

Domme kenmerken... complexe teksteigenschappen?

Een onderzoek naar cohesie- en interactiekenmerken van schoolteksten met behulp van de leesbaarheidstool T-Scan

Eindwerkstuk 2012/2013

MA Communicatiestudies

Naam: N.W. Dekker

Studentnummer: 3417077

Scriptiebegeleider: dhr. H. Pander Maat

Tweede beoordelaar: mw. J. Evers-Vermeul

Hij is een eenvoudig en naïef schrijver. Ik geloof niet dat hij een regel geschreven heeft die niet door een twaalfjarige begrepen kan worden. Je hoeft, om hem te kunnen waarderen, geen enkel ander boek gelezen te hebben behalve wat toevallige jeugdlectuur (Van het Reve, 1985, p. 415).

Karel van het Reve over Lev Tolstoj (1828-1910)

Inhoud

1. Inleiding	4
2. Theoretisch kader	5
2.1 T-Scan	5
2.2 Samenhang: cohesie en coherentie	6
2.3 Tekstrepresentatie: de complexe eigenschappen	6
2.4. En de domme kenmerken?	9
2.5 Narrative vs. expository texts	10
3. Methode	12
3.1 Materialen: Het tekstcorpus	12
3.2 Niveauverschillen	13
3.3 Onderzoeksopzet	13
3.4 Relevante tekstkenmerken in T-Scan	16
4. Hypothesen	19
5. Resultaten: Kwantitatieve analyse	20
5.1 Argumentoverlap	20
5.2 Connectieven	21
5.3 Informatiedichtheid	22
5.4 Eigennamen	23
6. Resultaten: Kwalitatieve analyse	24
6.1 Argumentoverlap	24
6.2 Connectieven	28
6.2.1 Causale connectieven	28
6.2.2 Alle connectieven	31
6.3 Informatiedichtheid	33
6.3.1 Type-token-ratio en rarity index	33
6.3.2 NP-dichtheid	36
6.3.3 Inhoudswoorden	38
6.4 Eigennamen	42
6.5 Overzicht van de tekstkenmerken	44
7. Conclusie	45
8. Discussie	47
9. Aangehaalde literatuur	50
Index van bijlagen	52

1. Inleiding

Wat maakt een tekst geschikt voor een bepaalde lezer? En, misschien nog belangrijker, hoe kom je daar achter? Je kunt een tekst natuurlijk laten beoordelen door een 'expert', maar dat is niet zonder nadelen. Ten eerste kost dat veel tijd en ten tweede zijn experts het niet altijd met elkaar eens over *wat* nou precies een moeilijke tekst is. De letterkundige Karel van het Reve meent dat Tolstoj gemakkelijk te begrijpen is, omdat 'zijn regels door een twaalfjarige kunnen worden begrepen'. Of dat een bewijs is dat romans als *Oorlog en Vrede* en *Anna Karenina* inderdaad door brugklassers begrepen *en* gewaardeerd worden... Daar zal menig collega-expert anders over denken.

Eén manier om dit probleem te omzeilen is het beoordelen van de tekst op meetbare oppervlaktekenmerken, zoals zinslengte, woordlengte, woordmoeilijkheid e.d. Die methode wordt al jaren toegepast: zo hebben de classici Owen & Goodspeed in 1906 een *Homeric Vocabulary* uitgegeven, waarin zij van elk woord, en per woordsoort in kaart brachten hoe vaak die in Homerus' Ilias en Odyssee voorkomt. Dit boek is later gebruikt als argument dat Homerus de plaats zou moeten innemen van Xenophon, die traditioneel gebruikt werd als schoolauteur, hoewel hij op basis van leesbaarheidsformules moeilijker zou moeten zijn (Pharr, 1985, pp. 28-29). Hoe groot dat karwei is geweest laat zich raden. Gelukkig is het sinds de jaren '50 mogelijk dit soort analyses geautomatiseerd uit te voeren; hoewel dat veel tijd bespaart, is de validiteit van deze methode omstreden (Jansen & Lentz, 2008; Kraf & Pander Maat, 2009).

Toch wordt er nog volop gewerkt aan nieuwe leesbaarheidstools: één daarvan is T-Scan. T-Scan is in een Utrecht ontwikkelde leesbaarheidstool, die in staat is meer dan 150 tekstkenmerken in kaart te brengen (Kraf, 2013). In dit onderzoek probeer ik er achter te komen in hoeverre deze tekstkenmerken in staat zijn twee complexe teksteigenschappen aan te tonen: *element interaction* en *inference calls*. De eerste eigenschap duidt op de mate waarin de tekst als een samenhangend geheel kan worden begrepen. De eigenschap *Inference calls* draait om de vraag in hoeverre lezers, om de tekst te begrijpen, zelfstandig inferenties (verbanden) moeten kunnen leggen tussen tekstdelen. Daarbij maak ik gebruik van een tekstcorpus, dat bestaat uit studieteksten voor de havo en het vmbo. De hoofdvraag die ik stel luidt:

In hoeverre bieden oppervlaktekenmerken uit T-Scan een goede afspiegeling van de mate van elementinteractie en de benodigde hoeveelheid inference calls in studieteksten?

In het theoretisch kader volgt een bespreking van T-Scan en een onderbouwing van de twee teksteigenschappen. Daarna volgt de methode, gevolgd door een kwantitatieve en kwalitatieve analyse. Ik sluit af met een conclusie en discussie.

2. Theoretisch kader

Zoals gezegd begin ik met een uitleg van de leesbaarheidstool T-Scan. De paragrafen 2.2 tot en met 2.5 zijn gewijd aan literatuur met betrekking tot de (complexe) teksteigenschappen die ik met T-Scan wil onderzoeken.

2.1 T-Scan

T-Scan is een softwareprogramma dat in staat is teksten op een groot aantal kenmerken 'door te lichten' (Kraf, 2013). Daarbij gaat het om kenmerken op oppervlakteniveau, zoals het aantal connectieven, woordlengte en werkwoordsvormen. Deze kunnen echter wel een aanduiding zijn voor dieperliggende teksteigenschappen zoals informatiedichtheid, coherentie, woord- en zinsmoeilijkheid. T-Scan kan bijvoorbeeld de *type-token-ratio* (TTR) berekenen; het aantal unieke woorden, gedeeld door het totaal aantal woorden. Een lage TTR geeft aan dat er relatief veel unieke woorden voorkomen, en er dus sprake is van weinig woordherhaling. Dat zou kunnen duiden op een moeilijke tekst, met veel verschillende begrippen. Een reeks plaatsnamen heeft echter een zeer lage TTR, maar is niet per se moeilijk te begrijpen. Het is daarom belangrijk je te realiseren dat T-Scan alleen in staat is 'domme kenmerken' boven water te halen; die moeten worden geïnterpreteerd en zo nodig worden gecontroleerd. Een score in T-Scan kan een aanduiding zijn voor een bepaalde teksteigenschap, maar er is geen causale relatie tussen T-Scan kenmerken en teksteigenschappen. Het is daarom beter te spreken van een *symptoomrelatie* (Kraf & Pander Maat, 2009).

Ik onderscheid drie typen van problemen die ik in T-Scan verwacht tegen te komen. Ten eerste kunnen bepaalde kenmerken dubbelzinnig zijn: 'daarnaast' kan een plaats aanduiden (*het huis daarnaast*), maar kan ook als connectief gebruikt worden (*daarnaast is het belangrijk...*). Bij het zoeken naar verbindingswoorden 'kijkt' T-Scan echter alleen of een woord in een bepaalde lijst voorkomt (Kraf, 2013, p. 52). Omdat 'daarnaast' in de lijst met opsommende verbindingswoorden staat, zal T-Scan dit woord als zodanig beschouwen, ongeacht de context. Dit probleem duid ik aan als *probleemtype 1*; hoewel er dubbelzinnigheid is, is één betekenis dominant, en zal het kenmerk meestal niet voor problemen zorgen. Met *probleemtype 2* duid ik gevallen aan waarin dat wél zo is. T-Scan herkent het woord *nu* bijvoorbeeld als causaal verbindingswoord (Kraf, 2013, p. 53). Of dat woord in de praktijk inderdaad *meestal* als causaal verbindingswoord voorkomt, is echter de vraag. Ten derde kunnen sommige tekstkenmerken per context verschillen. Eigennamen kunnen bijvoorbeeld duiden op een moeilijke tekst, met veel verwijzingen naar personen, geografische namen en andere eigennamen. Dat geldt echter niet altijd: de namen in een fictief verhaal maken de tekst meestal niet moeilijker. Een fictief

verhaal is meestal de enige bron van waaruit de lezer de personages moet leren kennen. In een geschiedenistekst hoeft dat niet zo te zijn: een auteur kan historische personen als bekend veronderstellen, zonder deze zelf besproken te hebben. Dat kan het leggen van verbanden (inferenties) lastig maken. De hoeveelheid eigennamen kan dus een aanwijzing zijn voor vereiste wereldkennis, maar dat hoeft niet. Dit 'contextprobleem' duid ik aan als *probleemtype 3*.

2.2 Samenhang: cohesie en coherentie

Om tot de complexe teksteigenschappen te komen die ik met T-Scan wil onderzoeken, bespreek ik eerst de kenmerken *cohesie* en *coherentie*. Het eerste begrip, cohesie, slaat op de samenhang *binnen* de tekst, op grond van tekstuele eigenschappen (Graesser, McNamara & Lauwerse, 2003). Denk hierbij aan verbindingswoorden en verwijzingen. Op grond van vooraf gestelde criteria kan de cohesie van een tekst worden vastgesteld (Graesser, McNamara, Lauwerse & Cai, 2004, pp 1-2). Een tekst met een hoge cohesie hoeft echter niet per se een begrijpelijke tekst te zijn; daarvoor is ook het tweede begrip *coherentie* belangrijk.

Coherentie is niet direct, objectief, uit de tekst af te leiden. Het gaat daarbij om de samenhang 'in het hoofd' van de lezer. Ten eerste moet het voor de lezer duidelijk zijn welke informatie aan welke referenten (personen of dingen) wordt toegekend en hoe in volgende zinnen naar deze of nieuwe referenten wordt verwezen: dit is referentiële coherentie (Noordman & Maes, 2000, pp. 34-35). Ten tweede moet de inhoudelijke samenhang tussen woorden en zinnen duidelijk zijn: dat is relationele coherentie. Tekstelementen (op oppervlakteniveau) zoals connectieven kunnen de lezer helpen bij het herkennen en begrijpen van coherentierelaties (p. 34).

Coherentie en cohesie kunnen dus samenhangen met oppervlaktekenmerken/ T-Scanmerken uit de tekst.

2.3 Tekstrepresentatie: de complexe eigenschappen

In hoeverre een tekst begrepen wordt, is afhankelijk van de *representatie* die de lezer zich van de tekst vormt. Een hoge coherentie (inhoudelijke samenhang) is daarbij belangrijk – maar geen sleutel tot succes, want de tekstrepresentatie is mede afhankelijk van de lezer.

Van Dijk & Kintsch (1983) onderscheiden drie niveaus van begrijpelijkheid. Op het eerste niveau, de *oppervlakterepresentatie*, worden formuleringen en grammaticale vormen herkend, maar ze worden slechts kort in het geheugen vastgehouden. Bij het tweede type, de *tekstbetekenisrepresentatie*, wordt de betekenis (niet de formulering) van de woorden en de samenhang tussen zinnen op tekstniveau herkend. Bij het derde representatietype, de *situatiemodelrepresentatie*, wordt de inhoud van de tekst *echt* begrepen; de lezer verbindt de elementen uit de tekst met zijn (wereld)kennis en is in staat zich een mentale representatie te

vormen: het *situatiemodel* (Van Dijk & Kintsch, 1983, pp. 11-2; McNamara e.a., 2004, p. 2; Kamalski, e.a., 2004, pp. 86-7).

Deze drie niveaus bouwen op elkaar voort: de lezer kan pas een situatiemodel vormen als de tekst op tekstbetekenis- en oppervlakteniveau begrepen is. Wanneer lezers daar moeite mee hebben - bijvoorbeeld door onvoldoende wereldkennis - kan de auteur een handje helpen:

Bijvoorbeeld:

- 1) *Voor Overijssel worden hagelbuien verwacht. Het wordt slecht weer in het Oosten. De voetbalwedstrijden in Deventer zijn afgelast.*

Deze drie zinnen hebben op oppervlakteniveau geen overeenkomsten; er is bijvoorbeeld geen sprake van woordherhaling. Toch is dit fragment wel te begrijpen wanneer de lezer zelf de volgende verbanden legt:

- 2) *Voor Overijssel [een provincie in het Oosten] worden hagelbuien [een vorm van slecht weer] verwacht. Het wordt slecht weer [want: hagelbuien] in het Oosten. De voetbalwedstrijden in Deventer [stad in Overijssel/ in het Oosten] zijn afgelast [de voetbalwedstrijden vinden buiten plaats en kunnen dus door het slechte weer niet doorgaan].*

De toevoegingen tussen haken moet de lezer zelf 'invullen' om een coherente representatie te kunnen vormen. Wanneer deze toevoegingen in de tekst zouden staan, zou het voor een onwetende lezer gemakkelijker zijn de tekst te begrijpen. Hij hoeft bijvoorbeeld, wanneer hij de tweede zin leest, niet zelf te bedenken dat Overijssel in het Oosten ligt. In de eerste regel is dat immers al expliciet genoemd. De ontbrekende verwijzingen (zoals in voorbeeld 1) noemen Britton & Gülgöz *inference calls* (1991, pp. 330-331). Zij stellen dat een tekst die *weinig* inference calls vereist beter begrepen en onthouden wordt: sommige lezers zijn niet in staat de vereiste inferenties te maken en zullen dus geen coherente representatie kunnen vormen. In mijn onderzoek maak ik ook gebruik van de term *inference calls*, hoewel ik dat begrip op een andere manier gebruik. Hoewel woordherhaling (zie tabel 1: *argumentoverlap*) zeker een rol speelt, denk ik niet dat dat de *enige* manier is om het aantal inference calls te verminderen. Dat blijkt ook uit voorbeeld 3:

- 3) *Voor Overijssel worden hagelbuien verwacht. [Dus] wordt het slecht weer in het Oosten. [Daarom] zijn de voetbalwedstrijden in Deventer afgelast.*

De connectieven (dus/daarom) maken het de lezer gemakkelijker het fragment te begrijpen, net als woordherhaling. Ik beschouw de verbanden in voorbeeld 3 daarom niet als inference calls. Britton & Gülgöz zouden dat – door het ontbreken van woordherhaling – wel doen.

Deze voorbeelden laten zien dat de lezer geholpen kan worden bij het vormen van een tekstbetekenisrepresentatie, bijvoorbeeld door connectieven. T-Scan is in staat connectieven te herkennen; dat is dus een van de ‘domme kenmerken’ die een aanduiding kan zijn voor dieper liggende eigenschappen.

Dat de lezer een tekstbetekenisrepresentatie vormt, betekent nog niet dat de tekst echt begrepen is. Om een tekst te kunnen begrijpen *en* onthouden, is het namelijk belangrijk dat de lezer zich een mentale voorstelling kan vormen van de wijze waarop de inhoud van de tekst samenhangt: een *schema* (Sweller, Van Merriënboer & Paas, 1998). In zo’n schema kunnen verschillende theoretische elementen worden ‘opgeslagen’, en wordt het *kortetermijngeheugen* (of werkgeheugen) ontlast. Dat kan namelijk maar zo’n zeven elementen tegelijk verwerken, en het schema kan als één onderdeel in het *langetermijngeheugen* worden opgenomen. Later kan het van daaruit weer worden opgeroepen en is het weer beschikbaar in het werkgeheugen. Een schema is te vergelijken met het hierboven besproken situatiemodel: ook bij het maken van een schema verbindt de lezer de inhoud van de tekst met zijn/ haar voorkennis. In hoeverre de lezer in staat is een schema of situatiemodel te vormen is echter ook afhankelijk van het onderwerp van de tekst. Mijn tweede ‘complexe’ tekstkenmerk *element interaction* speelt daarbij een grote rol.

Element interaction is de mate waarop elementen uit de tekst in relatie tot elkaar moeten worden begrepen (Sweller, & Chandler, 1994, pp. 188-189; Sweller e.a., 1998, pp. 263-4). Een opsomming van Amerikaanse presidenten in een geschiedenisboek heeft een zeer lage element interaction. Zo’n tekst is daarom bijna onmogelijk te schematiseren en dus lastig te onthouden. Bij zo’n zwak interacterende tekst is het markeren van coherentierelaties meestal geen oplossing; de onderdelen kunnen alleen los van elkaar worden onthouden. Een tekst over hoog- en laagconjunctuur in een economieboek daarentegen, kan er op het eerste gezicht moeilijk uitzien. Na het leggen van de benodigde inferenties kan de lezer hier echter *wel* een schema van vormen waardoor de theorie voor langere tijd onthouden wordt. De auteur kan de lezer daar eventueel nog bij helpen, bijvoorbeeld door coherentierelaties te markeren zoals in voorbeeld 3 (pag. 7): interacties worden dan verduidelijkt voor lezers met weinig voorkennis. Begrijpelijkheid is dus (behalve van lezerskenmerken) afhankelijk van twee tekstkenmerken:

1. **Inference calls (stijl)**: Hoeveel inference calls moet de lezer zelfstandig op basis van (wereld)kennis maken om de tekst te begrijpen?
2. **Element interaction (inhoud)**: In hoeverre is de inhoud van de tekst als een samenhangend geheel te begrijpen? Met andere woorden, hoe hoog is de element interaction?

De voorbeelden 1 t/m 3 (pag. 7) verschillen dus van elkaar op basis van *inference calls*; de lezer moet in tekst 1 zelf het verband leggen tussen het slechte weer, de ligging van de provincie Overijssel, de stad Deventer en het voetbal. In zin 3 zijn die relaties gemarkeerd, de lezer hoeft de inferenties daardoor niet helemaal op eigen kracht te maken. Het aantal connectieven zou dus, zoals ik heb laten zien, een aanduiding kunnen zijn voor het aantal inference calls. Qua element interaction verschillen de voorbeelden (pag. 7) niet van elkaar.

De eerste eigenschap, inference calls, lijkt op het eerder besproken begrip *cohesie*. Element interaction, de tweede eigenschap, is te beschouwen als *logische (relationele) coherentie*. Niet alleen moet de inhoud samenhangen (coherent zijn), de samenhang moet daarnaast zodanig zijn dat het vormen van een schema mogelijk is. Daarvoor is het belangrijk dat de tekst 'logisch' (of: causaal) samenhangt.

2.4. En de domme kenmerken?

De titel van dit onderzoek luidt: *Domme kenmerken... complexe eigenschappen?*. De 'complexe eigenschappen' heb ik zo-even besproken: dat zijn inference calls & element interaction. Aan de 'domme kenmerken' heb ik tot nu toe minder aandacht geschonken. Ik bespreek in deze paragraaf de belangrijkste typen 'domme kenmerken'. Zoals gezegd voer ik het onderzoek met T-Scan uit; de precieze definitie van de T-Scan kenmerken is te vinden in tabel 1 (pagina 16).

Laat ik beginnen met kenmerken die duiden op inference calls. Uit de voorbeelden op pagina 7 komt al een kenmerk naar voren: connectieven. Zoals uit de voorbeelden bleek, kunnen connectieven de relatie tussen tekstdelen expliciteren. De lezer heeft dan minder (wereld)kennis nodig om die relatie te leggen: veel connectieven zouden dus een aanduiding kunnen zijn van *weinig* inference calls. Naast connectieven kan ook een kenmerk als *woordherhaling* het leggen van verbanden vereenvoudigen. Veel woordherhaling zou immers kunnen duiden weinig besproken onderwerpen en weinig gedachtesprongen. Tot slot kan ook het gebruik van (eigen)namen duiden op veel inference calls. Vooral wanneer de auteur een bestaand persoon noemt, en deze bij de lezer bekend veronderstelt (zie ook hoofdstuk 2.1).

Oppervlaktekenmerken voor element interaction zijn iets lastiger te vinden. Connectieven kunnen echter ook bij deze eigenschap een rol spelen. Vooral het gebruik van causale

connectieven (dus, daardoor, daarom) duiden op een oorzakelijk opgebouwde tekst: een tekst dus, met een hoge interactie. Eigennamen zouden ook kunnen helpen bij het vaststellen van de interactie. Veel (eigen)namen zouden namelijk kunnen duiden op een verhalende of opsommende tekst: die hebben vaak een lage interactie (zie ook hoofdstuk 2.5). Wat voor inference calls geldt qua woordherhaling, zou ook voor element interaction kunnen gelden. Wanneer een tekst veel woordherhaling bevat, is de kans op lange opsommingen en uitweidingen mogelijk klein. Bovendien is de kans op een sterke causale samenhang wellicht groter wanneer er weinig (verschillende) onderwerpen besproken worden. Mogelijk hangt woordherhaling hierdoor samen met de interactie binnen de tekst. Er is natuurlijk nooit sprake van een *causaal* verband tussen een tekstkenmerk en een eigenschap zoals element interaction. Wel kan er, zoals gezegd, sprake zijn van een symptoomrelatie: teksten met een lage inhoudelijke samenhang bevatten *over het algemeen* minder causale connectieven dan hoog interacterende teksten. Wanneer dat inderdaad zo is, zou het mogelijk moeten zijn de complexe eigenschappen 'boven water te krijgen' door de teksten op (een combinatie van) 'domme' oppervlaktekenmerken te toetsen.

2.5 Narrative vs. expository texts

De twee complexe teksteigenschappen (element interaction & inference calls) staan niet helemaal los van elkaar. Ik stelde zo-even al, dat het markeren van coherentierelaties bij zwak interacterende teksten niet altijd zin heeft. Een reeks Amerikaanse presidenten is niet gemakkelijker te onthouden – laat staan schematiseren - door connectieven te gebruiken (daarna, daarna enz.). Het lijkt daarom aannemelijk dat juist hoog interacterende teksten in een stijl geschreven zijn die het maken van inferenties vergemakkelijkt: bijvoorbeeld door het expliciteren van relationele coherentie.

Een onderzoek van McNamara e.a. (2012) lijkt deze aanname te bevestigen. Zij onderzochten een groot corpus schoolteksten met *Coh-Matrix*, een leesbaarheidstool, vergelijkbaar met T-Scan. Daarbij onderscheidde zij twee teksttypen: *narrative texts* (verhalende teksten) en *expository texts* ('leerteksten').

Drie vakken stonden in hun onderzoek centraal:

- *Language arts*: Taal- en literatuurvakken.
- *Social studies*: Sociologie, geschiedenis, politicologie, maatschappijleer. (Grofweg alle alfavakken behalve de talen, samen met de gammavakken.)
- *Science texts*: Alle bètavakken zoals wiskunde en natuurkunde.

Het eerste vak bevat vooral verhalende teksten, terwijl in het laatste vak *expository texts* overheersen. Het genre *social studies* bevat beide teksttypen.

Uit de resultaten van Coh-Metrix bleek dat narratieve teksten een lagere referentiële en relationele cohesie hebben, en weinig causale betekenisrelaties bevatten. Verhalen zijn in hoge mate temporeel opgebouwd, en verbanden zijn veelal impliciet (2012, p.7). Mogelijk is de element interaction hierdoor zwak, en bevatten deze teksten veel inference calls. Het laatste vak (*science texts*) onderscheidt zich juist door een causale opbouw. Ook komen in dit vak veel gemarkeerde coherentierelaties voor: de lezer wordt dan geholpen bij het maken van inferenties. Mogelijk heeft dit schoolvak dus juist een hoge element interaction en weinig inference calls. Het tweede vak, *social studies*, combineert eigenschappen van *language arts* en *science texts*: *[S]ocial studies texts are more like narratives in the depiction of human experiences and commonly known concepts, but more like science texts in the use of cohesive elements to didactically explain the content* (2012, p.3).

Een ander onderzoek naar Coh-Metrix van Graesser, McNamara & Kulikowich (2011) laat niet alleen verschil zien tussen vakken, maar ook tussen niveaus. Graesser e.a. maakten daarbij gebruik van dezelfde drie vakken die ik hierboven besprak (*language arts*, *social studies* en *science texts*). Uit dit onderzoek blijkt dat teksten, naarmate het niveau stijgt, steeds minder narratief worden. Anders gezegd: de teksten krijgen steeds meer eigenschappen van *expository texts*. Dat geldt voor elk van de drie vakken. Wel was de referentiële coherentie hoger in de lagere niveaus (p. 230). Dat bleek onder andere uit de kenmerken argumentoverlap en type-token-ratio. De argumentoverlap was hoog in teksten voor een laag niveau; de type-token-ratio was juist laag (p. 229).

3. Methode

Om de hoofdvraag uit de inleiding te beantwoorden voer ik een corpusonderzoek uit. De teksten uit dit corpus (zie ook hoofdstuk 3.1) worden op basis van T-Scanresultaten aan een kwantitatieve analyse onderworpen. Vervolgens voer ik voor een aantal teksten een kwalitatieve analyse uit. Daarbij richt ik mij vanzelfsprekend vooral op de vraag in hoeverre oppervlaktekenmerken een afspiegeling zijn van 'complexe' eigenschappen. In dit hoofdstuk bespreek ik eerst de precieze opbouw van het tekstcorpus. Daarna ga ik verder met de onderzoeksopzet, waarin het kwantitatieve en het kwalitatieve onderzoek besproken worden.

3.1 Materialen: Het tekstcorpus

Het corpus bestaat uit geschiedenis- en economieteksten van de bovenbouw (klas 4-5) van de havo (34 geschiedenis teksten, 25 economie teksten) en de onderbouw (klas 1-2) van het vmbo (26 geschiedenis teksten, 25 economie teksten). Ik verwacht bij geschiedenis teksten vooral narratieve eigenschappen aan te treffen. In economieboeken verwacht ik daarentegen vooral veel expository texts te vinden. Geschiedenis teksten zijn immers vaak narratief; het Duitse woord voor verhaal is niet toevallig '*Geschichte*'. Economie teksten zijn naar mijn verwachting theoretischer en 'logischer' van aard: zeker in vergelijking met geschiedenis teksten. Deze aanname is vooral gebaseerd op het onderscheid tussen expository texts en narrative texts (McNamara e.a., 2012, pp. 5-7).

Het onderscheid in vak en niveau hebben wij gemaakt om twee redenen; ten eerste hadden wij simpelweg de beschikking over veel boeken van deze vakken en schooltypes. Ten tweede hebben deze boeken een totaal andere doelgroep, zowel qua leeftijd (onderbouw/ bovenbouw) als qua niveau (vmbo/ havo). Wij nemen daarom aan dat de teksten sterk verschillen op grond van 'leesniveau'. Wij hebben geen 'hybride boeken' gebruikt: boeken die bedoeld zijn voor twee schooltypen tegelijk, zoals havo én vwo, of vmbo én havo. Bij vmbo-boeken is enkel gebruik gemaakt van de drie hoogste niveaus, de theoretische, gemengde, en kaderberoepsgerichte leerweg. De verhouding tussen theorie en praktijk verschilt tussen deze stromingen, maar het niveau is gelijk: elk van deze niveaus geeft toegang tot een mbo-opleiding op niveau 4. Het laagste vmbo-niveau (basisberoepsgerichte leerweg) is niet gebruikt. De teksten voor dit niveau zijn zeer kort en onsamenhangend. Veel boeken bestaan slechts uit afbeeldingen met een korte toelichting. Hierdoor was het moeilijk - en vaak onmogelijk - geschikte teksten te vinden.

Een gedeelte van de benodigde boeken was aanwezig op de universiteit – het grootste gedeelte komt echter uit bibliotheken, zoals die van de Hogeschool Windesheim te Zwolle. Ook is er een flink aantal teksten uit online-inkijkexemplaren gehaald (zie ook bijlage 8).

Wij hebben een aantal eisen gesteld bij de aanleg van het corpus. Zo is de lengte van de teksten –een enkele uitzondering daargelaten – tussen de 300 en 500 woorden. Er is geprobeerd zo min mogelijk teksten op te nemen die bestaan uit veel puntsgewijze opsommingen of verwijzingen naar afbeeldingen. Hoewel een tekst soms uit meerdere paragrafen bestaat is er wel op de thematische coherentie gelet; de teksten zijn op zichzelf te lezen en te begrijpen. Er komen dus geen teksten voor die twee of meer heel verschillende onderwerpen behandelen. De teksten zijn met behoud van de opmaak in Word-bestanden bijeengebracht. Vetgedrukte, cursieve en onderstreepte woorden zijn als zodanig overgenomen. Dat geldt ook voor titels, tussenkopjes en verwijzingen tussen haakjes. Deze kenmerken zijn echter in T-Scan niet geanalyseerd. Een overzicht van de gebruikte teksten is te vinden in bijlage 8.

3.2 Niveauverschillen

Zoals gezegd gebruik ik teksten voor de havo en het vmbo. Naast verschillen tussen schoolvakken verwacht ik namelijk ook een verschil te vinden *tussen* deze niveaus. Dat effect is echter lastig te voorspellen. Enerzijds kan bij havo-leerlingen een hogere voorkennis worden verwacht: zij kunnen dus meer inferenties maken op basis van wereldkennis. Daarom zouden vmbo-teksten dus meer relationele/ referentiële coherentiesignalen moeten bevatten. Dat bleek ook uit het onderzoek van Graesser e.a. (2011), dat ik in hoofdstuk 2.5 besprak. Die aanname wordt echter door andere onderzoeken weerlegd. Omdat vmbo'ers moeite hebben met lezen, worden betekenisrelaties vaak juist *niet* geëxpliciteerd, om de zinnen niet te lang te maken. Deze teksten zijn *gefragmenteerd*: de zinnen zijn extreem kort, en er wordt nauwelijks gebruik gemaakt van connectieven (Land, 2009; Land, Sanders & van den Bergh, 2008). Dit teksttype komt echter vooral in het laagste niveau voor (basisberoepsgerichte leerweg). De teksten uit mijn corpus, daarentegen, zijn niet voor dit laagste niveau bedoeld, maar voor de 3 hogere niveaus (kaderberoepsgerichte, gemengde en theoretische leerweg). De vereenvoudiging op basis van zinslengte – met gefragmenteerde teksten tot gevolg- komt daarom niet veel voor.

3.3 Onderzoeksoptzet

Nadat de teksten met T-Scan zijn geanalyseerd, zijn de waarden in SPSS ingevoerd. Daarmee wordt vastgesteld in hoeverre de teksten verschillen op het gebied van verschillende T-Scan kenmerken (zie tabel 1). In het theoretisch kader stelde ik dat de score in T-Scan een aanduiding kan zijn voor de element interaction en de aanwezige inference calls van een tekst. In

hoeverre dat klopt probeer ik in een kwalitatief onderzoek vast te stellen. Daarbij ga ik als volgt te werk:

Stap 1: Wanneer één van de onafhankelijke variabelen (vak/ niveau) zorgt voor een significant effect op één van de tekstkenmerken (zoals argumentoverlap), zoek ik twee teksten die sterk verschillen op basis van *dat* kenmerk. Deze teksten neem ik de bijlagen 1-7 op.

Stap 2: Indien mogelijk wordt het tekstkenmerk gemarkeerd, connectieven worden bijvoorbeeld onderstreept. Sommige tekstkenmerken, zoals NP-dichtheid of argumentoverlap kunnen niet worden gemarkeerd. De tekst wordt dan wel in de bijlage opgenomen. Nu probeer ik na te gaan in hoeverre het tekstkenmerk het verwachte effect heeft: zorgen causale connectieven inderdaad voor causale betekenisrelaties? En heeft een tekst met opmerkelijk weinig causale connectieven inderdaad een zwakke causale opbouw?

Om de informatiedichtheid kwalitatief 'vast te stellen' geef ik elke 'betekeniseenheid' een nummer. In de tekst wordt deze als volgt genoteerd: [1] zin] . Als betekeniseenheid vat ik een zinsdeel op dat iets toevoegt aan een vorig zinsdeel, en kan worden verwijderd zonder dat de zin ongrammaticaal wordt. Bijvoorbeeld:

[1] Na de dood van keizer Theodosius] [2] in 395] [3] werd het Romeinse rijk verdeeld] [4] onder zijn zonen][3a] in een oostelijk en een westelijk deel]: [5] het Oost-Romeinse rijk zou nog tot 1453 blijven bestaan.] [6] Toen werd het veroverd door de Ottomaanse Turken.] (bijlage 5 regel 1-3).

Bepaling 2 markeer ik als zelfstandige betekeniseenheid, omdat deze kan worden verwijderd zonder dat dat zin ongrammaticaal wordt. De bepalingen 3 & 3a vormen echter samen één zinsdeel; 'verdelen' is een transitief werkwoord, en vereist dus een lijdend voorwerp. De zin: *[H]et Romeinse rijk werd verdeeld' is niet grammaticaal. '[H]et Romeinse rijk werd verdeeld in een oostelijk en een westelijk deel' is dat wel. De bepaling 'onder zijn zonen' kan wel worden verwijderd, en is dus een apart zinsdeel. De laatste zin kan niet verder worden gesplitst en vormt dus één betekeniseenheid.

Een hele zin is *in ieder geval* één betekeniseenheid, omdat een zin per definitie kan worden verwijderd zonder dat een andere zin ongrammaticaal wordt. Daarbij rijst één probleem: sommige teksten maken nu eenmaal gebruik van korte zinnen, en bevatten daardoor (ten onrechte) veel betekeniseenheden. Om dat probleem te omzeilen, maar teksten toch te kunnen onderscheiden qua informatiedichtheid, bereken ik zowel het aantal betekeniseenheden per 1000 woorden als het aantal betekeniseenheden per zin. Deze waarden geven samen een beeld van de informatiedichtheid van een tekst.

Stap 3: Nu leg ik de link met de teksteigenschappen inference calls en element interaction. Ik bekijk nu in hoeverre teksten –die verschillen in T-Scan kenmerken- ook van elkaar verschillen op een ‘dieper niveau’, namelijk qua element interaction *en* inference calls. Hoewel zo’n analyse handmatig is, probeer ik de eigenschappen door systematische redenering zo betrouwbaar mogelijk in kaart te brengen. Voor de element interaction maak ik onder andere gebruik van de argumentatieschema’s van Verhoeven & Schellens (2008). Hierin verbind ik verschillende betekenseenheden uit de tekst met elkaar, zodat de samenhang tussen deze eenheden inzichtelijk wordt. Wanneer er geen samenhang is tussen de onderdelen van de tekst, is het onmogelijk deze in zo’n schema samen te brengen. Inference calls probeer ik vooral door redeneren in kaart te brengen. Ik lees beide teksten zeer grondig, waarbij ik me richt op inferenties die noodzakelijk zijn om een tekst te begrijpen. Wanneer ik die betekenisvol acht neem ik de desbetreffende passage op in de lopende tekst. Ik maak daarbij inzichtelijk *welke* inferentie de leerling moet maken, *waarom* daar (wereld)kennis voor nodig is, en *waarom* dat problematisch kan zijn. Natuurlijk kan ik niet *alle* inferenties uit een bepaalde tekst bespreken. De lezer kan mijn analyse echter beoordelen door de tekst in de bijlage op te zoeken. In de kwalitatieve analyse wordt ook kritisch naar het programma T-Scan gekeken, met name naar de drie eerder genoemde probleemtypen (zie hoofdstuk 2.1).

3.4 Relevante tekstkenmerken in T-Scan

In tabel 1 staan de T-Scan kenmerken die relevant kunnen zijn voor het bepalen van de element interaction en inference calls van de tekst.

Tabel 1: Tekstkenmerken die mogelijk samenhangen met element interaction of inference calls

Tekstkenmerk	Definitie	Relatie met element interaction	Relatie met inference calls
<p>Argumentoverlap Met vorige zin</p> <p>Waarde in T-Scan: aantal overlappende argumenten met de vorige zin, berekend op basis van 1000 woorden.</p>	<p>Argumentoverlap met de vorige zin laat zien in hoeverre proposities (argumenten) uit de vorige zin herhaald worden. Als argument worden hier aangemerkt: voornaamwoorden (maar niet aanwijzende), namen, zelfstandige naamwoorden en hoofdwerkwoorden. Bijvoorbeeld:</p> <p>1). <i>Abraham Kuyper, greep naar een laatste redmiddel: een beroep op de Koning om de onderwijswet niet te ondertekenen. Vier dagen nadat de Tweede Kamer de wet had aangenomen (...)</i></p> <p>2). <i>Abraham Kuyper, greep naar een laatste redmiddel: een beroep op de Koning om de onderwijswet niet te ondertekenen. De koning ondertekende de onderwijswet echter wel, en vier dagen nadat de Tweede Kamer de wet had aangenomen (...)</i></p> <p>In het eerste voorbeeld worden argumenten uit de eerste zin niet herhaald. In de tweede zin is dat wel het geval, de woorden 'koning', 'ondertekende', en 'onderwijswet' worden herhaald (het voorbeeld is gebaseerd op tekst 2 uit bijlage 1. Deze tekst wordt besproken in paragraaf 6.1).</p> <p>Argumentoverlap is een goede indicator voor de referentiële coherentie van een tekst (Kintsch & van Dijk, 1978, p. 367).</p>	<p>Mogelijk duidt een hoge argumentoverlap op een hoge element interaction.</p> <p>Een tekst met een <i>lage</i> argumentoverlap behandelt immers veel argumenten, en de kans dat die vele argumenten een duidelijke interactie vormen is daardoor wellicht klein. In het voorbeeld hiernaast heeft argumentoverlap geen effect; de <i>element interaction</i> is in beide voorbeelden gelijk.</p>	<p>Omdat een hoge argumentoverlap kan duiden op een klein aantal besproken onderwerpen, ga ik er vanuit dat een hoge argumentoverlap duidt op weinig inference calls (en andersom). Naarmate er meer onderwerpen besproken worden, is de kans groter dat er meer inferenties gemaakt moeten worden.</p> <p>In het voorbeeld hiernaast hoeft in de tweede zin inderdaad een inferentie minder gemaakt te worden. De lezer hoeft niet zelf te bedenken dat het aannemen van de wet door de Tweede Kamer betekent dat de koning de wet wél ondertekend heeft.</p>
<p>Argumentoverlap Met voorgaande 50 woorden</p> <p>Waarde in T-Scan: aantal overlappende argumenten met de voorgaande 50 woorden, berekend op basis van 1000 woorden.</p>	<p>De argumentoverlap met de voorgaande 50 woorden laat zien in hoeverre argumenten in de vorige 50 woorden herhaald worden.</p> <p>Dit type argumentoverlap lijkt op de argumentoverlap met de vorige zin. Dit kenmerk meet de overlap echter op een minder lokaal niveau dan de argumentoverlap met de vorige zin.</p>	<p>Net als de argumentoverlap met de vorige zin, duidt een hoge argumentoverlap met de voorgaande 50 woorden mogelijk op een hoge <i>element interaction</i>.</p>	<p>Ook voor <i>inference calls</i> geldt hetzelfde als de argumentoverlap met de vorige zin: een hoge argumentoverlap duidt mogelijk op weinig inference calls.</p>

<p>Connectieven</p> <p>Alle connectieven</p> <p>Waarde in T-Scan: gemiddeld aantal connectieven per 1000 woorden.</p>	<p>Connectieven zijn verbindingswoorden die de relatie tussen tekstdelen (kunnen) expliciteren. Daardoor kunnen connectieven een tekst beter leesbaar maken (Kraf, 2013, p. 24). Het kenmerk 'alle connectieven' bestaat uit:</p> <ul style="list-style-type: none"> -Temporele connectieven (bijv. daarna, ooit, opeens) -Contrastieve connectieven (bijv. hoewel, behalve, maar) -Comparatieve connectieven (bijv. alsof, zoals) -Causale connectieven (bijv. doordat, omdat, nu, om) <p>Dit kenmerk wordt door T-Scan 'berekend' door de woorden op te zoeken in een viertal lijsten. Wanneer ik spreek over een 'temporeel verbindingswoord' bedoel ik dus een woord dat in de lijst connectieven' is opgenomen. Of dat juist is, is soms de vraag (zie par. 2.1).</p>	<p>Ik verwacht geen relatie tussen het totaal aantal connectieven en de elementinteractie. (Wel voor causale connectieven, zie daarvoor het volgende kenmerk.)</p>	<p>Connectieven helpen bij het leggen van verbanden; mogelijk duiden veel connectieven daarom op weinig inference calls.</p>
<p>Connectieven:</p> <p>Causale connectieven</p> <p>Waarde in T-Scan: gemiddeld aantal causale connectieven per 1000 woorden.</p>	<p>Connectieven zoals 'doordat', 'omdat' en 'om' zijn, zoals hierboven besproken, causaal. Anders gezegd: dit type connectieven duidt vaak een causale (oorzaak-gevolg) relatie aan tussen twee zinnen of zinsdelen.</p>	<p>Connectieven van causaliteit wijzen op een causale, en dus logisch opgebouwde tekst. Mogelijk duiden veel causale betekenisrelaties daarom op een hoge element interaction.</p>	<p>Ik verwacht geen relatie tussen het causale connectieven en het aantal inference calls. Wel verwacht ik een effect van het totale aantal connectieven (zie daarvoor het vorige kenmerk).</p>
<p>Informatiedichtheid:</p> <p>Type-Token-Ratio (TTR)</p> <p>Waarde in T-Scan: het aantal unieke woorden gedeeld door het totaal aantal woorden van de tekst.</p>	<p>De type-token-ratio wordt berekend door het aantal unieke woorden te delen door het totaal aantal woorden. Teksten met een hoge TTR hebben dus veel unieke woorden, en weinig woordherhaling. Alfavakken als geschiedenis zijn talig van aard, hebben over het algemeen een hoge TTR. Mogelijk duidt dat op een hoge informatiedichtheid. Een lage TTR duidt juist op een lage informatiedichtheid: er komen immers relatief weinig unieke woorden in de tekst voor, en de kans dat de tekst veel verschillende onderwerpen bespreekt is hierdoor klein.</p>	<p>Informatiedichtheid hangt mogelijk negatief samen met de element interaction: hoe lager de informatiedichtheid, hoe hoger de element interaction. Een lage TTR duidt dus mogelijk op een <i>hoge</i> element interaction.</p>	<p>Een lage TTR duidt op veel woordherhaling. Mogelijk is dat een aanwijzing voor weinig inference calls.</p>
<p>Informatiedichtheid</p> <p>Rarity index</p> <p>Waarde in T-Scan: het aantal woorden dat 4 keer of vaker herhaald wordt, gedeeld door het totaal aantal woorden van de tekst.</p>	<p>De <i>rarity index</i> wordt berekend door de lemma's die 4 keer of minder herhaald worden, te delen door het totaal aantal lemma's. Lidwoorden worden niet meegerekend, bij lemmaherhaling noch bij het totaal aantal lemma's. Dit kenmerk lijkt hiermee sterk op de TTR; een hoge rarity index duidt dus op een lage informatiedichtheid,</p>	<p>De rarity index is sterk verwant met de TTR. Een lage rarity index duidt dus (net als de TTR) op een hoge element interaction.</p>	<p>Net als de TTR duidt een lage rarity index op veel woordherhaling, en daarmee op weinig inference calls.</p>
<p>Informatiedichtheid:</p> <p>Inhoudswoorden</p>	<p>Deze maat telt het aantal inhoudswoorden per 1000 woorden. Inhoudswoorden zijn zelfstandig naamwoorden, adjectieven, bijwoorden en hoofdwerkwoorden. Veel inhoudswoorden duiden</p>	<p>Ik verwacht een lagere <i>element interaction</i> naarmate het aantal</p>	<p>Ik verwacht meer inference calls aan te treffen naarmate de tekst meer</p>

Waarde in T-Scan: het aantal inhoudswoorden per 1000 woorden.	op een hoge informatiedichtheid; een inhoudswoord voegt immers informatie toe.	inhoudswoorden- en daarmee de informatiedichtheid-toeneemt.	inhoudswoorden bevat. De kans dat er veel onderwerpen besproken worden (die een inferentie vereisen) neemt dan immers toe.
Informatiedichtheid: NP-dichtheid Waarde in T-Scan: het gemiddeld aantal NP's per zin gedeeld door het gemiddeld aantal woorden per zin.	Dit kenmerk berekent het gemiddeld aantal naamwoordgroepen (NP's: <i>noun phrases</i>) per zin. Bijvoorbeeld: 1. [[Die man] [met die hoed]] is [kunstenaar]. 2. [Die man] is [kunstenaar]. De eerste zin bevat 4 NP's: 1) die man, 2) die hoed, 3) die man met die hoed, 4) kunstenaar. De tweede zin bestaat uit twee NP's: 1) Die man, 2) kunstenaar. T-Scan deelt het aantal NP's door het aantal woorden in de zin; de NP-dichtheid voor de eerste zin is dus $4/7 = 0,57$, en voor de tweede zin $2/4 = 0,50$. Een hoge NP-density duidt mogelijk op een hoge informatiedichtheid; een NP voegt namelijk, zoals ook uit de voorbeeldzinnen blijkt, informatie toe.	Een hoge NP-dichtheid duidt op een hoge informatiedichtheid. Daarom verwacht een negatieve relatie tussen de NP-dichtheid en element interaction te vinden: hoe hoger de NP-dichtheid, hoe lager de element interaction.	Mogelijk duidt een hoge NP-dichtheid op veel inference calls. De NP-dichtheid hangt immers samen met de informatiedichtheid, net als de kenmerken 'inhoudswoorden', 'TTR' en inference calls.
Eigennamen Waarde in T-Scan het aantal eigennamen op 1000 woorden.	Onder eigennamen worden onder andere voor- en achternamen, bedrijfsnamen, geografische namen verstaan.	Veel eigennamen duiden op een tekst met een narratief karakter. Mogelijk is de element interaction dan laag.	Het gebruik van veel eigennamen duidt mogelijk op veel verwijzingen naar 'algemene kennis' en dus een groot aantal inference calls.

Zoals ik in hoofdstuk 2.4 al opmerkte, hangen de begrippen inference calls en element interaction samen. Dat is ook in de bovenstaande tabel te zien, meestal geldt: een hoge element interaction wijst op weinig inference calls (en andersom). Dat is niet verwonderlijk: juist in teksten met een hoge causale interactiviteit is het nuttig inferenties te expliciteren, bijvoorbeeld door connectieven te gebruiken, of de zin te herhalen in een andere formulering. Zo'n herhaling beïnvloedt tekstkenmerken als argumentoverlap of TTR. Dat betekent echter niet dat element interaction niet van inference calls niet te scheiden is; een tekst kan wel degelijk een lage element interaction bevatten, en tegelijkertijd weinig inference calls vereisen. Een aardrijkskundetext die de cultuur van een bepaald volk beschrijft zal bijvoorbeeld een lage element interaction hebben. Om zo'n tekst te begrijpen zijn niet per se veel inference calls nodig. Zie hierover ook hoofdstuk 8.

4. Hypothesen

De hypothesen die ik in dit onderzoek hanteer zijn vooral gebaseerd op het onderzoek van McNamara e.a. (2012). Zoals besproken in hoofdstuk 3.1 beschouw ik geschiedenisteksten in dit verband als narrative texts, en economieteksten als expository texts. McNamara e.a. (2012) vonden bij *science texts* een causale opbouw, tegenover een temporele opbouw in verhalende teksten (zie hoofdstuk 2.5). Ik verwacht dat het begrip causaliteit in dit verband sterk samenhangt met element interaction. Bovendien neem ik aan dat *science texts* (of: *expository texts*) coherentierelaties vaker markeren om het de lezer gemakkelijker te maken een tekstbetekenis- en daarmee situatiemodelrepresentatie te maken. De hypothesen zijn:

1. *Element interaction*

- a. De element interaction¹ is in economieteksten hoger dan in geschiedenisteksten;
- b. De element interaction is in schoolboeken voor lage(re) niveaus hoger dan voor hoge(re) niveaus.

2. *Inference calls*

- a. Economieteksten bevatten meer expliciete coherentierelaties, en vereisen dus minder wereldkennis bij het leggen van verbanden dan geschiedenisteksten
- b. Schoolboeken voor lagere niveaus bevatten meer gemarkeerde relationele en referentiële coherentierelaties, en vereisen dus minder wereldkennis bij het maken van inferenties dan schoolboeken voor hogere niveaus.

Vooraf de hypothesen 1b en 2b stel ik met enige voorzichtigheid. Ik verwacht echter dat samenstellers van schoolboeken die voor lage niveaus bedoeld zijn, hun teksten 'schematischer' opzetten. Teksten zijn zo weergegeven (vereenvoudigd) dat deze zo goed mogelijk 'in samenhang' te begrijpen zijn, en er dus zo min mogelijk 'gestampt' hoeft te worden. Hypothese 2b berust op het onderzoek van Graesser e.a. (2011), dat ik op pagina 11 besprak. Daaruit bleek teksten voor lagere niveaus een hogere referentiele coherentie hebben dan teksten voor hogere niveaus.

¹Om verwarring te voorkomen maak ik, net als in de voorgaande hoofdstukken, gebruik van de Engelse termen *element interaction* en *inference calls*.

5. Resultaten: Kwantitatieve analyse

5.1 Argumentoverlap

Tabel 2: Twee typen argumentoverlap; dichtheid per 1000 woorden.

Vak/ niveau	Met vorige zin M (SD)	Met voorgaande 50 woorden (gelemmatiseerd) M (SD)
ECONOMIE (n=50)	319,90 (30,68)	349,35 (276,93)
Vmbo (n=25)	334,12 (25,51)	485,91 (325,85)
Havo (n=25)	305,68 (29,15)	212,78 (107,39)
GESCHIEDENIS (n=61)	328,04 (29,89)	212,09 (123,20)
Vmbo (n=26)	348,70 (25,33)	241,38 (126,07)
Havo (n=35)	312,70 (23,15)	190,34 (118,13)
Totaal vmbo (n=51)	341,55 (26,22)	361,25 (272,32)
Totaal havo (n=60)	309,78 (25,82)	199,69 (113,39)

De argumentoverlap met de vorige zin is in vmbo-teksten hoger dan in havo-teksten ($F(1, 110) = 4,01, p < 0,001, \eta^2 = 0,29$). Verder is dit type argumentoverlap hoger in geschiedenis teksten dan in economie teksten ($F(1, 110) = 4,83, p = 0,03, \eta^2 = 0,04$).

Ook op basis van de voorgaande 50 woorden is de argumentoverlap hoger in vmbo-teksten ($F(1, 110) = 17,36, p < 0,001, \eta^2 = 0,14$). Anders dan argumentoverlap op basis van de vorige zin, is de argumentoverlap met de voorgaande 50 woorden hoger in economie teksten ($F(1, 110) = 14,03, p < 0,001, \eta^2 = 0,11$). Dit hoofdeffect bespreek ik verder niet: er is namelijk ook een interactie-effect voor vak en niveau ($F(1, 110) = 9,71, p = 0,002, \eta^2 = 0,08$). De argumentoverlap met de laatste 50 woorden is binnen havo-teksten voor beide vakken gelijk, maar bij vmbo-teksten niet; daar is de argumentoverlap binnen het vak economie veel groter. Het hoofdeffect van vak is dan ook enkel veroorzaakt door vmbo-teksten. Voor havo-teksten is er geen verschil tussen de vakken.

Tot slot lijkt het er op dat hoge lokale argumentoverlap op basis van de vorige zin, een teken is van *lage* argumentoverlap met de voorgaande 50 woorden. Er is inderdaad een negatieve correlatie tussen deze beide variabelen ($r = -0,21, p = 0,03$).

5.2 Connectieven

Tabel 3: frequentie van 4 typen connectieven + het totale aantal per 1000 woorden

Vak/ niveau	Temporele Connectieven <i>M (SD)</i>	Contrasterende Connectieven <i>M (SD)</i>	Comparatieve Connectieven <i>M (SD)</i>	Causale Connectieven <i>M (SD)</i>	Totaal <i>M (SD)</i>
ECONOMIE (<i>n</i> =50)	5,24 (5,43)	5,63 (4,94)	13,36 (7,17)	11,53 (6,64)	35,77 (13,11)
Vmbo (<i>n</i> =25)	6,23 (5,93)	5,91 (3,96)	14,17 (7,22)	10,73 (5,94)	37,03 (10,81)
Havo (<i>n</i> =25)	4,26 (4,80)	5,36 (5,82)	12,56 (7,19)	12,32 (7,31)	34,50 (15,18)
GESCHIEDENIS (<i>n</i> =61)	5,96 (5,07)	7,12 (3,71)	7,68 (5,19)	7,61 (5,16)	28,37 (8,84)
Vmbo (<i>n</i> =26)	4,79 (4,85)	6,76 (2,96)	8,10 (6,50)	6,32 (3,49)	25,96 (8,08)
Havo (<i>n</i> =35)	6,84 (5,12)	7,39 (4,20)	7,37 (4,03)	8,56 (5,98)	30,16 (9,06)
Totaal vmbo (<i>n</i> = 51)	5,49 (5,40)	6,34 (3,48)	11,07 (7,45)	8,48 (5,29)	31,39 (10,95)
Totaal havo (<i>n</i> =60)	5,76 (5,11)	6,54 (5,01)	9,53 (6,09)	10,13 (6,77)	31,97 (12,07)

Het totale aantal connectieven is in economieteksten hoger dan in geschiedenisteksten ($F(1, 110) = 13,55, p < 0,001, \eta^2 = 0,11$).

Causale connectieven komen in economieteksten meer voor dan in geschiedenisteksten ($F(1, 110) = 13,33, p < 0,001, \eta^2 = 0,11$). Dat geldt ook voor comparatieve connectieven ($F(1, 110) = 22,51, p < 0,001, \eta^2 = 0,17$).

Voor temporele connectieven is er een interactie-effect ($F(1, 110) = 4,10, p = 0,045, \eta^2 = 0,04$).

Binnen het vmbo is er qua temporele connectieven geen verschil, bij de havo bevatten geschiedenisteksten echter meer temporele connectieven dan economieteksten ($t(58) = 1,98, p = 0,046$).

5.3 Informatiedichtheid

Tabel 4: vier kenmerken m.b.t. informatiedichtheid

Vak/ niveau	Type-token-ratio <i>M (SD)</i>	Rarity index <i>M (SD)</i>	Inhoudswoorden <i>M (SD)</i>	NP-Dichtheid <i>M (SD)</i>
	Per tekst	Per tekst	Per 1000 Woorden	Per tekst
ECONOMIE (<i>n</i> =50)	0,493 (0,045)	0,904 (0,028)	370,39 (37,60)	0,304 (0,021)
Vmbo (<i>n</i> =25)	0,486 (0,037)	0,896 (0,026)	367,32 (33,44)	0,310 (0,022)
Havo (<i>n</i> =25)	0,501 (0,046)	0,912 (0,029)	373,46 (41,82)	0,299 (0,019)
GESCHIEDENIS (<i>n</i> =61)	0,549 (0,042)	0,926 (0,018)	344,44 (37,35)	0,315 (0,026)
Vmbo (<i>n</i> =26)	0,531 (0,036)	0,919 (0,018)	329,51 (35,62)	0,334 (0,021)
Havo (<i>n</i> =35)	0,562 (0,043)	0,931 (0,017)	355,53 (35,37)	0,300 (0,018)
Totaal vmbo (<i>n</i> =51)	0,509 (0,046)	0,908 (0,025)	348,04 (39,02)	0,322 (0,025)
Totaal havo (<i>n</i> =60)	0,536 (0,053)	0,923 (0,024)	362,00 (38,88)	0,300 (0,018)

De type-token-ratio (TTR) is in geschiedenis teksten hoger dan in economie teksten ($F(1, 110) = 43,17, p < 0,001, \eta^2 = 0,29$). Verder is de TTR in havo-teksten hoger dan in vmbo-teksten ($F(1, 110) = 7,96, p = 0,01, \eta^2 = 0,07$). Dat geldt ook voor de rarity index, zowel qua vak ($F(1, 110) = 23,73, p < 0,001, \eta^2 = 0,18$), als qua niveau ($F(1, 110) = 11,08, p < 0,001, \eta^2 = 0,09$).

Het aantal inhoudswoorden is binnen economie-teksten juist hoger dan binnen geschiedenis teksten ($F(1, 110) = 15,87, p < 0,001, \eta^2 = 0,13$). Verder bevatten havo-teksten meer inhoudswoorden dan vmbo-teksten ($F(1, 110) = 5,28, p = 0,02, \eta^2 = 0,05$).

Net als de TTR en de rarity index, is ook de NP-dichtheid is in geschiedenis teksten hoger dan in economie teksten ($F(1, 110) = 11,08, p < 0,001, \eta^2 = 0,09$). Verder is de NP-dichtheid in vmbo-teksten hoger dan in havo-teksten ($F(1, 110) = 34,35, p < 0,001, \eta^2 = 0,24$). Dit effect zal ik verder niet bespreken: er is namelijk ook een interactie-effect ($F(1, 110) = 10, p = 0,002, \eta^2 = 0,05$). Binnen de havo-teksten is er geen verschil qua NP-dichtheid; in het vmbo is dit verschil juist groot ($t(48) = 4,14, p < 0,001$). Het zo-even besproken hoofdeffect is dus enkel veroorzaakt door de vmbo-teksten.

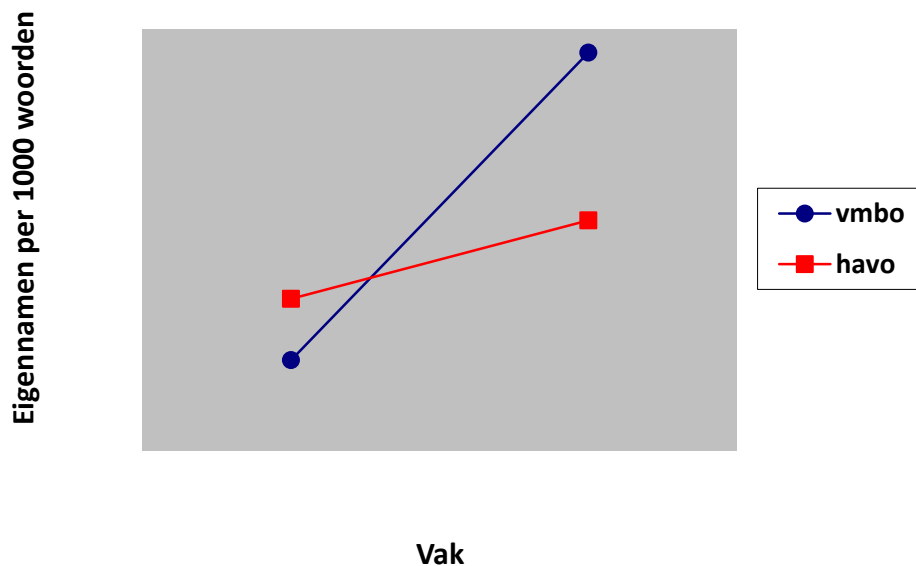
5.4 Eigennamen

Tabel 5: Aantal eigennamen per 1000 woorden; gesplitst op vak en niveau.

Niveau	Economie <i>M (SD)</i>	Geschiedenis <i>M (SD)</i>	Totaal
Vmbo (n=51)	17,16 (13,89)	72,57 (32,51)	45,41 (37,47)
Havo (n=60)	28,79 (25,25)	43,69 (29,81)	37,48 (28,74)
Totaal	22,98 (21,00)	56,01 (33,93)	40,47 (30,29)

Qua eigennamen is er zowel een hoofdeffect van vak, als een interactie-effect van de beide variabelen. Zoals verwacht bevatten geschiedenisteksten meer eigennamen dan economieteksten ($F(1, 110) = 46,99, p < 0,001, \eta^2 = 0,31$). Ten tweede het interactie-effect: bij economieteksten is het aantal eigennamen het grootst voor de havo, voor geschiedenisteksten is dat omgekeerd ($F(1, 110) = 15,60, p < 0,001, \eta^2 = 0,13$). Dat betekent echter niet dat het eerder genoemde hoofdeffect hiermee vervalt; zowel voor het vmbo ($t(34,11) = 7,86, p < 0,001$), als voor de havo ($t(58) = 2,03, p = 0,045$) is het aantal eigennamen het hoogst in geschiedenisteksten. Zie ook figuur 1 voor het interactie-effect:

Figuur 1: interactie-effect van vak/ niveau m.b.t. het aantal eigennamen



6. Resultaten: Kwalitatieve analyse

Ik bespreek de verschillen per kenmerk; tot slot zal de resultaten interpreteren in het licht van de hypothesen.

6.1 Argumentoverlap

Duidt een lage argumentoverlap inderdaad op weinig inference calls? Om die vraag te beantwoorden maak ik gebruik van twee geschiedenisteksten, één voor de havo (Feniks 5) en één voor het vmbo (Sfinx 3). De argumentoverlap op basis van de voorgaande zin is respectievelijk 256 en 404 (per 1000 woorden), en met de laatste 50 woorden 182 en 180 (gelemmatiseerd per 1000 woorden).

De argumentoverlap met de voorgaande zin lijkt in dit geval een redelijke aanduiding te geven van inference calls. Uit een kritische beschouwing van de tekst blijkt echter dat er niet één type 'inference call' is. Soms is er duidelijk sprake van een gedachtesprong tussen twee opeenvolgende zinnen. In andere gevallen is het lastig een specifieke plek in de tekst aan te wijzen, maar moet de lezer wel degelijk een aantal inferenties maken om de tekst *als geheel* te begrijpen. In het eerste geval spreek ik van *lokale* inference calls; in het tweede geval van *globale* inference calls. Ik zal het verschil tussen die twee typen hieronder uitleggen aan de hand van twee teksten. Ik begin met *lokale* inference calls.

De havo-tekst vereist meer lokale inference calls dan de vmbo-tekst. Dat blijkt onder andere uit het volgende fragment:

Fragment 1: Geschiedenis havo Feniks 5 r. 17-19 (zie bijlage 1).

De leider van de antirevolutionairen, Abraham Kuyper, greep naar een laatste redmiddel: een beroep op de Koning om de onderwijswet niet te ondertekenen.
Vier dagen nadat de Tweede Kamer de wet had aangenomen (...)

De stelling dat Kuyper 'naar een laatste redmiddel greep', namelijk 'een beroep op de koning,' betekent niet dat dit inderdaad de redding *was*. In de volgende zin is immers te lezen dat de wet *wel* aangenomen is. Kennelijk heeft de koning het verzoek van Kuyper niet gehonoreerd. De lezer moet dus op de hoogte zijn van het feit dat een wet pas geldig is wanneer de koning die ondertekend heeft. Ook moet de inferentie gemaakt worden dat Kuypers verzoek vergeefs was. Bij de argumentoverlap met de vorige zin zijn wel enige vraagtekens te plaatsen. Een gedachtesprong tussen twee zinnen (zoals in fragment 1) heeft zeker een lage argumentoverlap. Of de gemiddelde argumentoverlap met de vorige zin inderdaad een aanduiding geeft van

(lokale) inference calls van de tekst als geheel durf ik echter te betwijfelen. In de volgende paragrafen zal namelijk blijken dat dat meestal niet zo is. Ik richt mij daarom (tenzij anders aangegeven) enkel op globale inference calls. Het kenmerk argumentoverlap met de vorige zin wordt in de paragrafen 6.3.1, 6.3.3 & 6.4 verder besproken.

De havo-tekst vereist echter ook een aantal *globale* inference calls:

Fragment 2: Geschiedenis havo Feniks 5 r. 1-4 (zie bijlage 1).

Net zoals vrouwen, werden ook de katholieken in Nederland lange tijd achtergesteld. Pas in de herziene grondwet van 1848 werd 'vrijheid van godsdienst' als een belangrijk grondrecht opgenomen. Hiermee werd de weg vrijgemaakt voor emancipatie van de katholieken. De katholieken (en later ook protestanten) waren het niet eens met de manier waarop de liberalen steeds verder gingen met het loskoppelen van kerk en staat.

De lezer zou zich af kunnen vragen waarom het grondrecht 'vrijheid van godsdienst' alleen voor katholieken een emancipatie betekende, en niet voor protestanten. Die zijn toch ook 'godsdienstig'? En waarom waren de protestanten het pas *later* oneens met de manier waarop de liberalen kerk en staat loskoppelden? Om dat te begrijpen moet de lezer op de hoogte zijn van het feit dat Nederland tot 1848 een protestantse staatskerk had: katholieken waren daardoor sterk achtergesteld. Protestanten hadden de nieuwe wet dus niet nodig, zij waren immers al vrij om hun geloof te belijden. In 1848 werden kerk en staat grotendeels gescheiden. Katholieken en protestanten protesteerden dus niet om dezelfde reden 'tegen het loskoppelen van kerk en staat'. Ik denk wel dat het belangrijk is dat de lezer die inferentie maakt: het bijzondere van de Schoolstrijd was nu juist het feit dat het een strijd was van gelovigen tegen 'niet-gelovigen', in plaats van gelovigen tegen gelovigen. Een belangrijk verschil met de 'geloofsstrijd' tijdens de 80-jarige opstand, of de terechtstelling van de 'Martelaren van Gorcum' (Van der Horst, 2000, pp. 232-235 & pp. 301-308).

De vmbo-tekst lijkt geen relevante inference calls op lokaal niveau te bevatten. Dat wil zeggen; twee opeenvolgende zinnen die enkel door het maken van een inferentie (in samenhang) begrepen kunnen worden. Op globaal niveau (de tekst als geheel) vereist de tekst echter wel degelijk een aantal inferenties van de lezer. Zo heeft de vmbo-tekst een zeer lage inhoudelijke congruentie: in 301 woorden worden zowel de Romeinse limes, als de volksverhuizingen, kerstening van de Lage Landen, het ontstaan van het feodale systeem onder Karel de Grote en de standenvorming besproken (zie ook tabel 6). De verbanden tussen deze onderwerpen zijn dus niet alleen inhoudelijk, maar ook chronologisch zeer zwak. Dat wordt duidelijk wanneer de eerste en de laatste zin tegenover elkaar worden gesteld:

Fragment 3: Geschiedenis vmbo Sfinx 3, r.1 & r.19- 21 (zie bijlage 1).

Eerste zin (r.1): De Romeinen kozen de Rijn als natuurlijke grens.

(...)

Laatste zinnen (r.19-21): Het christendom was de enige godsdienst. De geestelijke, alle mensen die voor de kerk werkten, kregen steeds meer grond. De kerk werd hierdoor nog belangrijker en rijker. Zo ontstonden in de Middeleeuwen drie soorten beroepen. Je zegt ook wel drie standen: geestelijkheid, adel en boeren.

De gebeurtenis uit de eerste zin kan worden gedateerd rond 50 n.C., terwijl de gebeurtenis uit de laatste zin zich zo rond 1100 n.C. afspeelt. In 301 woorden wordt dus ruim 1000 jaar geschiedenis besproken. Hierdoor is het lastig te bepalen wat nu precies het *punt* van de tekst is. Waarom is de aanwezigheid -en het vertrek- van de Romeinen relevant bij het ontstaan van het feodale stelsel? En wat hebben Bonifatius en Willibrord (regel 9) te maken met de uiteindelijke kerstening van de Lage Landen (regel 19)? Verderop in de tekst (regel 12) wordt vermeld dat Karel de Grote wilde dat 'iedereen christelijk werd'. Waarom is dat, net als een nieuw muntstelsel en infrastructuur, een belangrijke verandering geweest? Deze verbanden moet de lezer zelf leggen. De lezer moet weten dat het feit dat het Karel de Grote lukte om zijn 'onderdanen' te bekeren, belangrijk was bij het ontstaan van eenheid in de Lage Landen. Van een sterk versnipperd gebied met veel verschillende stammen werd het min of meer een eenheid (in ieder geval religieus). Daardoor kon de uiteindelijke standensamenleving ontstaan. Omdat Willibrord en Bonifatius niet konden profiteren van deze eenheid, was het voor hen moeilijker de stammen te bekeren. Zonder deze verbanden te zien, is het lastig te tekst 'in zijn geheel' te begrijpen.

Qua inference calls is er dus weinig verschil. Dat geldt ook voor de element interaction: die is in beide teksten zwak. Zoals ook zal blijken uit andere teksten (zie bijvoorbeeld hoofdstuk 6.3.1) is de zwakke interactie vooral te wijten aan het grote aantal besproken onderwerpen/ thema's die slechts indirect met elkaar samenhangen. De Schoolstrijd, secularisering, staatskerk en onderwijswet krijgen elk (qua ruimte) evenveel aandacht in de havo-tekst (zie tabel 6 & bijlage 1). Het is echter, zeker wanneer je weinig voorkennis hebt, lastig de (complexe) verbanden tussen deze onderwerpen te begrijpen. In de onderstaande tabel is dat inzichtelijk gemaakt. In de eerste en derde kolom staan de besproken onderwerpen, in de tweede en vierde de woorden die eraan besteed worden:

Tabel 6: besproken onderwerpen voor een havo- en een vmbo-geschiedenistekst

Vmbo Sfinx 3		Havo Feniks 5	
Onderwerp	Woorden	Onderwerp	Woorden
Romeinse Limes	1-8	Achterstelling Katholieken op basis van godsdienst	1-12
Romeinse Forten	9-21	Gelijkheid door de Grondwet op basis van godsdienst	13-38
Friese terpen	22-45	Twisten tussen Katholieken en Liberalen	39-64
Romeinse stedenvorming	46-76	Twisten over onderwijs	65-106
Vertrek Romeinen	77-93	Nieuwe ideeën over onderwijs	107-130
Verdeling Lage Landen	94-118	Invloed van geestelijken op onderwijs	131-152
Goden Lage Landen	119-127	Samenwerking Katholieken en Protestanten	153-193
(Poging tot) kerstening Lage Landen	128-140	Tegenwerken van de onderwijswet van 1878	194-223
Karel de Grote	141-150	Redmiddel van de antirevolutionairen	223-246
Vernieuwingen Karel de Grote	151-179	Handtekeningenactie	247-315
Oorlogen Karel de Grote	180-191	Uitleg v.h. begrip Schoolstrijd	316-321
Ontstaan feodale systeem	192-224	Oplossing in 1917. Gelijkheid van christenen op basis van onderwijs	322-328
Definitie feodaal systeem	224-256		
Christendom en Kerk	257-283		
Ontstaan standensamenleving	284-301		

Uit de opeenvolging van onderwerpen –en de daaraan bestede ruimte - is te zien dat beide teksten een lage interactie hebben. Het verband tussen ‘Romeinse forten’ en ‘Friese terpen’ is bijvoorbeeld enkel temporeel: het verdwijnen van de Romeinen heeft met de Friezen weinig te maken. Evenzo heeft het ontstaan van het feodale systeem weinig te maken met de Romeinse stedenvorming of de pogingen de Friezen te bekeren. De (causale) samenhang tussen de onderwerpen is mede daardoor laag.

Beide typen argumentoverlap lijken in dit geval niet geschikt om element interaction boven water te krijgen. De argumentoverlap met de voorgaande 50 woorden verschilt niet, en zou dus een aanduiding kunnen zijn voor de element interaction, die immers *ook* niet verschilt. In de volgende onderdelen zal de validiteit van het kenmerk argumentoverlap met de voorgaande 50 woorden nog verder worden getoetst. Meer over dit kenmerk is te vinden in de paragrafen 6.3 en 6.4.

6.2 Connectieven

6.2.1 CAUSALE CONNECTIEVEN Op grond van het gebruik van causale connectieven (verbindingswoorden) is de hypothese bevestigd: economieteksten bevatten inderdaad meer connectieven van dit type (zie hoofdstuk 5.2). Maar in hoeverre zijn causale connectieven representatief voor element interaction? In bijlage 2 heb ik twee teksten opgenomen, één economie- en één geschiedenis tekst. De geschiedenis tekst (vmbo Pharos 1) bevat 3 causale connectieven, de economie tekst (vmbo Economisch Bekeken 4) bevat er 7. De connectieven lijken in dit geval inderdaad een signaal te zijn voor een hoge element interaction. Dat blijkt onder andere uit deze passage uit de economie tekst (r. 13-15):

Fragment 4: Economie vmbo Economisch bekeken 4, r. 14-17 (zie bijlage 2).

Inkomsten kunnen flink omlaag gaan, bijvoorbeeld DOORDAT je minder gaat werken of door ontslag. Bij een daling van de inkomsten is budgetteren lastig, WANT er moet bezuinigd worden op uitgaven waaraan je bent gewend. Bij een stijging van de inkomsten is budgetteren plezierig, WANT er is geld beschikbaar voor nieuwe uitgaven.

De geschiedenis tekst bevat volgens T-Scan minder causale verbindingswoorden. De element interaction lijkt in deze tekst inderdaad lager te zijn. Deze passage (r. 9-11) geeft daar een goed voorbeeld van:

Fragment 5: Geschiedenis vmbo Pharos 1, r. 9-11 (zie bijlage 2).

Alleen vrije, volwassen Atheense mannen mochten over de politiek beslissen. Dit waren de mannen die in het Atheense leger vochten. Zij betaalden zelf hun wapens. Een soldaat had veel nodig. Hij droeg een bronzen helm, een borstharnas, bronzen scheenbeschermers en een groot rond schild.

Er is hier enkel sprake van opsommende of verklarende betekenisrelaties. In het volgende onderdeel laat ik echter zien dat de aan/afwezigheid van causale connectieven niet per se een betrouwbare indicator is van element interaction.

Hoewel de geschiedenis tekst volgens T-Scan minder causale verbindingswoorden bevat zijn er wel degelijk een aantal causale relaties aan te wijzen. Eén daarvan is gemarkeerd met een causaal verbindingswoord (zie bijlage 2, r. 10):

Fragment 6: Geschiedenis vmbo Pharos 1, r. 11 (zie bijlage 2).

Soldaten waren heel belangrijk, DAAROM mochten zij stemmen.

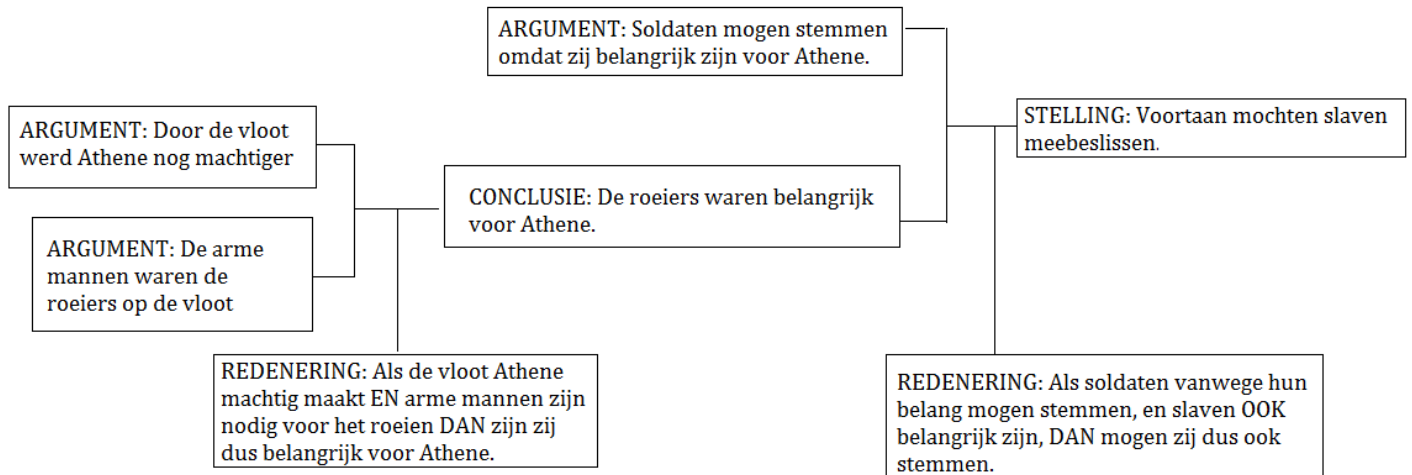
In regel 13-15 komt weer een causaal opgebouwde passage voor, dit keer zonder causale connectieven.

Fragment 7: Geschiedenis vmbo Pharos 1 , r. 14-16 (zie bijlage 2).

Door de vloot werd Athene nog machtiger. De arme mannen van Athene werden de roeiers op de oorlogsschepen. De roeiers waren ook belangrijk voor Athene. Voortaan mochten zij ook meebeslissen.

De interactiviteit tussen de betekenselementen is hier hoog, mede door het verband met regel 11 (fragment 6), dat de argumentatie versterkt. Dit schema (naar Verhoeven & Schellens, 2008) laat dat duidelijk zien:

Figuur 2: argumentatieschema geschiedenis vmbo Pharos 1 (zie bijlage 2)



Er zijn echter ook individuele teksten te vinden die de hypothese tegenspreken; een economietekst (havo In Balans II, tekst 2) zonder connectieven, tegenover een geschiedenistekst (havo Geschiedeniswerkplaats, tekst 2) met veel causale, contrastieve en vergelijkende connectieven. Deze teksten illustreren wederom het probleem dat ik hierboven besprak; de afwezigheid van causale connectieven betekent niet altijd dat de tekst niet causaal is opgebouwd.

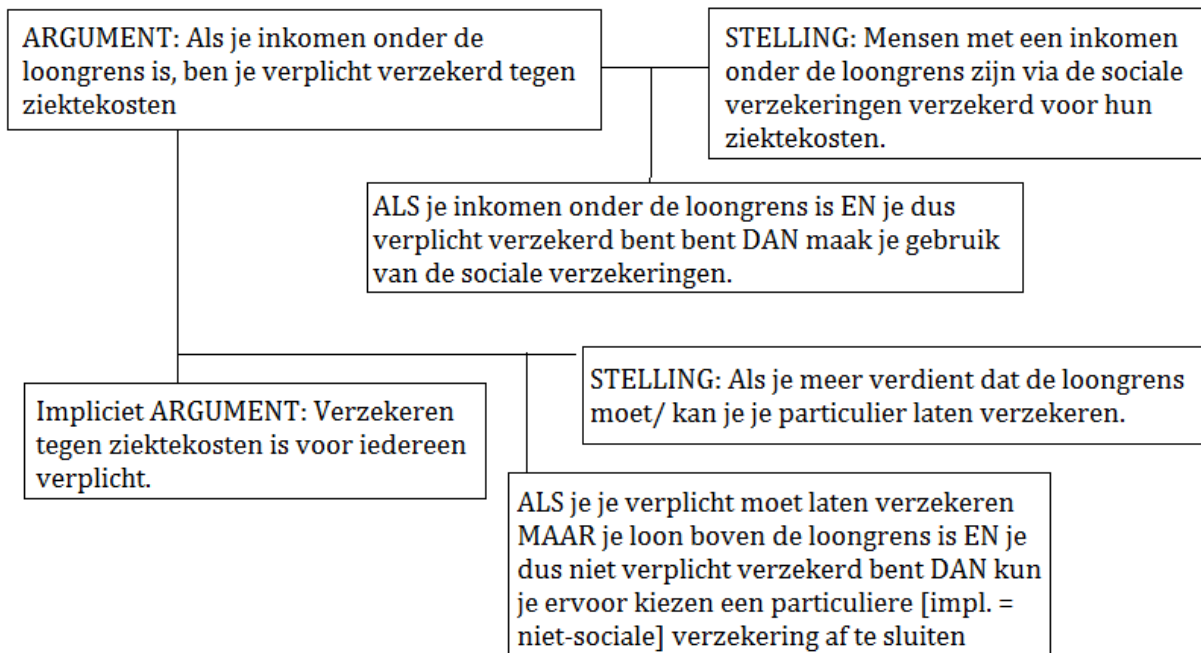
De economietekst is namelijk wel degelijk causaal opgebouwd (zie fragment 8 en figuur 3), terwijl de geschiedenistekst ondanks het gebruik van causale connectieven toch minder interactief lijkt te zijn (zie fragment 9).

Fragment 8: Economie vmbo havo In Balans II, 2 , r. 1-3 (zie bijlage 3).

Mensen met een inkomen onder de loongrens voor de ziekenfondsverzekering (...) zijn verplicht verzekerd tegen ziektekosten op grond van de ziekenfondswet (ZFW). Ze zijn DUS via de sociale verzekeringen tegen ziektekosten verzekerd. Werknemers met een inkomen boven de loongrens (...) moeten/kunnen zich particulier verzekeren tegen ziektekosten. (...) In dit geval is de sociale zekerheid DUS via de particuliere markt georganiseerd.

Deze inhoud van dit economiefragment vertoont een duidelijke interactie:

Figuur 3: argumentatieschema Economie havo In Balans II, 2 (zie bijlage 3)



In deze geschiedenis tekst heeft T-Scan causale verbindingswoorden gevonden: zoals gezegd zegt dat niet veel over de aanwezige element interaction:

Fragment 9: Geschiedenis havo Geschiedeniswerkplaats 1, r. 7-9 (zie bijlage 3).

Hitler bracht met zijn toespraken niet alleen het publiek dat op hem was afgestroomd in vervoering, **MAAR** bespeelde het hele volk DOORDAT hij overal voortdurend op de radio te horen en in de bioscoop was te zien.

Het verbindingswoord (doordat) voegt niet veel aan de causaliteit toe. Het connectief wordt vooral gebruikt als bepaling; het specificiert het gezegde uit de vorige zin (*bespelen; parafrase: 'Zo/ op die manier bespeelde Hitler...'*). Hoewel 'doordat' meestal een causale relatie aanduidt, en dus terecht als causaal connectief wordt herkend door T-Scan, wordt het in sommige contexten als middelaanduiding gebruikt, zoals in het bovenstaande fragment. Er is hier dus sprake van *probleemtype 1* (zie hoofdstuk 2.1). Daarbij moet wel worden opgemerkt dat een middel-doelrelatie wel degelijk een sterk 'causaal karakter' heeft.

Ook worden sommige woorden ten onrechte *niet* als connectief aangemerkt. In de economie tekst (havo In Balans II 2, bijlage 3) worden de woorden 'des te' bijvoorbeeld niet als connectief aangemerkt. Toch is het gebruik van 'des te' vaak een aanduiding van een vergelijkende relatie. Dat deze woorden niet herkend worden is niet aan T-Scan te wijten; de

woorden komen simpelweg niet voor in de lijst met connectieven (Kraf, 2013, pp. 52-53). Mogelijk kan 'des te' daaraan toegevoegd worden. Het gebruik van 'des te' is immers vergelijkbaar met die van het woord 'naarmate' dat *wel* als vergelijkend connectief in de connectievenlijst is opgenomen.

Conclusie: Causale verbindingswoorden lijken inderdaad een aanduiding te zijn voor een hoge element interaction. Aan de andere kant hoeft het uitblijven van causale connectieven niet te betekenen dat er een lage element interaction is. Hoewel de aanwezigheid van causale connectieven dus geen waterdicht middel is om de element interaction te voorspellen, is de hypothese op grond van deze kwalitatieve analyse wel bevestigd: economieteksten hebben inderdaad een hogere element interaction dan geschiedenisteksten.

6.2.2 ALLE CONNECTIEVEN In de tabel met tekstkenmerken stelde ik dat een grote hoeveelheid connectieven –ongeacht het type- zou kunnen duiden op weinig inference calls. De vmbo-teksten uit bijlage 2, die ik ook gebruikt heb om het gebruik van causale connectieven te bespreken, kan ik nu weer gebruiken. Ook op grond van het totaal aantal connectieven vertonen deze teksten een groot verschil. De geschiedenistekst (vmbo Pharos 1) bevat een gemiddelde van 46 connectieven per 1000 woorden, de economietekst (vmbo Economisch bekeken 4) heeft er maar liefst 83. Dat de teksten niet sterk verschillen qua element interaction heb ik laten zien in paragraaf 6.2.1. Nu is het dus de vraag in hoeverre de teksten inference calls vereisen. Moeten de inferenties in de geschiedenistekst vaker worden gemaakt met behulp van wereldkennis? In dit geval denk ik van niet, hoewel er weinig connectieven in de geschiedenistekst voorkomen. Qua inferenties is de economietekst wel problematisch, wat gedeeltelijk veroorzaakt wordt door de dubbelzinnigheid van de woorden 'budget' en 'budgetteren'. Budgetteren wordt als synoniem gebruikt van 'bezuinigen'. Wanneer je op bepaalde 'budgetten' 'budgetteert' – m.a.w.: op bepaalde uitgaven bezuinigt- houd je meer geld over voor andere 'budgetten'.

Fragment 10: Vmbo economie Economisch bekeken 4, r. 7-8 (zie bijlage 2).

ALS je een budget wilt verhogen, moet je budgetteren. WANT je houdt dan minder geld over voor de andere uitgaven.

Dit fragment bevat weliswaar veel connectieven, maar is daardoor zeker niet duidelijker. De leerling moet wel begrijpen dat het vierde woord (budget) en het laatste woord (andere uitgaven) eigenlijk als synoniem gebruikt worden. Als je EEN budget wilt verhogen, moet je bezuinigen op andere uitgaven (ANDERE budgetten). Daarvoor houd je dus minder geld over. Het woord 'als' herkent T-Scan als vergelijkend (comparatief) verbindingswoord. Dat lijkt mij – zeker in dit voorbeeld – onterecht. Het woord kan zonder problemen vervangen worden door

het connectief 'wanneer', dat T-Scan herkent als causaal connectief. In dit geval is er inderdaad sprake van een causale betekenisrelatie; het budgetteren wordt veroorzaakt door de wil een budget te verhogen. Dat T-Scan *als* herkent als vergelijkend verbindingswoord lijkt mij in dit geval onterecht. Het verbindingswoord 'als' kan zowel als causaal, als vergelijkend voegwoord gebruikt worden, waarbij er niet sprake is van één dominante gebruikswijze; er is in dit voorbeeld dus sprake van *probleemtype 2* (zie hoofdstuk 2.1).

De geschiedenis tekst bevat zeker niet meer inferenties dan de economie tekst; de totale hoeveelheid connectieven lijkt mij dus geen goede aanduiding voor het aantal inference calls (zie ook 6.2.1).

6.3 informatiedichtheid

Zoals ik in hoofdstuk 5,3 heb laten zien, zijn er inderdaad verschillen gevonden in tekstkenmerken die een aanduiding zouden moeten zijn voor de informatiedichtheid van de tekst. In deze paragraaf bespreek ik deze tekstkenmerken.

6.3.1 TTR & Rarity Index Qua Type-Token-Ratio en rarity index heb ik twee teksten genomen die het hoofdeffect van vak bevestigen: een havo-economietekst (Praktische economie 4, TTR= 0,40, RI= 0,85) en een havo-geschiedenistekst (Pharos 4, TTR= 0,66, RI= 0,96). De TTR deelt het aantal unieke woorden door het totaal aantal woorden. De economietekst bevat 301 woorden; het aantal unieke woorden is dus $(301 * 0,4) = 120$. In de geschiedenistekst is dat $(304 * 0,66) = 201$. Er zijn dus $(304 - 201) = 103$ woorden die *niet* uniek zijn, en dus vaker dan één keer in de tekst voorkomen. De rarity index deelt het aantal lemma's die minder dan 4 keer voorkomen door het totaal aantal woorden. In de economietekst komen dus $(301 * 0,85) = 256$ woorden minder dan 4 keer voor. Voor de geschiedenistekst zijn dat er $(304 * 0,96) = 292$.

Tabel 7: betekenisdichtheid in vergelijking met zins- en tekstlengte (zie bijlage 4)

	Havo Pr. Ec. 4	Havo Phar. 4
Aantal woorden	301	304
Aantal zinnen	22	15
Aantal betekenseenheden	30	34
Aantal betekenseenheden per 1000 woorden	100	112
Gemiddeld aantal betekenseenheden per zin	1,36	2,27

Zijn deze waarden inderdaad een aanduiding voor het aantal inference calls? Bij deze teksten lijkt dat verschil inderdaad relevant te zijn. De economietekst begint met een lange inleiding om het begrip informatie-asymmetrie bij het afsluiten van een verzekering te verklaren. Als voorbeeld wordt het afsluiten van een fietsverzekering genomen. Er wordt uitgelegd dat een verzekeraar op de hoogte is van het gemiddeld aantal keren dat hij uit moet keren, gebaseerd op ervaring (het aantal historische fietsdiefstallen). Daartegenover wordt de individuele verkeringsnemer gesteld die van zichzelf weet of zijn kans groter of kleiner is dan dat gemiddelde, afhankelijk van de manier waarop hij met zijn fiets omgaat. Om deze tekst te begrijpen moet de lezer één inferentie zelf leggen: het gemiddelde is belangrijk voor de verzekeraar, omdat hier de premie op gebaseerd is die hij zijn klanten vraagt (er moet immers aan verdiend worden). Hoe meer fietsen er echter gestolen worden, hoe kleiner de winst van de verzekeraar is, dus is het niet in zijn belang klanten te werven die slordig met hun verzekerde product omgaan. De conclusie is dat de verzekeringsnemer dat van zichzelf weet, maar de verzekeraar niet: de informatie-asymmetrie.

De geschiedenisstekst –die een hogere TTR en RI heeft- lijkt inderdaad meer inferenties van de lezer te vergen dan de economietekst. De geschiedenisstekst gaat over de gevaren aan boord van een 17^e-eeuws zeilschip. In zeer korte tijd worden 4 verschillende gevaren genoemd; bedorven voedsel (geen mogelijkheid tot conservatie), kou/ hitte, gebrek aan privacy, scheurbuik en slechte medische zorg. Dat deze onderwerpen met elkaar samenhangen blijft zeer impliciet. Het volgende fragment geeft daarvan een goed beeld:

Fragment 11: Geschiedenis havo Pharos 4, r. 7-10 (zie bijlage 4).

In een hoekje deelden de op elkaar gepropte zeevarenden de ruimte met ratten, wormen, spinnen en kakkerlakken. Vooral 's nachts kwamen die op het rottend voedsel af dat, hoewel het verboden was, door een kieskeurige matroos was weggegooid. De meest gevreesde ziekte die optrad door het slechte voedsel was scheurbuik.

Ten eerste is het op basis van het voorgaande stilistisch niet duidelijk dat de voorlaatste zin een *voorbeeld* is, en geen stelling. De inferentie dat ratten 's nachts op slecht opgeborgen voedsel afkomen –*bijvoorbeeld* voedsel dat door een matroos stiekem is weggegooid- moet de leerling zelf leggen. Dat het voedsel dat ratten aantrok 'rottend' was (regel 8), was niet de oorzaak van het probleem, maar het feit dat het voedsel niet goed opgeborgen werd. Kortom; er wordt ten onrechte gesuggereerd dat het voedselprobleem (m.b.t. conservering, bijlage 4, regels 2-4) samenhangt met het rattenprobleem. De laatste zin moet begrepen worden in dit licht: de meest gevreesde ziekte was scheurbuik, veroorzaakt door het slechte voedsel. Of het *eten* van slecht voedsel de enige oorzaak is, of dat de aanwezigheid van ratten (indirect ook veroorzaakt door het al dan niet weggegooid voedsel) dat ook was, blijft buiten beschouwing. Na dit fragment volgt een authentiek ooggetuigenverslag uit 1699, met behoud van archaische spelling en woordkeuze. Na dit ooggetuigenverslag eindigt de tekst als volgt:

Fragment 12: Geschiedenis havo Pharos 4, r. 22-24 (zie bijlage 4).

Dat amputatiezagen tot de standaarduitrusting van de scheepschirurgijn behoorden, zegt veel over de mogelijkheden (of beter gezegd: onmogelijkheden) van de medici aan boord van schepen.

Zelfs wanneer de lezer het ooggetuigenverslag heeft begrepen, is er wereldkennis nodig om te begrijpen *wat* de oorzaak van scheurbuik is. Die oorzaak lag weliswaar bij het voedsel, maar niet bij het eerder genoemde 'conserveringsprobleem', maar bij de *voedselkeuze*; dat bevatte immers niet genoeg vitamine C. De tekst van de scheepschirurgijn is dus niet bedoeld als verduidelijking van het begrip scheurbuik, maar eerder als illustratie, om de onwetendheid van medici te onderstrepen. Inhoudelijk is er immers wel het een en ander aan te merken op het ooggetuigenverslag, zoals de stelling dat brandewijn zou helpen tegen scheurbuik (bijlage 4, regel 21). De laatste, hierboven geciteerde zin maakt het nog lastiger. Want waarom zou de simpele aanwezigheid van een amputatiezaag iets zeggen over de onmogelijkheden van medici op zeilschepen? Tegenwoordig zullen scheepsartsen wellicht nog steeds zo'n zaag bij zich

hebben; maar dat zegt niets over de mogelijkheden die zij hebben. De leerling moet inzien dat amputeren niet helpt tegen een ziekte als scheurbuik *maar* dat de verwijzing naar de zaag gebruikt wordt om de onmacht van artsen aan te geven.

Deze twee teksten onderstrepen de aanname die ik deed op basis van TTR en RI: teksten met een hoge TTR en/of RI vereisen veel inferenties. Dat de geschiedenis tekst inderdaad meer inferenties vereist, wordt door sommige andere tekstkenmerken –die ook op inference calls duiden- bevestigd:

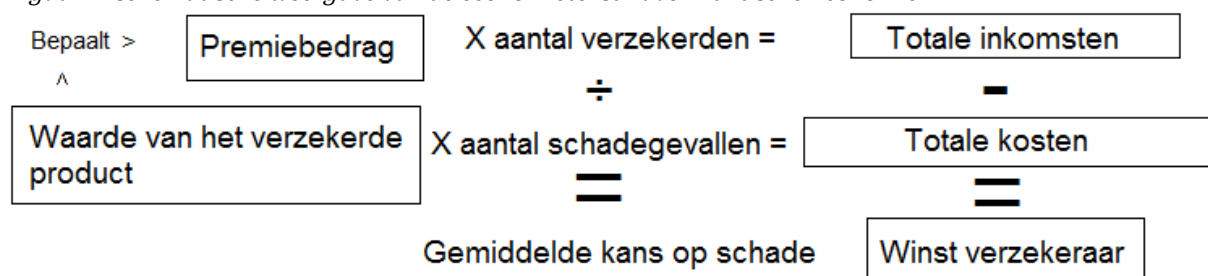
Tabel 8: De TTR/ RI in vergelijking met 4 andere tekstkenmerken m.b.t. inference calls

Vak	TTR	RI	Argumentoverlap met de vorige zin	Argumentoverlap (vrgd. 50 wrdn)	Connectieven	Eigennamen
	Per tekst	Per tekst	Per 1000 wrdn.	Per 1000 wrdn.	Per 1000 wrdn.	Per 1000 wrdn.
Havo Econ. Prakt. Econ. 4	0,40	0,85	292	336	66	10
Havo Gesch. Pharos 4	0,66	0,96	314	59	59	10

Het verschil in argumentoverlap wijst in dezelfde richting (meer inference calls in de geschiedenis tekst). Het verschil in connectieven is klein, en qua argumentoverlap met de vorige zin en eigennamen is er geen verschil. Argumentoverlap met de vorige zin is in dit geval geen geschikt middel om inference calls boven water te krijgen. Ik heb in ieder geval geen lokale inference calls kunnen ontdekken die een probleem zouden kunnen vormen. Het verschil in inference calls – dat ik hier aan de hand van de TTR en rarity index besprak – is vooral globaal.

Een tweede veronderstelling die ik deed was dat een *lage* TTR & RI duidt op een hoge element interaction. Bij het kenmerk ‘causale connectieven’ heb ik laten zien op welke manier ik de element interaction in teksten bepaal; ik zal er daarom nu iets minder uitvoerig op ingaan. De aanname lijkt me echter juist; de economie tekst (met een lage TTR & RI) heeft inderdaad een zeer hoge interactie.

Figuur 4: schematische weergave van de economie tekst havo Praktische Economie 4



In dit schema is te zien dat de 'hoofdlijn' van de tekst geheel in een causaal schema te vangen is.

Met de geschiedenis tekst is dat een stuk lastiger, hoewel er wel degelijk een samenhang aan te wijzen is. Ten eerste een thematische samenhang (gevaren op zee). Ten tweede is er sprake van een causale relatie:

1. Slecht eten >gevolg> scheurbuik
2. Ratten + slecht eten >gevolg> ziekte (o.a. scheurbuik)
3. Scheurbuik + slechte medische hulp >gevolg> dood (conclusie: slecht voedsel weinig privacy, ongedierte, geen conserveringsmiddelen, geen beschutting tegen weersomstandigheden en slechte medische hulp vormen gevaren op zee).

Het is dus niet gezegd dat deze geschiedenis tekst zich per uitstek onderscheidt door een lage inhoudelijke samenhang (lage element interaction). Net als in de teksten in hoofdstuk 6,1 is er voornamelijk sprake van lokale samenhang, wat vooral te wijten is aan het grote aantal besproken onderwerpen/ thema's. De aanwezige samenhang is dus van een lokaal niveau, terwijl de economie tekst in zijn geheel samenhangt (zie figuur 4). Dat die lokale samenhang voor bepaalde misschien lezers niet meteen duidelijk is, is een ander verhaal: dat heb ik besproken bij het eerste tekstkenmerk (inference calls, par. 6.3.1).

6.3.2 NP-DICHTHEID Het tweede kenmerk dat samenhangt met informatiedichtheid, en dus een aanduiding zou moeten zijn voor element interaction en inference calls is de NP-dichtheid. Ook voor dit kenmerk scoorde het vak geschiedenis hoger. Voor alle duidelijkheid; een hoge NP-dichtheid zou een aanduiding moeten zijn voor een hoge informatiedichtheid, en daarmee voor een *zwakke* element interaction en *veel* inference calls.

Om dit te illustreren leg ik wederom twee teksten naast elkaar. Beide teksten heb ik echter al eerder gebruikt bij de bespreking van de argumentoverlap: vmbo Sfinx 3 en havo Feniks 5. De eerste tekst heeft een zeer hoge NP-dichtheid (0,38), de tweede juist een zeer lage (0,29). Bij de bespreking van het kenmerk argumentoverlap liet ik al zien dat er –zeker op globaal niveau– geen verschil is tussen deze teksten. De NP-dichtheid is in dit geval dus niet geschikt inference calls te traceren.

Een tweede aanname was dat een hoge NP-dichtheid duidt op een lage element interaction. Zoals besproken bij het kenmerk argumentoverlap, verschillen de bovenstaande teksten (vmbo Sfinx 3 & havo Feniks 5, zie bijlage 5) niet van elkaar qua element interaction. In tabel 9 staan

vier eerder besproken teksten die *wel* sterk van elkaar verschillen op grond van element interaction. Die verschillen worden bevestigd door het kenmerk argumentoverlap met voorgaande 50 woorden, maar er is *geen* verschil in de NP-dichtheid:

Tabel 9: De NP-dichtheid in vergelijking met een ander kenmerk m.b.t. element interaction: argumentoverlap

Vak	Element-Interaction	Besproken bij tekstkenmerk	Bijlage	Argumentoverl. Voorg. 50 woorden	NP-Dichtheid
				Per 1000 wrdn	Per zin
Havo econ. In Balans II 2	Hoog	Causale connectieven	3	131	0,308
Havo gesch.-werkpl. 1	Laag	Causale connectieven	3	93	0,323
Havo econ. Prakt. Ec. 4	Hoog	TTR/ RI	4	336	0,316
Havo gesch. Phar. 4	Laag	TTR/ RI	4	59	0,297

De invloed van zinslengte is één van de valkuilen van dit tekstkenmerk; korte zinnen zorgen namelijk al snel voor een hoge NP-dichtheid. Er is inderdaad een sterke negatieve correlatie tussen zinslengte en NP-dichtheid ($r = -0,56$, $p < 0,001$). De volgende fragmenten geven daarvan een goed beeld:

Fragment 13: vmbo Sfinx 3, regel. 1, (zie bijlage 1)

Ze bouwden **forten** langs **de rivier**.

Fragment 14: havo Feniks 5, regel 4-5, (zie bijlage 1)

De katholieken (en later ook **protestanten**) waren het niet eens met **de manier** waarop **de liberalen** steeds verder gingen met **het loskoppelen van kerk en staat**.

Het eerste fragment (fragment 13) is afkomstig uit de eerder besproken vmbo-tekst (Sfinx 3, bijlage 1). In dit fragment komen 3 NP's voor: 'ze', 'forten' en 'de rivier'. De zin bestaat uit 6 woorden, de NP-dichtheid van dit fragment is dus $(3/6) = 0,5$. Het tweede fragment (fragment 14) is afkomstig uit de havo-tekst (Feniks 5, regel 4-5, bijlage 1). Deze zin bevat 8 NP's: 'de katholieken', 'protestanten', 'de manier', 'de liberalen', 'het loskoppelen van kerk en staat', 'kerk' en 'staat'. De zin bevat 26 woorden, dus is de NP-dichtheid $(8/26) = 0,31$. De grote valkuil van deze maat is dus dat niet alleen inhoudelijke, maar ook stilistische kenmerken een rol spelen. Een auteur die kiest voor lange, samengestelde zinnen, bereikt een lagere NP-dichtheid dan een auteur die gebruikt maakt van korte, enkelvoudige zinnen. Ook wanneer zij hetzelfde beschrijven verschilt de NP-dichtheid. Qua informatiedichtheid is er dan echter geen verschil.

Mogelijk is de tekst van de eerste auteur door de lange zinnen juist moeilijker te begrijpen, hoewel de NP-dichtheid lager is.

6.3.3 INHOUDSWOORDEN

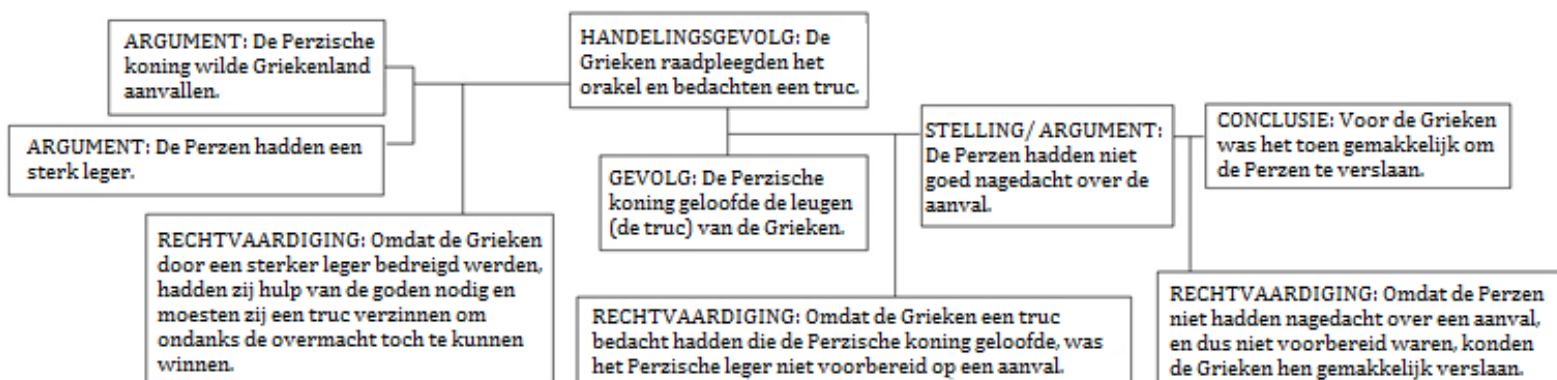
Tabel 10: betekenisdichtheid in vergelijking met . zins- en tekstlengte (zie bijlage 5)

	Havo Sprekend Verl. 2	Vmbo Sfinx 2
Aantal woorden	400	342
Aantal zinnen	31	43
Aantal betekenseenheden	56	45
Aantal betekenseenheden per 1000 woorden	140	132
Gemiddeld aantal betekenseenheden per zin	1,81	1,05

Op basis van de waarden in tabel 10 is de informatiedichtheid in de havo-tekst hoger dan die in de vmbo-tekst. Bovendien lijkt het aantal inhoudswoorden in dit geval een zeer goede aanduiding te zijn van element interaction (zie tabel 11). De vmbo-tekst bespreekt één gebeurtenis uit de klassieke Griekse geschiedenis, namelijk de slag bij Salamis. De tekst begint met de oorzaak: de binnenvallende Perzen, daarna wordt de list (het middel) besproken waarmee de Grieken de oorlog hebben gewonnen. De laatste alinea is een korte beschouwing over oorzaken en gevolgen in de geschiedenis. De havo-tekst begint met de splitsing en de val van het Romeinse rijk door de binnenvallende Germanen. De laatste 260 woorden van de tekst, zo'n 65%, worden aan dit volk gewijd; hun staatsvorm, gebruiken en godsdienst. Daarna worden enkele goden genoemd; elk met hun kenmerken en verschillende namen. Met betrekking tot godsdienst wordt ook ingegaan op verschillen tussen de Germaanse stammen. Tot slot wordt een overeenkomst met het christendom genoemd, namelijk het geloof in leven na de dood.

De element interaction lijkt in de vmbo-tekst inderdaad hoger te zijn:

Figuur 5: argumentatieschema vmbo geschiedenis Sfinx 2 (zie bijlage 5)



Zoals te zien is in tabel 11, verschillen ook de andere kenmerken met betrekking tot element interaction in de verwachte richting:

Tabel 11: Het aantal inhoudswoorden in vergelijking met 5 andere tekstkenmerken m.b.t. element interaction

Vak	Inhoudswrdrn Per 1000 woorden	TTR	RI	Argumentoverlap Met de vorige zin	Argumentoverlap Per 50 woorden	NP- dichtheid
Havo gesch. Spr.Verl. 2	418	0,525	0,933	358	160	0,348
Vmbo gesch. Sfinx 2	294	0,474	0,917	329	374	0,324

Afgezien van de *rarity index* laten alle variabelen een redelijk tot zeer groot verschil zien. Vooral de argumentoverlap met de voorgaande 50 woorden verschilt zeer sterk. De havo-tekst bevat dus niet alleen heel veel inhoudswoorden, maar ook veel *verschillende* inhoudswoorden. De inhoudswoorden in de vmbo-tekst worden vaker herhaald. Dat is niet vreemd; als de informatiedichtheid lager is – en er relatief weinig onderwerpen besproken worden – is de kans groter dat woorden herhaald worden die met die onderwerpen te maken hebben.

Bij het doorlezen van de teksten lijken deze weinig te verschillen qua inference calls. Beide teksten maken soms een ‘gedachtesprong’ waardoor een inferentie vereist is (zie ook fragment 15 & 16). Inderdaad vereist de havo-tekst niet veel meer inferenties van de lezer dan de vmbo-tekst. In beide teksten wordt soms een beroep gedaan op de wereldkennis van de lezer:

Fragment 15: Geschiedenis vmbo Sfinx 2, r. 4-7 (zie bijlage 5)

Ze gingen naar het orakel van Delphi. Dat was een heilige plaats. Daar spraken de goden met de mensen. Het orakel zei: “Vlucht naar het einde van de aarde. De god Zeus geeft jullie muren van hout. Deze muren zullen jullie redden.” De Grieken dachten lang na. Toen bedachten ze dat Zeus met de ‘muren van hout’ hun oorlogsschepen bedoelde.

De leerling moet, om deze passage te begrijpen, wel op de hoogte zijn van het feit dat een orakel per definitie in geheimtaal spreekt. Uit deze tekst wordt dat niet duidelijk, er wordt enkel gesproken over een heilige plaats waar goden met mensen spreken. Dat dat in geheimtaal gebeurt is niet vanzelfsprekend; denk aan Mozes in het Oude Testament, die ook naar een heilige plaats gaat (de Sinai), waar God duidelijk en ondubbelzinnig met hem spreekt (Exodus, hoofdstuk 19-31). Dit type inference calls komt echter ook in de havo-tekst voor:

Fragment 16: Geschiedenis havo Sprekend Verleden 2, r.7-8, 16-17 & 28-29 (zie bijlage 5)

(r. 7-8) Een deel van de Germanen die zich in het Romeinse rijk vestigden, romaniseerde. Het merendeel bleef de Germaanse cultuur geheel of gedeeltelijk trouw. (...) (r.16-17) De Germanen vereerden anders dan de christenen meer goden. (...) (r.28-29) Naast deze verschillen was er echter ook een overeenkomst: zowel christenen als Germanen geloofden in een leven na de dood.

Waarom de Germaanse godsdienst in dit verband meerdere keren vergeleken wordt met het christendom blijft zeer impliciet. Om te begrijpen dat de tegenstellingen tussen de Germaanse en de christelijke godsdienst tegelijkertijd een tegenstelling is tussen de Germanen en Romeinen moet de lezer op de hoogte zijn van het feit dat Theodosius, over wie in de eerste regel gesproken is, het christendom tot staatsgodsdienst gemaakt heeft in het Romeinse rijk. Ten tijde van de scheiding in een Oostelijk, en in een Westelijk deel (regel 1), was het Romeinse rijk dus christelijk. Dat zal wellicht niet voor iedere leerling duidelijk zijn; de Romeinen staan immers ook bekend om hun christenvervolgingen (bijv. in het Colosseum), en hun 'eigen' goden als Jupiter, Neptunus en Venus. Duidelijk is wel, dat het in beide voorbeelden om globale, en niet om lokale inference calls gaat (zie ook hoofdstuk 6.1).

Beide teksten doen buiten deze voorbeelden om geen groot beroep op de wereldkennis van de lezer bij het leggen van verbanden. Toch denk ik dat de bovenstaande 'inference call' in de havo-tekst een groter probleem vormt dan die in de vmbo-tekst. De vmbo-tekst is namelijk goed te begrijpen zonder precies te weten wat een orakel is. De inferentie in de havo-tekst is echter niet alleen heel complex, maar ook van groot belang bij het begrijpen van de tekst. Het startpunt van de tekst is het jaar 395. Het Romeinse rijk wordt dan gesplitst in een Oostelijk en een Westelijk deel dat 70 jaar later, in 476 zal vallen. Het feit dat het Romeinse rijk het christendom tot staatsgodsdienst had afgekondigd, en vreemde godsdiensten (zoals die van de Germanen) de laatste 70 jaar niet tolereerde, was een belangrijke oorzaak van het uiteindelijke verval van het Romeinse rijk "in een lappendeken van Germaanse staatjes" (Gibbon, 1994, pp. 117-119). Hoewel dat verband belangrijk is voor het begrijpen van de tekst, blijft het grotendeels impliciet. De tekst zou dit probleem gemakkelijk kunnen oplossen door expliciet aan te geven dat de Romeinen vanaf de vierde eeuw christelijk waren. De regels 16-17 & 28-29 zouden bijvoorbeeld zo kunnen luiden: *De Germanen vereerden anders de Romeinen, die het christendom aanhingen, meer goden. (...) Naast deze verschillen was er echter ook een overeenkomst: zowel de christelijke Romeinen als de Germanen geloofden in een leven na de dood.* De tekst zou dan minder inference calls vereisen, en de lezer zou de tekst op globaal niveau wellicht beter begrijpen.

Ik denk dat het aantal inhoudswoorden een betere aanduiding is voor informatiedichtheid dan de eerder besproken NP-dichtheid. Zoals gezegd valt de NP-dichtheid vaak hoog uit in teksten met korte zinnen; het aantal inhoudswoorden wordt echter berekend per 1000 woorden, en is

dus minder afhankelijk van stijlkenmerken zoals zinslengte. De teksten uit bijlage 1, die ik gebruikte bij de bespreking van de NP-dichtheid geven hiervan een goede illustratie. *Let op: de aanduidingen m.b.t. de woordsoort zijn van T-Scan, deze kloppen niet altijd:*

Fragment 17: vmbo Sfinx 3, regel. 1, (zie bijlage 1)

Ze bouwden **forten** langs de **rivier**.

Fragment 18: havo Feniks 5, regel 4-5, (zie bijlage 1)

De ~~katholieken~~ (en ~~later ook~~ **protestanten**) waren het niet eens met de **manier waarop** de **liberalen** ~~steeds verder~~ gingen met het loskoppelen van **kerk** en **staat**.

LEGENDA: Bijwoorden Adjectieven **Zelfstandige naamwoorden**

Fragment 17 heeft door de korte zinslengte een zeer hoge NP-dichtheid (0,50). De NP-dichtheid van fragment 18 is slechts 0,31. Qua inhoudswoorden is dit verschil echter omgekeerd; fragment 17 bevat 333 inhoudswoorden per 1000 woorden, fragment 18 bevat er maar liefst 500. Dat verschil geldt niet alleen voor deze teksten; er is namelijk een (licht) negatieve correlatie tussen de NP-dichtheid en het aantal inhoudswoorden ($r = -0,22$, $p = 0,03$). Hoewel het aantal inhoudswoorden in dit geval inderdaad een goede aanduiding geeft van de informatiedichtheid, is er ook het één en ander aan te merken op dit kenmerk. Ten eerste herkent T-Scan hoofdwerkwoorden niet; bij tellen van het aantal inhoudswoorden slaat T-Scan hierdoor alle werkwoorden over. Verder wordt een woord soms in een verkeerde categorie ingedeeld. In fragment 18 herkent T-Scan het zelfstandige naamwoord 'katholieken' bijvoorbeeld als adjectief. Hoewel dat voor het tellen van inhoudswoorden niet uitmaakt (zelfstandige naamwoorden en adjectieven zijn beide inhoudswoorden) kan dit toch een punt van aandacht zijn.

Ik denk dat het aantal inhoudswoorden een goede aanduiding kan geven van de element interaction. Andere kenmerken, zoals de argumentoverlap met de voorgaande 50 woorden bevestigen dit. Of dit kenmerk veel zegt over inference calls blijft echter onduidelijk. Zoals gezegd zijn de inference calls uit de havo-tekst naar mijn mening problematischer dan die uit de vmbo-tekst, maar dat lijkt grotendeels op toeval te berusten.

6.4 Eigennamen

In hoofdstuk 5.4 liet ik zien dat geschiedenisteksten meer eigennamen bevatten dan economieteksten. Als de aanwezigheid van eigennamen inderdaad negatief samenhangt met element interaction en de vereiste inference calls, zijn de hypothesen bevestigd; ik verwachtte immers dat geschiedenisteksten een lagere element interaction hebben en meer inference calls vereisen (hypothese 1a en 2a).

Voor de eerste analyse van eigennamen gebruik ik een geschiedenistekst (vmbo Pharos 5, eigennamen per 1000 woorden: 143) en een economietekst (vmbo Pharos 3, aantal eigennamen per 1000 woorden: 0). T-Scan is goed in staat eigennamen te onderscheiden, op een paar inconsequenties na. Zo wordt het woord 'Grieken' wel als eigennaam aangemerkt, maar 'Atheners' niet. Ook samengestelde eigennamen zoals 'Trojaanse Oorlog' worden niet herkend. Over het aantal inference calls lijkt dit kenmerk nu niet veel te zeggen. Dat is ook niet vreemd; eigenlijk gaat het hier gewoon om fictie (zie ook hoofdstuk 2.1). Het verhaal is bovendien niet complex, en het grote aantal eigennamen wordt vooral veroorzaakt door het vermijden van voornaamwoorden, wat leidt tot veel herhalende namen. Het aantal inference calls wordt hier dus juist *verminderd* door het grote aantal eigennamen: de lezer hoeft immers niet zelf het verband te leggen tussen een voornaamwoord en de referent. Ook de element interaction lijkt niet noemenswaardig te verschillen – in ieder geval niet door de aanwezige eigennamen.

En hoe zit dat met non-fictie? In bijlage 7 heb ik wederom twee teksten naast elkaar gezet, een havo-economietekst (Praktische Economie 5) en een havo-geschiedenistekst (Feniks 6). De eerste tekst bevat geen eigennamen, de tweede bevat er 44. T-Scan is weer goed in staat eigennamen te onderscheiden, maar weer zijn er een paar inconsequenties. De samenstellingen Eerste en Tweede Wereldoorlog worden wel herkend, maar het 'Verdrag van Versailles' niet. Het woord Versailles wordt echter wel afzonderlijk als eigennaam herkend. Namen van volken worden niet meegenomen wanneer deze bijvoeglijk gebruikt zijn (Duitse, Poolse etc.). Wanneer namen van volken zelfstandig worden gebruikt herkent T-Scan ze soms wel (Fransen r. 27), en soms niet (Engelsen, r. 27). Misschien wordt 'Fransen' als meervoud gerekend van de voornaam 'Frans'. In dat geval zou er sprake zijn van probleemtype 1; het woord 'Fransen' heeft immers twee betekenissen waarvan er één dominant is. 'Fransen' zal namelijk vaker gebruikt worden als aanduiding voor het volk, dan als meervoud van de voornaam 'Frans'.

Zorgt het grote aantal eigennamen in de geschiedenistekst inderdaad voor veel inference calls? Het lijkt er wel op, zie bijvoorbeeld dit fragment:

Fragment 19: Geschiedenis havo Feniks 7, r. 9-11 (zie bijlage 7).

DUITSLAND stapte uit de **VOLKENBOND** en **HITLER** verkondigde dat de **DUITSERS LEBENSRAUM** nodig hadden. Dat daarmee de bepalingen uit het Verdrag van **VERSAILLES** onder druk kwamen, lapte hij aan zijn laars. De Duitse economie werd zo snel mogelijk op wapenproductie ingesteld en het leger werd uitgebreid.

De lezer moet in dit fragment zelf invullen wat er bedoeld wordt met ‘Volkenbond’ en ‘Verdrag van Versailles’. Dat is wel op te maken uit de formulering: “daarmee kwamen de bepalingen uit het Verdrag van Versailles onder druk”. De leerling zal ook zonder kennis van de begrippen ‘Volkenbond’ en ‘Verdrag van Versailles’ de hoofdlijn uit het verhaal wel kunnen volgen: Hitler schond internationale afspraken en stevende daarmee op oorlog af. Dat het schenden van het Verdrag van Versailles samenhangt met de economie van Duitsland blijft echter volledig impliciet. Het schenden van het verdrag betekent immers dat Duitsland haar verplichte herstelbetalingen voor de schade van de Eerste Wereldoorlog niet nakomt, en daardoor in staat is wapens te produceren. Ook het verband tussen het begrip ‘*Lebensraum*’, de inval in het Rijnland en gebiedsuitbreiding naar het Oosten wordt niet nader toegelicht. Dat gebiedsuitbreiding in Oost-Europa bedoeld is als uitbreiding van het ‘*Lebensraum*’ en de inval in het Rijnland – dat in Franse handen was – om een andere reden gepleegd is, zal de leerling volstrekt onduidelijk blijven. Tot slot wordt vermeld dat in 1939 de inval in Polen begon: het begin van de Tweede Wereldoorlog. Ook dát moet de leerling op grond van wereldkennis begrijpen. De inval in het Rijnland (*ook* een inval) leidde immers *niet* tot een oorlog.

Deze resultaten worden door andere tekstkenmerken bevestigd:

Tabel 12: Het aantal eigennamen in vergelijking met 4 andere tekstkenmerken m.b.t. inference calls

Vak	Eigennamen Per 1000 wrdn.	Argumentoverlap vorige zin Per 1000 wrdn	Argumentoverlap voorgaande 50 woorden Per 1000 wrdn	Connectieven Per 1000 wrdn	TTR Per tekst
Havo Econ. Prakt. Econ. 5	0	324	321	81	0,531
Havo Gesch. Feniks 6	114	318	41	57	0,584

Net als eigennamen duidt ook het verschil in argumentoverlap, connectieven en type-token-ratio op een groter aantal inference calls in de geschiedenis tekst.

Het aantal eigennamen zegt veel over het aantal inference calls, zowel in deze teksten, als in de tekst die ik besprak in hoofdstuk 6,1. Wel denk ik dat eigennamen zoveel mogelijk in combinatie met andere tekstkenmerken moet worden beschouwd; de *afwezigheid* van eigennamen hoeft niet altijd te duiden op weinig inference calls. De resultaten op basis van het totaal aantal eigennamen zijn in ieder geval veelbelovend.

6.5 Overzicht van de tekstkenmerken

Tabel 13: Samenvatting v.d. resultaten tekstkenmerken T-Scan uit de kwalitatieve analyse

Kenmerk	Vergeleken Teksten (tekstparen)	Hfd.st	Bij- Lage	Elem. Int.	Inf. Calls.
Argumentoverlap vorige zin	Gesch. havo Feniks 5 gesch. vmbo Sfinx 3	6,1	1	-	+ (lokaal)
	Ec. havo Praktische Economie 4 Gesch. havo Pharos 4	6.3.1	4	-	-
	Gesch. havo Sprekend verl. 2 Gesch. vmbo Sfinx 2	6.3.3	5	-	∅
	Ec. havo Prakt. Ec. 5 Gesch. havo Feniks 6	6.4	7	-	-
	Argumentoverlap vorige 50 woorden	Gesch. havo Feniks 5 gesch. vmbo Sfinx 3	6,1	1	∅
	Ec. havo in Balans II 2 Gesch. havo Geschiedeniswerpl. 1	6.3.2	3	+	n.v.t.
	Ec. havo Praktische Economie 4 Gesch. havo Pharos 4	6.3.1	4	+	+
	Gesch. havo Sprekend verl. 2 Gesch. vmbo Sfinx 2	6.3.3	5	+	-
Causale connectieven	Ec. vmbo Economisch Bekeken Gesch. vmbo Pharos 1	6.2.1	2	+	n.v.t.
	Ec. havo in Balans II 2 Gesch. havo Geschiedeniswerpl. 1	6.2.1	3	-	n.v.t.
Alle connectieven	Ec. vmbo Economisch Bekeken Gesch. vmbo Pharos 1	6.2.2	2	n.v.t.	-
TTR & Rarity index	Ec. havo Praktische Economie 4 Gesch. havo Pharos 4	6.3.1	4	+	+
	Gesch. havo Sprekend Verl. 2 Gesch. vmbo Sfinx 2	6,3.3	5	∅	∅
	Ec. havo Prakt. Ec. 5 Gesch. havo Feniks 6	6.4	7	+	+
	NP-dichtheid	Gesch. havo Feniks 5 Gesch. vmbo Sfinx 3	6,3.2	1	-
	Ec. havo in Balans II 2 Gesch. havo Geschiedeniswerpl. 1	6,3.2	3	-	n.v.t.
	Ec. havo Prakt. Ec. 4 Gesch. havo Pharos 4	6,3.2	4	-	-
Inhoudswoorden	Gesch. havo Sprekend Verl. 2 Gesch. vmbo Sfinx 2	6,3.3	5	+	-
Eigennamen	Ec. vmbo Index 1 Gesch. vmbo Pharos 5	6.4	6	∅	∅
	Ec. havo Prakt. Ec. 5 Gesch. havo Feniks 6	6.4	7	+	+
	Ec. havo Prakt. Ec. 4 Gesch. havo Pharos 4	6.3.1	4	∅	∅

LEGENDA

- +** Het tekstkenmerk laat een daadwerkelijk aanwezig verschil in element interaction of inference calls zien.
- Het tekstkenmerk laat ten onrechte *geen* verschil zien, of laat wel een verschil zien in tegenovergestelde richting.
- ∅** Het tekstkenmerk verschilt niet, maar de tekst verschilt ook niet qua inference calls of element interaction.
- N.v.t.** Het kenmerk is voor deze tekst buiten beschouwing gelaten.

7. Conclusie

Ik heb mij in dit onderzoek de volgende hoofdvraag gesteld:

In hoeverre bieden oppervlaktekenmerken uit T-Scan een goede afspiegeling van de mate van elementinteractie en de benodigde hoeveelheid inference calls in studieteksten?

Gezien de schaalgrootte van het kwalitatieve onderzoek, beantwoord ik deze vraag met enige terughoudendheid. Sommige oppervlaktekenmerken lijken echter goede aanwijzingen te geven voor de 'diepere' eigenschappen element interaction en inference calls (zie tabel 13). Element interaction is vooral met de kenmerken argumentoverlap met de voorgaande 50 woorden, de type-token-ratio en de rarity index goed te voorspellen. Voor de verdere kenmerken zijn de resultaten wisselend; elk kenmerk heeft zijn eigen valkuilen. Enkel de kenmerken argumentoverlap met de vorige zin en NP-dichtheid vallen in negatieve zin op; deze lijken me dan ook niet geschikt voor het aantonen van 'mijn' teksteigenschappen. Ook voor de eigenschap inference calls zijn de resultaten wisselend. De argumentoverlap met de voorgaande 50 woorden, type-token-ratio, rarity index en het aantal eigennamen geven echter een redelijke aanduiding van deze eigenschap. Het valt op dat vooral 'globale' tekstkenmerken – die iets zeggen over de tekst als geheel – succes hebben, met name voor de eigenschap element interaction. De kenmerken NP-dichtheid en argumentoverlap met de vorige zin zijn toch lokaler: zij traceren een verschil binnen of tussen zinnen.

Uit de resultaten blijkt verder dat element interaction en inference calls, zoals verwacht (zie hoofdstuk 2.4 & 3), niet volledig los van elkaar staan. Meestal duidt een hoge interactie *ook* op weinig inference calls (zie ook hoofdstuk 8).

De drie probleemtypen die ik in hoofdstuk 2.1 besprak zijn allen voorbij gekomen. Soms is een kenmerk dubbelzinnig, maar is er sprake van een uitzondering ('doordat' in fragment 9, 'Fransen' in bijlage 7, r. 26). Soms zorgt dubbelzinnigheid voor grotere problemen; bijvoorbeeld omdat een woord niet één dominante betekenis heeft ('als' in fragment 10), of omdat een woord verkeerd wordt geïndexeerd (fragment 18). Tot slot zorgt de context soms voor een vertekend beeld, zoals het gebruik van eigennamen in de vmbo-geschiedenistekst Pharos 5 (bijlage 6). Door het fictieve karakter van deze tekst waren eigennamen geen indicator voor het aantal inference calls (zie ook hoofdstuk 2.1). Dat was in de havo-geschiedenistekst Feniks 6 (bijlage 7) echter wel het geval.

De kwalitatieve analyse – aan de hand waarvan de vier hypothesen getoetst worden – was niet uitputtend: van de 111 teksten uit het corpus heb ik er 14 aan een kwalitatieve analyse

onderworpen. Ik bespreek de hypothesen daarom met enige voorzichtigheid. Dat doe ik aan de hand van oppervlaktekenmerken die de eigenschappen met succes traceerden. Bij een 'bevestigde hypothese' staat er dus minstens één plusje in tabel 13: bovendien was dat verschil in de verwachte richting (voor vak of niveau). De hypothesen luiden als volgt:

1. *Element interaction*

- a. De element interaction is in economieteksten hoger dan in geschiedenisteksten;
- b. De element interaction is in schoolboeken voor lage(re) niveaus hoger dan voor hoge(re) niveaus.

2. *Inference calls*

- a. Economieteksten bevatten meer expliciete coherentierelaties, en vereisen dus minder wereldkennis bij het leggen van verbanden dan geschiedenisteksten
- b. Schoolboeken voor lagere niveaus bevatten meer gemarkeerde relationele en referentiële coherentierelaties, en vereisen dus minder wereldkennis bij het maken van inferenties dan schoolboeken voor hogere niveaus.

Hypothese 1a lijkt mij – met enige terughoudendheid- bevestigd; in economieteksten (*expository texts*) is de element interaction hoger dan in geschiedenisteksten (*narrative texts*). Dat bleek uit de tekstenparen in bijlage 3 & 4. Hypothese 1b is niet bevestigd: er is (voor element interaction) slechts twee maal een verschil gevonden tussen een havo- en een vmbo-tekst van hetzelfde vak. Eenmaal was dit verschil in lijn met de hypothese (teksten bijlage 5), eenmaal was er geen verschil (teksten bijlage 1).

Hypothese 2a is - eveneens met enige terughoudendheid - bevestigd; in twee gevallen vereiste de economietekst minder inference calls dan de geschiedenistekst (teksten bijlage 4 & 7), in één geval was er geen verschil (teksten bijlage 2). Er is geen ondersteuning voor hypothese 2b; twee keer is een vmbo-tekst vergeleken met een havo-tekst (teksten bijlage 1 & 5), tweemaal was er geen duidelijk verschil.

In de methode uitte ik de verwachting dat economieteksten vooral te vergelijken zijn met *expository texts*, en geschiedenisteksten met *narrative texts*. De verschillen die McNamara e.a. vonden tussen deze genres, verwachtte ik namelijk te vinden tussen de twee schoolvakken geschiedenis en economie. Die verwachting is –wederom met enige voorzichtigheid – bevestigd; zowel qua element interaction (hypothese 1a) als qua inference calls (hypothese 2a). Het verschil tussen de beide vakken is dus overtuigender aangetoond dan dat tussen de niveaus (hypothese 1b en 2b).

8. Discussie

Tabel 13 op pagina 44 geeft op het eerste gezicht niet bepaald een positief beeld; 12 keer voorspelde een tekstkenmerk een verschil, 15 keer lukte dat niet, en 9 keer bleef dat onduidelijk. Toch is dat resultaat geenszins teleurstellend. Sommige kenmerken, zoals de NP-dichtheid, argumentoverlap met de vorige zin, of 'alle connectieven' bleken geen goede voorspeller te zijn voor element interaction of inference calls (zie ook hoofdstuk 7). Wanneer die tekstkenmerken uit de tabel verwijderd worden, blijven er nog slechts 3 'minnetjes' over. Bovendien kan T-Scan niet één, maar 150 tekstkenmerken bepalen, en een *combinatie* van succesvolle tekstkenmerken geeft natuurlijk een betrouwbaarder beeld dan dat van één tekstkenmerk. De teksten uit bijlage 4 verschilden bijvoorbeeld zowel qua element interaction als inference calls. Dat verschil is met de kenmerken type-token-ratio, rarity index en argumentoverlap op basis van de laatste 50 woorden succesvol aangetoond. Ook het verschil tussen de teksten uit bijlage 5 is met twee kenmerken succesvol voorspeld; zowel het aantal inhoudswoorden als de argumentoverlap met de laatste 50 woorden toonden het verschil in element interaction aan. De 9 gevallen waarin er *geen* duidelijk verschil gevonden is (\emptyset in de tabel) betekenen niet dat T-Scan 'gefaald' heeft. Vaak is er sprake van een gemakkelijk op te lossen probleem. Het kenmerk 'eigennamen' zou bijvoorbeeld sterk verbeterd kunnen worden door onderscheid te maken tussen voor- en achternamen, persoonsnamen, topografische namen en andere eigennamen. Een verhaal wordt dan –door de vele voornamen– al snel herkend. Ook de mogelijkheid een type-token-ratio uit te voeren op het aantal eigennamen zou de betrouwbaarheid van dit kenmerk verbeteren. Wanneer dezelfde namen telkens herhaald worden zal dat de tekst immers niet moeilijker maken (zie ook hoofdstuk 6.4 en bijlage 6).

Ik merkte een paar keer op (hoofdstuk 2.4, 3 & 7) dat inference calls en element interaction grotendeels samenhangen: teksten met veel inference calls hebben vaak een lage element interaction en andersom. Dat heeft echter ook een nadeel: op basis van 'domme' T-Scan kenmerken zijn element interaction en inference calls moeilijk van elkaar te scheiden. Dat betekent overigens niet dat een tekst met een lage element interaction *altijd* veel inference calls vereist; de eerder genoemde lijst met Amerikaanse presidenten bevat een zeer lage interactie maar bevat *ook* maar weinig inference calls. Dat zal in T-Scan echter niet te zien zijn: zo'n lijst heeft namelijk een hoge TTR en rarity index, een lage argumentoverlap, veel eigennamen en een hoge informatiedichtheid.

Hoewel de kenmerken op conceptueel niveau dus duidelijk van elkaar te scheiden zijn, is dat op oppervlakteniveau niet altijd zonder problemen. Dat geldt ook voor een tekst met een hoge element interaction *en* veel inference calls. Bijvoorbeeld: *Socrates was niet ziek meer. Hij vroeg*

om een haan. Wanneer je niet op de hoogte bent van het feit dat Socrates een historisch persoon was die leefde in het klassieke Griekenland *en* dat het daar gewoonte was een haan te offeren wanneer je een ziekte overwonnen had, is de tekst niet te begrijpen. Er is echter wel een causale relatie tussen deze twee zinnen aanwezig. De T-Scanmerken zullen echter duiden op een lage element interactie *en* veel inference calls. In beide gevallen is er één kenmerk onzichtbaar: element interaction in het eerste voorbeeld, inference calls in het tweede voorbeeld. In de lijst Amerikaanse presidenten heeft T-Scan immers ‘gelijk’ wat betreft de element interaction, maar ongelijk wat betreft het aantal inference calls. Bij het bovenstaande voorbeeld over Socrates is dat andersom.

Bij het werken met T-Scan is het belangrijk stil te staan bij het tekstgenre. In dit onderzoek ging het over schoolteksten; die zijn per definitie bedoeld voor lezers met weinig voorkennis. Het is bij dit teksttype te verwachten dat de auteur aanwezige verbanden zal markeren zodat de tekst door de doelgroep te begrijpen is. Een auteur die voor experts schrijft zal dat misschien niet doen. Om de T-Scanmerken goed te kunnen interpreteren is het daarom belangrijk enigszins op de hoogte te zijn van het teksttype dat je onderzoekt.

Bij het kenmerk argumentoverlap (hoofdstuk 6.1) maak ik een tweedeling tussen lokale en globale inference calls. Het lijkt logisch dat argumentoverlap met de vorige zin vooral lokale, en argumentoverlap met de voorgaande 50 woorden vooral globale inference calls traceert. Die constatering gaat echter niet op: lokale inference calls zijn nauwelijks te onderscheiden. Natuurlijk zal een lokale ‘inference call’ meestal gekenmerkt worden door een lage argumentoverlap met de vorige zin. Er blijven dan echter nog tientallen gevallen over waarin een lage argumentoverlap met de vorige zin *niets* zegt over de (lokale) inference calls. En *dat* verschil kan T-Scan niet maken. Het verschil in lokale inference calls dat de argumentoverlap met de vorige zin laat zien in hoofdstuk 6.1 is dus toeval. Al met al denk ik dan ook niet dat de argumentoverlap met de vorige zin een relevant kenmerk is om inference calls te traceren: of die nu lokaal of globaal zijn. Gelukkig heeft het andere type argumentoverlap wel overtuigende resultaten laten zien.

Een laatste opmerking over inference calls. Zoals gezegd gebruik ik die term op een andere manier dan Britton & Gülgöz (1991), aan wiens onderzoek het begrip ontleend is. Dat heeft vooral te maken met hun onderzoeksdoel. Zij gebruiken het begrip met betrekking tot het *reading comprehension model* van Kintsch & Van Dijk (1978) en de manier waarop dit model wordt geïmplementeerd in Miller & Kintsch’ computerprogramma (1980). Miller & Kintsch proberen met dit programma de mentale representatie van de lezer na te bootsen. Omdat dit computerprogramma (uit 1980!) vooral gebruik maakt van oppervlaktekenmerken, hanteren

Britton & Gülgöz het begrip *inference calls* ook met betrekking tot oppervlaktekenmerken. In mijn onderzoek gaat het uiteindelijk echter juist om dieperliggende kenmerken: in tegenstelling tot de computer uit het onderzoek van Britton & Gülgöz kan een 'echte' lezer daar wel gebruik van maken. Daarom gebruik ik de term *inference calls* niet alleen voor verschillen op oppervlakteniveau (zoals het ontbreken van woordherhaling) maar ook op een dieper, conceptueel niveau. De eigenschap zou, wanneer ik dat niet zou doen, een vertekend beeld geven van de manier waarop mensen werkelijk lezen. Daarmee is de methode van Britton & Gülgöz niet 'verkeerd': hun onderzoek verschilt van het mijne, en de terminologie is daarop aangepast.

Tot slot noem ik ideeën voor vervolgonderzoek. Ten eerste zou het goed zijn *alle* teksten uit het corpus aan een kwalitatieve analyse te onderwerpen. Wanneer dat systematisch en voor elke tekst op gelijke wijze gebeurt, zou dat overtuigende ondersteuning kunnen opleveren voor de hypothesen. Een nadeel is dat dat heel veel tijd zou kosten. Ten tweede zou het interessant zijn eerst de kwalitatieve analyse uit te voeren, en pas *daarna* de kwantitatieve analyse. Dat wil zeggen: eerst worden er teksten uitgezocht die op basis van zorgvuldige lezing opvallen door hun (schijnbare) hoge of lage interactie of het aantal inferenties. Daarna worden deze teksten in T-Scan ingevoerd om te kijken of dat in de kenmerken terug is te zien. In feite wordt 'mijn' onderzoek dan omgekeerd uitgevoerd. Zo'n onderzoek zou verdere ondersteuning kunnen leveren voor de T-Scan kenmerken met betrekking tot element interaction en *inference calls*. Misschien komen op die manier meer kenmerken aan het licht waarmee het mogelijk is element interaction en *inference calls* van elkaar te kunnen onderscheiden met T-Scan.

De discussie over leesbaarheidsformules, die ik in de inleiding aanhaalde, is dus nog niet beslecht. Ik hoop echter dat mijn onderzoek heeft laten zien dat er een toekomst is voor leesbaarheidstools: en dat geldt zeker voor T-Scan!

9. Aangehaalde literatuur

Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83(3), 329.

Gibbon, E. (1994). *The Decline and Fall of the Roman Empire (Vol. IV)*, Everyman's Library, JM Dent and Sons, London (orginele uitgave: 1788).

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82-98). New York: Guilford.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods*, 36(2), 193-202.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, 40(5), 223-234.

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

Kamalski, J., Lentz, L., & Sanders, T. (2004). Coherentiemarkering in informerende en persuasieve teksten. *Tijdschrift voor Taalbeheersing*, 26(2), 85-103.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of tekst comprehension and production. *Psychological review*, 85(5), 363-394.

Kraf, R. & Pander Maat, H. (2009). Leesbaarheidsonderzoek: oude problemen, nieuwe kansen. *Tijdschrift voor Taalbeheersing*, 31(2), 97-123.

Kraf, R. (2013). *T-Scan Gebruikershandleiding 0.2*. Universiteit Utrecht: Faculteit Geesteswetenschappen

Land, J. (2009). *Zwakke lezers, sterke teksten?: effecten van tekst-en lezerskenmerken op het tekstbegrip en de tekstwaardering van VMBO-leerlingen*. Eburon Uitgeverij BV.

Land, J., Sanders, T., & Van den Bergh, H. (2008). Effectieve tekststructuur voor het vmbo Een corpus-analytisch en experimenteel onderzoek naar tekstbegrip en tekstwaardering van vmbo-leerlingen voor studieteksten. *Pedagogische Studiën*, 85(2), 76-94.

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. *Unpublished Grant proposal, University of Memphis, Memphis, Tennessee*.

McNamara, D. S., Floyd, R. G., Best, R., & Louwerse, M. (2004). World knowledge driving young readers' comprehension difficulties. In *Proceedings of the 6th international conference on Learning sciences* (pp. 326-333). International Society of the Learning Sciences.

McNamara, D. S., Graesser, A., & Louwerse, M. (2012). *Sources of Text Difficulty: Across Genres and Grades*. Measuring Up: Advances in How We Assess Reading Ability, 89.

Miller, J. R., & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory*, 6(4), 335-54.

Noordman, L.G.M., & Maes, A. (2000). Het verwerken van tekst. In A. Braet (Ed.), *Taalbeheersing als Communicatiewetenschap* (pp. 29-60). Bussum: Coutinho

Owen, W.B. & E.J. Goodspeed. (1906). *Homeric Vocabularies: Greek and English Wordlists for the Study of Homer*. Chicago: University of Chicago Press.

Pharr, C. (1985). *Homeric Greek: A book for beginners*. Oklahoma: University of Oklahoma Press. (originele uitgave: 1920).

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional science*, 32(1), 1-8.

Sweller, J. Van Merriënboer, J. van & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251 – 296.

Van het Reve, K. (1985). *Geschiedenis van de Russische literatuur: van Vladimir de Heilige tot Anton Tsjechov*. Amsterdam: Van Oorschot.

Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Horst, H. V. D. (2000). *Nederland. De vaderlandse geschiedenis van de prehistorie tot nu*. Amsterdam: Bert Bakker

Verhoeven, G. & Schellens, P.J. (2008). Argumentatieanalyse. In: Schellens, P.J. & Steehouder, M. (eds.), *Tekstanalyse: Methoden en toepassingen*. Assen: Van Gorcum..

Index van bijlagen

Bijlage 1: Geschiedenis vmbo Sfinx 3 & Geschiedenis havo Feniks 5

Bijlage 2: Economie vmbo Economisch Bekeken 4 & Geschiedenis vmbo Pharos 1

Bijlage 3: Economie havo In Balans II 2 & Geschiedenis havo Geschiedeniswerkplaats 1

Bijlage 4: Economie havo Praktische Economie 4 & Geschiedenis havo Pharos 4

Bijlage 5: Geschiedenis havo Sprekend Verleden 2 & Geschiedenis vmbo Sfinx 2

Bijlage 6: Economie vmbo Index 1 & Geschiedenis vmbo Pharos 5

Bijlage 7: Economie havo Praktische Economie 5 & Geschiedenis havo Feniks 6

Bijlage 8: Opbouw van het corpus

Bijlage 1: Geschiedenis vmbo Sfinx 3 & Geschiedenis havo Feniks 5

De Romeinen kozen de Rijn als natuurlijke grens. Ze bouwden forten langs de rivier. De grootste forten waren Nijmegen en Utrecht. Ten noorden van de Rijn woonden de Friezen op terpen. Terpen zijn zelf gemaakte heuvels. Op een terp bleef je droog bij een overstroming.

5 In de heuvels van Zuid-Limburg bouwden de Romeinen boerderijen. Ze noemden ze villa's, net zoals in Italië. Maastricht en Heerlen groeiden uit tot kleine steden met een fort en Romeinse badhuizen.

Rond het jaar 400 vertrokken de laatste Romeinen uit Nederland. Ze gingen Rome verdedigen tegen de vijanden. In het oosten van ons land kwamen de Saksen wonen.

10 In het zuiden bleven de Franken wonen. In het noorden bleven de Friezen de baas. Deze volken geloofden in goden als Wodan, Donar en Frya. De christenen Willibrord en Bonifatius probeerden de bevolking tot het christendom te bekeren.

Rond 800 was Karel de Grote koning van de Franken. Hij zorgde voor één munt, voor nieuwe wegen en een nieuw bestuur. Hij wilde dat alle mensen christen werden. Hij werd door de paus in Rome tot keizer gekroond.

15 Voor zijn oorlogen had Karel de Grote legeraanvoerders nodig en veel soldaten. Hij beloofde zijn aanvoerders met stukken land. Karel gaf de grond niet weg, maar hij leende het uit. Karel werd de leenheer en zijn legeraanvoerders werden leenmannen. Deze leenmannen noemen we ook adel. De leenmannen hadden boeren in dienst. Die werkten voor hen. Ze moesten altijd op de grond blijven. Ze hadden helemaal geen vrijheid. We noemen hen lijfeigenen. Het waren eigenlijk een soort slaven.

20 Het christendom was de enige godsdienst. De geestelijke, alle mensen die voor de kerk werkten, kregen steeds meer grond. De kerk werd hierdoor nog belangrijker en rijker. Zo ontstonden in de Middeleeuwen drie soorten beroepen. Je zegt ook wel drie standen: geestelijkheid, adel en boeren.

Net zoals vrouwen, werden ook de katholieken in Nederland lange tijd achtergesteld. Pas in de herziene grondwet van 1848 werd 'vrijheid van godsdienst' als een belangrijk grondrecht opgenomen. Hiermee werd de weg vrijgemaakt voor emancipatie van de katholieken.

5 De katholieken (en later ook protestanten) waren het niet eens met de manier waarop de liberalen steeds verder gingen met het loskoppelen van kerk en staat. Zo vonden ze het bijvoorbeeld verkeerd dat de kinderen op school niet volgens de katholieke of protestantse leer werden opgevoed. In de tweede helft van de negentiende eeuw groeide bij deze confessionelen de ergernis over de manier waarop het onderwijs georganiseerd was. In de grondwet van 1848 was weliswaar bepaald dat er 'vrijheid van

10 onderwijs' was, maar in de praktijk kwam dat neer op 'openbaar onderwijs'. De bisschoppen lieten vanaf 1870 steeds duidelijker merken dat ze het openbare onderwijs 'meer en meer onbruikbaar' en zelfs 'uiterst gevaarlijk' vonden. De katholieken en de protestanten – eeuwenlang gescheiden groepen die regelmatig vijandig tegenover elkaar stonden – werden op het punt van onderwijs bondgenoten. Zij wilden dat de overheid de bijzondere (katholieke en protestantse) scholen op dezelfde manier zou financieren als het openbaar onderwijs. Toen in 1878 een nieuwe onderwijswet werd behandeld – waarin

15 de financiële gelijkstelling van het openbaar en het bijzonder onderwijs niet was geregeld – stemde de katholieke en antirevolutionaire (protestantse) kamerleden tegen. De leider van de antirevolutionairen, Abraham Kuyper, greep naar een laatste redmiddel: een beroep op de Koning om de onderwijswet niet te ondertekenen.

20 Vier dagen nadat de Tweede Kamer de wet had aangenomen, lag op veel plaatsen een petitie ter tekening klaar. Tegenstanders van de Onderwijswet 1878 konden met het zetten van een handtekening kenbaar maken dat ze tegen deze wet waren. Nog eens vier dagen later hadden meer dan 300.000 Nederlandse protestanten hun handtekening gezet. Ook de katholieken organiseerden een handtekeningactie, maar bij hen bleef de score steken op 164.000 handtekeningen. De onderwijskwestie, bekend als de Schoolstrijd, werd uiteindelijk pas in 1917 definitief geregeld.

Bijlage 2: Economie vmbo Economisch Bekeken 4 Geschiedenis vmbo Pharos 1

Stephans grootste hobby is uitgaan. Zijn weekenden kosten bergen geld. Hij wil hier paal en perk aan stellen en neemt zich voor om zijn persoonlijke uitgaven te beperken tot € 200,— in de maand. Hij wil dit budget bewaken en houdt sinds 1 oktober al zijn persoonlijke uitgaven bij.

Een budget is bedoeld om aan vast te houden. ALS dat niet lukt en je gaat over je budget heen, moet je budgetteren. Anders kunnen er financiële problemen ontstaan. Je moet je uitgaven opnieuw afstemmen op je inkomsten.

ALS je een budget wilt verhogen, moet je budgetteren. WANT je houdt dan minder geld over voor de andere uitgaven. Je moet de andere budgetten verlagen, of minder gaan sparen. Budgetteren is ook nodig ALS je meer wilt gaan sparen, WANT er is dan minder geld beschikbaar voor je uitgaven.

ALS je een of meer budgetten verlaagt, moet je bezuinigen. Je kunt bezuinigen door minder te kopen, of door een uitgavenpost te schrappen. Bij de uitgaven voor de noodzakelijke behoeften gaat dat niet, WANT je kunt niet zonder dingen die nodig zijn om gezond te blijven. Hooguit kun je zoeken naar een goedkopere manier om deze behoeften te bevredigen.

Inkomsten kunnen flink omlaag gaan, bijvoorbeeld DOORDAT je minder gaat werken of door ontslag. Bij een daling van de inkomsten is budgetteren lastig, WANT er moet bezuinigd worden op uitgaven waaraan je bent gewend. Bij een stijging van de inkomsten is budgetteren plezierig, WANT er is geld beschikbaar voor nieuwe uitgaven.

Een map met bankafschriften is onderdeel van een administratie. Een administratie bestaat uit financiële overzichten, bijvoorbeeld van de inkomsten en uitgaven. Met een administratie is het gemakkelijker om de budgetten te bewaken. Je kunt HIERDOOR de werkelijke uitgaven vergelijken met de budgetten.

De Grieken noemden hun stad een polis. Zo'n polis was eigenlijk een klein staatje. Er waren honderden staatjes in Griekenland. Ze werden allemaal verschillende bestuurd. Het bestuur van de polis werd politiek genoemd. De Atheners hadden als enige een democratie. Democratie betekent dat het volk de beslissingen neemt. Op het grote plein van bron 8 kwamen de mannen van Athene bij elkaar. Hier namen ze samen beslissingen. In Athene mocht lang niet iedereen meebeslissen. De vrouwen mochten niet meedoen aan de democratie. Kinderen deden niet mee en ook buitenlanders niet. Tenslotte mochten slaven ook niet meedoen.

Alleen vrije, volwassen Atheense mannen mochten over de politiek beslissen. Dit waren de mannen die in het Atheense leger vochten. Zij betaalden zelf hun wapens. Een soldaat had veel nodig. Hij droeg een bronzen helm, een borstharnas, bronzen scheenbeschermers en een groot rond schild. Soldaten waren heel belangrijk, DAAROM mochten zij stemmen. Op bron 9 zie je hoe soldaten zich klaarmaken voor de strijd. Arme mannen telden eerst niet mee. Ze konden geen wapens betalen, en DUS konden ze niet voor de stad vechten. Dat veranderde TOEN de Atheense leider Themistokles een vloot liet bouwen. Door de vloot werd Athene nog machtiger. De arme mannen van Athene werden de roeiers op de oorlogsschepen. De roeiers waren ook belangrijk voor Athene. Voortaan mochten zij ook meebeslissen.

Veertig keer per jaar werden de vrije Atheense mannen uitgenodigd om naar de volksvergadering te komen. Daar beslisten de mannen hoe het verder moest met de stad. Bij een belangrijke vergadering waren er wel 6000 mannen aanwezig. Iedereen mocht een toespraak houden, MAAR dat deden maar een paar mannen. Zij werden de leiders van Athene. Een bekende leider was Perikles. Het was zijn idee om het Parthenon te bouwen. De Atheners waren bang dat iemand te veel macht kreeg. Dan was er geen democratie meer. DAAROM hielden de Atheense mannen elk jaar een speciale stemming. Iedereen kreeg een potscherf. Op die scherf schreef je de naam van iemand die volgens jou te veel macht had. De man die meer DAN de helft van de stemmen kreeg, mocht tien jaar lang niet meer in Athene komen. Dat was een hele zware straf, MAAR Athene bleef zo wel een democratie.

CAUSALE CONNECTIEVEN TEMPORELE CONNECTIEVEN

LEGENDA:

CONTRASTERENDE CONNECTIEVEN VERGELIJKENDE CONNECTIEVEN

Bijlage 3: Economie havo In Balans II 2 & Geschiedenis havo Geschiedeniswerkplaats 1

Mensen met een inkomen onder de loongrens voor de ziekenfondsverzekering (in 2002 € 30.700,- bruto per jaar) zijn verplicht verzekerd tegen ziektekosten op grond van de ziekenfondswet (ZFVV). Ze zijn DUS via de sociale verzekeringen tegen ziektekosten verzekerd.

5 Werknemers met een inkomen boven de loongrens voor de ziekenfondsverzekering moeten/kunnen zich particulier verzekeren tegen ziektekosten. Bekende particuliere ziektekostenverzekeraars zijn OHRA, FBTO en Zilveren Kruis. In dit geval is de sociale zekerheid DUS via de particuliere markt georganiseerd. Om de kosten van het stelsel van sociale zekerheid te verminderen, probeert de overheid de laatste jaren sociale verzekeringen over te hevelen naar de marktsector. We spreken DAN van privatisering. Een voorbeeld hiervan is de overheveling van de uitvoering van de ZW naar de marktsector in 1996. Sinds 10 1996 moeten werkgevers hun zieke werknemers op eigen kosten een in de cao overeengekomen percentage van hun brutoloon doorbetalen. Werkgevers kunnen zich hiertegen verzekeren bij een particuliere ziektekostenverzekeraar. Des te lager het ziekteverzuim in een bepaald bedrijf is, des te lager wordt de premie. Op die manier dalen de kosten van het stelsel van de sociale zekerheid. Bovendien worden werkgevers geprikkeld om de arbeidsomstandigheden voor hun personeel te verbeteren. Daar 15 staat tegenover dat werkgevers bij de selectie van nieuw personeel beter op de gezondheid van sollicitanten zullen letten.

BINNENKORT wordt de WAO waarschijnlijk geprivatiseerd. De regeling wordt DAN vergelijkbaar met de privatisering van de ZW.

20 De ANW is een voorbeeld van een gedeeltelijke overheveling van de sociale zekerheid van de collectieve sector naar de marktsector. Vroeger kregen alle nabestaanden van overleden kostwinners een sociale uitkering op basis van de AWW (Algemene weduwe- en wezenwet). Sinds 1997 geldt dat alleen voor een bepaalde groep nabestaanden. Degenen die niet meer in aanmerking komen voor een (volledige) ANW-uitkering, moeten tegenwoordig zelf beslissen of ze zich particulier verzekeren tegen de inkomensderving die ontstaat als de kostwinner overlijdt.

In de Sovjet-Unie en nazi-Duitsland was de propaganda overal. Alles werd versimpeld tot de tegenstelling 'goed-kwaad', 'vriend-vijand', 'wij-zij'. Tegenstanders werden met grove middelen verdacht gemaakt, de eigen maatschappij en beweging werden verheerlijkt, dezelfde boodschap werd eindeloos herhaald. De 5 propaganda was des te effectiever OMDAT hij werd gecombineerd met onderdrukking. Het uiten van afwijkende ideeën was gevaarlijk. Het was moeilijk of zelfs onmogelijk om onafhankelijke informatie te krijgen. Het effect van de propaganda werd verder vergroot door het gebruik van nieuwe communicatiemiddelen. Vooral radio en film bleken geschikt voor het beïnvloeden van de massa. Hitler bracht met zijn toespraken niet alleen het publiek dat op hem was afgestroomd in vervoering, **MAAR** 10 bespeelde het hele volk DOORDAT hij overal voortdurend op de radio te horen en in de bioscoop was te zien. De communisten en de nazi's maakten honderden propagandafilms. Een berucht voorbeeld is de nazifilm Der Ewige Jude. Die film vertelde dat joden in Duitsland normale mensen leken DOORDAT ze hun ware aard verborgen. Ze pasten zich aan het 'gastvolk' aan, ZODAT ze het in het geniep konden verzwakken en leegzuigen. Op een kaart met bewegende pijlen werd de verspreiding van joden 15 uitgebeeld, 'een wereld omspannend spinnenweb'. Volgens de commentaarstem was er een 'verbluffende parallel met de massale migratie van een al even rusteloos dier, de rat'. Minutenlang krioiden vervolgens ratten over het scherm. 'Waar ze ook opduiken, brengen ze vernieling, verderven ze voedsel, verspreiden ze ziektes. Ze zijn vals, laf en wreed, de verwezenlijking van het stiekeme... ZOALS de joden onder de mensen.' Niet alleen de film, alle kunst stond in het teken van de propaganda. In de Sovjet-Unie was de 20 enige toegestane stijl het 'socialistisch realisme'. Kunstenaars moesten de werkelijkheid 'realistisch' voor de massa begrijpelijk weergeven en tegelijk verwijzen naar de stralende toekomst. En dus toonden schilders lachende of vastberaden in de verte kijkende arbeiders en boerinnen met achter zich de opgaande zon of het vaderlijke portret van Stalin.

25 De leider werd eindeloos verheerlijkt. In de Sovjet-Unie hing Stalins portret overal. Het land stond vol reusachtige beelden van hem. Steeds weer werd hij geprezen ALS 'zon der mensheid, gids naar de toekomst, geniale geleerde of juist bescheiden man van het volk. In Duitsland werd Hitler als afgod vereerd.

LEGENDA:	<u>CAUSALE CONNECTIEVEN</u>	<u>TEMPORELE CONNECTIEVEN</u>
	<u>CONTRASTERENDE CONNECTIEVEN</u>	<u>VERGELIJKENDE CONNECTIEVEN</u>

Bijlage 4: Economie havo Praktische Economie 4 & Geschiedenis havo Pharos 4

5 [1] Verzekeringen zijn geen gewoon product][2] zoals een scooter of een paar schoenen.][3] Een verzekering betaalt een schade uit wanneer een schade zich voordoet.][4] Of die schade zich zal voordoen is niet zeker.][5] Alleen de kans op schade is bekend.][6] Hierin verschilt een verzekering van andere producten.][7] Als je een scooter koopt, weet je zeker dat je er op gaat rijden.][8] En een paar schoenen schaf je aan om te dragen.]

10 [9] Een verzekeraar kent de gemiddelde kans op een schade.][10] Deze gemiddelde kans is gebaseerd op het gedrag van heel veel consumenten.][11] Dat][12] in het fietsenvoorbeeld][11a] één op de tien fietsen gestolen wordt, is een ervaringsgegeven.][13] Als je op heel veel scholen vraagt hoeveel fietsen er zijn en hoeveel fietsen er gestolen worden, dan blijken dat er één op de tien te zijn.][14] Maar die gemiddelde kans zegt niet zoveel over een individuele leerling.][15] De ene leerling, Azra, zet haar fiets netjes op slot][16] en stalt haar fiets bij voorkeur in een bewaakte fietsenstalling.][17] De kans dat haar fiets gestolen wordt, is minder dan één op tien.][18] De andere leerling, Ingmar, zet zijn fiets nooit op slot][19] en zal zijn fiets niet stallen in een bewaakte fietsenstalling.][20] De kans dat zijn fiets gestolen wordt is groter dan één op tien.]

15 [21] Een verzekeraar heeft daardoor een informatieachterstand.][22] Hij kent alleen de gemiddelde kans op schade][23] en dus kent hij alleen het gemiddelde risico.][24] Hij weet niets over het risico van een individuele leerling.][25] Anders gezegd: er is een verschil in de hoeveelheid informatie die de verzekeraar heeft en die een individuele verzekeringsnemer heeft.][26] op het moment dat de verzekering wordt afgesloten.][27] Dit verschil heet informatie-asymmetrie.][28] de verzekeringsnemer heeft meer informatie over het te verzekeren risico dan de verzekeraar.][29] Ingmar weet van zichzelf dat hij slordig is,][30] de verzekeringsmaatschappij weet dat niet van hem.]

20

5 [1] Als het schip eenmaal was uitgevaren][2] loerden er allerlei levensbedreigende gevaren.][3] Het eerste gevaar was onvoldoende of slecht voedsel.][4] Op een schip moest voor maanden voedsel meegenomen worden,][5] zonder over andere conserveringsmiddelen dan het zoutvat te beschikken.][6] De voorraad eten en drinken werd opgeslagen in ongekoelde,][7] dikwijls broeiend hete][6a] ruimten. [8] Op een schip konden de opvarenden noch tegen strenge koude,][9] noch tegen felle warmte behoorlijk beschermd worden. [10] Het eten bedierf in de snikhete ruimten snel][11] en de stank die dan ontstond was niet te harden.][12] In de ruimten was bovendien geen privacy.][13] Zeilschepen waren klein][14] en de laadruimte moest maximaal benut worden.][15] In een hoekje deelden de op elkaar gepropte zeevarenden][16] de ruimte met ratten, wormen, spinnen en kakkerlakken]. [17] Vooral 's nachts kwamen die op het rottend voedsel af][18] dat, hoewel het verboden was, door een kieskeurige matroos was weggegooid.]

10 [19] De meest gevreesde ziekte die optrad][20] door het slechte voedsel][19a] was scheurbuik.][21] Uit het journaal van Abraham Bogaert,][22] die in 1699 als chirurgijn op 'de Peperboom' werkzaam was] , [21a] komt de volgende beschrijving van scheurbuik:

15 '[22] Deze ziekte, zo dodelijk voor ons, vertoonde zich eerst met bruinblauwe stippelen over de geheele Beenen], [23] daarna in blauwe plekken,][24] waar op verderving van het tandvlees, stinkende Adem, loomigheid, sterk bloeden uit de neus en slaperigheid volgde][25]; hier wam nog bij aamborstigheid (kortademigheid) tot stikkens toe, ongestadige ademhaaling, [26] ontspanning der Hooftdeelen, benauwtheid,][27] kugh, verstijving der gewrichten, paarsche knobbels op de beenen, geduurigeflauwtens en felle koortsen.][28] Om dezen gevaarlijke vyand in den beginne 't hoeft te bieden, oordeelde ik dat men een Medicijn][29] – Brandwijn][28a] behoorde te maken,][29] waar van niet alleen de scheurbuikigen,][30] maar 't geheele scheepsvolk][29a] zou drinken][31] 't geen de schipper goedkeurde.']

20 [32] Dat amputatiezagen tot de standaarduitrusting van de scheepschirurgijn behoorden, zegt veel over de mogelijkheden][33] (of beter gezegd: onmogelijkheden) van de medici][34] aan boord van schepen.

25

LEGENDA: [1] = scheiding tussen informatie-eenheden

Bijlage 5: Geschiedenis havo Sprekend Verleden 2 & Geschiedenis vmbo Sfinx 2

[1] Na de dood van keizer Theodosius][2] in 395][3] werd het Romeinse rijk][4] onder zijn zonen][3a] verdeeld in een oostelijk en een westelijk deel]: [5] het Oost-Romeinse rijk zou nog tot 1453 blijven bestaan.][6] Toen werd het veroverd door de Ottomaanse Turken.][7]Het West-Romeinse rijk zou in 476 uiteengevallen][8] in een lappendeken van Germaanse staten.][9]Tijdens de laatste eeuwen van het Romeinse rijk trokken de Germaanse volken het rijk binnen.][10]Sommige trokken zich weer terug][11] andere bleven.][12] Een groot deel van West-Europa kwam in handen van Germanen.][13] Een deel van de Germanen][14]die zich in het Romeinse rijk vestigden,][13a] romaniseerde.][14]Het merendeel bleef de Germaanse cultuur [15] geheel of gedeeltelijk][14a] trouw.][16] De Germaanse volken hadden veel gemeenschappelijk.][17] Hun talen leken sterk op elkaar.][18] Zij leefden vooral van landbouw][19] en woonden in dorpen.][20] Elk Germaans volk was verdeeld in verschillende stammen.][21] In iedere stam had de vergadering van vrije mannen de meeste macht.][22] Hun samenleving was gelaagd][23] met slaven als onderste laag,][24] vrijgelatenen als middelste laag][25] en vrije mannen als bovenlaag.][26] Vrouwen en kinderen moesten hun man of vader volledig gehoorzamen.][27] In oorlogstijd werd een aanvoerder gekozen.][28] Sommige aanvoerders behielden hun macht ook in vreedstijd][29] en werden koning.][30] De Germanen vereerden [31] anders dan de christenen][30a]meer goden.][32] Die goden verbeeldden vaak natuurkrachten.][33] zoals de vruchtbaarheid (Freya),][34] regen en onweer (Donar)][35]of het vuur (Loki).][36] Deze goden stonden nogal onverschillig tegenover de mensen][37] en moesten tot aandacht en welwillendheid gebracht worden][38] met brandoffers][39] van fruit, graan en vlees.][40] Kleinere goden of natuurkrachten huisden en bomen, rivieren en bergen.][41] Elke Germaanse stam gaf zijn goden een eigen naam][42] in de eigen taal][43] en bewaarde eigen verhalen en rituelen.][44]Tius, Zius, Thor, Wodan en Donar waren verschillende namen voor de oppergod.][45] De ene stam verbond de oppergod met de zon,][46] de andere met de maan, de bliksem of de donder.][47] Soms was de oppergod een herder, soms een jager, soms een krijger.][48] Freya, Nehalennia of Tanfara waren namen voor de hoogste godin.][49] Bij de ene stam was zij maangodin,][50] bij de ander godin van de schoonheid, van de vruchtbaarheid of van de lente.][51] Bij sommige stammen was zij echtgenote van de oppergod,][52] bij andere niet.][53] In ruil voor offers verwachtte iedere stam van deze goden begunstiging.][54] Naast deze verschillen was er echter ook een overeenkomst]: [55] zowel christenen als Germanen geloofden in een leven na de dood.][56] Over dat leven na de dood verschilden zij echter weer van opvatting.]

[1]In Turkije lagen veel Griekse kolonies.][2] De Perzische koning was toen de baas in Turkije.][3] Hij wilde ook de baas zijn in de Griekse kolonies.][4] De kolonies vroegen toen hulp aan de steden in Griekenland.][5] Zo begon een lange oorlog.][6] De mensen van Athene waren bang.][7] De Perzische koning had immers een sterk leger.][8]Ze gingen naar het orakel van Delphi.][9] Dat was een heilige plaats.][10] Daar spraken de goden met de mensen.][11] Het orakel zei: "Vlucht naar het einde van de aarde.][12] De god Zeus geeft jullie muren van hout.][13] Deze muren zullen jullie redden.]"[14] De Grieken dachten lang na.][15] Toen bedachten ze dat Zeus met de 'muren van hout' hun oorlogsschepen bedoelde.][16] De leider van de Atheners bedacht een truc.][17] Hij stuurde een knecht naar de Perzische koning.][18]De knecht vertelde de koning het volgende: "Ik ben gestuurd door de leider van Athene.][19] Hij wil dat u de oorlog wint.][20] De andere Grieken weten hier niets van.][21] Ik moet van hem zeggen dat de Grieken bang zijn.][22] Ze zijn van plan om te vluchten.][23]Ze zullen niet meer vechten.]"[24] De Perzische koning geloofde het verhaal.][25] Hij stuurde meteen oorlogsschepen naar Griekenland.][26] Maar de Grieken vluchtten natuurlijk niet.][27] Ze verstopten zich met hun oorlogsschepen][28]en wachtten tot de Perzen kwamen.][29] Toen de Perzen met hun schepen bij Griekenland kwamen, werden ze overvallen door de Grieken.][30] De Perzen hadden niet goed nagedacht over de aanval.][31] Ze schrokken van de Griekse schepen.][32] De voorste Perzische schepen draaiden om.][33] Ze wilden vluchten.][34] Maar achter hen lagen nog meer schepen.][35] En die wilden juist aanvallen.][36] Er ontstond paniek.][37] Voor de Grieken was het toen gemakkelijk om de Perzen te verslaan.][38] Oorzaken: Het verhaal van de truc is echt gebeurd.][39] Het verhaal is opgeschreven door Herodotus,][40] een Griekse schrijver.][41]Herodotus schrijft vooral over de oorzaken van de oorlog.][42]Hoe kwam het dat er oorlog kwam? [43] Als je over een gebeurtenis nadenkt, is het goed om eens na te denken over de oorzaken.][44] Hoe kwam het dat iets gebeurde?][45] En wat is de belangrijkste oorzaak geweest?]

LEGENDA: [1] = scheiding tussen informatie-eenheden

Bijlage 6: Economie vmbo Index 1 & Geschiedenis vmbo Pharos 5

Een belangrijke wedstrijd in de eredivisie kan tot heel wat problemen leiden. Er wordt extra politie ingezet om eventuele reischoppers op te pakken de politiemensen worden betaald door de overheid. Zo draait de overheid op voor de kosten van het handhaven van de openbare orde. Na afloop van de wedstrijd ligt er behoorlijk wat afval op straat en de gemeentereiniging komt het opruimen. Op die manier komen de kosten van het schoonmaken ook ten laste van de overheid. De overheid ontvangt zijn geld van de belastingbetalers. Omdat iedereen in de maatschappij belasting betaalt, noemen we deze kosten maatschappelijke kosten. De gehele maatschappij betaalt dus mee aan de kosten die maar door een kleine groep worden veroorzaakt.

Iedereen vindt een schoon milieu belangrijk. Maar niet iedereen is bereid om er ook wat voor te doen. De overheid probeert daarom via allerlei maatregelen het gedrag van de mensen te beïnvloeden. Zo kan de overheid bijvoorbeeld bepaalde zaken verbieden, zoals het gebruik van schadelijke stoffen. Dat is een behoorlijk strenge maatregel. Het kan ook anders. De overheid kan bedrijven ook extra belasting laten betalen voor het gebruik van schadelijke stoffen. De bedrijven willen deze extra kosten terugverdienen en zullen hun producten duurder maken. Als de producten duurder zijn, zullen de consumenten minder van deze producten kopen. De bedrijven hoeven dan ook minder van deze producten te maken, waardoor er minder schadelijke stoffen worden gebruikt. En dat is natuurlijk wat de overheid wil bereiken.

De extra belasting die de bedrijven moeten betalen voor het gebruik van schadelijke stoffen noemen we milieuheffingen. Zo is er bijvoorbeeld de ecotax, een milieuheffing op energieverbruik. Bedrijven die veel energie gebruiken, betalen meer ecotax dan bedrijven die zuiniger met energie omgaan. Op die manier probeert de overheid er voor te zorgen dat er zuiniger met energie wordt omgegaan. Dat is beter voor het milieu. Alle maatregelen die de overheid neemt voor een schoner milieu horen bij het milieubeleid van de overheid.

Hoe werd **ATHENA** de godin van **ATHENE**? De goden **ATHENA** en **POSEIDON** wilden allebei de god van **ATHENE** zijn. De Atheners mochten zelf kiezen wie hun god werd. **POSEIDON** en **ATHENA** gaven allebei een cadeau aan de Atheners. Poseidon gaf een paard en **ATHENA** gaf een olijfboom. Aan een paard had je niet zoveel in heuvels rond **ATHENE**. De olijfboom was veel nuttiger. De olijven kon je eten en je kon er olijfolie van maken. De mannen van **ATHENE** kozen **ATHENA** als godin. De vrouwen wilden liever **POSEIDON**. Wat nu? De goden waren het met de mannen eens. Zij beslisten dat **ATHENA** de godin werd van **ATHENE**. De vrouwen mochten nooit meer kiezen. Voor **ATHENA** werd een tempel gebouwd: het **PARTHENON**. Daarin kwam een groot beeld van **ATHENA**, zoals je ziet op bron 5. De Grieken hadden veel verhalen over goden en fantasiewezens. Die verhalen noemen we mythes.

Eén van de grote helden van **ATHENE** was **THESEUS**. Hij hielp **ATHENE** tegen koning **MINOS** van het eiland **KRETA**. Ieder jaar moest **ATHENE** zeven jongens en zeven meisjes naar **KRETA** sturen. Zij werden opgegeten door de **MINOTAUROS**. Dat was een monster met het lijf van een mens en de kop van een stier. De **MINOTAUROS** leefde in een groot doolhof zoals op bron 6. Niemand kon daar de weg vinden. **THESEUS** kreeg hulp van de dochter van koning **MINOS**: **ARIADNE**. **ARIADNE** gaf **THESEUS** een rode draad. Daarmee kon hij de weg in het doolhof zoeken. **THESEUS** doodde de **MINOTAUROS**. **ATHENE** was bevrijd!

Al eeuwenlang bestaat er een beroemd verhaal over de **GRIEKEN**. Het is het verhaal van de oorlog tegen **TROJE**. **TROJE** was een stad in **TURKIJE**. De **GRIEKEN** deden er tien jaar over om de stad te veroveren. De beroemdste held van de **GRIEKEN** was **ACHILLES**. Bijna niemand kon hem doden. Een andere Griekse held was **ODYSSEUS**. Hij is bekend geworden door zijn slimheid. De verhalen over de Trojaanse oorlog zijn opgeschreven door de dichter **HOMEROS**. De verhalen worden nu nog veel gelezen. De **GRIEKEN** beeldden de oorlog tegen **TROJE** af op het **PARTHENON**.

LEGENDA: Eigennamen zijn VETGEDRUKT

Bijlage 7: Economie havo Praktische Economie 5 & Geschiedenis havo Feniks 6

Door de economische conjunctuur wisselen goede tijden en slechte tijden elkaar af. Dat is niet zonder gevolgen. Het raakt alle onderdelen van de economie. In slechte tijden is de vraag laag waardoor bedrijven minder gaan produceren. Daardoor hebben ze minder arbeid nodig en verliezen mensen hun baan. Als mensen ontslagen onvrijwillig werkloos worden, is dat niet hun keuze. Ze worden onvrijwillig werkloos. Als mensen zelf ontslag nemen kiezen ze voor werkloosheid. Ze worden vrijwillig werkloos. Misschien heeft iemand genoeg gespaard om er een tijdje tussenuit te kunnen. Vrijwillig of onvrijwillig, in een periode van laagconjunctuur loopt het werkloosheidspercentage op. Een stijgende werkloosheid heeft twee gevolgen voor de begroting van de overheid. Ten eerste moet de overheid meer uitkeringen verstrekken, waardoor de overheidsuitgaven stijgen. Ten tweede komt er minder loonbelasting binnen, omdat er minder mensen aan het werk zijn. Hierdoor dalen de overheidsinkomsten. Door beide effecten verslechteren de overheidsfinanciën en neemt de kans op een financieringstekort toe. Een financieringstekort beperkt de keuzes die de overheid kan maken, net zoals een lager gezinsinkomen de keuzes van een gezin beperkt. Een belangrijke keuze die de overheid ieder jaar moet maken betreft de stijging van sociale uitkeringen. Door inflatie kun je met dezelfde nominale uitkering namelijk steeds minder kopen. In 1980 was je rijk als je iedere maand € 1.000 kreeg. Daar kon je toen veel voor kopen. Nu kun je met € 1.000 per maand nauwelijks rondkomen. De overheid bekijkt daarom ieder jaar of de sociale uitkeringen kunnen stijgen. Als de uitkeringen stijgen met hetzelfde percentage als de inflatie, is er sprake van een waardevervalende uitkering. Met de gestegen uitkering kun je dezelfde hoeveelheid producten blijven kopen. Wat je kocht is duurder geworden, maar je uitkering is met hetzelfde percentage gestegen. Ondanks de inflatie ga je er dus niet op achteruit. De koopkracht is gelijk gebleven. In tijden van laagconjunctuur kan de overheid waarschijnlijk geen waardevervalende uitkeringen garanderen. Daar is simpelweg geen geld voor.

De **EERSTE WERELDOORLOG** en de **TWEDE WERELDOORLOG** begonnen op verschillende wijze. Terwijl in 1914 de **EUROPEANEN** bijna enthousiast de oorlog in gingen, riepen de herinneringen aan de **EERSTE WERELDOORLOG** 22 jaar later gevoelens van angst op. De vier jaar lange strijd tussen 1914 en 1918, die grotendeels door soldaten in de loopgraven was gevoerd, had 8.500.000 doden opgeleverd. En dat lag in **SEPTEMBER** 1939, aan het begin van de **TWEDE WERELDOORLOG**, bij veel mensen nog vers in het geheugen. In de jaren die voorafgaan aan de **TWEDE WERELDOORLOG** hadden regeringsleiders en diplomaten uit met name **GROOT-BRITANNIË** en **FRANKRIJK** er alles aan gedaan om conflicten met **DUITSLAND** te voorkomen. **HITLER** stuurde na zijn machtsovername in 1933 vrijwel meteen aan op een oorlog. **DUITSLAND** stapte uit de **VOLKENBOND** en **HITLER** verkondigde dat de **DUITSLERSLEBENSRAUM** nodig hadden. Dat daarmee de bepalingen uit het Verdrag van **VERSAILLES** onder druk kwamen, laptte hij aan zijn laars. De Duitse economie werd zo snel mogelijk op wapenproductie ingesteld en het leger werd uitgebreid. Omdat de Duitse economie zichzelf niet volledig kon voorzien, streefde Duitsland naar gebiedsuitbreiding. De Europese landen protesteerden hiertegen, maar waren als de dood voor een nieuwe oorlog. Zo konden de Duitse troepen in 1936 ongestraft het **RIJNLAND** binnentrekken, terwijl **HITLER** verkondigde dat de Duitse bevolking in de toekomst gevoed zou moeten worden met voedsel van de Oekraïense graanvelden, in het **OOSTEN** van **EUROPA**. **EUROPA'S** machtigste land, **GROOT-BRITANNIË**, voerde in de jaren dertig een politiek van appeasement; elke internationale confrontatie werd gemedend. Van **FRANKRIJK**, dat de ene na de andere regeringscrisis beleefde, had **DUITSLAND** ook niets te vrezen. **HITLER** lijfde op 13 **MAART** 1938 zonder slag of stoot buurland **OOSTENRIJK** in. Weliswaar werd nog geprobeerd om tijdens de conferentie van **MÜNCHEN**, eind september 1938, **HITLER** een halt toe te roepen, maar hij kreeg in ruil voor valse beloften wel Sudetenland van **TSJECHOSLOWAKIJE**. In tegenstelling tot alle afspraken veroverde Hitler in het voorjaar van 1939 vervolgens heel **TSJECHOSLOWAKIJE** en toen was het **GROOT-BRITANNIË** en **FRANKRIJK** duidelijk dat met deze Duitse leider geen afspraken te maken vielen. Terwijl de Duitse minister van Buitenlandse Zaken **VONRIBBENTROP** met zijn Russische ambtsgenoot **MOLOTOV** onderhandelde over een niet-aanvalsverdrag (**AUGUSTUS** 1939), hoopten de **FRANSEN** en Engelsen dat er geen herhaling zou komen van de periode 1914-1918. Een week later begon de Duitse invasie in **POLEN**.

LEGENDA: Eigennamen zijn VETGEDRUKT

Bijlage 8:Opbouw van het corpus

Tabel A: havo-boeken

Vak	Auteur	Bron	Jaar	Lengte
EC	Haak, J.K. van den	Economie in Balans, 2e fase totaalvak havo, theorieboek 1, Nijgh Versluys, p.20-21	2002	330
EC	Haak, J.K. van den	Idem; p.60-61	2002	328
EC	Haak, J.K. van den	Idem; p.75	2002	308
EC	Haak, J.K. van den	Idem;, p.195-196	2002	306
EC	Haak, J.K. van den	Idem; p.268-269	2002	329
EC	Haak, J.K. van den	Idem; p.14-15	2002	298
EC	Haak, J.K. van den	Idem; p.56-57	2002	311
EC	Haak, J.K. van den	Idem; p.154-155	2002	302
EC	Haak, J.K. van den	Idem; p.177-179	2002	304
EC	Haak, J.K. van den	Idem; p.204-205	2002	334
EC	Hinloopen, J., Adriaansen, P., Zuiderwijk, A.	Praktische economie, havo voor de 2e fase, 5e druk, p.9-10	2009	359
EC	Hinloopen, J. e.a.	Idem; p.89	2009	383
EC	Hinloopen, J. e.a.	Idem; p.143-144	2009	375
EC	Hinloopen, J. e.a.	Idem; p.195	2009	301
EC	Hinloopen, J. e.a.	Idem; p.283	2009	321
EC	Kentson, J., Janssen, K., van de Donk, W.	Transactie havo voor de 2e fase deel 2, Meulenhoff Educatief, p.26	1998	321
EC	Kentson, J. e.a.	Idem; p.51	1998	415
EC	Kentson, J. e.a.	Idem;, p.62	1998	348
EC	Kentson, J. e.a.	Idem; p.66	1998	318
EC	Kentson, J. e.a.	Idem; p.75	1998	304
EC	Duijm, H., Gorter, G.F..	Percent havo, Totaalvak 1, Deel theorieboek, ThiemeMeulenhoff, p.99	2002	331
EC	Duijm, H., e.a.	Idem; p.126-127	2002	378
EC	Duijm, H., e.a.	Idem; p.173	2002	341
EC	Duijm, H., e.a.	Idem; p.188	2002	385
EC	Duijm, H., e.a.	Idem; p.193	2002	377
GS	Verkuil, D., Hijstek, B., van der Geugten, T. & Betten, E.	De Geschiedeniswerkplaats, Wolters-Noordhoff, p. 22	2006	300
GS	Verkuil, D. e.a.	Idem; p. 22	2006	339
GS	Verkuil, D. e.a.	Idem; p. 22	2006	299
GS	Verkuil, D. e.a.	Idem; p. 22	2006	561
GS	Verkuil, D. e.a.	Idem;, p. 22	2006	369
GS	Raymond de Kreek, Ludwich Verberne, Marian Veldkamp, Jos Venner, Albert Jan Bosch, Andre van Voorst, Jan van Oudheusden, Robert Boonstra, Cor van der Heijden, Maria van Haperen	Idem; p. 20	2007	280
GS	Raymond de Kreek, e.a.	Feniks, Havo tweede fase, overzicht v/d geschiedenis. p. 42	2007	278
GS	Raymond de Kreek, e.a.	Idem; p. 64	2007	384
GS	Raymond de Kreek, e.a.	Idem; p. 150	2007	377
GS	Raymond de Kreek, e.a.	Idem; p. 198	2007	334
GS	Raymond de Kreek, e.a.	Idem; p. 212	2007	393
GS	Raymond de Kreek, e.a.	Idem; p. 236	2007	362
GS		http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p. 68	2012	317
GS		http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p. 69	2012	586
GS		http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm pp. 70-71	2012	547
GS		http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p. 73	2012	331

GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.74		2012	635
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.76		2012	322
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.78		2012	323
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.79		2012	526
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm pp. 80-81		2012	378
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.83		2012	285
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.84		2012	464
GS	http://www.malmberg.nl/Voortgezet-onderwijs/Methodes/Mens-Maatschappij/geschiedenis.htm p.85		2012	615
GS	Hageraats, B, van der Heyden, C., van Oudheusden, J, van de Pol, L., Raaijmakers, J., Rongen, W., Salman, J., Schuitemaker, P., van Voorst, A & van der Kaap, A.	Pharos themaboek voor de tweede fase (druk 1). Meulenhoff Educatief, pp. 12-13	1998	344
GS	Hageraats, B, e.a.	Idem;, pp. 44-45	1998	341
GS	Hageraats, B, e.a.	Idem;, pp. 55	1998	375
GS	Hageraats, B, e.a.	Idem; pp. 81	1998	303
GS	Hageraats, B, e.a.	Idem; pp. 162	1998	321
GS	Buskop, H., Dalhuisen, L., van der Geest, R & Steegh, F., Bastiaans, C. & de Waal, C.	Sprekend verleden 2e fase Handboek A (druk 2/ oplage 1) Nijgh Versluys, pp. 16-17	1998	374
GS	Buskop, H., e.a.	Idem; pp. 37	1998	400
GS	Buskop, H., e.a.	Idem; pp. 66-67	1998	353
GS	Buskop, H., e.a.	Idem; pp. 208	1998	326
GS	Buskop, H., e.a.	Idem; pp. 244-245	1998	347

Tabel B: vmbo-boeken

Vak	auteur	Bron	Jaar	Lengte
EC	Adriaansen, P.	Praktische Economie 2 vmbo kgt, Malmberg, p.16-17	2004	339
EC	Adriaansen, P.	Idem; , p.34-35	2004	335
EC	Adriaansen, P.	Idem; p.52-53	2004	343
EC	Adriaansen, P.	Idem; p.66-67	2004	326
EC	Adriaansen, P.	Idem; p.68-69	2004	326
EC	Adriaansen, P.	Idem; p.98-99	2004	327
EC	Huitema, J., Peters, L., Vaart, I. van der	Economisch bekeken, basisvorming vmbo-kgt handboek, zesde druk, Malmberg, p.6-7	2004	278
EC	Huitema, J. e.a.	Idem; p.44-45	2004	300
EC	Huitema, J. e.a.	Idem; p.62-63	2004	319
EC	Huitema, J. e.a.	Idem; p.76-77	2004	291
EC	Huitema, J. e.a.	Idem; p.80-81	2004	308
EC	Huitema, J. e.a.	Idem; p.90-91	2004	305
EC	Huitema, J. e.a.	Idem; p.92-93	2004	299
EC	Kruis, M.	Pincode onderbouw vmbo kgt leerboek, Noordhoff, p.12-13	2006	341
EC	Kruis, M.	Idem; p.42-43	2006	334
EC	Kruis, M.	Idem; p.54-55	2006	398
EC	Kruis, M.	Idem; p.110-111	2006	341
EC	Kruis, M.	Idem; p.144-145	2006	318
EC	Kruis, M.	Idem; p.200-201	2006	299
EC	Kruis, M.	Idem; p.204-205	2006	323
EC	Scholte, P., Janssen, K., Kuijpers, M., Voorend, P., Wevers, F.	Index vmbo kgt voor de basisvorming, eerste druk, ThiemeMeulenhoff, p.70-71	2003	322
EC	Scholte, P., e.a.	Idem; p.101	2003	325
EC	Scholte, P., e.a.	Idem; p.106	2003	333
EC	Scholte, P., e.a.	Idem; p.110-111	2003	327
EC	Scholte, P., e.a.	Idem; p.122	2003	328
GS	Inkijkexemplaar van de nieuwste druk 2012/ 2013 www.noordhoffuitgevers.nl	De Geschiedeniswerkplaats informatieboek deel 1	2012	338
GS	Idem	Idem	2012	346
GS	Idem	Idem	2012	314
GS	Idem	Idem	2012	395
GS	Idem	Idem	2012	414
GS	Inkijkexemplaar van de nieuwste druk 2012/ 2013 www.noordhoffuitgevers.nl	Memo handboek KGT deel 2	2012	415
GS	Idem	idem	2012	385
GS	Idem	idem	2012	423
GS	Idem	idem	2012	324
GS	Idem	idem	2012	307
GS	Idem	idem	2012	341
GS	van Bon, J., Hendriks, S., Lelieveld, J., Spree, M., Stroo, R., Veldkamp, M., Venner, J., Voogt, A., & van Voorst, A.	Pharos kgt deel 1 Thieme-Meulenhoff pp.36-37	2003	382
GS	van Bon, e.a.	Idem; pp.46-47	2003	338
GS	van Bon, e.a.	Idem; pp.122-123	2003	336
GS	van Bon, e.a.	Idem; pp.142-143	2003	420
GS	van Bon, e.a.	Idem; pp.34-35	2003	355
GS	Berents, D., Bos, J., Rombouts, G. & Veldkamp, M.	Sfinx leesboek 1 vbo (druk 1) Thieme, pp. 14-15	2000	357
GS	Berents, D, e.a.	Idem; pp. 28-29	2000	370
GS	Berents, D, e.a.	Idem; pp. 44-45	2000	333
GS	Berents, D, e.a.	Idem; pp. 64-65	2000	370
GS	Berents, D, e.a.	Idem; pp. 104-105	2000	323
GS	Buskop, H., Dalhuisen, L., van der	Sprekend verleden handboek 1 (druk 4)	2004	516

	Geest, R., Steegh, F. & van der Waal, C.	Nijgh Versluys pp. 42-43		
GS	Buskop, H., e.a.	Idem; 19-20	2004	300
GS	Buskop, H., e.a.	Idem; pp. 76-77	2004	378
GS	Buskop, H., e.a.	Idem; pp. 72-73	2004	320
GS	Buskop, H., e.a.	Idem; pp. 24-25	2004	330