

# VU Research Portal

## Advanced Learner Writing

Springer, P.E.

2012

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Springer, P. E. (2012). *Advanced Learner Writing: a corpus-based study of the discourse competence of Dutch writers of English in the light of the C1/C2 levels of the CEFR*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

## Samenvatting in het Nederlands (Summary in Dutch)

### Inleiding en achtergrond

Nederlandse studenten staan over het algemeen bekend om hun goede beheersing van het Engels en worden geacht om aan de universiteit een *near-native* niveau te halen tijdens hun opleiding. Tegelijkertijd groeit het besef dat op het gebied van academisch schrijven er nog een wereld te winnen is. Bovendien zijn er hoe langer hoe meer Nederlandse universiteiten die academisch schrijven in het Engels aanbieden, hetzij in de master opleiding, hetzij als onderdeel van de bachelor. Het is daarom van belang om uit te vinden wat Nederlandse leerders schrijven en hoe dat zich verhoudt tot teksten geschreven door moedertaalsprekers van het Engels om vervolgens uitspraken te kunnen doen met betrekking tot de implicaties voor het Nederlandse onderwijs in academisch Engels schrijven.

Tot voor kort ging de aandacht in tweedetaalverwervingsonderzoek uit naar beginnende leerders. Sinds de introductie van computer leerdercorpora begin jaren 90, is er meer onderzoek gekomen naar vergevorderde leerders. Er is echter nog relatief weinig corpusonderzoek gedaan naar de taal van vergevorderde leerders van het Engels met Nederlands als eerste taal (L1).

Het doel van het onderzoek zoals dit is beschreven in dit proefschrift is het achterhalen welke lexico-grammaticale middelen een schrijver zich eigen moet hebben gemaakt om op het hoogste niveau te schrijven en in welke mate studenten Engels deze middelen inzetten, al dan niet op juiste wijze, in verschillende stadia van hun studie, om vervolgens vast te kunnen stellen welke aanpassingen nodig zijn in de schrijfvaardigheidscurricula om de ontwikkeling van de discourse competentie van de leerders in de richting van het hoogst haalbare niveau te sturen. Hiervoor is een stapsgewijze aanpak nodig. Allereerst moet worden vastgesteld wat het betekent om een vergevorderde leerder te zijn. Met andere woorden, welke definitie van vergevorderd wordt er gehanteerd en welke competenties horen daar bij? De descriptoren voor de twee hoogste niveaus, C1 en C2, van het alomtegenwoordige en breed gedragen Gemeenschappelijk Europees Referentiekader voor Moderne Vreemde Talen: Leren, Onderwijzen, Beoordelen, kortweg Europees Referentiekader, (hierna: ERK) zijn gekozen als uitgangspunt voor deze omschrijving.

Vervolgens moet er worden vastgesteld wat een vergevorderde leerder moet kunnen volgens het ERK met betrekking tot schrijfvaardigheid. Dit lijkt eenvoudiger dan het is. Het ERK is niet een eenvoudig en transparant document en bevat veel inconsistenties met betrekking tot terminologie en omschrijvingen. Het ERK formuleert taalvaardigheid in zogeheten '*can-do statements*', zoals de volgende voor C2 schrijfvaardigheid:

Kan heldere, soepel lopende complexe teksten schrijven in een gepaste en doelmatige stijl en met een logische structuur die de lezer helpt om belangrijke punten te herkennen. (Nederlandse Taalunie 2008: 59)

Echter, het ERK geeft niet aan welke lexico-grammatica de leerder moet beheersen om dit voor elkaar te krijgen. Deze lacune maakt de volgende stap nodig, namelijk het vertalen van deze omschrijving naar specifieke lexico-grammaticale middelen om een systematische corpusanalyse mogelijk te maken van teksten geschreven door de Nederlandse leerders.

## Methode

Allereerst is er vastgesteld wat de hoofdkenmerken van het schrijven op C1/C2 niveau zijn en vervolgens werd geconstateerd dat hetgeen de C1/C2 niveaus onderscheidt van de lageregelegen niveaus een goed ontwikkelde discourse competentie is. Discoursecompetentie werd gedefinieerd als: ‘de kennis en kunde om de lexico-grammaticale middelen die beschikbaar zijn in een taal te gebruiken om soepel lopende, samenhangende (*cohesive* en *coherent*) tekst te produceren in een gegeven genre door zinnen logisch te ordenen en het gebruik van de juiste verbindingswoorden’. Om deze definitie te operationaliseren in het corpusonderzoek zijn er lexico-grammaticale constructies geannoteerd in de volgende vier hoofdcategorieën: *attention getting / focusing devices* (1), *backgrounding / clause combining devices* (2), *cohesive devices* (3), en *stance* (4). Die vervolgens weer uiteenvallen in vele subcategorieën. Uit elk van de vier hoofdcategorieën volgt hieronder een voorbeeld:

- 1) ***On board were a doctor and two nurses.***
- 2) ***Although Peter Jackson made a very good trilogy, the time-limit does not allow the films to represent the novels fairly.***
- 3) ***Last but not least, television is a great thing to watch.***
- 4) ***The internet, undoubtedly, can have negative effects on children’s development.***

Alle geannoteerde categorieën zijn ook geannoteerd met betrekking tot hun positie in de zin, bijv. vooraan, in het midden, of aan het eind van de zin, en daar waar relevant ook onderverdeeld op basis van semantische kenmerken. Een volledig overzicht van alle annotatie is te vinden in Appendix IV.

De corpusanalyse van leerderteksten en teksten geschreven door professionele moedertaalsprekers is het centrale en meest uitgebreide onderdeel van het onderzoek. Ten behoeve van het onderzoek zijn er twee corpora samengesteld en geannoteerd (van in totaal bijna 300.000 woorden). Alle teksten in de twee hoofdcorpora, VULCAN (VU Learner Corpus of Advanced writing) en VUNSPRAC (VU Native Speaker Published Research Article Corpus) zijn in zijn geheel toegevoegd (geen fragmenten) en handmatig geannoteerd. Vervolgens is alle annotatie gekwantificeerd en met behulp van de statistische software SPSS zijn deze resultaten geanalyseerd. Elk corpus bestaat uit twee subcorpora: het moedertaalspreker corpus bestaat uit een taalkundig- en een letterkundig subcorpus; het leerdercorpus bestaat uit een eerstejaars- en een derdejaars subcorpus. Een uitgebreid overzicht van de opbouw van de corpora is te vinden in sectie 3.3.3.

Het hoofddoel van het onderzoek is om op systematische wijze –met gebruik van authentieke taaldata uit representatieve corpora, niet op basis van intuïtie– te beschrijven waarin het Engels van Nederlandse vergevorderde leerders afwijkt van dat van professionele moedertaalsprekers.

Aannemend dat vergevorderde leerders zich de regels van de Engelse morphosyntaxis eigen hebben gemaakt en het dus niet fouten zijn of idiosyncratisch gebruik van het Engels is dat problemen oplevert, wordt de analyse gemaakt op basis van het onder- en overgebruik van bepaalde lexico-grammaticale middelen van de leerders ten opzichte van de moedertaalsprekers. De analyse is dus voornamelijk kwantitatief van aard, maar ook, waar nodig, kwalitatief. Daarnaast bevat het proefschrift een vervolgonderzoek op basis van een analyse van bestaande leerboeken op het gebied van academisch Engels (Hoofdstuk 6).

## Bevindingen uit het corpusonderzoek

Een belangrijke bevinding van dit onderzoek is dat Nederlandse leerders op veel fronten zich dicht in de buurt van de moedertaalsprekers bevinden. Bijvoorbeeld op het gebied van het

gebruik van adverbiale connectoren, zoals *however*, *firstly*, en *for example*, is er geen verschil wat betreft de relatieve frequentie en positie in de zin tussen de leerders en de moedertaalsprekers. Hierin onderscheiden de Nederlandse leerders zich van vergevorderde leerders met een andere eerste taal dan het Nederlands, blijkt uit de literatuur. Maar ook op andere gebieden was een grondige kwantitatieve analyse nodig –vaak gekoppeld met een kwalitatieve analyse– voordat er verschillen aan het licht kwamen.

De corpusanalyse bracht verschillende gebieden van over- en ondergebruik aan het licht. Sommige subtiel, maar statistisch significant, zoals het feit dat de leerders over het algemeen voltooide niet-finiete bijwoordelijke bijzinnen (*When consulted in their conventional printed form, parallel texts also present a number of pitfalls*) over het geheel genomen ondergebruiken, maar daarentegen weer overgebruiken aan het begin van de zin. Andere verschillen zijn overduidelijk, zoals het overgebruik van bijzinnen die beginnen met het onderschikkend voegwoord *because* en het overgebruik van de causale betekenisrelatie in het algemeen. Hoofdstuk 4 geeft een uitgebreide kwantitatieve analyse van de corpusanalyse en deze wordt samengevat in Table 101 op bladzijde 158. Er worden een aantal trends geconstateerd met betrekking tot semantische categorieën, positie en syntactische complexiteit, waarvan hieronder een korte samenvatting.

Een van de duidelijkste uitkomsten van de kwantitatieve analyse was dat de leerders een aantal lexico-grammaticale middelen overgebruiken die uitdrukking geven aan een causale betekenisrelatie, met name adverbiale bijzinnen van reden en oorzaak en adverbiale connectoren van reden of gevolg(trekking), zoals *therefore* en *because of X*. Ook andere constructies die beginnen met *because of* worden overgebruikt, net als lexicale formules met een vergelijkbare betekenis, zoals *this means that*. Een ander zeer duidelijk geval van overgebruik betreft indicatoren van opeenvolging, zoals bijvoorbeeld bijwoorden als *firstly*, *finally*, maar ook bijwoordelijke bijzinnen van tijd (zowel finiete als non-finiete) en het overgebruik van de determinator *another*, in terugverwijzende naamwoordgroepen, de zogenoemde *retrospective labels*.

Met betrekking tot positie is het duidelijk dat de meeste leerders een voorkeur hebben voor het begin van de zin als de mogelijkheid zich aandient om een constituent toe te voegen aan de zin. Dit blijkt onder andere door het overduidelijke overgebruik van beperkende betrekkelijke bijzinnen en nominale apposities in het gedeelte van de zin voor de persoonsvorm en het overgebruik van non-finiete bijzinnen aan het begin van de zin, relatief gezien (hiermee wordt bedoeld dat deze minder ondergebruikt worden aan het begin van de zin dan in het midden of aan het eind van de zin). Het midden van de zin wordt door de leerders veel minder gebruikt en deze negatieve voorkeur om de zin te onderbreken om (achtergrond)informatie toe te voegen wordt weerspiegeld door het ondergebruik van parenthetische zinnen.

Ook wat betreft zinscomplexiteit zijn er belangrijke verschillen te constateren tussen de leerders en de professionals. Zinnen in het moedertaalcorpus zijn over het algemeen langer en complexer dan de zinnen in het leerdercorpus. Van de eerstejaars naar de derdejaars naar het taalkundig subcorpus (VUNSPRAC\_ling) naar het letterkundig subcorpus (VUNSPRAC\_lit) neemt de gemiddelde lengte van de bijzinnen toe en daarmee de gemiddelde zinslengte. Dezelfde 'route' laat een toename zien van simpele zinnen, d.w.z. zinnen die geen bijzinnen of apposities bevatten en een toename van zinnen met drie of meer van deze lexico-grammaticale constructies. Deze volgorde is zichtbaar in veel van de geobserveerde frequenties van de onderzochte lexico-grammaticale constructies. Dit houdt ook in dat er dus verschillen zijn tussen VUNSPRAC\_ling en VUNSPRAC\_lit. VUNSPRAC\_lit bevat twee keer zo veel zinnen die met *and* of *but* beginnen dan VUNSPRAC\_ling. Echter, het werd ook vastgesteld dat hoewel het belangrijk is om op te merken dat de twee subcorpora op belangrijke punten significant van elkaar verschillen (genoeg om niet uit het oog te verliezen

dat het hier om twee verschillende subgenres van academisch Engels gaat), dit verschil geen fundamenteel probleem vormde voor de kwantitatieve analyse van het over- en ondergebruik van de leerders.

Hoofdstuk 5 biedt een bespreking van een selectie van de resultaten uit Hoofdstuk 4 alsmede een kwalitatieve analyse. De uitkomsten van deze analyse vormen de basis voor een systematische groepering van het over- en ondergebruik in vier dimensies van discourcecompetentie. Deze vier dimensies zijn:

- 1) Beknoptheid (*concision*): van een overgebruik van lange constructies die de bedoeling van de schrijver expliciteren, naar een balans tussen lange en compacte constructies zonder aan duidelijkheid in te boeten;
- 2) Zelfvertrouwen (*authorial confidence*): van algemeenheid naar interactiviteit en specificiteit, door de discourse waar nodig te onderbreken en/of aan te vullen, daarbij gebruikmakend van alle mogelijke posities in de zin;
- 3) Complexiteit (*sophistication*): van minder naar meer complex, zowel cognitief, d.w.z. dat schrijvers aan het complexere uiteinde van deze dimensie een groter bereik hebben van semantische betekenisrelaties –en daarmee een grotere mate van semantische precisie bereiken– dan leerders aan het ‘simpelere’ uiteinde van deze dimensie, als tekstueel, d.w.z. dat bepaalde teksttypen meer gelegenheden bieden tot het gebruik van specifieke lexico-grammaticale constructies;
- 4) Mate van formaliteit (*degree of formality*): van Engels dat lijkt op (informele) spreektaal naar tekst die geschikt is in een formele geschreven context.

Dimensie 1, beknoptheid, heeft betrekking op het vermogen van een schrijver om doelmatige teksten te schrijven zonder alles te expliciteren en wordt bij de leerders gekenmerkt door het overgebruik van finiete bijwoordelijke bijzinnen (de lange variant) gekoppeld aan het ondergebruik van non-finiete bijwoordelijke bijzinnen (de compacte variant), alsook het overgebruik van beperkende betrekkelijke bijzinnen (de lange variant) gekoppeld aan het ondergebruik van nominale apposities en ascriptieve naamwoordgroepen (*ascriptive NP NCDUs*) (de compacte variant).

Dimensie 2, zelfvertrouwen, heeft betrekking op het vermogen van de schrijver zich te presenteren expert als de schrijver van zijn tekst op een zodanige wijze dat hij het ‘aandurft’ om blijk te geven van betrokkenheid bij de tekst door de interactie aan te gaan met de lezer door middel van het toevoegen van details en anderszins de stroom van de discourse te onderbreken met behulp van specifieke constructies. Deze dimensie wordt voor de leerders gekenmerkt door relatief overgebruik van non-finiete bijzinnen voorafgaand aan de hoofdzin en het ondergebruik van het midden van de zin, parenthetische zinnen, bijzinnen die verwijzen naar de bron van informatie (zowel inter- als intratekstueel), bijzinnen die met *as* beginnen (intratekstuele bronverwijzing), constructies die de mogelijkheid bieden om uit te weiden of informatie in andere bewoordingen te brengen, zoals *in other words*, en ascriptieve naamwoordgroepen (*ascriptive NP NCDUs*).

Dimensie 3, complexiteit, heeft betrekking op twee verschillende (maar tot op zekere hoogte overlappende) componenten: cognitieve complexiteit en complexiteit van tekst type en onderwerp. In sommige gevallen is het vrijwel onmogelijk om vast te stellen of een bepaald geconstateerd ondergebruik te wijten is aan de relatief lage complexiteit van het schrijven of aan het feit dat het teksttype weinig aanleiding geeft tot het gebruik van deze constructie. Deze dimensie wordt voor de leerders gekenmerkt door het ondergebruik van toegevend bijzinnen (bijv. bijzinnen die beginnen met het voegwoord *although* of *whereas*), en vergelijkbare adverbiale connectoren (bijv. *though*), het ondergebruik van bijzinnen die verwijzen naar de bron van informatie en een beperkt scala van *stance* bijwoorden. Het

overgebruik dat typerend is voor deze dimensie heeft betrekking op het overgebruik van indicatoren van opeenvolging, zoals bijzinnen van tijd en adverbiale connectoren van toevoeging (*secondly, additionally, etc.*), constructies die een betekenisrelatie van oorzaak of reden aangeven, maar ook het overgebruik van *focusing devices*, het overgebruik van voorwaardelijke en doelaanwijzende bijzinnen en het overgebruik van vage zelfstandige naamwoorden in verwijzende naamwoordgroepen.

Dimensie 4, mate van formaliteit, heeft betrekking op het vermogen van de schrijver om zich een formele stijl aan te meten die past bij het formele register waarin hij behoort schrijft. Hoewel het niet een hoofddoel was van het onderzoek en er ook niet bijzonder veel voorbeelden van te vinden waren, is het wel een kenmerk van leerdertaal waar in de literatuur vaak melding van wordt gemaakt (bijv. De Haan 2007; Hinkel 2002, 2003; Hoogesteger 1998; Wiktorsson 2003: 110, 121) Het vermogen om tekst te hanteren in termen van stijl en register staat ook vermeld als onderdeel van discoursecompetentie<sup>74</sup> in het ERK (Nederlandse Taalunie 2008: 114) en is ook opgenomen in de definitie van discoursecompetentie die hierboven gegeven is. Hoewel dus niet talrijk, zijn er wel degelijk kenmerken te vinden in deze dimensie die voor de leerders van belang zijn. Deze dimensie wordt voor de leerders gekenmerkt door het ondergebruik van de formele appositie signalerende afkortingen *e.g.* (bijv.) en *i.e.* (d.w.z.) gekoppeld aan het overgebruik van *like* in dezelfde functie. Ook het overgebruik van *And* en *But* aan het begin van de zin, het overgebruik van *even though, while* als voegwoord voor bijzinnen van tijd (niet voor toegevend bijzinnen), van *wh-clefts* (talrijker in gesproken Engels dan in geschreven academisch Engels), van indicatoren voor causale betekenisrelaties en informele woordgroepen zoals *a lot of* en *lots of* zijn kenmerkend voor deze dimensie.

### **Bevindingen uit het vervolgonderzoek**

Aangezien veel van de gevonden verschillen tussen leerderteksten en teksten geschreven door moedertaalsprekers zo subtiel zijn dat ze alleen door corpusonderzoek aan het licht komen en door veel docenten niet bewust worden opgemerkt, kan er niet van leerders worden verwacht dat zij deze verschillen vanzelf oppikken door blootgesteld te worden aan grote hoeveelheden academische teksten. Sterker nog, onderzoeken tonen aan dat leerders zelfs veel meer in het oog springende kenmerken van academisch schrijven niet opmerken puur door middel van het lezen van deze teksten (Cobb 2003; Hinkel 2003, 2004). Expliciete instructie is derhalve nodig. Het is aangetoond dat expliciete instructie bij tweedetaalverwerving een positief effect heeft (zie o.a. Norris & Ortega 2000 voor een overzicht). Om die reden is het centrale corpusonderzoek uitgebreid met een analyse van gangbare leerboeken op het gebied van academisch schrijven in het Engels –om vast te stellen of en hoe deze voorzien in de behoeften van de leerders– op basis van de hierboven beschreven vier dimensies van discoursecompetentie.

Deze vier dimensies zijn uiteraard een uitvloeisel van het huidige onderzoek en komen daarom natuurlijk ook niet als zodanig in de leerboeken voor. Maar door middel van het inventariseren van de lexico-grammaticale middelen die behandeld worden in deze leerboeken is het toch mogelijk een uitspraak te doen over welke dimensies behandeld worden en hoe dit gebeurt.

Zoals verwacht is de meest voor de hand liggende dimensie, Dimensie 4 (formaliteit) de enige die als zodanig herkenbaar is in de geanalyseerde leerboeken. Alle leerboeken besteden er enige aandacht aan. Voor de hand liggend, omdat het feit dat academisch

---

<sup>74</sup> Vreemd genoeg wordt de oorspronkelijke term *discourse competence* op dezelfde pagina (114) en in dezelfde context eerst vertaald met *discoursecompetentie* en later met *gesprekscompetentie*. Ik houd de eerste vertaling aan daar deze hier van toepassing is.

schrijven formeel van aard is, alsook het feit dat beginnende schrijvers (zowel L1 als L2) hier moeite mee hebben, algemeen bekend en breed gedocumenteerd is. Maar ook hierbij kan de corpusanalyse relevante informatie toevoegen, als het gaat om het precies vaststellen van de verschillen tussen leerders en ervaren schrijvers. De meeste leerboeken lijken geen gebruik te hebben gemaakt van corpusonderzoek om hun beweringen te onderbouwen en in die gevallen waarin dat wel is gebeurd, is dat op basis van L1 corpora. Hoewel de meeste leerboeken dus aandacht besteden aan de verschillen tussen schrijftaal en spreektaal en/of tussen informeel en formeel taalgebruik en relevante kenmerken van beide noemen, behandelen zij niet allemaal alle lexico-grammaticale constructies in dit onderzoek aan het licht kwamen in het kader van Dimensie 4.

De andere drie dimensies kwamen veel minder aan bod in de geanalyseerde leerboeken. Wat betreft Dimensie 1, beknoptheid, komen de lexico-grammaticale middelen die de syntactische efficiency van de leerders zouden kunnen bevorderen, zoals beknopte bijzinnen, apposities en andere werkwoordloze constructies (*NCDUS*) nauwelijks aan bod. De meeste leerboeken besteden geen aandacht aan non-finiete bijzinnen, hooguit om de leerders te waarschuwen tegen de zogeheten ‘dangling participle’.

Wat betreft Dimensie 2, zelfvertrouwen, worden de leerders niet geattendeerd op de mogelijkheden die er bestaan om hun discourse te onderbreken, specificeren, of uit te breiden door middel van het strategisch benutten van posities in het midden en het eind van de zin om onderschikkende constructies te plaatsen, om daarmee hun autoriteit en betrokkenheid te bevestigen. M.a.w. de meeste leerboeken laten het discoursepotentieel van het variëren van de positie van lexico-grammaticale constructies in de zin onderbelicht. Ook het gebruik van parenthetische zinnen (die ook dienen als een instrument om een zin te onderbreken of er informatie aan toe te voegen) wordt door geen van de leerboeken behandeld. Bijzinnen die de bron van informatie vermelden (*reporting clauses* en *as-clauses*) daarentegen worden wel behandeld, hoewel niet overal even grondig of nauwkeurig. Adverbiale constructies die de mogelijkheid bieden om uit te weiden of informatie in andere bewoordingen te brengen, zoals *in other words*, worden niet uitgebreid behandeld, ondanks het feit dat adverbiale connectoren in alle leerboeken aan bod komen, ofschoon met weinig uitleg of context.

Wat Dimensie 3, complexiteit, betreft blijken de meeste leerboeken samengestelde zinnen niet te behandelen, laat staan de verschillende semantische categorieën van bijzinnen die onder- of overgebruikt worden door de leerders. Die leerboeken die wel verschillende semantische categorieën behandelen doen dat zonder enige verwijzing naar de frequentie waarmee ze voorkomen in authentiek geschreven academisch Engels of adviezen op het gebied van over- en ondergebruik. Slechts één leerboek beschrijft terugverwijzende naamwoordgroepen (*retrospective labels*), maar zonder daarbij een onderscheid te maken tussen technische hoofdnaamwoorden (*analysis, dichotomy, aggregation*) en vage hoofdnaamwoorden (*problem, advantage, thing*). Ten slotte, het overgebruik van adverbiale connectoren van toevoeging, een kenmerk van Dimensie 3, lijkt mogelijk versterkt te worden door de volgorde waarin adverbiale connectoren worden behandeld in de leerboeken: connectoren als *firstly, secondly* en *finally* worden vaak als eerste aangeboden, vaak zonder context en zonder dat de verschillende connectoren worden geordend naar frequentie, waardoor deze onbedoeld en onterecht aan belang winnen.

### **Terugkoppeling naar het ERK**

Ten slotte wordt er een terugkoppeling gemaakt naar het beginpunt van dit onderzoek, namelijk het ERK. In het ERK vinden we dat:

leerders, om volledig doeltreffend te kunnen deelnemen aan communicatiegebeurtenissen, het volgende moeten hebben geleerd of verworven:

- de noodzakelijke competenties, ... ;
- het vermogen om deze competenties om te zetten in handelingen...;
- het vermogen om de strategieën toe te passen die noodzakelijk zijn om de competenties om te zetten in handelingen. (Nederlandse Taalunie 2008: 121)

De focus van dit onderzoek is doeltreffend academisch schrijven en in hoofdstuk 2 worden er twee belangrijke conclusies getrokken ten aanzien van de noodzakelijke competenties: om op het hoogste niveau te kunnen schrijven, moeten leerders een goed ontwikkelde discoursecompetentie bezitten en binnen het concept van communicatieve competentie is zijn de linguïstische competenties centraal en staan aan de basis van de discoursecompetentie. In het ERK staat ook: '[d]e ontwikkeling van de linguïstische competenties van de leerder is een centraal en onontbeerlijk aspect van het leren van talen' (Nederlandse Taalunie 2008: 135) en 'Grammaticale competentie, het vermogen om zinnen zo te ordenen dat zij een betekenis uitdragen, is natuurlijk een centraal aspect van communicatieve competentie' (Nederlandse Taalunie 2008: 136).

Linguïstische competentie, waarvan grammaticale competentie een onderdeel vormt, is de centrale competentie, daar leerders een breed repertoire aan lexico-grammaticale middelen met een grote mate van semantische diepgang nodig hebben om in staat te zijn 'heldere, soepel lopende, complexe teksten [te] schrijven in een gepaste en doelmatige stijl en met een logische structuur die de lezer helpt belangrijke punten te herkennen' (Nederlandse Taalunie 2008: 59).

Het is mogelijk om de vier dimensies van discoursecompetentie op één lijn te brengen met over de descriptor voor het C2 niveau van algemene schriftelijke productie heen te leggen. Zoals schematisch is weergegeven in Tabel 1:

**Tabel 1: De vier dimensies van discoursecompetentie en de C2 descriptor het ERK**

<i>Dimensie</i>	<i>naam</i>	<i>onderdeel omschrijving C2 niveau schrijven ERK</i>
3	<i>complexiteit</i>	'Kan heldere, soepel lopende, complexe teksten schrijven'
4	<i>formaliteit</i>	'in een gepaste'
1	<i>beknoptheid</i>	'en doelmatige stijl'
2	<i>zelfvertrouwen</i> <sup>74</sup>	'men met een logische structuur die de lezer helpt belangrijke punten te herkennen'

Het is aan diegenen die onderwijsmateriaal ontwikkelen om in overweging te nemen welke lexico-grammatical middelen aan leerders aangeboden zouden moeten worden (en uiteindelijk door hen beheerst zouden moeten worden), opdat hun linguïstisch repertoire zodanig wordt verbreed en verdiept dat zij linguïstisch bekwaam genoeg zijn om tekstueel doeltreffend te zijn. De vier dimensies uit dit onderzoek zouden daarbij als een leidraad kunnen fungeren.

## **Conclusie**

Het hoofddoel van deze studie was het in kaart brengen van de verschillen tussen teksten geschreven door Nederlandse vergevorderde leerders van het Engels en moedertaalsprekers om aanbevelingen te kunnen doen voor het schrijfvaardigheidsonderwijs van het Engels op universitair niveau. Door zijn brede opzet heeft deze studie getracht bij te dragen aan de

<sup>74</sup> Een van de bepalende kenmerken van deze dimensie is het gebruiken van verschillende posities in de zin om informatie te ordenen.



onderzoeksvelden van schrijfvaardigheidstheorievorming (*writing theory*), leerdercorpus-onderzoek (CLC), tweedetaalverwerving (SLA) en Engels voor Academische Doeleinden (EAP).

Op het gebied van theorievorming heeft dit onderzoek getracht een bijdrage te leveren aan de interpretatie van het begrip *discoursecompetentie*. Deze studie heeft *discoursecompetentie* op zodanige wijze gedefinieerd dat het een gedetailleerdere en geactualiseerde versie is van die van Canale (1983) en tegelijkertijd in overeenstemming is met de beschrijving van het ERK. Bovendien is het concept vertaald naar verschillende typen van discourse kenmerken en daarna naar een grote reeks specifieke lexico-grammaticale constructies. Omdat dit corpusonderzoek zich niet beperkt tot één kenmerk, zoals vaak gebruikelijk, maar zich daarentegen richt op een reeks van kenmerken, ontstaat er een completer beeld van hoe een theoretisch concept als *discoursecompetentie* vertaald kan worden in daadwerkelijk taalgebruik, in dit geval geschreven academisch Engels. Door het alomtegenwoordige ERK als uitgangspunt te nemen en een aantal concepten ervan te verhelderen heeft dit onderzoek bijgedragen aan een beter begrip van wat het voor een leerder betekent om in staat te zijn op C1 en C2 niveaus te kunnen schrijven. Dit is nuttig voor taalprofessionals die zich inspannen om hun onderwijsmateriaal op één lijn te brengen met de descriptoren van het ERK. En hoewel het niet het uitgangspunt van deze studie was, zou deze informatie ook een toepassing kunnen vinden in toetsingsonderzoek.

De bijdrage op het gebied van leerdercorpora-onderzoek is drieledig. Ten eerste heeft het onderzoek een model opgeleverd voor het onderzoeken van *discoursecompetentie* in CLC en een annotatieschema verschaft. Ten tweede zijn er veel gevallen van onder- en overgebruik in het Engels van Nederlandse vergevorderde leerders aan het licht gekomen die tot nu toe niet bekend waren in dit onderzoeksgebied. Dit kan ook van nut zijn voor onderzoekers die Engelse leerdertaal onderzoeken van leerders met een andere L1 dan het Nederlands, omdat dit mogelijkheden biedt tot een contrastieve intertaalanalyse. Tot slot, CLC onderzoek streeft er o.a. naar om inzichten te verschaffen die ten goede komen aan onderzoek op het gebied van tweedetaalverwerving en taalvaardigheidsonderwijs. Vooralsnog zijn er vooral studies geweest naar beginnende leerders en weinig naar vergevorderde leerders. Door de omschrijving van de *discoursecompetentie* van vergevorderde leerders levert dit onderzoek ook een bijdrage aan deze zich nog ontwikkelende tak van onderzoek naar leerdertaal.

Op het gebied van onderzoek naar Engels voor academische doeleinden (EAP) heeft dit onderzoek vooral een bijdrage geleverd door inzichten te verschaffen in de verschillen tussen het academisch Engels van moedertaalsprekers en dat van vergevorderde leerders, *in casu* Nederlandse leerders. De corpusanalyse heeft een aantal vermoedens die er bestonden bij taalonderwijsprofessionals bevestigd, zoals o.a. het ondergebruik van niet-finiete bijzinnen, een aantal weerlegd, zoals o.a. het ondergebruik van beperkende betrekkelijke bijzinnen (ze blijken namelijk óvergebruikt te worden), en een aantal aan het licht gebracht dat tot op heden nog niet in de literatuur beschreven zijn, zoals o.a. het ondergebruik van nominale apposities. Ook zijn er door de corpusanalyse en de kwantitatieve analyse een aantal kenmerken aan het licht gebracht die met het blote oog niet te ontwaren zijn, zoals bijvoorbeeld het feit dat leerders weliswaar niet-finiete bijzinnen ondergebruiken, maar dat alleen doen in het midden en aan het eind van de zin, maar niet aan het begin van de zin. Bovendien is er een voorstel gedaan om al deze voorbeelden van onder- en overgebruik te groeperen in systematische categorieën, ongeacht hun syntactische verwezenlijking, die gezien kunnen worden als dimensies waarlangs de *discoursecompetentie* van vergevorderde leerders zich ontwikkeld. Dat betekent dat toekomstig onderwijsmateriaal zich op deze gebieden zou kunnen richten om het verschil te overbruggen tussen wat dit type leerder al kan en de uiteindelijk beoogde bekwaamheid.

De inzichten die deze studie heeft opgeleverd kunnen, tot slot, ook hun toepassing vinden in andere deelgebieden van EAP. Veel corpusonderzoeken gebruiken academisch Engels corpora die zijn opgebouwd uit teksten afkomstig uit zeer veel verschillende academische disciplines, om uitspraken te kunnen doen over gemeenschappelijke kenmerken van academisch schrijven in het Engels (o.a. Coxhead 1998; Biber 1999). Het controlecorpus in het huidige onderzoek bestaat uit teksten van twee disciplines: taalkunde en letterkunde. Het gebruik van zo een gespecialiseerd corpus in plaats van een algemeen academisch corpus biedt de mogelijkheid om verschillen tussen het schrijven in deze twee disciplines voor het voetlicht te brengen. Op deze manier levert het onderzoek ook een bijdrage aan de inzichten die er bestaan over genre-specificiteit en de mate van specialisatie die er eventueel nodig is in het academisch schrijfvaardigheidsonderwijs.

Natuurlijk kennen het onderzoek en de conclusies die eraan verbonden kunnen worden ook hun beperkingen. De methodologische beperkingen zijn in feite allen direct of indirect terug te voeren op de tijdsintensieve aard van de samenstelling, handmatige annotatie, en analyse van de corpora. Idealiter zouden de corpora groter zijn dan ze nu zijn. Hoewel ze representatief zijn voor de onderzochte taalvariëteiten en groter dan de meeste corpora die gebruikt zijn in de in dit proefschrift aangehaalde onderzoeken, bleken zij toch te klein voor een accurate analyse van sommige laagfrequente lexico-grammaticale constructies. Ook zijn niet alle mogelijk relevante lexico-grammaticale middelen geannoteerd die een rol spelen in de discoursecompetentie van de leerder. Echter, die welke wel zijn geannoteerd vertegenwoordigen zo een breed spectrum, dat een relevante analyse van de discoursecompetentie van Nederlandse vergevorderde leerders goed mogelijk is. Het grote aantal verschillende lexico-grammaticale constructies dat geanalyseerd is heeft tot gevolg gehad dat dit onderzoek zich voornamelijk heeft gericht op kwantitatieve analyses; slechts een beperkt aantal resultaten is in meer detail –kwitatief– geanalyseerd. Op het gebied van de interpretatie van de resultaten zijn ook nog twee opmerkingen te maken. Ten eerste is het van belang om voor ogen te houden welke taalvariëteit een corpus vertegenwoordigt en conclusies n.a.v. de bevindingen ook daartoe te beperken. Conclusies gebaseerd op resultaten uit het leerdercorpus zouden voorzichtig doorgetrokken kunnen worden naar alle Nederlandse leerders van het Engels, maar niet naar alle leerders van het Engels in het algemeen. Tot slot, zoals de meeste corpusstudies houdt het huidige onderzoek geen rekening, statistisch gezien, met de verschillen tussen teksten binnen het corpus. Hoewel dit geen probleem hoeft te zijn, aangezien het corpus is samengesteld op basis van zeer strenge objectieve selectiecriteria, zou er eventueel een grotere accuratesse behaald kunnen worden door het gebruik van bijvoorbeeld een Multilevel Modeling analyse (Mulder 2008: 43-56).

### **Tot slot**

In 2001 verzuchtte Flowerdew (2001: 366) dat maar weinig bevindingen uit leerdercorpusonderzoek hun toepassing vinden in het taalonderwijs en blijven hangen op het niveau van aanbevelingen. Als er tien jaar na dato uitgevers zijn die de uitdaging aandurven om onderwijsmateriaal te ontwikkelen dat is toegesneden op leerders met een specifieke taalachtergrond en dat gebaseerd is op empirisch onderzoek van zowel genre-specifieke corpora van moedertaalsprekers als leerdercorpora, dan zullen zij de vier dimensies van discoursecompetentie die in dit onderzoek naar voren zijn gekomen als een nuttige leidraad kunnen gebruiken voor de selectie en ordening van hun onderwijsmateriaal. Dit onderzoek heeft laten zien dat Dimensie 1 (beknoptheid), Dimensie 2 (zelfvertrouwen en betrokkenheid) en Dimensie 3 (complexiteit) minstens zo relevant zijn voor vergevorderde leerders als de in leerboeken vaak genoemde Dimensie 4 (formaliteit). Het gebruik van alle vier de dimensies als richtsnoer is een manier om coherentie te verschaffen aan toekomstig onderwijsmateriaal op het gebied van academische schrijfvaardigheid in het Engels.