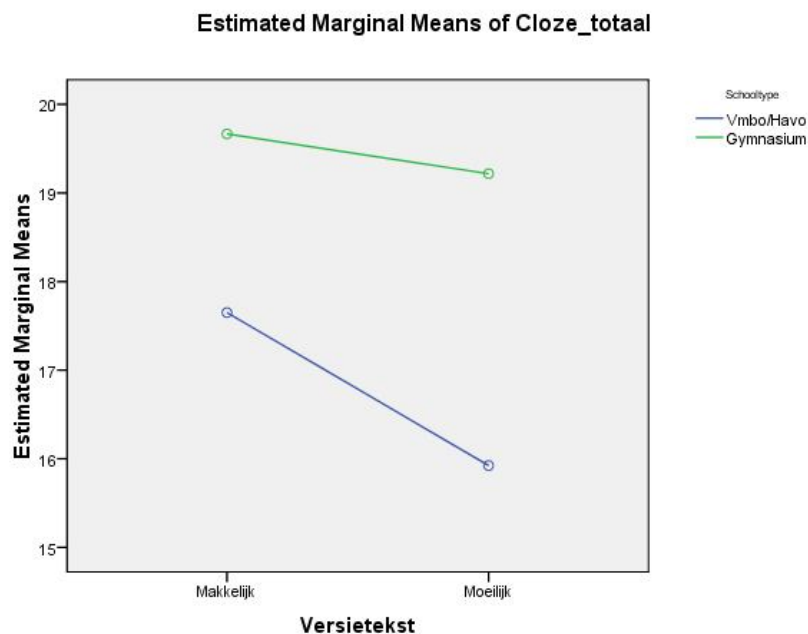
testantwoorden kan verklaard worden door schooltype, wat een redelijk effect is. Het Gymnasium heeft, ongeacht de tekstversie, dus een significant beter tekstbegrip dan de Vmbo/Havo-leerlingen.

Hypothese 2: 'Gymnasiumleerlingen scoren gemiddeld beter op beide tekstversies dan Vmbo/Havo leerlingen' kan hierdoor bevestigd worden.

Tabel 6: Gemiddeldes en std. van de cloze-test per schoolniveau en tekstversie

Schooltype	Versietekst	Gem.	Std. Deviatie	Aantal leerlingen
Vmbo/Havo	Makkelijk	17,65	3,13	20
	Moeilijk	15,92	2,68	26
	Totaal	16,67	2,98	46
Gymnasium	Makkelijk	19,67	1,77	21
	Moeilijk	19,22	1,51	23
	Totaal	19,43	1,63	44
Totaal	Makkelijk	18,68	2,70	41
	Moeilijk	17,47	2,75	49
	Totaal	18,02	2,78	90

Figuur 1: Tekstbegrip per niveau per tekstversie



Aangezien de cloze-testen gemanipuleerd zijn op moeilijkheid zou de makkelijke versie beter gemaakt moeten zijn dan de moeilijke versie. Beide schoolniveaus scoren hoger op de makkelijke versie. Het verschil tussen de twee tekstversies blijkt inderdaad significant te zijn ($F(1)=4.74$; $p=0.032$). Alleen een klein percentage, 5,2%, kan verklaard worden door van de verschillende versies ($\eta^2=0.052$).



De Vmbo/Havo-leerlingen hebben de moeilijke versie bijna 2 punten slechter gemaakt dan de makkelijke versie. Bij de Gymnasiumleerlingen zijn de gemiddeldes van de moeilijke en makkelijke versie bijna gelijk aan elkaar. Uit een T-toets blijkt dan ook dat de moeilijkheidsgraad van de tekst bij de Vmbo/Havo-leerlingen significante invloed heeft op tekstbegrip: $t(44)=2.01$; $p=0.050$. Bij het Gymnasium is geen significant verschil tussen de moeilijke en makkelijke versie: $t(42)=0.91$; $p=0.368$. Dit zou het kleine percentage van 5,2% van de tekstversie-variantie kunnen verklaren. Dit betekent ook dat de woordmoeilijkheid van een tekst grotere invloed heeft op tekstbegrip bij zwakkere lezers dan bij sterke lezers.

De resultaten uit de ANOVA ondersteunen dit echter niet. Al blijkt uit de T-toets dat de woordmoeilijkheid meer effect heeft gehad bij de Vmbo/Havo'ers dan bij de Gymnasiumleerlingen, van een significant interactie-effect tussen schooltype en tekstversie is volgens een meervoudige ANOVA geen sprake ($F(1,1)=1.63$; $p=0.205$). Dit zou betekenen dat de moeilijkheid niet meer of minder effect heeft gehad bij de verschillende schooltypes.

Afzonderlijke T-toetsen per schooltype hebben minder waarnemingen dan de ANOVA, waardoor ze minder snel een significant effect aan tonen. In dat geval zou ik resultaten uit de ANOVA beter kunnen interpreteren. Toch lijkt het resultaat op de T-toets zeer aannemelijk vanwege het kleine verschil van 0,45 tussen de moeilijke en makkelijke versie. Het effect van woordmoeilijkheid op de Gymnasiumleerlingen is dus twijfelachtig.

Hypothese 3: 'Zowel de Gymnasiumleerlingen als de Vmbo/Havo-leerlingen scoren gemiddeld lager op de moeilijke versie' kan deels aangenomen worden: inderdaad scoren ze beide lager op de moeilijke versie. Bij de Vmbo/Havo-leerlingen is dit verschil significant, dus kan de hypothese bevestigd worden. Of dat ook voor de Gymnasiumleerlingen geldt is dubieus. Zij scoren inderdaad iets lager op de makkelijke versie, maar of dit significant is of niet, valt niet te zeggen vanwege de twee tegenstrijdige toetsresultaten. Op basis van de T-toets, die mij het meest aannemelijk lijkt, zou ik zeggen dat deze hypothese voor het Gymnasium verworpen moet worden. Zwakkere lezers hebben in dat geval meer moeite hebben met de woordmoeilijkheid van een tekst dan sterke lezers. Dit blijkt ook uit de gemiddeldes. Maar zoals ik al zei, omdat de significantie twijfelachtig is, mag ik dit eigenlijk niet zo concluderen en rapporteren.

Hypothese 4: 'Het verschil tussen de scores op de moeilijke en makkelijke versie zal bij de Vmbo/Havo leerlingen groter zijn dan bij de Gymnasiumleerlingen' kan wel geheel bevestigd worden.

3.3 Schooltype-effect nader bekeken

Uit de vorige analyses blijkt dat schooltype een significante invloed heeft op de resultaten van alle testen. In dit deel ga ik kijken of ik het schooltype-effect op tekstbegrip nader kan verklaren aan de hand van het niveau van de leesvaardigheid en woordenschat. Leesvaardigheid, woordenschat en schooltype hangen namelijk samen: leesvaardigere lezers hebben vaak ook een hogere woordenschat en zitten op het Gymnasium. Hoe hoger hun leesvaardigheid en woordenschat is, hoe beter ze zullen scoren op de cloze-testen. Als de ene cloze-versie echter leesvaardigere lezers heeft dan de andere versie zal deze daarom beter gemaakt worden. In dit deel ga ik de leesvaardigheid en woordenschat

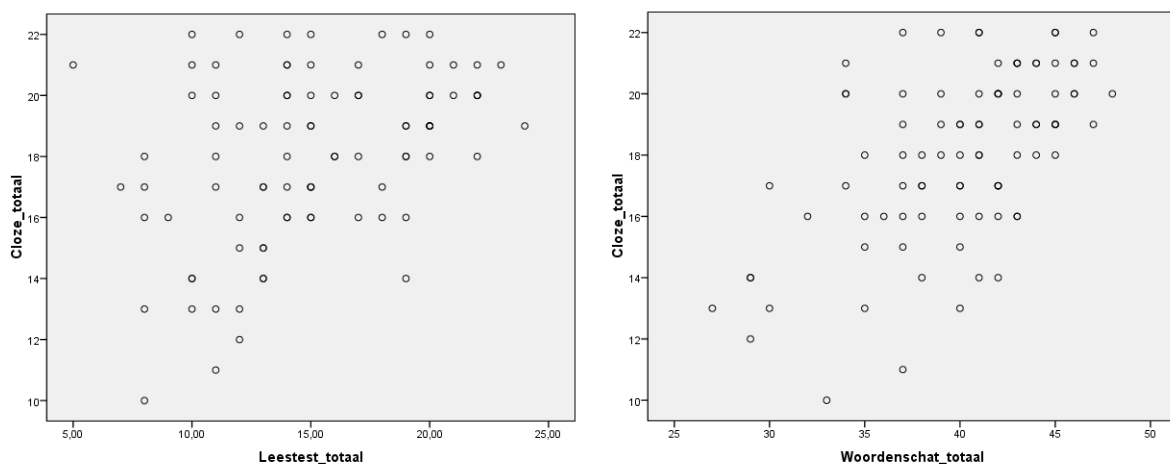


gelijk trekken voor alle leerlingen waardoor ik vervolgens objectief kan bekijken of de makkelijke versie beter is gemaakt dan de moeilijke versie.

3.3.1 Correlaties

Uit een correlatietest blijkt eveneens dat zowel de leesvaardigheidtest, woordenschattoets en cloze-test significant positief met elkaar correleren. De leesvaardigheidtoets en cloze-toets correleren een beetje ($r=0.39$; $p<0.001$), de woordenschattoets en cloze-test correleren een stuk meer ($r=0.59$; $p<0.001$). De leestest en woordenschattoets correleren redelijk ($r=0.48$; $p<0.001$). Uit figuur 2 blijkt de woordenschattoets, zoals de correlaties ook al aangeven, beter correleert met de cloze-test dan de leesvaardigheidtoets, maar ook de leesvaardigheidtoets geeft een redelijk lineair verband aan.

Figuur 2: Correlatie tussen leestest en cloze-test en woordenschattoets en cloze-test



3.3.2 Co-varianten

Ondanks dat de leesvaardigheidtoets minder correleert dan de woordenschattoets neem ik beide uiteindelijk mee als co-variant omdat deze wel degelijk invloed hebben op de scores van de cloze-test, oftewel op het tekstbegrip. Hoe hoger bijvoorbeeld iemands woordenschat is, hoe hoger de scores op cloze-test uit zullen vallen. Door woordenschat in te voeren als co-variant neem ik het effect van de woordenschat weg uit de scores op de cloze-test. Bij de leesvaardigheid gebeurt uiteraard hetzelfde.

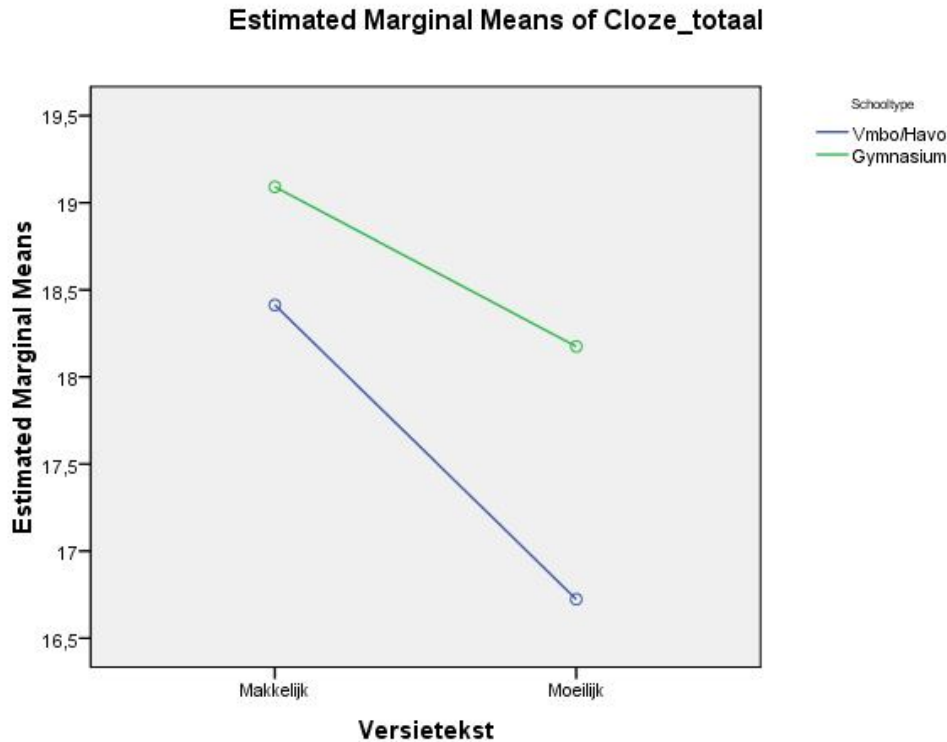
Als eerste stap heb ik alleen het effect van leesvaardigheid op tekstbegrip gecorrigeerd door leesvaardigheid als co-variant in te voeren bij een meervoudige ANOVA. Hieruit blijkt dat leesvaardigheid een significante invloed heeft op tekstbegrip ($F(1)=6,67$; $p=0.011$). Ook schooltype ($F(1)=12,50$; $p=0.001$) en tekstversie ($F(1)=8,69$; $p=0.004$) blijven een significante invloed hebben op tekstbegrip. Als alleen woordenschat als co-variant meegenomen wordt, komt naar voren dat de woordenschat een nog grotere significante invloed heeft op tekstbegrip dan de leesvaardigheid ($F(1)=20,73$; $p<0.001$). Schooltype ($F(1)=5,48$; $p=0.022$) en tekstversie ($F(1)=0,89$; $p=0.017$) blijven eveneens significante invloed uitoefenen.

De leesvaardigheid en woordenschat beïnvloeden natuurlijk samen het tekstbegrip dus moeten tegelijkertijd als co-variant meegenomen worden. Het tekstbegrip per niveau en tekstversie, waarvoor



leesvaardigheid en woordenschat zijn gecorrigeerd, zijn visueel weergegeven in figuur 3. In bijlage 7 staan de precieze gemiddeldes.

Figuur 3: Tekstbegrip per niveau en tekstversie, gecorrigeerd voor leesvaardigheid en woordenschat



Als woordenschat en leesvaardigheid tegelijkertijd gecorrigeerd worden blijkt dat de leesvaardigheid geen significant effect meer heeft op tekstbegrip ($F(1)=1.78$; $p=0.186$), maar dat dit effect weggenomen wordt door de woordenschat ($F(1)=14.83$; $p<0.001$). Ook het significante effect van schooltype valt weg ($F(1)=3.58$; $p=0.062$). Dit betekent dat de variantie die eerst aan het schoolniveau werd toegewezen, nu wordt toegewezen aan het hoge of lagere niveau van leesvaardigheid en woordenschat van de leerling, waarbij woordenschat een sterkere invloed heeft op de cloze-resultaten dan de leesvaardigheid. Wat tevens een grote significante invloed heeft op tekstbegrip is de woordmoeilijkheid van de tekst, oftewel de tekstversie ($F(1)=7,51$; $p=0.007$). Uit het figuur en de gemiddeldes komt naar voren dat de moeilijke versie (significant) een stuk slechter gemaakt is dan de makkelijke versie, ook nu het niveau van woordenschat en leesvaardigheid van alle leerlingen gelijk is getrokken.

Hypothese 1: 'De makkelijke tekstversie wordt beter begrepen dan de moeilijke tekstversie' kan hierdoor bevestigd worden.

Al lijkt in figuur een interactie-effect te bestaan tussen schoolniveau en tekstversie, dit effect is niet significant ($F(1)=0.62$; $p=0.432$).



3.4 Cloze-items nader bekeken

Tot slot wil ik de cloze-items beter bekijken. Aan welke andere factoren kan het liggen dat de items goed of juist heel slecht zijn gemaakt? Als er een significant verschil in antwoord bestaat op een moeilijk en makkelijk cloze-item, ligt dat dan wel echt aan de moeilijkheid? Om een inzicht te krijgen in waar de verschillen precies aan kunnen liggen tussen beide versies, is een T-toets uitgevoerd waarin de antwoorden op beide versies van de verschillende cloze-items zijn vergeleken. In tabel 7 staan de drie moeilijkste en makkelijkste items en de drie items waarbij de verschillen tussen de versies significant zijn. De gehele tabel staat in bijlage 7.

Tabel 7: Resultaten per cloze-item per tekstversie

		Versietekst	Aantal	Gem.	Std. Deviatie
Moeilijk in te vullen items	Cloze_item3	Makkelijk	41	0,46	0,51
		Moeilijk	49	0,39	0,49
	Cloze_item4	Makkelijk	41	0,61	0,49
		Moeilijk	49	0,51	0,51
	Cloze_item16	Makkelijk	41	0,61	0,49
		Moeilijk	49	0,45	0,50
Makkelijk in te vullen items	Cloze_item10	Makkelijk	41	0,98	0,16
		Moeilijk	49	1,00	0,00
	Cloze_item15	Makkelijk	41	1,00	0,00
		Moeilijk	49	0,98	0,14
	Cloze_item17	Makkelijk	41	1,00	0,00
		Moeilijk	49	0,98	0,14
Significante verschillen tussen tekstversies	Cloze_item8	Makkelijk	41	0,80	0,40
		Moeilijk	49	0,53	0,50
	Cloze_item11	Makkelijk	41	1,00	0,00
		Moeilijk	49	0,84	0,37
	Cloze_item21	Makkelijk	41	0,95	0,22
		Moeilijk	49	0,80	0,41

3.4.1 Moeilijke cloze-items

Item 3, 4 en 16 blijken in beide versies zeer moeilijk te zijn geweest.

Item 3 werd bij de helft van de gatenteksten open gelaten.

Veel Duitsers waren nog boos over het feit dat(2)..... de Eerste Wereldoorlog had verloren, maar ook over de manier waarop de(3)..... (Frankrijk en Engeland) omgingen met Duitsland. (winnaars)

Veel Duitsers waren nog furieus over de nederlaag die(2)..... in de Eerste Wereldoorlog had geleden, maar ook over de manier waarop de(3)..... (Frankrijk en Engeland) Duitsland behandelen. (winnaars)

Leerlingen die nummer 3 foutief ingevuld hadden dachten vooral aan de woorden 'landen', 'mensen', 'koninkrijken'. Dit betekent dat ze waarschijnlijk vooral lokaal hebben gelezen en geen verband



hebben gelegd met de zin ervoor: in de zin ervoor stond dat Duitsland de Eerste Wereldoorlog verloren had. Iets in het antwoord op item 3 zou erop moeten duiden dat Frankrijk en Engeland de oorlog hadden gewonnen of in ieder geval tegenstanders van Duitsland waren, ondanks dat dit niet letterlijk in de tekst stond. Woorden als 'Tegenstanders' en 'vijanden' zijn dan ook goed gerekend. Een groot verschil tussen de moeilijke versie en makkelijke versie is er echter niet: maar 7% van de leerlingen heeft de makkelijke versie beter gemaakt.

Item 4 is door beide versies ook zeer slecht gemaakt.

Dit was één van de redenen waarom Hitlers partij zo snel(4)..... (groeide)

Dit was één van de motieven waarom Hitlers partij zo snel(4)..... (groeide)

Om item 4 in te kunnen vullen is zinsverband nodig of eventueel voorkennis. Item 4 kan niet door lokaal tekstbegrip ingevuld worden, maar kan alleen voortkomen uit de informatie uit de vorige twee zinnen. Het beoogde woord is 'groeide'. Foutief ingevulde antwoorden waren veelal het tegenovergestelde: 'verloor', 'afhaakte', 'opgaf' en 'aanviel', of het item werd blanco gelaten. Blijkbaar zijn de zinnen ervoor niet goed begrepen in zowel de makkelijke als de moeilijke versie of hebben ze het verband niet gelegd en alleen naar de zin op zichzelf gekeken. Hierbij geldt weer dat dit woord geïnterpreteerd moet worden uit de vorige zinnen en niet al ergens letterlijk is genoemd. De Vmbo/Havo-leeringen hebben dit item in beide versies slecht gemaakt, maar de Gymnasiumleeringen hebben de makkelijke versie (gem 0.81) significant beter gemaakt dan de moeilijke versie (gem 0.52). Dit kan betekenen dat woordmoeilijkheid van de zinnen ervoor wel degelijk invloed heeft gehad.

Het derde item dat het slecht gemaakt is, is item 16.

Er werden zomaar mensen(16)..... in gevangenissen en concentratiekampen, vooral veel Joden.
(opgesloten)

Er werden klakkeloos mensen(16)..... in gevangenissen en concentratiekampen, vooral veel Joden.
(opgesloten)

Hiervoor geldt hetzelfde als bij de vorige 2 slecht gemaakte items: waarschijnlijk is alleen lokaal naar de zin gekeken. Het woord dat ingevuld had moeten worden is 'opgesloten'. Bijna de helft van de leerlingen heeft hier 'gedood' of een ander synoniem voor vermoorden ingevuld. Dit is fout aangezien twee zinnen verderop staat dat de gevangenen/Joden vermoord worden. Ze hebben dus niet eerst verder gelezen of het naderhand veranderd toen ze verder lazen. 16% van de leerlingen heeft de moeilijke versie van dit item slechter gemaakt. Dit zou kunnen betekenen dat ze het woord 'geëxecuteerd' niet hebben begrepen en deze vraag slechter hebben gemaakt vanwege de woordmoeilijkheid. Dit verschil is echter niet significant.

Hieruit kan geconcludeerd worden dat leerlingen moeite hebben woorden in te vullen die niet letterlijk in de tekst of in dezelfde zin terug te vinden zijn. Die ze dus moeten interpreteren. Het lijkt hierbij geen verschil te maken of de benodigde informatie voor de zin of na de zin staat. Dit duidt erop dat ze lokaal lezen om de woorden in te vullen of moeite hebben om verbanden te leggen. Woordmoeilijkheid lijkt daar bij deze items geen grote invloed op te hebben.



3.4.2 Makkelijke cloze-items

Item 10 is door praktisch iedereen goed ingevuld.

De(10)..... bombardeerden het centrum van Rotterdam. (Duitsers)

De(10)..... bombardeerden het centrum van Rotterdam. (Duitsers)

Dit zou kunnen komen doordat de Duitsers in de zin ervoor al genoemd worden. Dit woord staat zelfs achteraan in de zin waardoor het nog in hun hoofd hangt als ze aan de zin met cloze-item 10 beginnen. Er is hier dus wel sprake van het leggen van zinsverband, maar het woord dat ingevuld moet worden ('Duitsers') is al letterlijk genoemd in de zin ervoor.

Ook item 15 kent bijna geen foutief ingevulde antwoorden. Dit valt misschien meer te wijten aan het gekozen gat:

Nederlandse mannen moesten in Duitse fabrieken gaan(15)..... . (werken)

Nederlandse mannen werden verplicht in Duitse fabrieken te(15)..... . (werken)

Met een beetje algemene kennis kan bedacht worden wat mannen in fabrieken doen, zonder dat het ergens in de tekst voorkomt. Doordat dit gat lokaal ingevuld kan worden, wordt zins- of tekstverband met deze vraag niet getest.

Item 17 is eigenlijk ook een weggevertje:

De meeste(17)..... werden gedood. (Joden)

De meeste(17)..... werden geëxecuteerd. (Joden)

Deze zin is zonder context niet voor 100% zekerheid in te vullen, maar in de zinnen erboven wordt beschreven dat mensen, vooral Joden, in gevangenen worden opgesloten. Een link naar 'gevangenen' of 'Joden' is dus niet moeilijk te maken. Net als bij item 10 is het antwoord letterlijk twee zinnen terug te vinden.

Uit deze gegevens valt niet heel veel te concluderen. Het enige wat uit deze minimale voorbeelden blijkt is dat leerlingen minder moeite hebben een gat in te vullen als het antwoord letterlijk in de tekst genoemd worden. Dan is het minder moeilijk zinsverbanden te leggen en het gat in te vullen.

3.4.3 Verschil in resultaten van beide versies

Het grootste verschil tussen de antwoorden in de moeilijke versie en de makkelijke versie valt te vinden in item 8:

Bij het aanvallen van(8)..... werden ook Nederland en België ingenomen. (Frankrijk)

In het offensief tegen(8)..... werden ook Nederland en België bezet. (Frankrijk)

80% van de leerlingen heeft een goed antwoord ingevuld in de makkelijke versie, tegenover 53% in de moeilijke versie. Foutief ingevulde antwoorden waren vooral 'Duitsland' of 'Duitsers'. Daarnaast kwamen 'Joden', 'Engeland' en 'Europa' voorbij of werd het item leeg gelaten. De aanwijzing voor 'Frankrijk' staat in de zin ervoor:

Daarna wilde hij Duitslands grote(7)..... uitschakelen: Frankrijk. (vijand)

Daarna wilde hij Duitslands grote(7)..... elimineren: Frankrijk. (vijand)

Dit gat vereist tekstbegrip en zinsintegratie. Ook hiervoor geldt, de aanwijzing staat letterlijk in de zin ervoor genoemd. De makkelijke versie wist Frankrijk er als juiste antwoord dan ook zo uit te pikken.



Schijnbaar waren de woorden 'eliminieren' en 'offensief' bij de moeilijke versie te moeilijk om de context nog te begrijpen.

Ondanks dat item 11 in beide versies goed is gemaakt, ken deze een significant verschil tussen de antwoorden.

Ze wilden ook nog andere steden gaan(11)..... . (bombarderen)

Ze dreigden om ook andere nederzettingen te(11)..... . (bombarderen)

Om dit gat in te vullen moet weer naar de zin ervoor gekeken worden, waarin letterlijk staat dat de [gat= Duitsers] het centrum van Rotterdam bombardeerden. In de makkelijke versie wist iedereen dit eruit te halen. 'Plaatsen' of 'plegen' waren antwoorden in de moeilijke versie. Misschien kennen ze het woord 'nederzetting' niet als 'stad', maar kennen ze alleen uit de context van geschiedenisboeken dat er nederzettingen gesticht werden door bijvoorbeeld de Romeinen. De structuuraanduiding 'ook' geeft echter aan dat er toch steden vernield zullen worden. De zin erna ziet er zo uit:

Toen besloot de leiding van het(12)..... leger zich over te geven. (Nederlandse)

Toen besloot de leiding van het(12)..... leger te capituleren. (Nederlandse)

Ook hier had uit gehaald kunnen worden dat het leger niet blij was met de actie die Duitsland uit zou halen, het bombarderen dus. Als leerlingen het woord capituleren niet kennen, in combinatie met nederzettingen, blijkt dat woordmoeilijkheid wel degelijk een rol heeft gespeeld.

Item 21 is ook door beide tekstversies goed gemaakt, maar verschilt significant:

Meer dan 20.000 mensen stierven van(21)..... . (honger)

Meer dan 20.000 mensen bezweken aan(21)..... . (honger)

Honger is een woord dat uit de context gehaald moet worden, maar waarvoor geen grote tekstinterpretatie nodig is. De twee zinnen ervoor luiden:

Zij maakten een verschrikkelijke 'Hongerwinter' mee. Er was bijna geen(20)..... meer. Mensen aten bijvoorbeeld tulpenbollen om in leven te blijven. (eten)

Zij hadden te lijden onder een erbarmelijke 'Hongerwinter'. Er was grote schaarste aan(20)..... .

Mensen aten bijvoorbeeld tulpenbollen om te overleven. (eten)

Uit deze zinnen blijkt dat mensen honger hadden. In de moeilijke versie werden ook woorden als 'ziektes', 'tulpenbollen' en 'voedselvergiftiging' ingevuld. Dit zijn in principe helemaal geen gekke gedachtegangen na het eten van tulpenbollen. Hier kan het ook zo zijn dat leerlingen 'bezweken aan' alleen in de context van ziektes kennen en daarom de associatie met ziektes leggen na het eten van tulpenbollen. Op deze manier speelt de woordmoeilijkheid wel een rol.

De antwoorden op deze cloze-items zijn in te vullen door zinsverband te leggen. De in te vullen woorden zijn letterlijk of bijna letterlijk terug te vinden in enkele zinnen voor de zin waarin het cloze-item voorkomt. Op die manier zijn de cloze-items vergelijkbaar met de items die door beide versies goed gemaakt zijn (paragraaf 3.4.2). In de makkelijke versie zijn de items ook zeer goed gemaakt. In de moeilijke versie significant minder. Het enige waarop de zinnen verschillen is woordmoeilijkheid. Dit speelt dus zeker een rol in het tekstbegrip van de leerlingen. Een reden zou kunnen zijn dat leerlingen een woord niet, of maar in één bepaalde context kennen en het daarom moeilijk vinden dit woord in een andere context te gebruiken.



4. Conclusie

De betrouwbaarheid van de cloze-test is aangetoond met een α van 0.675. Alle items correleren met elkaar. Dit betekent dat de gevonden effecten in meerdere items verschijnen. Uit een item-analyse blijkt eveneens dat cloze-items de invloed van woordmoeilijkheid op tekstbegrip kunnen meten.

Uit de resultaten komt naar voren dat de woordmoeilijkheid van teksten wel degelijk invloed heeft op tekstbegrip. Het tekstbegrip bij de tekst met makkelijke woorden is significant hoger dan bij de tekst met moeilijke woorden. Als de leesvaardigheid en woordenschat als co-variant meegenomen worden, valt het schooltype-effect weg. De verschillen tussen schoolniveaus kunnen dus volledig worden toegeschreven aan de verschillen in leesvaardigheid en woordenschat. Door deze twee factoren te corrigeren, is er geen sprake van een interactie-effect tussen tekstversie en schooltype. Dit betekent dat de moeilijkheid niet meer of minder effect heeft gehad bij de Vmbo/Havo- en de Gymnasiumleerlingen, het schooltype-effect viel namelijk weg. De woordenschat van de leerlingen blijkt een grotere invloed te hebben op tekstbegrip dan de leesvaardigheid. Omdat het in dit onderzoek gaat over de moeilijkheidsgraad van woorden in een tekst, is het logisch dat leerlingen met een grotere woordenschat de tekst met moeilijkere woorden beter begrijpen. Hypothese 1 'De makkelijke tekstversie wordt beter begrepen dan de moeilijke tekstversie' kan bevestigd worden.

Om te bekijken of woordmoeilijkheid ook invloed heeft op tekstbegrip bij sterke en zwakke lezers, kan er niet alleen gekeken worden naar de resultaten waarbij de leesvaardigheid en woordenschat zijn gecorrigeerd. Hierdoor wordt namelijk het schooltype-effect, en dus het verschil tussen goede en zwakke lezers, weg genomen. Dat Gymnasiumleerlingen betere lezers zijn dan de Vmbo/Havo-leerlingen blijkt uit het feit dat zij een significant grotere woordenschat en hogere leesvaardigheid hebben dan de Vmbo/Havo-leerlingen. Gymnasiumleerlingen scoren daarom waarschijnlijk significant hoger op zowel de moeilijke als de makkelijke tekstversie van de begripstoets. Hypothese 2 'Gymnasiumleerlingen scoren gemiddeld beter op beide tekstversies dan Vmbo/Havo-leerlingen' kan ook bevestigd worden.

Tot slot blijken de resultaten op beide tekstversies ook zonder co-varianten significant te verschillen. De moeilijke versie is significant slechter gemaakt ten opzichte van de makkelijke versie. Hierbij is echter een duidelijk verschil op te merken tussen de beide schooltypes. De Vmbo/Havo-leerlingen hebben de moeilijke versie significant slechter gemaakt dan de makkelijke versie. Het verschil is maar liefst bijna 2 punten. Bij de Gymnasiumleerlingen is de moeilijke versie wel ietsje slechter gemaakt dan de makkelijke versie (nog geen 0.5 punt), maar vanwege tegenstrijdige uitkomsten op een T-toets en ANOVA is het niet duidelijk of dit verschil significant is. Het lijkt erop dat de woordmoeilijkheid van een tekst een grotere invloed heeft op tekstbegrip bij zwakkere lezers dan bij sterke lezers, maar dit mag niet zo als feit gerapporteerd worden. Hypothese 3 'Zowel de Gymnasiumleerlingen als de Vmbo/Havo-leerlingen scoren gemiddeld lager op de moeilijke versie' kan daardoor deels of geheel bevestigd worden en hypothese 4 'Het verschil tussen de scores op de moeilijke en makkelijke versie zal bij de Vmbo/Havo leerlingen groter zijn dan bij de Gymnasiumleerlingen' kan in zijn geheel bevestigd worden. Woordmoeilijkheid heeft dus wel degelijk invloed op tekstbegrip, alleen lijkt deze invloed groter te zijn bij zwakke lezers dan bij sterke lezers.



5. Discussie

De meeste hypothesen zijn uitgekomen. Het lijkt erop dat woordmoeilijkheid daadwerkelijk invloed heeft op tekstbegrip van leerlingen. Om de conclusies te kunnen generaliseren is dit onderzoek echter te kleinschalig. Met 90 proefpersonen zou dit onderzoek meer als een pilotstudy gezien kunnen worden dan als een echt onderzoek. Daarnaast is het onderzoek afgenomen op leerlingen uit tweede klassen. Voor leerlingen uit vierde klassen kunnen de resultaten heel anders zijn, aangezien zij over het algemeen een grotere woordenschat en leesvaardigheid hebben. Ook het schoolniveau is een twistbaar punt. Hebben VWO-leerlingen hetzelfde leesniveau als Gymnasiumleerlingen? En gelden de resultaten van Vmbo/Havo-leerlingen ook voor de leerlingen van het Vmbo-kadergericht? Waarschijnlijk niet. Misschien zijn de conclusies wel generaliseerbaar als niet naar het schoolniveau gekeken wordt, maar alleen naar de resultaten van de leesvaardigheid en de woordenschat. Dus als alleen gekeken wordt naar zwakke lezers of sterke lezers, ongeacht in welke klas of welk schoolniveau ze zitten. Dan geldt: zwakkere lezers hebben meer moeite met woordmoeilijkheid dan sterke lezers. Alleen is het moeilijk om op deze basis een praktisch advies uit te brengen aan schrijvers van schoolboekteksten. Je kunt namelijk moeilijk in één klas verschillende schoolboeken uitdelen die voor zwakke of sterke lezers geschreven zijn.

Wat met het uitvoeren van een grotere studie naar dit onderwerp anders gedaan moet worden is het verdelen van de tekstversies over de klassen. Omdat in één klas leerlingen zitten met gradaties in niveau, is één tekstversie van de cloze-toets in één hele klas uitgedeeld. De verschillen tussen leerlingen zouden ondervangen worden door het klassenaspect. Achteraf was het toch beter geweest de versies random uit te delen in klassen. De moeilijke versie is nu bijvoorbeeld door 49 leerlingen gemaakt, de makkelijke versie maar door 41. De ene klas bevatte meer jongens dan de andere klas, waardoor de makkelijke versie door 25 jongens is gemaakt en door 14 meisjes, terwijl de verdeling van het geslacht over de moeilijke versie wel gelijk is verdeeld. Voor de resultaten op de cloze-test heeft dit geen effect gehad, maar wel voor de resultaten op de leesvaardigheidstoets. Uit de analyses is gebleken dat jongens meer moeite hebben met leesvaardigheid dan meisjes. Echter, leesvaardigheid was de laatste toets die ingevuld moest worden. Eén gymnasiumklas, waarin 3 meisjes zaten en de rest jongens, heeft de testen de dag voor hun herfstvakantie ingevuld. Het kan zijn dat het merendeel geen zin meer had in de laatste toets omdat het bijna herfstvakantie was en de leesvaardigheidstoets daarom slechter hebben gemaakt. Omdat er meer jongens in deze klas zaten en meer jongens de test waarschijnlijk slechter hebben gemaakt, kan het zijn dat het verschil daardoor aan het geslacht is toegeschreven. Bij de woordenschattoets en cloze-toets is immers geen geslachtsverschil waarneembaar.

Een ander vreemd punt wat uit de resultaten naar voren komt, is het resultaat op de cloze-test wanneer niet is gecorrigeerd voor leesvaardigheid en woordenschat. Uit een T-test blijkt dat de scores van de Vmbo/Havo-leerlingen op de makkelijke en moeilijke versie significant verschillen. De T-test wijst eveneens uit dat de Gymnasiumleerlingen de makkelijke versie niet significant beter maken dan de moeilijke versie. Ook in de gemiddeldes is nauwelijks een groot verschil te zien (19.76 op makkelijk en 19.22 op moeilijk). Toch wijst de ANOVA uit dat er geen interactieverschil bestaat tussen schooltype en tekstversie, dus dat de woordmoeilijkheid van de tekst niet meer of minder effect heeft



gehad bij de schooltypes. De T-test wijst dit toch wel degelijk uit. Het is moeilijk te bepalen welke resultaten, of de ANOVA of de T-test, geïnterpreteerd moeten worden om hypothese 3 geheel aan te nemen of gedeeltelijk te verwerpen.

Ook over de manipulatie van de cloze-test valt iets op te merken. De tekst van de cloze-test is gemanipuleerd op woordmoeilijkheid. Nog steeds is woordmoeilijkheid een lastige factor om te bepalen. Wat zijn immers moeilijke woorden voor leerlingen van 13 jaar? Daar bestaan geen bepaalde richtlijnen voor. Is de makkelijke versie niet te eenvoudig gemaakt of de moeilijke versie te moeilijk, of juist nog steeds te makkelijk voor een moeilijke versie?

Tot slot, als ik advies uit zou moeten brengen aan schoolboekschrijvers op basis van de bevindingen van dit onderzoek, zou ik zeggen dat zwakkere lezers baat hebben bij makkelijker taalgebruik en dat sterke lezers moeilijker taalgebruik goed aan kunnen. In dit geval zouden de schoolboeken voor de Vmbo/Havo-leerlingen makkelijker woorden moeten bevatten dan de schoolboeken voor de Gymnasiumleerlingen.

Dit onderzoek kan als op zichzelf staand beschouwd worden, maar kan beter gezien worden in een groter geheel met andere onderzoeken naar tekstbegrip. Woordmoeilijkheid blijkt wel degelijk een bepalende factor te zijn voor tekstbegrip, alleen spelen ook andere factoren als zinslengte, hoeveelheid verbindingswoorden of structuuraanduidingen, de plek van de clozes, de frequentie waarin het weggelaten woord voorkomt en de hoeveelheid context die nodig is om het gat in te vullen, een rol. Om een goed advies uit te brengen om schoolboekteksten op maat te schrijven, kan dit onderzoek dus het beste gecombineerd worden met andere onderzoek die zich onder andere richten op bovengenoemde factoren.

Als ik dit onderzoek bijvoorbeeld combineer met 'Vocabulary and readability: how knowing word meanings affects comprehension' van Stahl (2003) kan ik al een beter en concreter advies geven. Uit mijn onderzoek blijkt dat iemands woordenschat sterk gerelateerd is aan tekstbegrip. Dit komt ook naar voren in het onderzoek van Stahl. Wat verder uit Stahl's onderzoek komt is dat kennis van woordbetekenissen tekstbegrip kan vergroten, maar de vocabulaire die de lezer kent moet hij/zij wel tot op zekere hoogte weten om de tekst ook echt te begrijpen. Hiermee bedoelt Stahl dat woorden meerdere betekenissen hebben in verschillende contexten. Pas als je deze woorden ook in verschillende contexten kan gebruiken, zul je de tekst begrijpen. Hij stelt dat hoe meer woorden kinderen leren, hoe beter ze leren op een wereldwijze of hoogontwikkelde manier over de wereld na te gaan denken. Juist dit leidt tot een groter begrip. Volgens Stahl kan nieuwe woorden leren dan ook het beste gedaan worden door ze zowel de definitie van het woord te leren en daarbij, zeker zo belangrijk, dit woord in verschillende contexten te laten zien. Hierdoor leren kinderen niet alleen de woordbetekenis, maar kunnen ze het woord ook in een grotere context plaatsen. Daardoor leren ze het kennisdomein waar het woord aan gerelateerd wordt kennen. Op deze manier leert het kind meerdere categorieën kennen, waardoor ook kennis van de wereld uitgebreid wordt.

Voor de gymnasiumleerlingen begrijpen de tekst al erg goed als er moeilijke woorden in voor komen. De teksten voor sterke lezers kunnen dus moeilijke woorden bevatten zonder dat het tekstbegrip er sterk op achteruit gaat. Uit paragraaf 3.4.3 in dit onderzoek komt wel naar voren dat leerlingen sommige moeilijke woorden inderdaad waarschijnlijk maar in één bepaalde context kennen,



en daarom de betekenis van het woord waarbij deze in een andere context in de tekst staat niet geheel snappen. Om deze leerlingen een grotere woordenschat aan te reiken, zouden teksten geschreven moeten worden waarin wel degelijk moeilijke woorden voorkomen, maar daarbij ook in de context duidelijk moet worden wat deze woorden betekenen. Misschien kennen ze al één definitie van het woord en leren ze er zo andere definities bij, of kunnen ze uit de context raden wat het moeilijke woord betekent. Op deze manier vergroten de Gymnasiumleerlingen, of sterke lezers, al lezende hun woordenschat, wat in het vervolg weer zorgt voor een groter tekstbegrip. Voor de Vmbo/Havo-leerlingen, of zwakkere lezers in het algemeen, zou ik het advies uitbrengen voorzichtiger om te gaan met het gebruik van moeilijke woorden. Hun tekstbegrip gaat er daardoor sterk op achteruit. Natuurlijk geldt ook hier dat zij kunnen leren van contextbeschrijvingen, maar dit zou misschien toegepast moeten worden op makkelijkere woorden. Op deze manier kunnen ook zij hun woordenschat vergroten. Omdat grotere woordenschat leidt tot beter tekstbegrip zullen ook zij daardoor uiteindelijk teksten beter begrijpen.

Door op deze manier onderzoeksresultaten over verschillende tekstkenmerken te combineren kan uiteindelijk een concreet en veelomvattend advies uitgebracht worden aan schoolboektekstschrijvers.



Literatuurlijst

- Abraham, R.G., Chapelle, C.A. (1992). "The meaning of cloze test scores: an item difficulty perspective." *Modern Language Journal* 76 (4): 468-479.
- Alderson, J.C. (1979b). "The effect on the cloze test of changes in deletion frequency." *Journal of Research in Reading* 2: 108-119.
- Campbell, A.C. (1961). "Some Determinants of the Difficulty of Non-Verbal Classification Items." *Educational and Psychological Measurement* 21: 899-913.
- Carroll, J. (1989). "Intellectual Abilities and Aptitudes." *Foundations for a Psychology of Education*. Ed. Alan Lesgold & Robert Glaser. Hillsdale, NJ: Erlbaum. 137-197
- Chall, J. (1958). *Readability: An appraisal of research and application*. Columbus, OH: Ohio State University.
- Chall, J.S. (1995). *Readability revisited: The new Dale-Chall readability formula*. Cambridge, MA: Brookline Books.
- Endsley, M., Georgia, M. (2000). *Situation Models: An Avenue To The Modeling Of Mental Models*. Proceedings 14th Triennial Congress Of The International Ergonomics Association And The 44th Annual Meeting Of The Human Factors And Ergonomics Society.
- Freebody, R., Anderson, R.C. (1983a). "Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension." *Journal of Reading Behavior* 15: 19-40.
- Hacquebord, T.R., Linthorst, B., Stellingwerf (2004). *Vocabulairverwerving bij taalzwakke leerlingen in het voortgezet onderwijs*. [PDF] ORDPaper.
- Harrison, C. (1980). *Readability in the classroom*. Cambridge: Cambridge University Press.
- Jansen, C., Woudstra, E. (1979). "Theorie en praktijk van het Nederlandse leesbaarheidsonderzoek. Een analyse van twee formules." *Tijdschrift voor Taalbeheersing* 1: 43-60.
- Kamalski, J. (2007). *Coherence marking, comprehension and persuasion. On the processing and representation of discourse*. Dissertatie. Utrecht: Universiteit Utrecht.
- Kintsch, W., van Dijk, T. A. (1978). "Toward a model of text comprehension and production." *Psychological Review* 85: 363-394.



Klare, G.R. (1984). "Readability", in P.D. Pearson (red.), *Handbook of Reading Research*, New York et Londres: Longman. 681-744.

Kobayashi, M. (2002b). "Cloze tests revisited: Exploring item characteristics with special attention to scoring methods." *Modern Language Journal* 86: 571-586.

Omanson, R.C. (1985) "Knowing words and understanding texts." In T.H. Carr (ed) *The Development of Reading Skills* (New Directions for Child Development) 27: 35-53. San Francisco, CA: Jossey-Bass.

Sachs, J., Tung, P., Lam, R. (1997). "How to Construct a Cloze Test: Lessons from Testing Measurement Theory Models." *Perspectives* (City University of Hong Kong) 9: 145-160

Stahl, S.A., Jacobson, M.G., Davis, C.E., Davis, R.L (1989). "Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text." *Reading research Quarterly* 24 (1): 27-43.

Staphorsius, G. (1994). *Leesbaarheid en leesvaardigheid. De ontwikkeling van een domeingericht meetinstrument*. Arnhem: Cito

Staphorsius, M., Sanders, T. (2008). "Leesbaarheid en tekststructuur in basisschoolteksten. Een exploratief onderzoek." *Tijdschrift voor Taalbeheersing* 30 (2): 174-197.

Taylor, W.L. (1953). "Cloze procedure: A new tool for measuring readability." *Journalism Quarterly* 30: 415-433.

Zakaluk, B. L., & Samuels, S.J. (1988). *Readability: Its past, present, & future*. Newark, DE: International Reading Association.



Bijlagen

Bijlage 1: Originele tekst Tweede Wereldoorlog

Bijlage 2: Regels voor cloze-ontwerp

Bijlage 3: Gemanipuleerde teksten inclusief cloze-items

Bijlage 4: Moeilijke en makkelijke tekstversie inclusief gaten

Bijlage 5: Verdeling teksten en versies per schooltype

Bijlage 6: Goed gerekende niet-letterlijke antwoorden bij cloze-test

Bijlage 7: SPSS uitdraai



Bijlage 1: Originale tekst 'Tweede Wereldoorlog'

De Tweede Wereldoorlog (298 woorden)

In 1933 kwam in Duitsland Hitler aan de macht. Hij was de leider van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij. Een partij die Joden haatte en iedereen vertelde dat ook te doen.

Veel Duitsers waren nog kwaad over het feit dat Duitsland de Eerste Wereldoorlog had verloren. Maar ook over de manier waarop de winnaars (Frankrijk en Engeland) Duitsland behandelden. Dit was één van de redenen waarom Hitlers partij zo snel groeide.

Hitler wilde van Duitsland het machtigste land van Europa maken. Eerst viel hij met zijn leger de buurlanden van Duitsland aan. Daarna wilde hij Duitslands grote vijand uitschakelen: Frankrijk. In de aanval op Frankrijk werden ook Nederland en België bezet.

Het Nederlandse leger was veel te zwak om de Duitse aanval tegen te houden. De Duitsers bombardeerden het centrum van Rotterdam. Ze dreigden om ook andere steden te bombarderen. Toen besloot de leiding van het Nederlandse leger zich over te geven. De regering en de koningin waren toen al gevlucht naar Engeland.

Eerst leek de bezetting mee te vallen. Maar al snel werd duidelijk wat het betekende om niet vrij te zijn. Nederlandse mannen moesten in Duitse fabrieken gaan werken. Er werden zomaar mensen opgesloten in gevangnissen en concentratiekampen, vooral veel Joden. Ze werden vervoerd in goederentreinen. De meeste Joden werden vermoord.

In het najaar van 1944 werd het zuiden van Nederland bevrijd door het Engelse, Amerikaanse en Canadese leger. Het gebied boven de grote rivieren (vooral de grote steden in het westen) was nog niet bevrijd. Zij hadden te lijden onder een verschrikkelijke 'Hongerwinter'. Er was bijna geen eten meer. Mensen aten bijvoorbeeld tulpenbollen om in leven te blijven. Meer dan 20.000 mensen stierven van honger.

Op 5 mei 1945 gaf het Duitse leger zich over en was heel Nederland bevrijd.



Bijlage 2: Regels voor Cloze-ontwerp

Regels voor cloze-ontwerp

Henk Pander Maat & Rogier Kraf

Inleiding

De mechanische cloze (ieder n-de woord weglaten) is eenvoudig te maken, maar heeft tot nadeel dat hij een aantal gaten oplevert die vooral grammaticale en lexicale kennis aftappen. Een zogenaamde 'rationele' cloze-test worden gaten zo gekozen dat het vullen ervan een beroep doet op de context, en als het even kan niet alleen op de context binnen de zin, maar ook die buiten de zin (Abraham & Chapelle 1992). Het blijkt dat zo'n rationele cloze zich voorspelbaarder gedraagt dan een mechanische cloze.

Het probleem is natuurlijk dat er een procedure nodig is om een rationele cloze te ontwikkelen. Hoe komen we tot een cloze met gaten die tekst begrip vergen, maar dan op een intersubjectief betrouwbare manier, in die zin dat twee constructeurs met dezelfde tekst tot dezelfde cloze-items komen?

Enkele algemene regels

- a *Woorden die veranderd zijn als onderdeel van de tekstmanipulatie zijn natuurlijk geen geldige opties voor cloze gaten!*
- b *Op een cloze gat mag slechts 1 woord ingevuld worden!*
- c *Geen gaten maken om de lezer erin te laten komen:*
 - Eerste zin intact laten
 - Naam of vakterm die de eerste keer valt (bv. 'palts') niet weglaten
 - Getallen niet weglaten
- d *Nooit een gat van maken, omdat er weinig tekstbegrip voor nodig is:*
 - Functiewoorden (lidwoorden, voorzetsels, voegwoorden, voornaamwoorden)

Het argument tegen veel functie woorden geldt eigenlijk ook voor:

- Hulpwerkwoorden (eigenlijk functiewoorden, ook gesloten klassen)
 - Koppelwerkwoorden (ook in de praktijk een heel beperkte set)
- e *Nooit doen wegens lokale lexicale voorspelbaarheid:*
 - Delen van Phrasal verbs (staat X toe, staat X bij, enz.); het ene deel voorspelt het ander sterk



- Delen van vaste uitdrukkingen (houdt zich bezig met, raad en daad)

f Niet doen wegens bijkomstigheid van de informatie

- Bijvoeglijke naamwoorden
- Bijwoorden

De intentie hier is dat zelfstandige naamwoorden en werkwoorden in het algemeen vaker centrale informatie bevatten dan bijwoorden.

g Herhaling van woorden op cloze gaten

- Een cloze gat mag in principe meerdere meerdere malen op hetzelfde woord vallen, mits dit overeenkomt met de frequentie van dit woord in de tekst. Een voorbeeld: in een tekst waarin 40 cloze gaten geplaatst moeten worden, komen 80 naamwoorden voor die allen kandidaat cloze gat zijn. Het woord "trein" komt 8 keer voor. Dat is 10% van 80, wat betekent dat het toegestaan is 4 van de cloze gaten op het woord "trein" te laten vallen.

h Spreiding van gaten

- Je kunt mechanisch spreiden, dus een bepaalde afstand tussen de gaten in woorden aanhouden, maar het klinkt redelijk om gaten in principe linguïstisch te spreiden, dus bv. een gat per clause. Een clause is een stap in het verhaal of betoog van de tekst. Wij zullen dan ook proberen het aantal gaten per clause constant te houden

i Wat te doen wanneer er meer clauses zijn dan gaten?

Welke clauses kunnen afvallen?

- Clauses die geen kandidaat-gat hebben
- Non-finiëte clauses zijn van minder belang ('om zijn rijk te kunnen besturen') dan finiete
- Er is een korte matrixzin ('verondersteld wordt ...') die op zich incompleet is (subjectloos).
- Daarna zouden we naar de kortste en minst informatierijke bijzinnen kunnen kijken.

Methode bij het maken van een cloze, fase 1

- We willen uiteindelijk een gat op 12 woorden, dus in deze tekst 21 gaten.
- Streef eerst naar een gat in iedere deelzin (clause: zinsdeel met werkwoord) of er manipulaties mogelijk zijn. Later kun je nog selecteren.
- In aanmerking komen alleen zelfstandige naamwoorden of naamwoordelijk delen van het gezegde
- Als er meerdere ZN in aanmerking komen, kies je op basis van een syntactische ranglijst: subject, object, indirect object, bepalingen

Extra regels:

- blijft een nevenschikking over, dan kiezen we daarvan het eerste lid

Fase 2



Als fase 1 meer gaten oplevert dan we nodig hebben:

- schrappen een paar gaten uit samengestelde zinnen, beginnend met de kortste deelzinnen

Makkelijke versie	Moeilijke versie	Clauses	Gaten	Commentaar
<i>De eerste spoorlijn (269 woorden)</i>	<i>De eerste spoorlijn (264 woorden)</i>			
Op 20 september 1839 werd de eerste spoorlijn in Nederland met een groots feest geopend.	Op 20 september 1839 werd de eerste spoorlijn in Nederland <u>in een grootse plechtigheid</u> geopend.			
De stoomlocomotief 'De Arend' deed 25 minuten over het stuk van Amsterdam naar Haarlem.	De stoomlocomotief 'De Arend' <u>legde</u> in 25 minuten <u>het traject af</u> van Amsterdam naar Haarlem.	1	1	Enige optie
Veel reizigers waren niet erg tevreden: het ging veel te hard, en met veel te veel lawaai.	Veel reizigers waren niet erg enthousiast: het ging veel te hard, en met veel te veel <u>geruis</u> .	2	1	Enige optie in eerste deelzinnen; tweede deelzin heeft geen geschikt zelfstandig naamwoord of predicatsnomen.
En was de veiligheid wel verzekerd?	En was de veiligheid wel <u>gewaarborgd</u> ?	1	1	Enige optie
Eerder dat jaar was bij Gent nog de stoomketel van een vertrekkende trein ontploft.	Eerder dat jaar was bij Gent nog de stoomketel van een vertrekkende trein <u>geëxplodeerd</u> .	1	1	Enige optie; <i>stoomketel</i> is vakterm
Ondanks deze twijfels in het begin begon er met de eerste trein een tijdperk van enorme verandering.	Ondanks <u>die aanvankelijke scepsis</u> <u>luidde</u> de eerste trein een tijdperk <u>in</u> van enorme verandering.	1	1	Opties <i>trein</i> en <i>tijdperk</i> . Subject gekozen.
De lijn Amsterdam -Haarlem werd al snel uitgebreid met de lijn tussen Amsterdam naar Rotterdam.	<u>Aan</u> het <u>tracee</u> Amsterdam -Haarlem werd al snel de lijn tussen Amsterdam naar Rotterdam <u>toegevoegd</u> .	1	1	Herhaalde naam, mag.
Meer lijnen volgden, allemaal beheerd door verschillende spoorwegmaatschappijen.	Meer lijnen volgden, allemaal <u>geëxploiteerd</u> door verschillende spoorwegmaatschappijen.	1	1	<i>Lijnen</i> en <i>spoorweg-maatschappijen</i> zijn opties. <i>Lijnen</i> is subject.
Rond 1900 was de trein het belangrijkste vervoermiddel in Nederland.	Rond 1900 was de trein het voornaamste <u>transportmiddel</u> in Nederland.	1	1	Enige optie



Men beseft tegenwoordig niet meer welke verandering het spoor heeft veroorzaakt in de Nederlandse samenleving .	Men <u>onderkent</u> tegenwoordig niet meer welke <u>transformatie</u> het spoor heeft <u>bewerkstelligd</u> in de Nederlandse samenleving .	2	1	Opties zijn <i>spoor</i> en <i>samenleving</i> . <i>Spoor</i> is co-referentieel met het vorige gat. Niet doen.
Voor de komst van het spoor duurde reizen bijzonder lang; voor de meeste mensen was het duur; en soms was het zelfs een gevaarlijke onderneming .	Voor de komst van het spoor <u>was</u> reizen bijzonder <u>tijdrovend</u> ; voor de meeste mensen was het <u>kostbaar</u> ; en soms was het zelfs een <u>hachelijke</u> onderneming .	3	3	Opties zijn <i>spoor</i> , <i>reizen</i> , <i>mensen</i> en <i>onderneming</i> . In de eerste deelzin kiezen we het onderwerp; in de andere deelzinnen hebben we geen keus. Tweede deelzin is kort, kan later waarschijnlijk weg.
De betere verbindingen en het grotere reisgemak hebben meegeholpen aan de eenwording van Nederland: mensen uit verschillende streken kregen meer contact en de Nederlandse staat kon zijn gebied beter organiseren.	De <u>verbeterde</u> verbindingen en het <u>toegenomen</u> reisgemak hebben <u>bijgedragen tot</u> de eenwording van Nederland: mensen uit verschillende <u>regio's</u> kregen meer contact en de Nederlandse staat kon zijn <u>territorium</u> beter organiseren.	3	2	Opties in de eerste deelzin zijn <i>verbindingen</i> , <i>reisgemak</i> en <i>eenwording</i> . Dat laatste is wel een erg moeilijk woord. Dan maar het eerste lid van de nevenschikking. In de tweede deelzin nemen we het subject <i>mensen</i> . In de derde deelzin is er geen keuze. Naderhand is de tweede deelzin gesneuveld wegens kortheid.
Het spoorwegennet was hard nodig voor de industrialisatie van Nederland, die pas na 1870 goed begon.	Het spoorwegennet was <u>een</u> <u>voorwaarde</u> voor de industrialisatie van Nederland, die pas na 1870 goed <u>intrad</u> .	2	1	Opties zijn <i>spoorwegennet</i> en <i>industrialisatie</i> . Het subject van de zin is gekozen. Check of dit kan qua tekstfrequentie van de items <i>spoor</i> , <i>trein</i> , <i>spoornet</i> enz.
Grondstoffen , producten en ook arbeiders moesten immers vervoerd worden.	Grondstoffen , producten en ook <u>werklieden</u> moesten immers <u>getransporteerd</u> worden.	1	1	Opties zijn <i>grondstoffen</i> en <i>producten</i> . Eerste lid is gekozen.
En de industrialisatie hielp op haar beurt weer bij de uitbouw van het spoorwegennet.	En de industrialisatie <u>droeg</u> op haar beurt weer <u>bij tot</u> de <u>expansie</u> van het spoorwegennet.	1	1	Opties zijn <i>industrialisatie</i> en <i>spoorwegennet</i> . Het subject is gekozen.
Alle spoorlijnen kwamen in 1938 terecht bij de NV Nederlandse Spoorwegen (NS), een bedrijf dat	Alle spoorlijnen <u>werden</u> in 1938 <u>ondergebracht</u> bij de NV Nederlandse Spoorwegen (NS),	2	2	In de eerste deelzin is er geen keus. In de tweede deelzin is het onderwerp gekozen.



tot 1995 in handen bleef van de staat.	een bedrijf dat tot 1995 <u>eigendom</u> bleef van de overheid.			
Nu is de NS een gewoon bedrijf; maar dan wel een bedrijf dat heel belangrijk is in onze samenleving.	Nu is de NS een <u>particulier</u> bedrijf; maar dan wel een bedrijf dat heel belangrijk is in onze samenleving.	2	1	Onderwerp gekozen, in de tweede deelzin predicaatsnomen. Naderhand is eerste korte deelzin gesneuveld.
De politiek en allerlei belangenorganisaties kijken voortdurend of het treinvervoer goed blijft.	De politiek en allerlei belangenorganisaties <u>volgen</u> <u>continu</u> de <u>kwaliteit</u> van het treinvervoer.	1	1	Eerste lid van het subject gekozen.
		26	21	



Bijlage 3: Gemanipuleerde teksten inclusief uitleg gekozen cloze-items

Makkelijke versie	Moeilijke versie	Clauses	Gaten	Commentaar
<i>De Tweede Wereldoorlog (295 woorden)</i>	<i>De Tweede Wereldoorlog (291 woorden)</i>			Dik gedrukte woorden zijn cloze-gaten. <u>Onderstreepte woorden</u> zijn manipulaties.
In 1933 kwam in Duitsland Hitler aan de macht.	In 1933 <u>kreeg Hitler in</u> Duitsland <u>heerschappij</u> .			
Hij was de <u>baas</u> van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij.	Hij was de <u>voorman</u> van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij.			
Een partij die Joden haatte en iedereen vertelde dat ook te doen.	Een partij die Joden <u>verafschuwde</u> en iedereen vertelde <u>daarin mee te gaan</u> .	2	1	Opties: 'partij' en 'vertelde'. Hier gekozen voor subject omdat bij 'vertelde' veel werkwoorden ingevuld kunnen worden zoals beveelde, aanraadde etc. Moeilijk een grens te trekken tussen wat goed en fout is.
Veel Duitsers waren nog <u>boos</u> over het feit dat Duitsland de Eerste Wereldoorlog had verloren,	Veel Duitsers waren nog <u>furieus</u> over <u>de nederlaag die Duitsland in</u> de Eerste Wereldoorlog <u>had geleden</u> ,	1	1	Opties: 'Duitsers', 'Duitsland' en 'Eerste Wereldoorlog'. Eerste Wereldoorlog is een nieuw begrip, dus kan niet. 'Duitsland' test meer tekstbegrip dan 'Duitsers' omdat 'Duitsland' weer terugkomt in de volgende zin.
maar ook over de manier waarop de winnaars (Frankrijk en Engeland) <u>omgingen met</u> Duitsland.	maar ook over de manier waarop de winnaars (Frankrijk en Engeland) Duitsland behandelen.	1	1	Opties: 'Manier', 'Winnaars', 'Frankrijk', 'Engeland' en 'Duitsland'. Engeland en Frankrijk zijn uitgesloten omdat dit nieuwe begrippen zijn. 'De manier waarop' is een vaste uitdrukking, dus makkelijk te raden. 'Duitsland' is hiervoor al geweest, dus heb daarom voor het subject gekozen.
Dit was één van de redenen waarom Hitler's partij zo snel groeide .	Dit was één van de <u>motieven</u> waarom Hitlers partij zo snel groeide .	1	1	Opties: 'Partij' en 'groeide'. Partij is al eerder aan bod geweest, vandaar 'groeide'.
Hitler wilde van Duitsland het sterkste land van Europa maken.	Hitler wilde van Duitsland het <u>machtigste</u> land van Europa maken.	1	1	Opties: 'Hitler', 'land' en 'Europa'. Europa kan niet omdat het een nieuw begrip is. Gekozen voor subject 'Hitler' omdat dit tekstbegrip meet en 'land' makkelijker te raden valt.
Eerst viel hij met zijn leger de <u>landen naast Duitsland</u> aan.	Eerst viel hij met zijn <u>krijgsmacht</u> de buurlanden van Duitsland aan.	1	1	Opties: 'viel' en 'Duitsland'. 'Viel' is Phrasal verb, dus blijft 'Duitsland' over.
Daarna wilde hij Duitslands grote vijand uitschakelen: Frankrijk.	Daarna wilde hij Duitslands grote vijand <u>eliminieren</u> : Frankrijk.	1	1	Opties: 'Duitslands', 'partij' en 'Frankrijk'. 'Duitsland' is al een aantal keer aan bod geweest en 'Frankrijk' is een nieuw begrip. Omdat in de volgende zin over



				offensief/aanval wordt gesproken test 'vijand' het best tekstbegrip.
Bij het aanvallen van Frankrijk werden ook Nederland en België <u>ingenomen</u> .	In het offensief tegen Frankrijk werden ook Nederland en België bezet.	1	1	Opties: 'Frankrijk', 'Nederland' en 'Belgie'. 'Nederland' en 'Belgie' zouden geen tekstbegrip maar kennis testen, dus blijft 'Frankrijk' over.
Het Nederlandse leger was veel te zwak om de Duitse aanval te <u>stoppen</u> .	Het Nederlandse leger was veel te zwak om de Duitse aanval te <u>pareren</u> .	1	1	Opties: 'Leger', 'zwak', 'aanval'. Gekozen voor subject.
De Duitsers bombardeerden het centrum van Rotterdam.	De Duitsers bombardeerden het centrum van Rotterdam.	1	1	Opties: 'Duitsers', 'bombardeerden', 'centrum' en 'Rotterdam'. 'Centrum' en 'Rotterdam' zijn kennisvragen omdat het nieuwe begrippen zijn. 'Bombardeerden' is de enige cloze-optie in de volgende zin, dus kan hier geen gat vormen. 'Duitsers' blijft over.
Ze <u>wilden</u> ook <u>nog</u> andere steden <u>gaan bombarderen</u> .	Ze dreigden om ook andere <u>nederzettingen</u> te bombarderen .	1	1	Enige optie.
Toen besloot de leiding van het Nederlandse leger zich over te geven.	Toen besloot de leiding van het Nederlandse leger te <u>capituleren</u> .	1	1	Opties: 'Besloot', 'leiding', 'Nederlandse' en 'leger'. Ik denk dat 'Nederlandse' hier het beste tekstbegrip meet omdat het ervoor over Nederland gaat. 'Leger' is al eerder een cloze geweest en de andere twee opties zijn moeilijk te raden woorden als het een gat is.
De regering en de koningin waren toen al gevlucht naar Engeland.	De regering en de <u>vorstin</u> waren toen al gevlucht naar Engeland.	1	1	Opties: 'regering', 'gevlucht' en 'Engeland'. 'Regering' en 'Engeland' zijn kennisvragen, nieuwe begrippen. 'Gevlucht' blijft over.
Eerst leek de <u>overheersing</u> mee te vallen, maar al snel werd duidelijk wat het <u>was</u> om niet vrij te zijn.	Eerst leek de bezetting <u>gunstig</u> te <u>verlopen</u> , maar al snel werd duidelijk wat het betekende om <u>onderdrukt</u> te zijn.	2	2	Enige opties.
Nederlandse mannen moesten in Duitse fabrieken gaan werken .	Nederlandse mannen <u>werden verplicht</u> in Duitse fabrieken te <u>werken</u> .	1	1	Opties: 'Nederlandse', 'mannen', 'Duitse', 'fabrieken' en 'werken'. 'Nederlandse' en 'Duitse' zijn bijvoeglijke naamwoorden. 'Fabrieken' test kennis en geen tekstbegrip. Bij 'mannen' kan veel passends ingevuld worden (bijvoorbeeld soldaten, mensen, gevangenen) zonder tekstbegrip te meten. 'Werken' blijft over.
Er werden zomaar mensen opgesloten in gevangnissen en concentratiekampen, vooral veel Joden.	Er werden <u>klakkeloos</u> mensen opgesloten in gevangnissen en concentratiekampen, vooral veel Joden.	1	1	Opties: 'Mensen', 'opgesloten', 'gevangnissen', 'concentratiekampen' en 'Joden'. 'Gevangnissen', 'concentratiekampen' en 'Joden' zijn nieuwe begrippen die niet uit de tekst zijn te halen. Bij 'mensen' kan weer veel passends ingevuld worden. 'Opgesloten' blijft over.



Ze werden afgevoerd in goederentreinen. De meeste Joden werden <u>gedood</u> .	Ze werden <u>gedeporteed</u> in goederentreinen. De meeste Joden werden <u>geëxecuteerd</u> .	2	1	Enige optie.
In het najaar van 1944 werd het zuiden van Nederland bevrijd door het Engelse, Amerikaanse en Canadese leger.	In het najaar van 1944 werd het zuiden van Nederland bevrijd door <u>de</u> Engelse, Amerikaanse en Canadese <u>krijgsmacht</u> .	1	1	Opties: 'Najaar', 'zuiden', 'bevrijd', 'Nederland', 'Engelse', 'Amerikaanse' en 'Canadese'. 'Engelse', 'Amerikaanse' en 'Canadese' zijn bijvoeglijke naamwoorden, dus vallen af. 'Najaar' is een bijwoord, dus valt ook af. Zuiden is een kennisvraag in plaats van tekstbegrip. 'Bevrijd' is een cloze in de volgende zin, dus 'Nederland' blijft over.
Het gebied boven de grote rivieren (vooral de grote steden in het westen) was nog niet bevrijd .	Het <u>domein</u> boven de grote rivieren (<u>overwegend</u> de grote steden in het westen) was nog niet bevrijd .	1	1	Opties: 'Rivieren', 'steden', 'westen' en 'bevrijd'. 'Rivieren', 'steden' en 'westen' is nieuw, valt niet uit de tekst te halen, dus blijft 'bevrijd' over.
Zij <u>maakten</u> een verschrikkelijke 'Hongerwinter' <u>mee</u> .	Zij hadden te lijden onder een <u>erbarmelijke</u> 'Hongerwinter'.	1	0	Hongerwinter is geen optie omdat dit een nieuw begrip is, dus geen mogelijkheden voor cloze.
Er was bijna geen eten meer.	Er was <u>grote schaarste aan eten</u> .	1	1	Enige optie.
Mensen aten bijvoorbeeld tulpenbollen om in leven te blijven.	Mensen aten bijvoorbeeld tulpenbollen om <u>te overleven</u> .	1	0	Opties: 'mensen', 'tulpenbollen'. Tulpenbollen is een nieuw begrip en mensen kan niet in verband met spreiding omdat dan 2 gaten achter elkaar vallen.
Meer dan 20.000 mensen stierven van honger .	Meer dan 20.000 mensen <u>bezweken aan</u> honger .	1	1	Opties: 'Mensen' en 'honger'. Honger test beter tekstbegrip omdat het daarvoor over eten gaat. 'Mensen' is makkelijker te raden.
Op 5 mei 1945 gaf het Duitse leger zich over en was heel Nederland <u>vrij</u> .	Op 5 mei 1945 <u>capituleerde</u> het Duitse leger en was heel Nederland bevrijd.	2	1	Opties: 'Leger' en 'Nederland'. Met 'Nederland' als cloze test je precies hetzelfde als 5 zinnen terug. Heeft geen zin dus. Vandaar 'leger'.



Bijlage 4: Moeilijke en makkelijke tekstversie inclusief gaten

De Tweede Wereldoorlog (makkelijke versie)

In 1933 kwam in Duitsland Hitler aan de macht. Hij was de baas van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij. Een(1)..... die Joden haatte en iedereen vertelde dat ook te doen.

Veel Duitsers waren nog boos over het feit dat(2)..... de Eerste Wereldoorlog had verloren, maar ook over de manier waarop de(3)..... (Frankrijk en Engeland) omgingen met Duitsland. Dit was één van de redenen waarom Hitlers partij zo snel(4)..... .

.....(5)..... wilde van Duitsland het sterkste land van Europa maken. Eerst viel hij met zijn leger de landen naast(6)..... aan. Daarna wilde hij Duitslands grote(7)..... uitschakelen: Frankrijk. Bij het aanvallen van(8)..... werden ook Nederland en België ingenomen.

Het Nederlandse(9)..... was veel te zwak om de Duitse aanval te stoppen. De(10)..... bombardeerden het centrum van Rotterdam. Ze wilden ook nog andere steden gaan(11)..... . Toen besloot de leiding van het(12)..... leger zich over te geven. De regering en de koningin waren toen al(13)..... naar Engeland.

Eerst leek de overheersing mee te vallen, maar al snel werd(14)..... wat het was om niet vrij te zijn. Nederlandse mannen moesten in Duitse fabrieken gaan(15)..... . Er werden zomaar mensen(16)..... in gevangenissen en concentratiekampen, vooral veel Joden. Ze werden afgevoerd in goederentreinen. De meeste(17)..... werden gedood.

In het najaar van 1944 werd het zuiden van(18)..... bevrijd door het Engelse, Amerikaanse en Canadese leger. Het gebied boven de grote rivieren (vooral de grote steden in het westen) was nog niet(19)..... . Zij maakten een verschrikkelijke 'Hongerwinter' mee. Er was bijna geen(20)..... meer. Mensen aten bijvoorbeeld tulpenbollen om in leven te blijven. Meer dan 20.000 mensen stierven van(21)..... . Op 5 mei 1945 gaf het Duitse(22)..... zich over en was heel Nederland vrij.

De Tweede Wereldoorlog (moeilijke versie)

In 1933 kreeg Hitler in Duitsland heerschappij. Hij was de voorman van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij. Een(1)..... die Joden verafschuwde en iedereen vertelde daarin mee te gaan.

Veel Duitsers waren nog furieus over de nederlaag die(2)..... in de Eerste Wereldoorlog had geleden, maar ook over de manier waarop de(3)..... (Frankrijk en Engeland) Duitsland behandelen. Dit was één van de motieven waarom Hitlers partij zo snel(4)..... .

.....(5)..... wilde van Duitsland het machtigste land van Europa maken. Eerst viel hij met zijn krijgsmacht de buurlanden van(6)..... aan. Daarna wilde hij Duitslands grote(7)..... elimineren: Frankrijk. In het offensief tegen(8)..... werden ook Nederland en België bezet. Het Nederlandse(9)..... was veel te zwak om de Duitse aanval te pareren. De(10)..... bombardeerden het centrum van Rotterdam. Ze dreigden om ook andere nederzettingen te(11)..... . Toen besloot de leiding van het(12)..... leger te capituleren. De regering en de vorstin waren toen al(13)..... naar Engeland.

Eerst leek de bezetting gunstig te verlopen, maar al snel werd(14)..... wat het betekende om onderdrukt te zijn. Nederlandse mannen werden verplicht in Duitse fabrieken te(15)..... . Er werden klakkeloos mensen(16)..... in gevangenissen en concentratiekampen, vooral veel Joden. Ze werden gedeporteerd in goederentreinen. De meeste(17)..... werden geëxecuteerd.

In het najaar van 1944 werd het zuiden van(18)..... bevrijd door de Engelse, Amerikaanse en Canadese krijgsmacht. Het domein boven de grote rivieren (overwegend de grote steden in het westen) was nog niet(19)..... . Zij hadden te lijden onder een erbarmelijke 'Hongerwinter'. Er was grote schaarste aan(20)..... . Mensen aten bijvoorbeeld tulpenbollen om te overleven. Meer dan 20.000 mensen bezweken aan(21)..... . Op 5 mei 1945 capituleerde het Duitse(22)..... en was heel Nederland bevrijd.



Bijlage 5: Verdeling teksten en versies per schooltype

Verdeling teksten en versies over Gymnasium en Vmbo/Havo niveaus

Gymnasium		Vmbo/Havo	
Klas 1	Klas 2	Klas 1	Klas 2
Tekst 1, versie 2	Tekst 1, versie 1	Tekst 1, versie 2	Tekst 1, versie 1
Tekst 2, versie 1	Tekst 2, versie 2	Tekst 2, versie 1	Tekst 2, versie 2

Tekst 1: Tweede wereldoorlog, tekst 2: Eerste spoorlijn. Versie 1: makkelijk, versie 2: moeilijk

Aantal ingevulde teksten + tekstversies per niveau

		Tekst 1, versie 1	Tekst 1, versie 2	Tekst 2, versie 1	Tekst 2, versie 2	Totaal leerlingen
Gymnasium	Klas 1	0	23	23	0	23
	Klas 2	21	0	0	21	21
Totaal Gym						44
Vmbo/Havo	Klas 1	20	0	0	20	20
	Klas 2	0	26	26		26
Totaal V/H						46
Totaal Gym + V/H		41	49	41	49	90

Tekst 1: Tweede wereldoorlog, tekst 2: Eerste spoorlijn. Versie 1: makkelijk, versie 2: moeilijk



Bijlage 6: Goed gerekende niet-letterlijke antwoorden bij cloze-test

Makkelijke versie	Moeilijke versie		Goede gerekende antwoorden op de clozes
<i>De Tweede Wereldoorlog (295 woorden)</i>	<i>De Tweede Wereldoorlog (291 woorden)</i>		
In 1933 kwam in Duitsland Hitler aan de macht.	In 1933 <u>kreeg Hitler in</u> Duitsland <u>heerschappij</u> .		
Hij was de <u>baas</u> van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij.	Hij was de <u>voorman</u> van de Duitse partij NSDAP, de Nationaal Socialistische Duitse Arbeiders Partij.		
Een partij die Joden haatte en iedereen vertelde dat ook te doen.	Een partij die Joden <u>verafschuwde</u> en iedereen vertelde <u>daarin mee te gaan</u> .		Man, nazi, racist, dictator, groep, beweging
Veel Duitsers waren nog <u>boos</u> over het feit dat Duitsland de Eerste Wereldoorlog had verloren,	Veel Duitsers waren nog <u>furieus</u> over <u>de nederlaag die Duitsland in</u> de Eerste Wereldoorlog <u>had geleden</u> ,		Ze, zij
maar ook over de manier waarop de winnaars (Frankrijk en Engeland) <u>omgingen met</u> Duitsland.	maar ook over de manier waarop de winnaars (Frankrijk en Engeland) Duitsland behandelen.		Fransen en Engelsen, tegenstanders, geallieerden, buurlanden, vijanden, overwinnaars
Dit was één van de redenen waarom Hitler's partij zo snel groeide .	Dit was één van de <u>motieven</u> waarom Hitlers partij zo snel groeide .		Uitbreidde, aan macht kwam, steeg, verspreidde
Hitler wilde van Duitsland het sterkste land van Europa maken.	Hitler wilde van Duitsland het <u>machtigste</u> land van Europa maken.		hij
Eerst viel hij met zijn leger de <u>landen naast Duitsland</u> aan.	Eerst viel hij met zijn <u>krijgsmacht</u> de buurlanden van Duitsland aan.		
Daarna wilde hij Duitslands grote vijand uitschakelen: Frankrijk.	Daarna wilde hij Duitslands grote vijand <u>eliminieren</u> : Frankrijk.		Rivaal, concurrent, tegenstander
<u>Bij het aanvallen van</u> Frankrijk werden ook Nederland en België <u>ingenomen</u> .	In <u>het offensief tegen</u> Frankrijk werden ook Nederland en België bezet.		
Het Nederlandse leger was veel te zwak om de Duitse aanval te <u>stoppen</u> .	Het Nederlandse leger was veel te zwak om de Duitse aanval <u>te pareren</u> .		
De Duitsers bombardeerden het centrum van Rotterdam.	De Duitsers bombardeerden het centrum van Rotterdam.		Gevechtsvliegtuigen, nazi's, Duitse luchtmacht, Duitse vliegtuigen, Moffen
Ze <u>wilden</u> ook <u>nog</u> andere steden <u>gaan bombarderen</u> .	Ze dreigden om ook andere <u>nederzettingen</u> te bombarderen .		Elimineren, verwoesten, vernielen, aanvallen, veroveren
Toen besloot de leiding van het Nederlandse	Toen besloot de leiding van het Nederlandse		



leger zich over te geven.	leger te <u>capituleren</u> .		
De regering en de koningin waren toen al gevlucht naar Engeland.	De regering en de <u>vorstin</u> waren toen al gevlucht naar Engeland.		Vertrokken, opweg, lang, snel, weg
Eerst leek de <u>overheersing</u> mee te vallen, maar al snel werd duidelijk wat het <u>was</u> om niet vrij te zijn.	Eerst leek de bezetting <u>gunstig te verlopen</u> , maar al snel werd duidelijk wat het betekende om <u>onderdrukt</u> te zijn.		Geopenbaard, gezien, bekend, bedacht, gemerkt, gevoeld
Nederlandse mannen moesten in Duitse fabrieken gaan werken .	Nederlandse mannen <u>werden verplicht</u> in Duitse fabrieken <u>te werken</u> .		
Er werden zomaar mensen opgesloten in gevangenissen en concentratiekampen, vooral veel Joden.	Er werden <u>klakkeloos</u> mensen opgesloten in gevangenissen en concentratiekampen, vooral veel Joden.		gevangen, vastgezet, gegooid, gedumpte, gestopt, geplaatst, gevangen genomen, vastgehouden
Ze werden afgevoerd in goederentreinen. De meeste Joden werden <u>gedood</u> .	Ze werden <u>gedeporteerd</u> in goederentreinen. De meeste Joden werden <u>geëxecuteerd</u> .		Mensen, gevangenen, daarvan
In het najaar van 1944 werd het zuiden van Nederland bevrijd door het Engelse, Amerikaanse en Canadese leger.	In het najaar van 1944 werd het zuiden van Nederland bevrijd door <u>de</u> Engelse, Amerikaanse en Canadese <u>krijgsmacht</u> .		
Het gebied boven de grote rivieren (vooral de grote steden in het westen) was nog niet bevrijd .	Het <u>domein</u> boven de grote rivieren (<u>overwegend</u> de grote steden in het westen) was nog niet bevrijd .		vrij
Zij <u>maakten</u> een verschrikkelijke 'Hongerwinter' <u>mee</u> .	Zij hadden te lijden onder een <u>erbarmelijke</u> 'Hongerwinter'.		
Er was bijna geen eten meer.	Er was <u>grote schaarste aan eten</u> .		Voedsel, voedingsmiddelen, levensmiddelen
Mensen aten bijvoorbeeld tulpenbollen om in leven te blijven.	Mensen aten bijvoorbeeld tulpenbollen om <u>te overleven</u> .		
Meer dan 20.000 mensen stierven van honger .	Meer dan 20.000 mensen <u>bezweken aan</u> honger .		Hongersnood, ondervoeding, de Hongerwinter, voedselgebrek, verhongering
Op 5 mei 1945 gaf het Duitse leger zich over en was heel Nederland <u>vrij</u> .	Op 5 mei 1945 <u>capituleerde</u> het Duitse leger en was heel Nederland bevrijd.		



Bijlage 7: SPSS uitdraai

Betrouwbaarheidsanalyse voor de cloze-toets

Case Processing Summary

		N	%
Cases	Valid	90	100,0
	Excluded ^a	0	,0
	Total	90	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
,675	22

Item Statistics

	Mean	Std. Deviation	N
Cloze_item1	,92	,269	90
Cloze_item2	,77	,425	90
Cloze_item3	,42	,497	90
Cloze_item4	,56	,500	90
Cloze_item5	,97	,181	90
Cloze_item6	,84	,364	90
Cloze_item7	,76	,432	90
Cloze_item8	,66	,478	90
Cloze_item9	,90	,302	90
Cloze_item10	,99	,105	90
Cloze_item11	,91	,286	90
Cloze_item12	,77	,425	90
Cloze_item13	,92	,269	90
Cloze_item14	,69	,466	90
Cloze_item15	,99	,105	90



Cloze_item16	,52	,502	90
Cloze_item17	,99	,105	90
Cloze_item18	,84	,364	90
Cloze_item19	,91	,286	90
Cloze_item20	,93	,251	90
Cloze_item21	,87	,342	90
Cloze_item22	,90	,302	90

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Cronbach's Alpha if Item Deleted
Cloze_item1	17,10	7,507	,087	,676
Cloze_item2	17,26	6,956	,254	,663
Cloze_item3	17,60	6,715	,290	,659
Cloze_item4	17,47	6,948	,193	,673
Cloze_item5	17,06	7,626	,049	,677
Cloze_item6	17,18	7,204	,189	,669
Cloze_item7	17,27	6,378	,524	,629
Cloze_item8	17,37	6,415	,440	,638
Cloze_item9	17,12	7,255	,223	,666
Cloze_item10	17,03	7,561	,234	,670
Cloze_item11	17,11	7,111	,337	,657
Cloze_item12	17,26	7,249	,121	,679
Cloze_item13	17,10	7,417	,149	,672
Cloze_item14	17,33	6,607	,369	,648
Cloze_item15	17,03	7,651	,078	,675
Cloze_item16	17,50	6,612	,326	,654
Cloze_item17	17,03	7,606	,156	,673
Cloze_item18	17,18	7,092	,249	,663
Cloze_item19	17,11	7,426	,128	,673
Cloze_item20	17,09	7,475	,123	,673
Cloze_item21	17,16	6,829	,426	,646
Cloze_item22	17,12	7,255	,223	,666



Variantie-analyse van niveau en tekstversie op totale cloze-resultaten

Between-Subjects Factors

		Value Label	N
Schooltype	0	Vmbo/Havo	46
	1	Gymnasium	44
Versietekst	0	Makkelijk	41
	1	Moeilijk	49

Descriptive Statistics

Dependent Variable: Cloze_totaal

Schooltype	Versietekst	Mean	Std. Deviation	N
Vmbo/Havo	Makkelijk	17,65	3,133	20
	Moeilijk	15,92	2,682	26
	Total	16,67	2,982	46
Gymnasium	Makkelijk	19,67	1,770	21
	Moeilijk	19,22	1,506	23
	Total	19,43	1,634	44
Total	Makkelijk	18,68	2,697	41
	Moeilijk	17,47	2,747	49
	Total	18,02	2,776	90

Tests of Between-Subjects Effects

Dependent Variable: Cloze_totaal

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	206,980 ^a	3	68,993	12,388	,000	,302
Intercept	29238,528	1	29238,528	5249,771	,000	,984
Schooltype	157,088	1	157,088	28,205	,000	,247
Versietekst	26,375	1	26,375	4,736	,032	,052
Schooltype * Versietekst	9,091	1	9,091	1,632	,205	,019
Error	478,976	86	5,569			
Total	29918,000	90				
Corrected Total	685,956	89				

a. R Squared = ,302 (Adjusted R Squared = ,277)



Variantie-analyse van niveau en tekstversie op totale clozeresultaten met leesvaardigheid en woordenschat als covariaat

Between-Subjects Factors

		Value Label	N
Schooltype	0	Vmbo/Havo	46
	1	Gymnasium	44
Versietekst	0	Makkelijk	41
	1	Moeilijk	49

Tests of Between-Subjects Effects

Dependent Variable: Cloze_totaal

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	308,864 ^a	5	61,773	13,760	,000
Intercept	42,986	1	42,986	9,576	,003
Woordenschat_totaal	66,576	1	66,576	14,830	,000
Leestest_totaal	7,988	1	7,988	1,779	,186
Schooltype	16,055	1	16,055	3,576	,062
Versietekst	33,722	1	33,722	7,512	,007
Schooltype * Versietekst	2,796	1	2,796	,623	,432
Error	377,092	84	4,489		
Total	29918,000	90			
Corrected Total	685,956	89			

a. R Squared = ,450 (Adjusted R Squared = ,418)

Schooltype * Versietekst

Dependent Variable: Cloze_totaal

Schooltype	Versietekst	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Vmbo/Havo	Makkelijk	18,413 ^a	,500	17,418	19,408
	Moeilijk	16,723 ^a	,451	15,826	17,621
Gymnasium	Makkelijk	19,091 ^a	,498	18,102	20,081
	Moeilijk	18,175 ^a	,532	17,117	19,233

a. Covariates appearing in the model are evaluated at the following values:

Woordenschat_totaal = 39,99, Leestest_totaal = 15,0000.



T-toets van de scores op de cloze-items per tekstversie

Group Statistics

	Versietekst	N	Mean	Std. Deviation	Std. Error Mean
Cloze_item1	Makkelijk	41	,95	,218	,034
	Moeilijk	49	,90	,306	,044
Cloze_item2	Makkelijk	41	,85	,358	,056
	Moeilijk	49	,69	,466	,067
Cloze_item3	Makkelijk	41	,46	,505	,079
	Moeilijk	49	,39	,492	,070
Cloze_item4	Makkelijk	41	,61	,494	,077
	Moeilijk	49	,51	,505	,072
Cloze_item5	Makkelijk	41	,98	,156	,024
	Moeilijk	49	,96	,200	,029
Cloze_item6	Makkelijk	41	,85	,358	,056
	Moeilijk	49	,84	,373	,053
Cloze_item7	Makkelijk	41	,78	,419	,065
	Moeilijk	49	,73	,446	,064
Cloze_item8	Makkelijk	41	,80	,401	,063
	Moeilijk	49	,53	,504	,072
Cloze_item9	Makkelijk	41	,88	,331	,052
	Moeilijk	49	,92	,277	,040
Cloze_item10	Makkelijk	41	,98	,156	,024
	Moeilijk	49	1,00	,000	,000
Cloze_item11	Makkelijk	41	1,00	,000	,000
	Moeilijk	49	,84	,373	,053
Cloze_item12	Makkelijk	41	,80	,401	,063
	Moeilijk	49	,73	,446	,064
Cloze_item13	Makkelijk	41	,93	,264	,041
	Moeilijk	49	,92	,277	,040
Cloze_item14	Makkelijk	41	,61	,494	,077
	Moeilijk	49	,76	,434	,062
Cloze_item15	Makkelijk	41	1,00	,000	,000
	Moeilijk	49	,98	,143	,020



Cloze_item16	Makkelijk	41	,61	,494	,077
	Moeilijk	49	,45	,503	,072
Cloze_item17	Makkelijk	41	1,00	,000	,000
	Moeilijk	49	,98	,143	,020
Cloze_item18	Makkelijk	41	,88	,331	,052
	Moeilijk	49	,82	,391	,056
Cloze_item19	Makkelijk	41	,88	,331	,052
	Moeilijk	49	,94	,242	,035
Cloze_item20	Makkelijk	41	,98	,156	,024
	Moeilijk	49	,90	,306	,044
Cloze_item21	Makkelijk	41	,95	,218	,034
	Moeilijk	49	,80	,407	,058
Cloze_item22	Makkelijk	41	,90	,300	,047
	Moeilijk	49	,90	,306	,044