

Hoe groot is het verschil tussen B1 en B2?

Verschillen in kennis van woordenschat en grammatica tussen NT2-leerders op B1-B2-niveau van spreekvaardigheid

Jan H. Hulstijn (Universiteit van Amsterdam)

Niuja H. de Jong (Universiteit Utrecht)

Margarita P. Steinel (Universiteit van Amsterdam)

Arjen Florijn (Universiteit van Amsterdam)

Rob Schoonen (Universiteit van Amsterdam)

This study examined lexical and grammatical knowledge of adult learners of Dutch as a second language (L2) at the B1 and B2 speaking-proficiency levels of the Common European Framework of Reference for Languages. In a sample of 208 Dutch L2 learners, 80 and 30 participants were found to be proficient in speaking at the B1 and B2 levels respectively, as assessed in eight computer-administered speaking tasks. Participants also performed paper-and-pencil tests of vocabulary and grammar knowledge. Average vocabulary sizes were obtained of 4000 and 7000 words (with standard deviations of 1623 and 1456) in the B1 and B2 groups, respectively. Overall performance on the grammar test, which assessed knowledge in ten grammatical domains, was significantly higher in the B2 group than in the B1 group, with average correct scores of 86% and 71%, respectively. For each of the ten domains of grammar, examples are given of features that B1 and B2 learners did, or did not yet, control.¹

Inleiding

Onder docenten Nederlands buiten Nederland en Vlaanderen, ook als ze buiten Europa werkzaam zijn, is het Europees Referentiekader (ERK; Nederlandse Taalunie 2006; Council of Europe 2001)² gemeengoed geworden, al is het alleen maar omdat de examens van het Certificaat Nederlands als Vreemde Taal (CNaVT)³ en de staatsexamens NT2⁴ voorzien zijn van ERK-labels. Het ERK onderscheidt zes taalvaardigheidsniveaus (van laag naar hoog: A1, A2, B1, B2, C1 en C2) en de CNaVT- en de staatsexamens meten taalvaardigheid op de niveaus A2 (alleen

CNaVT), B1 en B2, al naar gelang het certificaat of examen. Dat gebeurt met zogenaamde integratieve toetsen, wat betekent dat de kandidaat luister-, spreek-, lees- en schrijfopdrachten uitvoert (soms aan elkaar gekoppeld, bijvoorbeeld een schrijfopdracht over een leestekst) waarin linguïstische deelvaardigheden geïntegreerd zijn. Kennis van woordenschat, grammatica, uitspraak, spelling en dergelijke worden in integratieve toetsen dus niet afzonderlijk getoetst. Een belangrijke vraag echter voor zowel kandidaten als docenten is hoeveel linguïstische kennis iemand eigenlijk moet bezitten om luister-, spreek-, lees-, of schrijfvaardig te zijn op een van de zes ERK-niveaus. Dat antwoord wordt niet gegeven door het ERK zelf, want de ERK-schalen van de deelvaardigheden, zoals woordenschat, grammatica en uitspraak, zijn in zeer algemene termen geformuleerd, zoals de volgende citaten laten zien (Nederlandse Taalunie 2006).

Bereik van woordenschat (p. 104):

B2. Beschikt over een goede woordenschat voor zaken die verband houden met zijn of haar vakgebied en de meeste algemene onderwerpen. Kan variatie aanbrengen in formuleringen om te veel herhaling te voorkomen, al kunnen hiaten in de woordenschat nog wel tot aarzeling en omschrijving leiden.

B1. Beschikt over voldoende woorden om zich, met enige omhaal van woorden, te uiten over de meeste onderwerpen die betrekking hebben op het dagelijks leven, zoals familie, hobby's en interesses, werk, reizen en actualiteiten.

Grammaticale correctheid (p. 106):

B2. Goede beheersing van de grammatica; incidentele vergissingen, niet-stelsmatige fouten en kleine onvolkomenheden in de zinsstructuur kunnen nog voorkomen maar zijn zeldzaam en worden meestal achteraf verbeterd. Toont een betrekkelijk grote beheersing van de grammatica. Maakt geen vergissingen die tot misverstanden leiden.

B1. Communiceert redelijk correct in vertrouwde omstandigheden; vertoont over het algemeen een goede grammaticale beheersing maar met merkbare invloed vanuit de moedertaal. Fouten komen voor, maar het is altijd duidelijk wat hij of zij probeert uit te drukken. Maakt met een redelijke mate van nauwkeurigheid gebruik van een repertoire van veelgebruikte 'routines' en patronen die bekend zijn van meer voorspelbare situaties.

Om tegemoet te komen aan de wens van het onderwijsveld zijn nu voor ettelijke Europese talen meer specifieke niveaubeschrijvingen ontwikkeld of in ontwikkeling,⁵ maar niet voor het Nederlands.⁶ Deze meer specifieke invullingen roepen echter de vraag op welke specifieke niveaubeschrijving, gemaakt door welke instelling, voor welk examen als maat in onderwijs en politiek gehanteerd zal worden. Sinds het verschijnen van het ERK en door de erkenning ervan door steeds

meer instanties in Europa is er een levendig, soms zelfs fel debat ontstaan over de vermeende deugden en nadelen van het ERK (zie bijvoorbeeld de bijdragen van negen deskundigen hierover in *The Modern Language Journal* 91 [4], pp. 641-685; Alderson, Figueras, Kuijper, Nold, Takala & Tardieu 2006; Davidson & Fulcher 2007; Hulstijn 2011; Hulstijn, Alderson & Schoonen 2010; Weir 2005). In Nederland leverden onder meer Janssen-Van Dielen (2000, 2001, 2006), Kwakernaak (2010) en Van Hest (2011) een bijdrage aan dit debat.

Hoewel de discussie rond het ERK nog wel langere tijd zal voortduren, is het intussen interessant om te onderzoeken wat precies de beheersing van deelvaardigheden is van kandidaten die geslaagd zijn voor een CNaVT- of staatsexamen. Voor zover wij weten, is hiernaar nog geen onderzoek gedaan. Maar de gegevens die wij verzameld hebben in een onderzoek dat eigenlijk niet direct met het ERK te maken heeft, zijn in dit verband wellicht relevant. In dit artikel rapporteren wij over de opzet en resultaten van dit onderzoek. De onderzoeksvraag die wij in dit artikel willen beantwoorden, luidt: *Welke kennis van woordenschat en grammatica treffen we aan bij twee groepen NT2-leerders die spreekvaardig zijn op, respectievelijk, de functionele B1- en B2-niveaus van het ERK?*

Methode

Het onderzoek

Het onderzoek, uitgevoerd aan het *Amsterdam Center for Language and Communication* (ACLIC) van de Universiteit van Amsterdam, met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek, had primair tot doel te onderzoeken welke deelvaardigheden bepalend zijn voor spreekvaardigheid. De werktitel van het project luidde dan ook *What is Speaking Proficiency?* (kortweg WiSP). We geven hier een korte beschrijving van het hoofdonderzoek; meer informatie is te vinden in De Jong, Steinel, Florijn, Schoonen en Hulstijn (2012a en 2012b) en Hulstijn, Schoonen, De Jong, Steinel en Florijn (2012).

Aan het hoofdonderzoek werd deelgenomen door 208 volwassen NT2-leerders en 58 volwassen moedertaalsprekers van het Nederlands. Bijna alle NT2-leerders volgden NT2-onderwijs op halfgevoerd of gevorderd niveau om zich voor te bereiden op een studie aan de Universiteit van Amsterdam. Alle proefpersonen werden individueel getest. In twee afnamesessies die samen 2.5 tot 3 uur duurden, deden ze acht computergestuurde spreektaken, vier andere computergestuurde taken (twee taken waarin ze plaatjes moesten benoemen, een taak waarin ze zinnen moesten maken, en een uitspraaktaak), een woordenschat- en een grammaticatest, en ze vulden een vragenformulier in met achtergrondgegevens. In dit artikel zullen we ons beperken tot de spreektaken en de woordenschat- en grammaticatest. We beschrijven nu eerst hoe we uit de groep van 208 NT2-leerders een groep van 80 proefpersonen spreekvaardig op B1-niveau en een groep

van 30 spreekvaardig op B2-niveau vonden. Daarna beschrijven we woordenschat- en grammaticatest.

Selectie van personen op B1- en B2-niveau van functionele spreekvaardigheid

De spreektaken in ons onderzoek waren zo opgezet dat de proefpersonen eerst op het scherm van de computer informatie kregen over de spreeksituatie, hun rol daarin, het gespreksonderwerp en de concrete opdracht. Terwijl deze informatie op het scherm zichtbaar bleef, kregen de deelnemers eerst een minuut om na te denken over wat ze zouden zeggen, en daarna twee minuten tijd om te spreken. De onderwerpen van de spreektaken betroffen vier beschrijvingen (het interieur van een woning, een ongeval, de taken van een verpleegkundige, en de ontwikkeling van de werkloosheid in de afgelopen tien jaar) en vier argumentatieve betogen (over de beste locatie voor een speelplaats voor kinderen, de oplossing van de fileproblemen in Nederland, de bouw van een parkeergarage, en het beste advies aan een vrouw met kinderen over de combinatie van zorg voor de kinderen, studie en werk). In de spreektaken vertolkten de proefpersonen soms de rol van een vriend of familielid, en soms vervulden ze een rol in een meer formele situatie, bijvoorbeeld als getuige van een ongeval, in de rechtszaal.

De spreekprestaties van de deelnemers werden opgenomen en later door een panel van twaalf niet-taalkundig geschoolde personen (studenten psychologie en rechten) beoordeeld aan de hand van een dertigpuntsschaal die verdeeld was in zes hoofdniveaus, die elk weer in vijf subniveaus waren onderverdeeld. Er werd gekozen voor niet-taalkundig geschoolde beoordelaars omdat de spreekprestaties alleen op inhoud en niet op taalfouten beoordeeld moesten worden. Het criterium van deze beoordelingsschaal had *geen* betrekking op vormelijke aspecten van de spreekprestaties (zoals uitspraak, woordenschat of grammaticale correctheid) maar *alleen* op de communicatieve adequaatheid (de inhoud) ervan. Bijvoorbeeld, in de taak waarin de deelnemer voor de rechter als getuige van een ongeval optreedt, had de beoordelingsschaal primair betrekking op de vraag in hoeverre de rechter zich op basis van de spreekprestatie een precies beeld kon vormen van het ongeval.

Door een panel van zes NT2-deskundigen die veel ervaring hadden met het ERK werden de acht spreektaken ingedeeld naar ERK-niveau volgens criteria ontleend aan de functionele beschrijvingen van het ERK. Eén spreektaak werd door de deskundigen ingeschaald op A2-niveau, drie taken op B1-niveau en vier taken op B2-niveau. Drie van de zes deskundigen beoordeelden bovendien een steekproef van de spreekprestaties van de niet-Nederlandstalige deelnemers. Op basis van een regressiemethode die gedetailleerd beschreven wordt in Hulstijn et al. (2012) kon vervolgens de beoordeling van de spreekpresentaties in deze steekproef geëxtrapoleerd worden naar alle spreekprestaties. Zo kon berekend worden welke van de 208 niet-Nederlandstalige deelnemers geacht konden worden functioneel spreekbaar-

dig te zijn op B1-niveau en welke op B2-niveau. Het resultaat van de berekeningen was dat 80 proefpersonen geacht konden worden zich op het B1-niveau van functionele spreekvaardigheid te bevinden en dat 30 personen zich op B2-niveau bevonden. We zullen deze groepen hierna aanduiden als B1'ers en B2'ers.

Onder de 80 B1'ers waren 34 moedertalen vertegenwoordigd, waarvan achttien Europese talen. De 30 B2'ers spraken dertien moedertalen, waarvan elf Europese. De leeftijd liep uiteen van 20 tot 56 jaar in de B1-groep en van 21 tot 43 jaar in de B2-groep, met in beide groepen een gemiddelde van 29 jaar. De verblijfsduur in Nederland varieerde van een maand tot 17 jaar in de B1-groep (gemiddeld 3,5 jaar) en van twee maanden tot elf jaar in de B2-groep (gemiddeld 3 jaar en 10 maanden). Het opleidingsniveau in beide groepen was overwegend voortgezet onderwijs (vwo) of tertiair onderwijs, waarbij negen B1'ers en twee B2'ers geen informatie opgegeven hebben over hun vooropleiding.

Resultaten

Woordenschat

Productieve woordenschat werd getest met een test bestaande uit 116 items, verdeeld over twee delen. In deel 1 werd de productieve kennis van 90 losse woorden bevraagd, in deel 2 kennis van 26 collocaties. Elk item bestond uit een zin waaruit het te toetsen woord was weggelaten, op de eerste letter(s) na. De deelnemers moesten het weggelaten woord invullen. De draagzinnen bevatten alleen frequente woorden, dat wil zeggen alleen woorden die tot de 2000 meest frequente woorden van het Nederlands behoren. De 90 doelwoorden van deel 1 waren als volgt gekozen. Eerst werden de 10.000 meest voorkomende woorden in het *Corpus Gesproken Nederlands* (CGN; Nederlandse Taalunie 2004) geselecteerd. Die woordenlijst werd opgedeeld in tien stukken: van 1–1000, 2000–3000 enzovoort tot 9000–10.000. Uit elk duizendtal werden negen inhoudswoorden (zelfstandige naamwoorden, werkwoorden, bijvoeglijke naamwoorden of bijwoorden) geselecteerd. Tot het eerste duizendtal behoorden testwoorden zoals *vandaag*, *idee*, *straks*, *helft* en *gang*; tot het tiende duizendtal behoorden testwoorden zoals *moedig*, *streepje*, *beangstigend*, *levering*, en *tussenstand*. Bij elk van de 90 woorden werd een draagzin gemaakt. De 26 collocaties kozen wij, bij gebrek aan frequentiegegevens, op basis van onze intuïties, waarbij we zorgden voor opname van wat wij beschouwden als frequente en minder frequente vaste combinaties met een voorzetsel (14 items) of combinaties van een werkwoord met een zelfstandig naamwoord (12 items). Hieronder staan enkele zinnen uit de woordenschattoets.

Voorbeelditems uit deel 1

1. Als je een vreemde taal spreekt, dan zeg je ook dat je die taal beh____. (Antwoord: *beheerst*)

2. Het spel was heel moeilijk en niet te snappen zonder uitleg. (Antwoord: uitleg)

Voorbeelditems uit deel 2

3. Heb je gevraagd of de reiskosten voor vergoeding in aanmerking komen? (Antwoord: komen)
4. Ze wilde niet ingaan op mijn vraag. (Antwoord: op)

Elk goed antwoord leverde één punt op. Niet-relevante spellingfouten werden genegeerd. De maximale score was dus 116.

De interne consistentie (een vorm van betrouwbaarheid) van de woordenschat-test als geheel (116) was hoog voor zowel de 80 B1'ers ($\alpha=.95$) als de 30 B2'ers (.95). Zoals we verwacht hadden, presteerden de B1'ers significant minder goed dan de B2'ers (tabel 1). Variantieanalyses lieten zien dat de B1'ers op elk onderdeel van de woordenschattoets significant lager presteerden dan de B2'ers (zie Hulstijn et al. 2012 voor details). De prestaties van zowel B1'ers als B2'ers gingen omlaag met de daling van woordfrequentie. Toch is het niet zo dat de B1'ers helemaal geen laagfrequente woorden kenden. Als we de kennis van de losse woorden naar frequentie in drie gelijke delen van 30 items splitsen, dan zien we de kennis dalen van 63% via 38% naar 18% bij de B1'ers en van 87% via 71% naar 51% bij de B2'ers.

Een van de moeilijkste opgaven betrof het testwoord *binnenin*, dat slechts door één van de 80 B1'ers en door twee van de 30 B2'ers goed werd ingevuld in de zin 'Van de buitenkant ziet een telefoon er simpel uit, maar binnen zitten allemaal kleine onderdelen'. Een van de makkelijkste opgaven betrof het woord *maag*, dat door 70 B1'ers en 29 B2'ers goed werd ingevuld in de zin 'Als je honger hebt, dan heb je waarschijnlijk een lege maag'.

Tabel 1. Prestaties van B1'ers en B2'ers op de productieve woordenschat-test*

Testdeel	ERK	N	k	M	Ondergrens**	Bovengrens
Losse woorden	B1	80	90	36	33	39
	B2	30	90	63	58	68
Collocaties	B1	80	26	9	8	10
	B2	30	26	18	16	20

* N = aantal proefpersonen; k = aantal items; M = gemiddelde goedscore (afgerond).

** Gemiddelden hebben altijd een zekere foutenmarge omdat ze gebaseerd zijn op steekproeven uit veronderstelde populaties van B1'ers en B2'ers. Daarom is het gebruikelijk om ook de onder- en bovengrens van het interval te rapporteren waarbinnen de gemiddelden met 95% zekerheid liggen, rekening houdend met individuele variatie en steekproefgrootte.

Als we de prestaties van de B1'ers en de B2'ers op deel 1 van de test extrapoleren naar kennis van de 10.000 meest frequente woorden in het CGN komen we uit op bijna 4000 woorden ($35.8/90 \cdot 10000 = 3977$) voor de B1'ers (met een 95% betrouwbaarheidsinterval⁷ van 3622 tot 4333 woorden en een standaarddeviatie⁸ van 1623 woorden) en bijna 7000 woorden ($62.7/90 \cdot 10000 = 6966$) voor de B2'ers (met een betrouwbaarheidsinterval van 6422 tot 7511 woorden en een standaarddeviatie van 1456 woorden).

Zijn de getallen van 4000 en 7000 (respectievelijk voor B1 en B2) aan de hoge kant? Dat is niet onmogelijk. Het ERK stelt dat ongelijke profielen heel gewoon zijn. Het is dus niet ondenkbaar dat onze proefpersonen (gemiddeld) een grotere woordenschat hadden dan wat minimaal nodig is voor spreekvaardigheid op B1- en B2-niveau, maar gezegd moet worden dat er onder de B1'ers en B2'ers ook mensen waren met een kleine woordenschat. Trekken we van de gemiddelde prestaties één standaarddeviatie af, dan weten we dat 68% van de deelnemers in elke groep een score behaalde van ten minste 2354 (B1'ers) en 5510 (B2'ers) woorden. Veel gebruikte NT2-leergangen die taalvaardigheid tot B1-niveau onderwijzen, zoals *Code* (Boers, Olijhoek, Van der Voort, Heijne & Hidma 2004), bevatten een woordenlijst van ongeveer 3000 woorden. Daar ligt het door ons gevonden gemiddelde van 4000 ruim boven. Maar cijfers in de woordenschatliteratuur moeten om uiteenlopende redenen met de nodige voorzichtigheid genomen worden, redenen die onder andere te maken hebben met de keuze en validiteit van corpora, de trekking van steekproeven, en de definitie van wat een woord of woordfamilie is.

Grammatica

De grammaticatest bestond uit 142 items verdeeld over de volgende domeinen: buiging van sterke en zwakke werkwoorden (19 items), korte of lange vorm van bijvoeglijke naamwoorden (dat wil zeggen zonder of met een *e*-uitgang; 19 items), woordvolgorde in hoofd- en bijzinnen (25 items), de splitsing van de twee delen van scheidbaar samengestelde werkwoorden (8 items), *het* en *er* als voorlopig onderwerp of voorlopig voorwerp (26 items), de volgorde van meerdere modale partikels (bijwoorden) in één zin (10 items), betrekkelijke voornaamwoorden (15 items), bezittelijke voornaamwoorden (5 items), keuze van hulpwerkwoord (*hebben* of *zijn*; 10 items), en constructie van passieve zinnen (5 items). De antwoorden werden gescoord als goed (1 punt) of fout (0 punten). Niet relevante spellingfouten werden genegeerd (bijvoorbeeld *geantwoort* in het onderdeel waarin kennis van het verleden deelwoord werd bevestigd). Voorbeelden van de manier waarop grammaticakennis in deze domeinen bevestigd werd, zijn te vinden in de appendix. Bij sommige items werden meerdere antwoorden goed gerekend. Elk onderdeel bevatte opgaven waarvan wij vermoedden dat ze (heel) makkelijk en (heel) moeilijk zouden zijn. Dat vermoeden is voornamelijk ingegeven door de (on)re-

gelmatigheid en de ingeschatte frequentie van de geteste verschijnselen, maar wij beschikten niet over betrouwbare gegevens omtrent de frequentie waarmee bepaalde constructies of verbogen vormen van bijvoorbeeld werkwoorden in het Nederlands voorkomen.

De grammaticatest (142 items) was betrouwbaar voor de 80 B1'ers ($\alpha=.90$), en voor de 30 B2'ers ($\alpha=.77$). De B1'ers hadden een gemiddelde goedscore van 101 (71%; standaarddeviatie=15); de B2'ers hadden een gemiddelde goedscore van 122 (86%; standaarddeviatie=8); 27 items werden door alle B2'ers foutloos beantwoord. De verschillen tussen B1'ers en B2'ers waren significant op alle onderdelen van de test; het sterkst (effectgrootte) waren de verschillen op de onderdelen 1) *het en er*, als voorlopig onderwerp of voorlopig voorwerp, 2) de volgorde van meerdere modale partikels in één zin, en 3) betrekkelijke voornaamwoorden. Tabel 2 geeft de prestaties van de B1'ers en B2'ers per onderdeel.

Tabel 2. Prestaties (in % correct) van B1'ers en B2'ers op de grammaticatest, per linguïstisch domein*

Linguïstisch domein	k	B1'ers (n = 80)		B2'ers (n = 30)		Verschil
		M	SD	M	SD	
Werkwoord	19	80	17	94	7	14
Bijvoeglijk naamwoord	19	76	14	84	12	8
Woordvolgorde	25	81	11	91	8	10
Scheidbare werkwoorden	8	66	29	83	14	17
<i>Het en er</i>	26	69	10	81	7	12
Volgorde van modale partikels	10	57	19	84	16	27
Betrekkelijke voornaamwoorden	15	56	22	83	12	27
Bezittelijke voornaamwoorden	5	86	17	98	6	12
Hulpwerkwoorden <i>hebben en zijn</i>	10	66	20	84	14	18
Passiefconstructies	5	59	32	81	29	22

* k = aantal items; M = gemiddelde score; SD = standaarddeviatie; Verschil = verschil tussen de twee gemiddelden.

Hieronder geven we per onderdeel aan wat de B1'ers wel en niet beheersten en wat de meest opvallende verschillen met de B2'ers waren.

Buiging van sterke en zwakke werkwoorden

Met de vier opgaven waarin een vorm van de tegenwoordige tijd moest worden ingevuld (ik *loop*; hij *kookt*; zij *wonen*, jullie *denken*) hadden de B1'ers geen moeite:

ten minste 75 van de 80 B1'ers gaven het goede antwoord. Met de onvoltooid verleden tijd van de gevraagde sterke werkwoorden had meer dan de helft van de B1'ers geen moeite (de cijfers tussen haakjes geven aan hoeveel deelnemers het goede antwoord gaven): hij *ging* (69), ik *vond* (67) hij *vertelde* (59), jullie *kwamen* (58), wij *lagen* (55), de twee zusjes *liepen* (55), hij *zei* (53). Maar bij de zwakke werkwoorden presteerden ze nog beter: de burens *maakten* (72), jij *belde* (74). Ook met drie van de vier verleden deelwoorden wisten de meeste B1'ers wel raad: *gekeken* (70), *gevraagd* (63), *gewacht* (62). Maar dat wisten ze niet met *geantwoord* (33).

Korte of lange vorm van bijvoeglijke naamwoorden

De korte vorm van het bijvoeglijk naamwoord in predicatieve positie werd door vrijwel iedereen goed ingevuld, bijvoorbeeld in: 'De planten zijn bijna dood'. Iets meer dan de helft van de B1'ers had kennis van het feit dat materiaal-aanduidende bijvoeglijke naamwoorden en de verleden deelwoorden van sterke werkwoorden helemaal geen lange vorm kennen: 'deze houten stoel' (56), 'afgelopen jaar' (56), 'een gesloten deur' (52). De lange vorm van het bijvoeglijke naamwoord (in attributieve positie) werd meestal goed ingevuld: 'een gekke man' (78), 'de rode auto' (78), 'heel bijzondere bloemen' (71), 'een heel leuke avond' (66), 'een iets oudere vrouw' (61), 'het prachtige schilderij' (57); maar moeilijk waren: 'ons nieuwe huis' (47) en vooral 'een heel vriendelijk iemand' (29). De B2'ers presteerden beter dan de B1'ers, maar ook voor hen waren 'ons nieuwe huis' (in Vlaanderen is 'ons nieuw huis' overigens de gewone vorm; 16/30) en 'een heel vriendelijk iemand' (18/30) de moeilijkste items.

Woordvolgorde in hoofd- en bijzinnen

De 25 items in deze rubriek waren zo opgebouwd dat het begin van elke zin gegeven was en dat alle volgende woorden tussen haakjes in willekeurige volgorde stonden. De deelnemers moesten de tussen haakjes geplaatste woorden in de juiste volgorde zetten. De hoofdzinopgaven vormden geen probleem, voor B1'ers noch voor B2'ers, ook niet de opgaven waarin het onderwerp na de persoonsvorm staat, zoals in 'Misschien kunnen jullie ons bellen'. De opgave met *want* was voor de B2'ers geen probleem (28/30) maar voor sommige B1'ers wel (63/80): 'Ze is vroeg naar bed gegaan, want ze was heel moe.' De B1'ers en B2'ers hadden geen moeite met de opgaven met bijzinsvolgorde, behalve twee extra moeilijke: 'Wie ze het ook vroegen, niemand wist het' (6/80 en 12/30) en 'Ze zegt dat in het algemeen mensen steeds dikker worden' (35/80 en 19/30).

De splitsing van de twee delen van scheidbaar samengestelde werkwoorden

Het wel of niet moeten splitsen van scheidbaar samengestelde werkwoorden ging de 30 B2'ers goed af, maar met twee items hadden ze moeite: 'Wij steken zo meeteen de straat *over*' (20/30) en 'Het was veel beter geweest als het feest was *doorgegaan*' (18/30). Bij de meeste van de 80 Br'ers echter was van beheersing nog geen sprake: de scores varieerden van 39 tot 66 goede antwoorden per item.

Het en er als voorlopig onderwerp of voorlopig voorwerp

Elk van de 25 opgaven in deze rubriek bestond uit twee minimaal verschillende zinnen waarvan er één goed en één fout was, zoals bij: (A) 'Zij vindt dat het niet hoort om iemand heel lang aan te kijken' en (B) '*Zij vindt dat niet hoort om iemand heel lang aan te kijken'. De raadkans lag dus bij 50%. Zeven opgaven werden goed beantwoord door 65 of meer van de 80 Br'ers; 16 opgaven werden goed beantwoord door 27 of meer van de B2'ers. De moeilijkste opgaven waren: 'Zijn er nog koekjes? Ja, er zijn (*er*) nog maar twee' (23/80 en 15/30), 'Toen ik klein was wilde ik goed kunnen zingen, maar dat is (*er*) tot nu toe nog niet van gekomen' (32/80 en 18/30), 'Meestal valt *het/er* tegen om kaartjes voor een voetbalwedstrijd te krijgen' (27/80 en 2/30).

De volgorde van meerdere modale partikels in één zin

De opgaven in dit onderdeel hadden een meerkeuzeformaat met twee, drie of vier antwoordalternatieven. De B2'ers presteerden tamelijk goed met itemscores die varieerden van 20/30 tot 30/30; de Br'ers hadden met dit onderdeel moeite, met itemscores die varieerden van 68/80 ('Maar hoe zit het ook al *weer* met je nieuwe baan?') tot 29/80 ('Ze moet dit boek *alvast maar eens even* lezen').

Betrekkelijke voornaamwoorden

Het invullen van het betrekkelijk voornaamwoord *die* (drie opgaven) viel de Br'ers niet zwaar (77, 70 en 63 goed), maar het invullen van *dat* was al moeilijker: 'Het brood *dat* ze hier verkopen, is niet lekker' (57). 'De vriendin *van wie* ...' (56) en 'De vrouw *van wie* ...' (54) werd ook nog door meer dan de helft van de Br'ers goed gemaakt. De opgaven met zinnen waarin andere betrekkelijke voornaamwoorden ingevuld moesten worden, werden slecht gemaakt. *Wat/hetgeen*, betrekking hebbend op een predikaat (vijf opgaven), werd door Br'ers niet beheerst (scores 34, 29, 29, 26, 11), evenmin als *waarmee* in 'De pen *waarmee* we het formulier ingevuld hebben ...' (29) en *wie* in 'Wie kookt, hoeft niet af te wassen' (39). De meeste B2'ers hadden met al deze opgaven geen moeite (scores tussen 23/30 en 30/30) met één uitzondering: 'Dit is het enige *dat/wat* ik weet' bleek ook voor bijna de

helpt van de B2'ers te moeilijk (17/30). (In de laatst genoemde zin hebben we zowel *wat* als *dat* goed gerekend.)

Bezittelijke voornaamwoorden

De vijf opgaven waarin een bezittelijk voornaamwoord moest worden ingevuld, vormden voor verreweg de meeste B1'ers geen probleem: 'onze tafel' (75), 'zijn kinderen' (74), 'jouw ouders' (71), 'haar tas' (67), en 'hun problemen' (55). De B2'ers maakten deze opgaven vrijwel foutloos.

Hulpwerkwoord *hebben* of *zijn*

In combinaties met slechts één hulpwerkwoord gaven de meeste B1'ers het goede antwoord: 'is verhuisd' (69), 'is gestorven' (66), 'hebben gezeten' (57), 'zijn geslaagd' (56), en 'is gesprongen' (48). Maar als het predicaat bestond uit meerdere werkwoorden, dan was het voor B1'ers soms lastig te kiezen tussen *hebben* en *zijn*: 'heeft laten verhuizen' (65), 'hebben hem niet kunnen zien fietsen' (60), 'heb moeten blijven wachten' (47), 'zijn gaan helpen' (39), en 'Hij was verkouden en heeft niet kunnen zingen' (20). De B2'ers hadden met deze opgaven geen moeite (ten minste 24/30 goede antwoorden per item) op één uitzondering na: slechts 16 B2'ers gaven het goede antwoord in 'heeft niet kunnen zingen'.

Passiefconstructies

De vijf opgaven in deze rubriek bestonden uit een actieve zin, gevolgd door het begin van een passieve zin. De tweede zin moest afgemaakt worden met de informatie uit de actieve zinnen. Meer dan de helft van de B1'ers had geen moeite met vier van de vijf opgaven (61, 55, 53, en 43 goede antwoorden). Alleen de omzetting van 'Men zegt soms dat ...' in 'Er wordt soms gezegd dat ...' lukte slechts 20 B1'ers. De B2'ers presteerden in de passiefrubriek tussen 27/30 en 21/30 (21/30 in de opgave met 'Er wordt soms gezegd dat ...').

Discussie

Ons onderzoek geeft een indruk van de verschillen in woordenschat en grammaticakennis tussen 80 en 30 NT2-leerders die zich wat betreft spreekvaardigheid op respectievelijk de niveaus B1 en B2 van het ERK bevonden. Men moet voorzichtig zijn bij het trekken van conclusies op basis van een onderzoek dat zijn beperkingen kent wat betreft 1) proefpersonen (aantal, selectiecriteria), 2) de beoordeling van spreekvaardigheid en de vertaling ervan naar ERK-niveaus, en 3) de meting van kennis van woordenschat en grammatica (toetsconstructie). Het is daarom met enige voorzichtigheid dat we stellen dat de productieve woordenschat van de

B1'ers en B2'ers in ons onderzoek ligt op respectievelijk 4000 en 7000 (met een betrouwbaarheidsmarge van ruim 400 aan beide zijden van deze getallen en een standaarddeviatie van respectievelijk 1623 en 1455 woorden). Zoals we hierboven al vermeldden, lijken dit betrekkelijk hoge cijfers. Misschien is het mogelijk voor NT2-leerders met een veel geringere woordenschat toch spreekvaardig te zijn (in communicatief opzicht) op B1- en B2-niveau. In zulke gevallen zou het oordeel over spreekvaardigheidsniveau af kunnen hangen van een relatief goede uitspraak of vloeiendheid van de kandidaat.⁹ Maar belangrijker is het om vast te stellen dat de vraag welke kennis noodzakelijk is voor een bepaald ERK-niveau, altijd lastig te beantwoorden zal zijn omdat de niveaubeschrijvingen van het ERK tamelijk vaag zijn. Het is aan de politiek en het onderwijsveld om al dan niet de wens uit te spreken (en te besluiten) dat de ERK-niveaus preciezer beschreven moeten worden dan nu het geval is (zie ook de in de Inleiding genoemde literatuur).

Bij de interpretatie van de verschillen in prestaties op de grammaticatoets tussen B1'ers en B2'ers moet volgens ons een onderscheid gemaakt worden tussen onderdelen die minder en onderdelen die meer verbonden zijn met woordkennis. In de onderdelen woordvolgorde, passiefconstructies en het en *er* als voorlopig onderwerp of voorwerp wordt kennis van algemene regelmatigheden gemeten, maar in de meeste items van de overige onderdelen zou het geven van het goede antwoord meer een kwestie kunnen zijn van kennis van 'lexicale feitjes' dan van kennis van syntactische of morfologische regels. Het is bijvoorbeeld moeilijk uit te maken of iemand die het goede antwoord geeft in de opgave 'Heeft ze ooit _____ of ze mee mag komen? (vragen)' dit doet op basis van kennis van een lexicaal feit (het deelwoord van *vragen* is *gevraagd*) of op basis van een regel uit de werkwoordmorfologie. Het kan best zijn dat de goede vorm van het bijvoeglijke naamwoord in 'de rode auto' en 'het prachtige schilderij' op basis van regelkennis gegeven wordt maar niet in 'een heel vriendelijk iemand', omdat 'iemand' geen zelfstandig naamwoord is waarvan je kunt zeggen of het een *de-* of een *het-*woord is. Misschien is het wel zo dat men nu eenmaal als los feit weet dat het bijvoeglijk naamwoord voor *iemand* geen lange vorm heeft. De toets, bestaande uit 142 items verdeeld over tien terreinen, was in ons onderzoek vooral bedoeld om een globaal, breed beeld van grammaticale kennis (in de ruime zin des woords) te krijgen. Om verschillen in grammaticakennis tussen NT2-leerders op halfgevoerd en gevorderd niveau zichtbaar te maken zijn we zeker geslaagd (De Jong et al., 2012). Om kennisverschillen tussen deelnemers zichtbaar te maken moest de toets dus zowel makkelijke als moeilijke items bevatten en moeilijke items betreffen dikwijls onregelmatige gevallen zoals 'een heel vriendelijk iemand'.

De grens tussen woordenschat en grammatica is in de taalwetenschap soms moeilijk te trekken. Woorden hebben immers grammaticale eigenschappen, bijvoorbeeld dat 'bier' een niet-telbaar zelfstandig naamwoord is en dat 'niezen' een onovergankelijk werkwoord is. In elementen die door Goldberg (1995) 'constructies' genoemd worden, vormen lexicon en syntaxis samen iets wat meer is dan de

som der delen. Er zijn volgens Goldberg bijvoorbeeld constructies van het volgende type: ‘Onderwerp dat een levend wezen aanduidt + overgankelijk werkwoord dat een bewegingshandeling uitdrukt + lijdend voorwerp + richtingsbepaling’, zoals in *De vrouw veegde het schuim van het bier*. Omdat deze abstracte constructies deel uitmaken van de taalkennis van *native speakers* (en vergevorderde NT2-leerders), kunnen zij ook een grappige zin begrijpen als *De vrouw niesde het schuim van het bier*, hoewel deze zin strikt genomen ongrammaticaal is – het werkwoord *niesen* is immers onovergankelijk. (Voor een Nederlandstalige inleiding in de constructiegrammatica zie Verhagen 2005.) Dat kennis van woordenschat en grammatica samengaan in de T2-verwerving, bleek ook uit ons onderzoek. De prestaties op de woordenschat- en grammaticatest van alle 181 NT2-ers van wie we complete gegevens hebben, waren sterk gecorreleerd ($r=.84$) en deze twee vormen van kennis waren allebei sterk gerelateerd aan de functionele spreekvaardigheid in de acht spreektaken: $r=.79$ en $.75$ tussen enerzijds (respectievelijk) woordenschat en grammatica en anderzijds functionele spreekvaardigheid.¹⁰

Zoals we in het begin van dit artikel al opmerkten, had ons onderzoek niet tot doel een precies beeld te krijgen van grammaticakennis in de tien opgenomen terreinen. Toch geven de resultaten onzes inziens een aardig idee van wat de meeste B1'ers al goed lijken te beheersen. Dat is vooral kennis van syntactische, morfologische en lexicaal-morfologische verschijnselen die *frequent* in het Nederlands voorkomen, ongeacht of ze regelmatig of onregelmatig zijn. De hierboven genoemde voorbeelden (met de percentages goede antwoorden) illustreren deze interpretatie, die we niet hard kunnen onderbouwen omdat we, zoals eerder opgemerkt, niet beschikten over frequentiegegevens van constructies of verbogen vormen van bijvoorbeeld werkwoorden. Al weer met enige flinke slagen om de arm, durven we te stellen dat deze resultaten steun geven aan de stelling dat wat T2-leerders in de beginfase van de T2-verwerving leren, vooral de taalelementen zijn die zowel veel voorkomen (frequentie) als betekenisvol zijn (semantische opvallendheid; zie bijvoorbeeld Ellis 2006a, 2006b; Ellis & Collins 2009).

Noten

1. Dit onderzoek werd mogelijk gemaakt door een subsidie van de Nederlandse organisatie van Wetenschappelijk Onderzoek (NWO; project 254-70-030) aan Hulstijn en Schoonen. We danken onze onderzoeksassistenten Renske Berns, Andrea Friedrich en Kimberley Mulder. We danken Ton Wempe en Rob van Son voor hun technische hulp en advies.

Correspondentie over dit artikel kan gericht worden aan Jan H. Hulstijn, ACLC, Universiteit van Amsterdam, Spuistraat 134, 1012 VB Amsterdam. E-mail: j.h.hulstijn@uva.nl.

2. Het Nederlandstalige document, in 2006 onder auspiciën van de Nederlandse Taalunie en de Raad van Europa uitgegeven, is een vertaling van het oorspronkelijke Engelstalige document (Council of Europe 2001), dat op zijn beurt het voorlopige eindpunt vormt van een reeks documenten waarvan Van Ek (1975) een van de eerste documenten was (zie ook Van Ek & Trim 1991). Vóór de publicatie van het ERK verschenen al Neder-

- landse uitwerkingen van Van Ek (1975), namelijk *Drempelniveau* (Wynants 1985) en *Over de drempel naar sociale redzaamheid* (Coumou, Michorius, Van de Wouw, Bronzwaer & Geertsen 1987).
3. Zie <http://www.cnavt.org/main.asp>, laatst gezien op 31 juli 2012.
 4. Zie http://www.ib-groep.nl/particulieren/examens/NT2/s10_nederlands_als_tweede_taal.asp, laatst geraadpleegd op 31 juli 2012.
 5. Zie http://www.coe.int/t/dg4/linguistic/dnr_EN.asp, laatst geraadpleegd op 31 juli 2012.
 6. Liemberg en Meijer (2004) geven een uitwerking van het ERK in de vorm van taalprofielen, met zogenaamde *can do statements*, voor elke vaardigheid en elk niveau. Deze uitwerking is echter taalneutraal, net als het ERK zelf.
 7. Zie de tweede noot onder tabel 1 voor een uitleg van betrouwbaarheidsinterval.
 8. De standaarddeviatie (SD) geeft een indruk van de spreiding van de scores rond het gemiddelde (M). Ongeveer 68% van de scores ligt tussen het gemiddelde min één SD en het gemiddelde plus één SD; ongeveer 95% van de scores ligt tussen het gemiddelde min twee SD en het gemiddelde plus twee SD. Bijvoorbeeld, in het geval van de woordenschat van de Br'ers, 68% van de 80 deelnemers heeft een score gehaald tussen de 2354 (3977-1623) en 5600 (3977+1623).
 9. Uitspraak, intonatie en vloeiendheid hebben we in ons onderzoek ook gemeten (zie De Jong et al. 2011 en 2012) maar voor de rapportage daarvan is in dit artikel geen ruimte.
 10. Het gaat om latente variabelen, dus de correlaties (de waardes achter het symbool r) zijn niet verzwakt door ruis/onbetrouwbaarheid.

Bibliografie

- Alderson, J.C., N. Figueras, H. Kuijper, G. Nold, S. Takala & C. Tardieu, 'Analysing tests of reading and listening in relation to the Common European Framework of Reference: The experience of The Dutch CEFR Construct Project'. *Language Assessment Quarterly* 3, 2006, 3-30.
- Boers, T., V. Olijhoek, C. van der Voort, N. Heijne & M. Hidma, *Code 1. Basisleergang Nederlands voor anderstaligen*. Utrecht/Zutphen, 2004.
- Coumou, W., J. Michorius, M. van de Wouw, W. Bronzwaer & M. Geertsen, *Over de drempel naar sociale redzaamheid: Omschrijving van een drempelniveau en een niveau ten behoeve van het onderwijs van het Nederlands als tweede taal aan volwassen buitenlanders*. Utrecht, 1987.
- Council of Europe, *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, 2001.
- Davidson, F. & G. Fulcher, 'The Common European Framework of Reference (CEFR) and the design of language tests: A matter of effect'. *Language Teaching* 40, 2007, 231-241.
- Ek, J.A. van, *Systems development in adult language learning: the threshold level in a European unit/credit system for modern language*. Strasbourg, 1975.
- Ek, J.A. van & J.L.M. Trim, *Threshold 1990*. Strasbourg, 1991.
- Ellis, N.C., 'Language acquisition as rational contingency learning'. *Applied Linguistics* 27, 2006a, 1-24.
- Ellis, N.C., 'Selective attention and transfer phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning'. *Applied Linguistics* 27, 2006b, 164-194.
- Ellis, N. & L. Collins, 'Input and second language acquisition: The roles of frequency, form, and function'. *The Modern Language Journal* 93, 2009, 329-335.

- Goldberg, A.E., *Constructions. A construction grammar approach to argument structure*. Chicago, 1995.
- Hest, E. van, 'De invloed van het ERK op de toetsing van vreemde talen'. *Levende Talen Tijdschrift* 12, (2) 2011, 25-32.
- Hulstijn, J.H.. 'Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment'. *Language Assessment Quarterly* 8, 2011, 229-249.
- Hulstijn, J.H., J.C. Alderson & R. Schoonen, Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? I. Bartning, M. Martin & I. Vedder (red.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research*. Eurosla Monographs Series 1, 11-20. Laatst geraadpleegd op 31 juli 2012 op <http://eurosla.org/monographs/EM01/EM01-home.html>.
- Hulstijn, J.H., R. Schoonen, N.H. de Jong, M.P. Steinel & A. Florijn, 'Linguistic competences of Dutch second-language learners at the B1 and B2 levels of the Common European Framework of Reference for Languages (CEFR)'. *Language Testing* 29, Online gepubliceerd 28 november 2011. DOI: 10.1177/0265532211419826.
- Janssen-Van Dieten, A.-M, 'Hoe "gemeenschappelijk" is het gemeenschappelijk Europees referentiekader?' *Toegepaste Taalwetenschap in Artikelen* 60, 2000, 59-70.
- Janssen-Van Dieten, A.-M., "Nee, ik ben niet van de trap gevallen." Hoe geschikt zijn de niveaudecriptoren van het CEF voor NT2?' *Levende Talen Tijdschrift* 2, (3) 2001, 20-26.
- Janssen-Van Dieten, A.-M., CEFR and L2 learners with a low level of education. I. van de Craats (red.), *Low-educated adult second language and literacy acquisition*. Utrecht, 2006, 143-150.
- Jong, N.H. de, M.P. Steinel, A.F. Florijn, R. Schoonen & J.H. Hulstijn, 'Facets of speaking proficiency'. *Studies in Second Language Acquisition* 34, 2012a, 4-34.
- Jong, N.H. de, M.P. Steinel, A.F. Florijn, R. Schoonen & J.H. Hulstijn, The effect of task complexity on native and nonnative speakers' functional adequacy, aspects of fluency, and lexical diversity. A. Housen, F. Kuiken, & I. Vedder (red.), *Dimensions of L2 performance and proficiency. Investigating complexity, accuracy and fluency in SLA*, Amsterdam, te verschijnen.
- Kwakernaak, E., 'Het ERK hinkt'. *Levende Talen Magazine* 97, (5) 2010, 4-9.
- Liemberg, E. & D. Meijer, *Taalprofielen*. Enschede, 2004.
- Nederlandse Taalunie, *Corpus of spoken Dutch*. 2004. Laatst geraadpleegd op 31 juli 2012 op <http://lands.let.ru.nl/cgn/ehome.htm>.
- Nederlandse Taalunie, *Gemeenschappelijk Europees referentiekader voor moderne vreemde talen: Leren, onderwijs, beoordelen*. Nederlandse Taalunie, 2006.
- Verhagen, A., 'Constructiegrammatica en "usage based" taalkunde'. *Nederlandse Taalkunde* 10, 2005, 197-222.
- Weir, C.J., 'Limitations of the Common European Framework for developing comparable examinations and tests'. *Language Testing* 22, 2005, 301-320.
- Wynants, A., *Drempelniveau: Nederlands als vreemde taal*. Strasbourg, 1985.

APPENDIX

Voorbeelden van de vraagtypen in de grammaticatest

Hieronder staan voorbeelden van de manier waarop kennis in verschillende grammaticale domeinen werd getoetst. De acht kopjes, die de domeinen en de aantallen items vermelden, zijn hier toegevoegd; ze kwamen niet voor in de test zelf.

1. Buiging van sterke en zwakke werkwoorden (19 items) en korte of lange vorm van bijvoeglijke naamwoorden (19 items)

Instructie: Vul de correcte vorm in van het woord tussen haakjes.

Voorbeeld:

(gaan) Anna gaat morgen naar het strand.
(verschillend) Hij had veel verschillende flessen staan naast zijn koelkast.

2. Woordvolgorde in hoofd- en bijzinnen (25 items)

Instructie: Gebruik alle woorden om een zin te maken. Het begin van de zin is al aangegeven. *Je hebt geen extra woorden nodig, gebruik alleen de woorden die er staan.* Schrijf de zin voluit. Soms zijn er twee correcte mogelijkheden, dan is het genoeg om er één op te schrijven.

Voorbeeld:

Je moet dit boek lezen . (lezen dit boek moet)

3. De splitsing van de twee delen van scheidbaar samengestelde werkwoorden (8 items)

Instructie: Gebruik het woord tussen haakjes om de zin af te maken. Soms moet je het werkwoord opsplitsen. De plaats waar het werkwoord moet staan is hier niet aangegeven, dat moet u zelf uitzoeken. Schrijf de zin voluit.

Voorbeeld:

(opstaan) Hoe laat moeten we? → *Hoe laat moeten we opstaan?*

4. Het en er als voorlopig onderwerp of voorlopig voorwerp (26 items)

Instructie: Geef aan welke zin (per paar) je beter vindt door de letter daarvoor te omcirkelen.

1. A/ Zij vindt dat het niet hoort om iemand heel lang aan te kijken.
B/ Zij vindt dat niet hoort om iemand heel lang aan te kijken.

5. De volgorde van meerdere modale partikels (bijwoorden) in één zin (10 items)

Instructie: Kies de beste optie.

1. Anne wil (eens ook)een keer mee.
 A/ eens ook
 B/ ook eens

6. Betrekkelijke voornaamwoorden (15 items); keuze van hulpwerkwoord (hebben of zijn) (10 items)

Instructie: Maak de zin compleet door één geschikt woord in te vullen. Vul een antwoord in dat correct is volgens de regels van het Nederlands.

Voorbeeld:

- Het heeft gisteren de hele dag geregend.
 Dit zijn de mooiste oorbellen die ik ooit heb gehad.

7. Bezittelijke voornaamwoorden (5 items)

Instructie: Maak de zin af door een woord in te vullen dat hetzelfde betekent als wat in de haakjes na het '=' teken is omschreven.

Voorbeeld:

- Deze tas is niet van mij. Is het misschien uw tas? (... = van u)

8. Constructie van passieve zinnen (5 items)

Instructie: Schrijf de zin opnieuw op. Begin zoals aangegeven en maak de zin af zonder te vermelden door wie iets gedaan wordt.

Voorbeeld:

- Ze hebben het huis vorig jaar geschilderd.
 Het huis is vorig jaar geschilderd .