

- Spivey, N., & King, J. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, 24(1), 7–26.
- Verhavert, S., De Maeyer, S., Donche, V., & Coertjens, L. (2017). Scale Separation Reliability: What does it mean in the context of comparative judgement? *Applied Psychological Measurement*, 42(6), 428–445.
- Ward-Penny, R. (2010). *Cross-curricular teaching and learning in the secondary school... Mathematics*. Routledge.
- Winder, R., Kathpalia, S. S., & Koo, S. L. (2016). Writing centre tutoring sessions: Addressing students' concerns. *Educational Studies*, 42(4), 323–339.
- Wubalem, A. Y. (2021). Assessing learning transfer and constraining issues in EAP writing practices. *Asian-Pacific Journal of Second and Foreign Language Education*, 6(17), 1–22.
- Yancey, K. B. (2017). Writing, transfer, and ePortfolios: A possible trifecta in supporting student learning. In J. Moore & R. Bass (Eds.), *Understanding Writing Transfer* (pp. 39–48). Routledge.

EMMA HILBRINK is religiewetenschapper en neerlandica. Ze is recent afgestudeerd aan de Radboud Universiteit en nu werkzaam als docent Nederlands in de bovenbouw van de havo en het vwo op het Pius X College in Almelo.
E-mail: <emma.hilbrink@ru.nl>

JOCHEM ABEN is universitair docent Taalbeheersing van het Nederlands aan de Radboud Universiteit Nijmegen. Hij doet onderzoek naar onder andere schrijfvaardigheidsonderwijs.
E-mail: <jochem.aben@ru.nl>

Bouwen aan bewustetaalvaardigheidsmetingen in het grammaticaonderwijs:

Een instrumentatieonderzoek

GIJS LEENDERS, RICK DE GRAAFF & MARJO VAN KOPPEN

Dit artikel behandelt de totstandkoming van een instrument voor het evalueren van bewuste taalvaardigheid in het grammaticaonderwijs in de L1 (Nederlands) en de L2 (Engels en Duits) in vwo 4. Hiertoe zijn items ontworpen bij drie grammaticale concepten: grammaticale functie, woordvolgorde en congruentie. Deze items richten zich ofwel op het verkrijgen van het 'juiste antwoord' (taalvaardigheid), ofwel op de redenering die aan dat antwoord voorafgaat (taalbewustzijn). We beschrijven hoe het taalvaardigheidsniveau momenteel wordt vastgesteld in het grammaticaonderwijs, hoe onze items tot stand zijn gekomen, op basis waarvan die items zijn geselecteerd, en hoe daarmee drie gelijkwaardige toetsversies zijn samengesteld die gebruikt kunnen worden om het bewustetaalvaardigheidsniveau van leerlingen te evalueren.

Huidige taalvaardigheidsmetingen in het grammaticaonderwijs

In het Nederlandse talenonderwijs wordt de grammaticale kennis van leerlingen zowel in het schoolvak Nederlands als in de moderne vreemde talen in de onderbouw, hoofdzakelijk geoefend en getoetst door middel van

conventionele focus on forms-methoden (o.a. Bonset & Hoogeveen, 2010; Graus & Coppen, 2016; Long, 1991; Rouffet et al., 2022; 2023; Van Rijt & Coppen, 2017; 2021). Deze methoden omvatten het invullen van contextloze zinnen, het reproduceren van allerhande rijtjes vervoegingen en het ontleden van op zichzelf staande zinnen met behulp van beschrijvende concepten, zoals 'onderwerp' en 'lijdend voorwerp'.

De vormgerichte manier van toetsen en de daaraan voorafgaande didactiek richt zich vrijwel uitsluitend op het benoemingsresultaat als opbrengst van diverse, grotendeels taalspecifieke vuistregels (o.a. Coppen, 2011; Graus & Coppen, 2016; Long, 1991) en stelt leerlingen niet in staat om te redeneren over grammaticale vraagstukken (Coppen, 2011; Van Rijt et al., 2020). Zo leveren de geleerde regels problemen op wanneer ze worden toegepast in andere talen en op zinnen die niet specifiek voor het ontleden zijn geconstrueerd (Coppen, 2010; Leenders, 2020; Leenders et al., 2021). Bovendien wordt op geen enkele manier het inzicht gestimuleerd dat er meerdere goede antwoorden kunnen bestaan en de context bepaalt welke daarvan het beste past (Coppen, 2011; Dielemans & Coppen, 2021; Wijnands et al., 2021).

Leerlingen richten zich mede hierdoor vrijwel uitsluitend op het vinden van ‘het juiste antwoord’ en besteden weinig tot geen aandacht aan het doorgronden van de relatie tussen vorm, functie en betekenis (o.a. Coppen, 2010; Wijnands et al., 2021).

De in de onderbouw gebruikelijke *focus on forms*-benadering vertegenwoordigt een structuralistische kijk op taal die zich voornamelijk richt op vorm en veel minder op betekenis, terwijl een *focus on form*-benadering (in Nederland gebruikelijker in de bovenbouw) zich meer richt op taalvaardigheid en impliceert dat leerlingen grammaticale constructies succesvol verwerven binnen een betekenisvolle, communicatieve context (Ellis, 2006; Funk, 2014; Long, 1991; Newby, 2014). Wanneer het grammaticaonderwijs gericht zou zijn op een combinatie van deze benaderingen door het begrip van de relatie tussen vorm, functie en betekenis centraal te stellen (zoals beschreven in o.a. Coppen, 2010 en Wijnands et al., 2021), zal de kennis van taal als systeem onder leerlingen toenemen (Fontich, 2014; Leenders, 2020; Leenders et al., 2021; Dielemans & Coppen, 2021, Van Rijt et al., 2019a). Die toename zou hen in staat kunnen stellen om verschillende grammaticale structuren te verbinden aan de betekenissen die ze uitdrukken, waarmee ze enerzijds hun eigen taal nauwkeuriger, expressiever en doelgerichter leren gebruiken en anderzijds de taal van anderen accurater leren analyseren (Ellis, 2006; Van den Broek & Dielemans, 2017).

Focus op bewuste taalvaardigheid

Met de introductie van ‘bewuste taalvaardigheid’ als beoogd doel van het Nederlandse talenonderwijs (Neijt et al., 2015) en als bouwsteen in de actualisatie van de examenprogramma’s van zowel Nederlands (Herder et al., 2021), als de moderne vreem-

de talen (Fasoglio & Tammenga-Helmantel, 2021) ontstaat er een discrepantie tussen de didactische verrijkingen in het grammaticaonderwijs enerzijds en de conventionele manier van toetsen anderzijds (Van Rijt et al., 2019b; Wijnands et al., 2021). Die vormgerichte aanpak leidt niet tot enig inzicht in taal, laat staan tot bewuste taalvaardigheid (Coppen, 2010; 2011; Van Rijt & Coppen, 2017; Van Rijt et al., 2019a; 2020), terwijl uit onderzoek blijkt dat juist taalbewustzijn de communicatieve vaardigheden kan versterken (o.a. Bialystok, 1988; Garrett & Cots, 2018).

Hoewel de verschillende beschrijvingen van bewuste taalvaardigheid op het eerste gezicht uiteenlopend en complex lijken, is er consensus onder alle auteurs dat het gaat om een vorm van begrip, zij het dat dit begrip wordt aangeduid als ‘inzicht’ (Bax et al., 2021), ‘bewustzijn’ (Carter, 2003) of ‘kennis van het waarom’ (Bolitho et al., 2003). Wat begrepen dient te worden, omvat zowel de structuur van taal (*content knowledge*; taalbewustzijn) als het praktische gebruik ervan (*procedural knowledge*; taalvaardigheid) (Garrett & James, 2000; Van den Broek, 2020; Van Rijt et al., 2020), waarbij impliciete kennis wordt aangevuld met expliciete kennis (Svalberg, 2007) als middel om logisch denken te bevorderen (o.a. Benjamin & Oliva, 2007). Bewuste taalvaardigheid in de context van het grammaticaonderwijs impliceert dus een dieper begrip waarbij het vaak impliciete ‘hoe’ (taalvaardigheid) wordt aangevuld met het explicietere ‘waarom’ (taalbewustzijn).

Een benadering waarin meer aandacht is voor de ontwikkeling van taalbewustzijn, gericht op de taalkundige aspecten van taal en het analyseren van taal(gebruik), zou kunnen bijdragen aan de vakinhoudelijke verdieping van het talenonderwijs van de toekomst (Van den Broek & Dielemans, 2017) en sluit aan bij de vraag naar uitdagender, creatiever

en toekomstgericht talenonderwijs (Fasoglio & Tammenga-Helmantel, 2021; Herder et al., 2021). Leerlingen doen kennis op over taal, en ontwikkelen inzicht in het taalsysteem en (hun eigen) taalgebruik.

In breder kader benadrukt Biggs (1996) het belang van samenhang tussen leerdoelen, leeractiviteiten en toetsing in een curriculum, en concluderen Biggs et al. (2022) dat toetsing en leerdoelen op elkaar afgestemd moeten worden om het leren en de prestaties van leerlingen te bevorderen. Zowel in Nederland als in onderwijscontexten daarbuiten blijken docenten open te staan voor verrijkingen van grammaticatoetsing, maar worden grote klassen en het gebrek aan kennis en tijd om zelf materiaal te ontwikkelen als belangrijkste oorzaken van de huidige toetsing (en didactiek) genoemd (zoals beschreven in Graus & Coppen, 2016; Leenders et al., 2021; Rouffet et al., 2022; 2023).

In een eerdere studie hebben we een voorstel gedaan voor de herziening van leerdoelen en lesinhoud binnen het grammaticaonderwijs met betrekking tot de concepten *grammaticale functie*, *woordvolgorde* en *congruentie* en lieten we zien hoe we in drie verschillende

ontwerpstadia een taalcontrastieve lessenreeks ontwikkelden (Leenders et al., 2024). In de voorliggende studie bouwen we hierop voort en ontwikkelen we een taalcontrastief meetinstrument dat beter aansluit op het vernieuwde doel van het talenonderwijs. Met deze aanvulling hopen we in het kader van *constructive alignment* (Biggs, 1996) een positieve *washback effect* (Alderson & Wall, 1993; Green, 2007) te bewerkstelligen, waarin de toetsing een positieve invloed uitoefent op zowel de didactiek als op het leren. Daarom luidt de onderzoeksvraag:

Welke itemkenmerken zijn essentieel voor het evalueren van bewuste taalvaardigheid in het grammaticaonderwijs in vwo 4?

Vervaardiging van items

In totaal zijn er 936 items ontwikkeld, verdeeld over drie talen, twee bewustetaalvaardigheidstypen (taalvaardigheid/taalbewustzijn), drie concepten en vier itemtypen (zie Tabel 1). Hieronder worden de verschillende itemtypen beschreven en in Tabel 2 t/m 5 is er van elk type een voorbeeld opgenomen.

		Grammaticale functie				Woordvolgorde		Congruentie		Totaal
		I	II	III	IV	I	II	I	II	
NL	Taalvaardigheid	30	11	8	24	60	20	66	24	259
	Taalbewustzijn		11	8	12		20		24	59
EN	Taalvaardigheid	30	11	8	24	60	20	60	20	249
	Taalbewustzijn		11	8	12		20		20	55
DU ¹	Taalvaardigheid	30	11	8	24	60	20	66	22	257
	Taalbewustzijn		11	8	12		20		22	57
Totaal		90	66	48	108	180	120	192	132	936

Tabel 1. Verdeling van de vier verschillende itemtypen over de concepten, talen en taalvaardigheidstypen

Itemtype I

In dit type taalvaardigheidsitems (zie Tabel 2), ook wel bekend onder de naam grammaticaliteitsbeoordelingen², beoordelen de leerlingen de grammaticaliteit van zinnen door per zin ‘correct’ of ‘incorrect’ aan te kruisen. Veel onderzoekers zijn het erover eens dat grammaticaliteitsbeoordelingsitems een prestatie maatstaf vormen voor het taalkundig vermogen van leerlingen (Loewen et al., 2009). Al decennialang worden ze gebruikt in tweedetaalverwervingsonderzoek, maar in het schoolvak Nederlands worden ze tot op heden nog maar weinig ingezet, terwijl ze gemakkelijk te beoordelen zijn voor grote aantallen leerlingen. Bovendien kunnen ze inzicht verschaffen in het beheersingsniveau van struc-

turen die in hun eigen (spontane) productie maar zelden voorkomen (Ellis, 1991).

De door ons ontwikkelde type I-items zijn geïnspireerd op de materialen bij Ammar et al. (2010) en Derwing et al. (2002) en zijn gericht op de impliciete taalkennis en het taalgevoel van leerlingen (Ellis, 1991). Bij het beantwoorden ervan worden zodoende de volgende verwerkingshandelingen van de leerlingen verwacht: observeren en snappen wat er staat (*semantic processing*), vaststellen of de gekozen vorm past bij wat er bedoeld wordt en beoordelen of de zin grammaticaal correct is of niet (*noticing*, Ellis, 2004).

Itemtype II

Dit itemtype is gebaseerd op de *Error Correction*

Itemtype	Taalvaardigheid	Taalbewustzijn	Voorbeelditem
I	x		<p>Beoordeel de grammaticaliteit van de onderstaande zin.</p> <p>I assure I will not tell anyone.</p> <p><input type="checkbox"/> correct <input type="checkbox"/> incorrect</p>

Tabel 2. Voorbeeld itemtype I

Itemtype	Taalvaardigheid	Taalbewustzijn	Voorbeelditem
II		x	<p>Leg uit waarom de onderstaande zin grammaticaal incorrect is.</p> <p>*Der Lehrer wollte die Schüler nicht lassen gehen, als die Glocke läutete.</p>
	x		<p>Kopieer en verbeter de gegeven zin.</p> <p>Zet de verbeterde fout IN HOOFDLETTERS in jouw zin.</p>

Tabel 3. Voorbeeld itemtype II

Test bij Spada et al. (2014) en de eerste sectie van de *Test of metalinguistic knowledge – L2 German* bij Roehr en Gánem-Gutiérrez (2009). In het verlengde van hun werkwijzen bestaan ook onze type II-items uit twee delen (zie Tabel 3). In het eerste deel leggen de leerlingen van grammaticaal incorrecte zinnen uit waarom ze grammaticaal incorrect zijn (taalbewustzijn) en in het tweede deel verbeteren de leerlingen de zin en zetten ze de verbeterde fout in hoofdletters (taalvaardigheid). Het eerste deel vereist naast *semantic processing* en *noticing* nog een derde verwerkingshandeling die Ellis (2004) *metalinguistic reflection* noemt: het identificeren van dat wat onjuist is en uitleggen waarom het onjuist is. Het moedigt leerlingen aan om te redeneren op basis van expliciete kennis van het taalsysteem in plaats van op gevoel of intuïtie en biedt docenten inzicht in de aard van de gebruikte taalkennis, het redeneerproces en het algehele taalbewustzijnsniveau (Loewen et al., 2009). De gevraagde analyse bevordert zodoende het doorgronden van de relatie tussen vorm, functie en betekenis (o.a. Coppen, 2010; Wijnands et al., 2021). In het tweede deel

van dit itemtype kunnen leerlingen zich, bij het verbeteren, zowel baseren op hun expliciete kennis als, bijvoorbeeld bij een gebrek daaraan, uitsluitend op hun impliciete kennis en intuïties. We testen met dit deel of ze de analyse ook kunnen toepassen (productief) op een specifieke situatie.


Voor het concept *grammaticale functie* zijn er naast type I- en II-items ook type III- en IV-items ontwikkeld (zie Tabel 4 en 5). Deze itemtypen zullen hieronder worden beschreven.

Itemtype III

In de geest van itemtype II bestaan ook type III-items uit twee delen (zie Tabel 4), waarin de leerlingen in het eerste deel van verschillende incorrecte grammaticale relaties uitleggen waarom ze grammaticaal incorrect zijn (taalbewustzijn) om ze vervolgens in het tweede deel te verbeteren (taalvaardigheid). Bij het beantwoorden van dit type items worden leerlingen eveneens geacht gebruik te maken van de drie verwerkingshandelingen *semantic processing*, *noticing* en *metalinguistic reflection* (Ellis, 2004).

Itemtype	Taalvaardigheid	Taalbewustzijn	Voorbeelditem
III		x	<p>Leg uit waarom de onderstaande grammaticale relatie incorrect is.</p> <p>Door onze hond werd hij in zijn gezicht gelikt.</p> <p>hij = *LV (lijdend voorwerp)</p>
	x		<p>Kopieer en verbeter de gegeven grammaticale relatie.</p>

Tabel 4. Voorbeeld itemtype III

Itemtype	Taalvaardigheid	Voorbeelditem
	Taalbewustzijn	
IV	x	<p>Zet een kruisje voor alle afbeeldingen die volgens jou een incorrecte interpretatie van de onderstaande zin weergeven.</p> <p>In meinem Schlafanzug habe ich einen Käfer gefunden.</p> 
	x	<p>Leg uit waarom afbeelding 1 volgens jou een correcte/incorrecte interpretatie van de gegeven zin weergeeft.</p> <p>Kopieer de onderstaande zin en vul aan:</p> <p>Afbeelding 1 is (correct/incorrect), omdat...</p>

Tabel 5. Voorbeeld itemtype IV

Itemtype IV

Geïnspireerd op een opgave op bladzijde vijf van de *Test of metalinguistic knowledge* bij Tellier en Roehr-Brackin (2017) zijn er tot slot ook type IV-items ontwikkeld (zie Tabel 5). Hierin wordt de leerlingen gevraagd om in het eerste deel de juistheid te beoordelen van vier visueel weergegeven interpretaties van een ambigue zin (taalvaardigheid) en in het tweede deel voor twee van die interpretaties te onderbouwen waarom ze een (in) correcte weergave vormen van de gegeven zin (taalbewustzijn). De afbeeldingen helpen de leerlingen hierbij om de verschillende betekenissen van de bevroegde zin te begrijpen (*semantic processing*) en te beoordelen (*noticing*) en het onderbouwen van hun oordeel vereist vooral *metalinguistic reflection* (Ellis, 2004).

Deelnemers, analyse en interbeoordelaarsbetrouwbaarheid

In totaal hebben 293 vwo 4-leerlingen (twee scholen, tien klassen, 102 jongens en 191 meisjes) deelgenomen aan deze studie. De drie concepten (*grammaticale functie, woordvolgorde en congruentie*) werden willekeurig aan de leerlingen toegewezen (zie Tabel 6). Na een herhalingsles over het toegewezen concept, maakten ze er in de daaropvolgende les een toets over. Het onderzoek werd goedgekeurd door de Ethische Toetsingscommissie Linguïstiek van de Universiteit Utrecht.

Voor de analyse van de taalbewustzijnsitems werd na een eerste verkenning een analysemodel opgesteld waarin voor elke vraag voorbeelden van vijf antwoordcategorieën werden verzameld, bedoeld om verschillen

	Grammaticale functie	Woordvolgorde	Congruentie	Totaal
Nederlands	36	35	33	104
Engels	32	30	32	94
Duits	31	32	32	95
Totaal	99	97	97	293

Tabel 6. Verdeling van de vwo 4-leerlingen over de verschillende concepten en talen

de gradaties van begrip te weerspiegelen in lijn met literatuur waarin wordt gesteld dat begrip gradueel is (o.a. Baumberger, 2019; De Regt, 2009). Indien de redenering ontbrak of als deze onjuist was, ontvingen de leerlingen nul punten (categorie a en b). Voor onvolledige of slechts gedeeltelijk correcte redeneringen ontvingen ze één punt (categorie c en d) en voor redeneringen die zowel volledig als correct waren twee punten (categorie e). Het 'blind' gebruik van terminologie leverde in lijn met Van Rijj et al. (2020) en Van Rijj en Coppen (2021) ook in onze analyse automatisch nul punten op. In Tabel 7 hebben we een analysevoorbeeld opgenomen van een type III-taalbewustzijnsitem (onderdeel A) waarin leerlingen moesten uitleggen waarom de gegeven incorrecte grammaticale relatie incorrect was. Deze werkwijze hebben we ook voor de type II- en IV-taalbewustzijnsitems gevolgd.

Bij het analyseren van de taalvaardigheidsitems werden twee categorieën gebruikt: incorrecte antwoorden leverden nul punten op en correcte antwoorden één punt. In Tabel 7 hebben we een voorbeeld opgenomen van een type III-taalvaardigheidsitem (onderdeel B) waarin leerlingen de gegeven incorrecte grammaticale relatie moesten verbeteren. Dit type taalvaardigheidsitems kent maar één correct antwoord (net als de

type I- en IV-taalvaardigheidsitems). Bij het verbeteren van de grammaticaal incorrecte zinnen bij de type II-items zijn er echter vaak meerdere goede verbeteringen mogelijk en kunnen leerlingen dus met verschillende antwoorden één punt verdienen (bijvoorbeeld door in het geval incongruentie het getal van ofwel het onderwerp ofwel de persoonsvorm te verbeteren).

Tien procent van de toetsen werd nagekeken door zowel de eerste auteur, als een docent Nederlands, Engels of Duits. De interbeoordelaarsbetrouwbaarheid werd geëvalueerd met behulp van Krippendorffs alpha (Hayes & Krippendorff, 2007) en bleek goed voor de taalbewustzijnsitems (drie categorieën, $\alpha = 0,84 - 0,89$) en uitstekend voor de taalvaardigheidsitems (twee categorieën, $\alpha = 0,96 - 0,98$). De verschillen werden besproken tot er overeenstemming werd bereikt. Gezien de sterke initiële overeenstemming, werd besloten om de eerste auteur de overige toetsen te laten beoordelen.

Selectie van items

Ter selectie van geschikte bewustetaalvaardigheidsitems hebben we een itemanalyse uitgevoerd waarbij we per item hebben gekeken naar de discriminatiewaarde (rit-waarde)

Incorrecte grammaticale relaties uitleggen en verbeteren.
Er zijn dit jaar maar weinig huizen gebouwd.
*Zijn gebouwd = *NG (naamwoordelijk gezegde)*

A. Leg uit waarom deze grammaticale relatie incorrect is. (taalbewustzijn)

Punten	Typering van de antwoorden	Voorbeeldantwoorden van leerlingen
0	a. De redenering <u>ontbreekt</u> . ³	'zijn gebouwd is het WG' 'Is geen naamwoordelijk, maar werkwoordelijk gezegde'
	b. De redenering is <u>incorrect</u> .	'zijn & gebouwd = staan in verschillende tijden' 'een NG heeft nooit datgene in wat is.'
1	c. De redenering is <u>onvolledig</u> .	'Er is geen bijv. naamwoord' 'Zijn gebouwd is het werkwoordelijk gezegde, naamwoordelijk gezegde is wat anders. Voorbeeld: De boom IS DIK.' 'zijn gebouwd zijn allebei werkwoorden'
	d. De redenering is <u>slechts gedeeltelijk correct</u> .	'dit is fout omdat dit gaat over een actie, en niet over een eigenschap van het werkwoord' 'naamwoordelijk gezegde moet ook een stukje zin met bijvoegelijk naamwoord dat wat zegt over het onderwerp'
2	e. De redenering is <u>volledig en correct</u> .	'zijn = hulpwerkwoord: gebouwd is een handeling geen eigenschap, dus het kan geen ng zijn' 'Bij een naamwoordelijk gezegde horen alle werkwoorden + het deel wat iets zegt over het onderwerp. Zijn gebouwd is dus geen naamwoordelijk gezegde, omdat er alleen de werkwoorden in zitten.'

B. Kopieer en verbeter de grammaticale relatie. (taalvaardigheid)

Punten	Typering van de antwoorden	Voorbeeldantwoorden van leerlingen
0	Het antwoord is <u>incorrect</u> .	?
		zijn gebouwd= hulpwerkwoord zijn gebouwd = werkwoord
1	Het antwoord is <u>correct</u> .	zijn gebouwd = WG (werkwoordelijk gezegde)

Tabel 7. Analysevoorbeeld van een type III-taalbewustzijnsitem en een type III-taalvaardigheidsitem

en de moeilijkheidsgraad (p-waarde) om de validiteit en betrouwbaarheid te waarborgen. Als eerste stap hebben we alle te makkelijke en te moeilijke items verwijderd. Deze items zijn niet in staat om onderscheid te maken tussen verschillende bewustetaalvaardigheidsniveaus.

Als tweede stap hebben we alle items ingedeeld per bewustetaalvaardigheidstype. Daarbij hebben we onderscheid gemaakt tussen items die voornamelijk gericht zijn op het evalueren van het taalvaardigheidsniveau en items die voornamelijk gericht zijn op het evalueren van het taalbewustzijnsniveau. Per bewustetaalvaardigheidstype hebben we vervolgens een exploratieve factoranalyse uitgevoerd en hebben we de items geselecteerd die zowel positief met elkaar samenhangen als het merendeel van de variantie in de data vertegenwoordigen.

Als derde stap hebben we een betrouw-

baarheidsanalyse (GLB) uitgevoerd (Sijtsma, 2009) om de interne consistentie en daarmee de betrouwbaarheid van de geselecteerde items te kunnen waarborgen. Deze drie stappen hebben bijgedragen aan de selectie van een robuuste en betrouwbare verzameling items.

Samenstelling van het meetinstrument

Op basis van de geselecteerde items, variërend qua bewustetaalvaardigheidstype (taalvaardigheid/taalbewustzijn), concept en taal, konden nu drie gelijkwaardige toetsversies worden samengesteld. Hierbij hebben we gelet op de gemiddelde moeilijkheidsgraad en discriminatiewaarde van de items volgens het principe van *matched random subtests* (Gulliksen, 1950). In Tabel 8 hebben we

Item		Moeilijkheidsgraad (p-waarde)	Discriminatiewaarde (rit-waarde)	Standaarddeviatie
1.2 Wer singst das Lied "Hallelujah" im Original?	Deze zin is grammaticaal incorrect vanwege de <u>incongruentie</u> tussen het onderwerp en de persoonsvorm.	0,58	0,53	0,51
5.3 Steht ihr sein Geld?	Deze zin is grammaticaal correct vanwege de <u>congruentie</u> tussen het onderwerp en de persoonsvorm.	0,59	0,53	0,51
5.4 Die Frau da draußen hat heute Geburtstag.	Deze zin is grammaticaal correct vanwege de <u>congruentie</u> tussen het onderwerp en de persoonsvorm.	0,59	0,53	0,51

Tabel 8. Voorbeeldset van drie vergelijkbare type I-taalvaardigheidsitems voor het schoolvak Duits bij het concept congruentie

		Grammaticale functie				Woordvolgorde		Congruentie	
		I	II	III	IV	I	II	I	II
NL	Taalvaardigheid	3	1	2	2	2	2	2	2
	Taalbewustzijn		1	2	1		2		2
EN	Taalvaardigheid	3	1	2	2	2	2	2	2
	Taalbewustzijn		1	2	1		2		2
DU	Taalvaardigheid	3	1	2	2	2	2	2	2
	Taalbewustzijn		1	2	1		2		2

Tabel 9. Verdeling van het aantal geselecteerde items over de concepten, itemtypen, talen en taalvaardigheidstypen per toetsversie

een voorbeeld opgenomen van een set type I-taalvaardigheidstypen die zowel qua moeilijkheidsgraad (p-waarde) als discriminatiewaarde (rit-waarde) vergelijkbaar bleken en zodoende verdeeld konden worden over de verschillende toetsversies.

Zoals in Tabel 9 te zien is, bevat elke toets evenveel items van elke taal, bewustetaalvaardigheidstype (taalvaardigheid/taalbewustzijn) en concept. Vooruitlopend op toekomstig effectonderzoek naar de taalcontrastieve lessenserie die we eerder ontwikkelden (Leenders et al., 2024) hebben we de nadruk op grammaticale functie wat betreft het aantal lessen, vertaald naar het aantal items per concept. Het aantal punten voor zowel taalvaardigheid als taalbewustzijn is op elkaar afgestemd om een eerlijke vergelijking mogelijk te maken. Zo kunnen leerlingen per taal in totaal 16 punten verdienen voor taalvaardigheid (1 punt per item) en 16 punten voor taalbewustzijn (2 punten per item).

Conclusie en discussie

Voor het evalueren van bewuste taalvaardigheid in het grammaticaonderwijs in vwo

4 bleken de gemiddelde moeilijkheidsgraad (p-waarde) en discriminatiewaarde (rit-waarde) essentieel. Op basis van deze itemkenmerken kon een meetinstrument samengesteld worden voor het evalueren van bewuste taalvaardigheid in het grammaticaonderwijs in zowel de L1 (Nederlands) als de L2 (Engels en Duits) in vwo 4. Het meetinstrument, bestaande uit drie verschillende toetsversies, is door de manier waarop het is samengesteld in principe even moeilijk en even betrouwbaar; verschillen tussen meetmomenten kunnen dus bij toekomstig gebruik geïnterpreteerd worden als verschillen in het bewustetaalvaardigheidsniveau van leerlingen en niet als verschillen tussen de toetsversies. Hieronder zullen we enkele inhoudelijke kenmerken van het meetinstrument opsommen.

Voor de type I-taalvaardigheidstypen hebben we ervoor gekozen om ook authentieke zinnen te bevragen, afkomstig uit de schrijfproducten van leerlingen in de drie schoolvakken. Zo leren ze ook de grammaticaliteit van zinnen te beoordelen die niet speciaal voor dit doel geconstrueerd zijn (Coppen, 2010; Leenders, 2020; Leenders et al., 2021). Het gebruik van dit type taalvaardigheidstypen stelt ons bovendien in staat om meer correcte

en incorrecte items in de toets op te nemen, wat de betrouwbaarheid ten goede komt.

Voor de type II- en III-items kozen we bewust niet voor grammaticaal correcte zinnen of correcte grammaticale relaties, omdat uit een pilot bleek dat die de leerlingen de mogelijkheid boden om bij verschillende items dezelfde uitleg te geven (o.a. “Het getal van het onderwerp en de persoonsvorm zijn goed op elkaar afgestemd.”) en we daarmee dus onvoldoende inzicht verkregen in hun taalbewustzijnsniveau met betrekking tot de verschillende concepten. Ook een tweede pilotvariant, waarin het bedoelde concept in alle zinnen onderstreept werd, bood geen uitkomst. Hoewel het eerder geconstateerde gebrek aan sturing werd verholpen, leverde deze werkwijze een nieuw probleem op. De fout in de grammaticaal incorrecte items was nu veel gemakkelijker te herkennen, omdat het onderstreepte zinsdeel als aanwijzing diende voor het gemanipuleerde concept. Daarom werd besloten bij deze itemtypes uitsluitend grammaticaal incorrecte zinnen te bevragen. De items bleken goed in staat om inzicht te bieden in de verschillen in het beheersingsniveau van taalvaardigheid en taalbewustzijn. Bovendien stelden ze leerlingen in staat om te redeneren over grammaticale vraagstukken waar in eerder onderzoek voor gepleit werd (Coppen, 2011; Van Rijt et al., 2020). Het viel op dat leerlingen over het algemeen goed in staat waren om de grammaticale incorrecte zinnen/relaties te verbeteren, maar dat hun taalbewustzijn vaak tekortschoot bij het uitleggen waarom de zin of relatie incorrect was.

Met de type IV-items konden we het inzicht stimuleren dat er meerdere goede antwoorden kunnen bestaan en de context bepaalt welke daarvan het beste past. Daarmee voorzien deze items in de tekortkoming die eerder werd geconstateerd in de toetsing in het grammaticaonderwijs (Coppen, 2011; Dielemans & Coppen, 2021;

Wijnands et al., 2021). Omdat we in de type IV-items gebruikmaakten van ambigue zinnen, bleek het voor dit itemtype wel mogelijk om zowel correcte als incorrecte interpretaties van grammaticaal correcte zinnen te bevragen. Het vermogen om zinnen zoals “In meinem Schlafanzug habe ich einen Käfer gefunden” op verschillende manieren te kunnen interpreteren, vormt een cruciaal onderdeel van taalbewustzijn. Het ondersteunt leerlingen bij het begrijpen van complexe taalkundige structuren, helpt hen om zich bewust te worden van de diverse interpretaties van woorden en zinnen, ontwikkelt hun gevoel voor taalnuances en verfijnt hun eigen taalgebruik. Bovendien stimuleert deze vorm van taalkundige analyse het kritisch denken van leerlingen. Dit is van onschatbare waarde in diverse contexten, waaronder academisch schrijven, vertalen en communicatie in professionele omgevingen.

Sommige type IV-items werden echter door leerlingen heel anders geïnterpreteerd dan bedoeld en leverden als zodanig uitdagingen op bij het nakijken. Bij de selectie van items van dit type zijn er dus voornamelijk items afgevallen waarin de afbeeldingen door leerlingen op alternatieve manieren werden geïnterpreteerd en ze daarmee ruis opleverden in de analyses. Wat in positieve zin opviel, was dat de geselecteerde items goed in staat waren om de verschillende beheersingsniveaus van bewuste taalvaardigheid te onderscheiden.

Hoe dan ook dienen we de resultaten met zorg te interpreteren. In de eerste plaats omdat het onderzoek werd afgenomen bij een relatief kleine groep vwo 4-leerlingen. Hoewel het totale aantal leerlingen aanzienlijk is, varieert het aantal leerlingen dat de items bij een specifiek concept en taal heeft gemaakt van dertig tot zesendertig. Bovendien is niet uit te sluiten dat de itemkenmerken beïnvloed werden door de omringende items. Zo is het mogelijk dat een

relatief moeilijk item volgend op een vergelijkbaar makkelijker item, beter gemaakt wordt dan wanneer het in de toekomst in een andere volgorde wordt bevestigd. We hebben dit deels kunnen ondervangen door gebruik te maken van twee versies van elke toets. In versie A werden de verschillende itemtypen sequentieel bevestigd (I-IV) en in versie B in omgekeerde volgorde (IV-I). Deze werkwijze had als bijkomend voordeel dat we van elk item data konden verzamelen, ook wanneer een deel van de leerlingen de toets niet af zou krijgen. In deze studie kregen alle leerlingen de toets binnen de tijd af, wat resulteerde in een complete dataset. Verder hebben de leerlingen in deze studie slechts items beantwoord bij één concept in één taal dat ze de vorige les nog bestudeerden. Het ontwikkelde instrument zal echter in vervolg op deze studie door ons gebruikt worden voor de evaluatie van een taalcontrastieve lessenserie die we eerder ontwikkelden (Leenders et al., 2024). In deze lessenserie (CLAP), bestaande uit animaties en zowel meerkeuze- als open vragen⁴, zijn alle drie de concepten (*grammaticale functie*, *woordvolgorde*, *congruentie*) en talen opgenomen en zodoende zullen de leerlingen ook in de toetsen items beantwoorden bij alle drie de concepten en talen. Daarbij zullen ze dus, anders dan in de voorliggende studie, ook moeten beoordelen om welk van de drie concepten het gaat. Wellicht heeft dit consequenties voor de moeilijkheidsgraad van de items.

Implicaties voor de praktijk

Al met al onderschrijft deze studie de meerwaarde van bewuste taalvaardigheid als doel van het grammaticaonderwijs: inzicht in de structuur van taal (*content knowledge*; taalbewustzijn) en het praktische gebruik ervan (*procedural knowledge*; taalvaardigheid) (Garrett & James, 2000; Van den Broek, 2020; Van

Rijt, 2020). We bepleiten *constructive alignment* (Biggs, 1996) tussen leerdoelen, leeractiviteiten en toetsing om een positief *washback-effect* (Alderson & Wall, 1993; Green, 2007) te bewerkstelligen. Dit impliceert dat de toetsing niet alleen dient voor het meten van kennis, maar ook voor het bevorderen van begrip en bewustzijn van grammaticale structuren.

Het ontwikkelde instrument richt zich daarom op het taalcontrastief evalueren van zowel taalvaardigheid als taalbewustzijn en impliceert daarmee een verschuiving van puur regelgerichte benaderingen, waarin docenten vaak te maken hebben met ondoorzichtige situaties en moeten raden hoe leerlingen tot hun antwoord gekomen zijn, naar een benadering die docenten inzicht biedt in de denkstappen van hun leerlingen betreffende de relatie tussen taalvorm, functie en betekenis (o.a. Coppen, 2010; Wijnands et al., 2021). Hiervan werd al eerder aangegevoerd dat het de groei van de kennis van taal als systeem kan bevorderen (Fontich, 2014; Leenders, 2020; Leenders et al., 2021; Dielemans & Coppen, 2021; Van Rijt et al., 2019a). De verschuiving zou ervoor kunnen zorgen dat leerlingen voortaan diverse grammaticale vormen kunnen verbinden aan de betekenissen die ze overbrengen, en hen zodoende in staat stellen om hun eigen taal nauwkeuriger, expressiever en doelgerichter te leren gebruiken en de taal van anderen accurater te leren analyseren (Ellis, 2006; Van den Broek & Dielemans, 2017). Zo is taal niet alleen relevant binnen school, maar wordt het nadenken over taal ook interessant buiten school (Van den Broek & Dielemans, 2017). Vermoedelijk is het geïsoleerd bevestigen van redeneringen ook goed toe te passen in de didactiek en toetsing binnen andere domeinen van het (talen-)onderwijs.

De gebruikte methode, bekend als *matched random subtests* (Gulliksen, 1950), waarbij items op basis van hun gemiddelde moeilijkheidsgraad (p-waarde) en discriminatie-

waarde (rit-waarde) willekeurig worden verdeeld over verschillende toetsversies, kan zowel in (onderwijs-)onderzoek als in het (talen-)onderwijs breder worden toegepast: in (onderwijs-)onderzoek voor het samenstellen van betrouwbare toetsversies waar groei mee kan worden vastgesteld door middel van voor- en nametingen, en in het (talen-)onderwijs voor het parallel inzetten van verschillende toetsversies om afkijken te voorkomen of voor het samenstellen van een gelijkwaardige set van een toets, een inhaaltoets en een herkansing.

Het zou mooi zijn als het ontwikkelen en afnemen van toetsen in de toekomst niet alleen een verplichting is, maar ook kansen biedt voor groei, begrip en echt leren. Educatieve uitgeverijen zouden hierin een cruciale rol kunnen spelen door inzicht te bieden in de moeilijkheidsgraden en discriminatiewaarden van hun items. Dit zou docenten in staat stellen om zelf robuuste en betrouwbare toetsen samen te stellen die naadloos aansluiten op hun leerdoelen. In het verlengde daarvan is het van essentieel belang dat lerarenopleidingen toekomstige docenten hiertoe opleiden en directies van middelbare scholen hun docenten de nodige tijd geven om betrouwbare toetsen samen te stellen en om zich op dit gebied te kunnen professionaliseren.⁵

NOTEN

- Omdat we in de toetsen de mate waarin leerlingen hun grammaticale kennis konden toepassen evalueerden, werd voor het schoolvak Duits een overzicht van de naamvallen uitgedeeld dat leerlingen als naslagwerk mochten gebruiken tijdens de toets.
- Mackey and Gass (2005) stellen dat L2-leerders hun oordeel bij dit type items baseren op de mate waarin ze het item accepteren als welgevormd. Daarom zou volgens hen een *Grammaticality Judgment*

Task (GJT) moeten worden aangeduid als een *acceptability judgement task*. Hoewel we ons bewust zijn van deze nuance, zullen we in dit artikel 'grammaticaliteitsbeoordeling' als vertaling van de meer gangbare term (GJT) gebruiken.

- Voor het correcte antwoord kunnen de leerlingen in het tweede deel van de vraag een punt verdienen. In dit deel wordt specifiek naar een redenering gevraagd en is een antwoord zonder redenering dus geen punten waard.
- De animaties die de kern vormen van deze lessenserie zijn te bekijken op de pagina 'Nederlands' op www.clap4education.nl.
- Hierbij willen we Dudoc-Alfa en het Amadeus Lyceum bedanken voor het faciliteren van dit onderzoek. Ook zijn we dankbaar voor de medewerking van alle taalkundigen, vakdidactici, docenten en leerlingen en de feedback van de redactie op een eerdere versie van het manuscript. Jullie input heeft ons geholpen om onze keuzes beter te begrijpen en te presenteren, wat heeft bijgedragen aan de verbetering van dit artikel.

LITERATUUR

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Ammar, A., Lightbown, P. M., & Spada, N. (2010). Awareness of L1/L2 differences: Does it matter? *Language Awareness*, 19(2), 129–146.
- Baumberger, C. (2019). Explicating objectual understanding: Taking degrees seriously. *Journal for General Philosophy of Science*, 50(3), 367–388.
- Bax, S., Coppen, P.-A., Evers-Vermeul, J., Gloppe, K. de, Herten, M. van, Mantingh, E., Oostendorp, M. van, & Voorst, S. van (2021). *Bewuste geleterdheid in perspectief: kennis, vaardigheden en inzichten*. Meesterschapsteam Nederlands. Vakdidactiek Geesteswetenschappen.

- Benjamin, A., & Oliva, T. (2007). *Engaging grammar. Practical advice for real classrooms*. NCTE.
- Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, 24(4), 560–567.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, 32(3), 347–364.
- Biggs, J., Tang, C., & Kennedy, G. (2022). *Teaching for quality learning at university* (5e editie). McGraw Hill Education & Open University Press.
- Bolitho, R., Carter, R., Hughes, R., Ivanič, R., Masuhara, H., & Tomlinson, B. (2003). Ten questions about language awareness. *ELT journal*, 57(3), 251–259.
- Bonset, H., & Hoogeveen, M. (2010). *Taalbeschouwing: Een inventarisatie van empirisch onderzoek in basis- en voortgezet onderwijs*. SLO.
- Broek, E. W. R. van den (2020). Language awareness in foreign language education. Exploring teachers' beliefs, practices and change processes. [Dissertatie, Radboud Universiteit Nijmegen].
- Broek, E. van den, & Dielemans, R. (2017). Bewuste taalvaardigheid in de klas. *Levende Talen Magazine*, 104(4), 12–16.
- Carter, R. (2003). Language awareness. *ELT Journal*, 57(1), 64–65.
- Coppen, P.-A. (2010). De taal is een rommeltje. In H. Hulshof & T. Hendrix (Eds.), *Taalkunde en het schoolvak Nederlands* (pp. 26–29). Special Levende Talen Magazine, VLLT.
- Coppen, P.-A. (2011). Grammatica is een werkwoord. In A. Mottart & S. Vanhooren (Eds.), *Vijfentigste conferentie Het Schoolvak Nederlands* (pp. 222–228). Skribis.
- Derwing, T. M., Rossiter, M. J., & Ehrensberger-Dow, M. (2002). 'They spoke and wrote real good': Judgements of non-native and native grammar. *Language Awareness*, 11(2), 84–99.
- Dielemans, R., & Coppen, P.-A. (2021). Defining linguistic reasoning: Transposing and grounding a model for historical reasoning to the linguistic domain. *Dutch Journal of Applied Linguistics*, 9(1-2), 182–206.
- Ellis, R. (1991). *Instructed second language acquisition: Learning in the classroom*. Wiley-Blackwell.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54(2), 227–275.
- Ellis R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83–107.
- Fasoglio, D., & Tammenga-Helmantel, M. (2021). *Startnotitie Moderne vreemde talen: Bovenbouw voortgezet onderwijs*. SLO.
- Fontich, X. (2014). Grammar and language reflection at school: Checking out the whats and the hows of grammar instruction. In T. Ribas, X. Fontich, & O. Guasch (Eds.), *Grammar at school. Research on metalinguistic activity in language education* (pp. 255–283). Peter Lang.
- Funk, H. (2014). Übungsformen im fremdsprachlichen Grammatikunterricht. In S. Dengscherz, M. Businger, J. Taraskina (Eds.), *Grammatikunterricht zwischen Linguistik und Didaktik: DaF/DaZ lernen und lehren im Spannungsfeld von Sprachwissenschaft, empirischer Unterrichtsforschung und Vermittlungskonzepten* (pp. 183–197). Narr.
- Garrett, O., & Cots, J.M. (2018). *The Routledge Handbook of Language Awareness*. Routledge.
- Garrett, P., & James, C. (2000). Language awareness. In M. Byram, (Ed.), *Routledge Encyclopedia of Language Teaching and Learning* (pp. 330–332). Routledge.
- Graus, J., & Coppen, P. A. (2016). Student teacher beliefs on grammar instruction. *Language Teaching Research*, 20(5), 571–599.
- Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39–51.
- Gulliksen, M. (1950). *Theory of mental tests*. Wiley.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.
- Herder, A, Silfhout, G. van, & Jansen, I. (2021). *Startnotitie Nederlands: Bovenbouw voortgezet onderwijs*. SLO.
- Leenders, G. (2020). Grammaticaal taalbewustzijn in de vakken Nederlands, Engels en Duits: Wie/wat+ waar ben ik nou eigenlijk mee bezig? *Levende Talen Magazine*, 107(3), 14–20.
- Leenders, G. P. M., Graaff, H. C. J. de, & Koppen, J. M. van (2021). Hoe meet je bewuste taalvaardigheid? Grammaticaal redeneren in de vakken Nederlands, Engels en Duits. *Pedagogische Studiën*, 98(1), 67–93.
- Leenders, G. P. M., Amstel, M. O. van, Graaff, H.C.J. de, & Koppen, J.M. van (2024). *Het didactiseren van bewuste taalvaardigheid in het grammaticaonderwijs: Grammaticale concepten ingebed in hun grammaticale en taalcontrastieve context* [Manuscript in voorbereiding]. Faculteit Geesteswetenschappen, Universiteit Utrecht.
- Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukasa, K., Ahn, S., & Chen, X. (2009). Second language learners' beliefs about grammar instruction and error correction. *The Modern Language Journal*, 93(1), 91–104.
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *English Language Teaching in its social context* (pp. 180–190). Routledge.
- Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design*. Erlbaum.
- Neijt, A., Coppen, P.-A., & Glopper, K. de (2015). Bewuste taalvaardigheid als eindterm voor het schoolvak Nederlands in havo en vwo: visie en voorbeelden. In A. Mottart, & S. Vanhooren (Eds.), *Negenentwintigste Conferentie Onderwijs Nederlands* (pp. 198–202). Academia Press.
- Newby D. (2014). Do grammar exercises help? Assessing the effectiveness of grammar pedagogy. In C. Haase & N. Orlova (Eds.), *ELT: Harmony and diversity* (pp. 3–16). Cambridge Scholars.
- Regt, H. de (2009). The epistemic value of understanding. *Philosophy of Science*, 76(5), 585–597.
- Roehr, K., & Gánem-Gutiérrez, G. A. (2009). The status of metalinguistic knowledge in instructed adult L2 learning. *Language Awareness*, 18(2), 165–181.
- Rouffet, C., Beuningen, C. van, & Graaff, R.de (2022). Samenhang in vreemdtaalen curricula; Een verkenning van onderwijs- en toetspraktijken in het voortgezet onderwijs. *Levende Talen Tijdschrift*, 23(3), 3–17.
- Rouffet, C., Beuningen, C. van, & Graaff, R. de (2023) Constructive alignment in foreign language curricula: an exploration of teaching and assessment practices in Dutch secondary education, *The Language Learning Journal*, 51(3), 344–358.
- Rijt, J. van, & Coppen, P.-A. (2017). Bridging the gap between linguistic theory and L1 grammar education—experts' views on essential linguistic concepts. *Language awareness*, 26(4), 360–380.
- Rijt, J. H. van, & Coppen, P.-A. (2021). The conceptual importance of grammar: Knowledge-related rationales for grammar teaching. *Pedagogical Linguistics*, 2(2), 175–199.
- Rijt, J. van, Swart, P. de, Wijnands, A., & Coppen, P.-A. (2019a). When students tackle grammatical problems: Exploring linguistic reasoning with linguistic metacognition. *Linguistics and Education*, 52, 78–88.
- Rijt, J. van, Hulshof, H., & Coppen, P.-A. (2021). 'X is the odd one out, because the other two are more about the farmland' - Dutch L1 student teachers' struggles to reason about grammar in odd one out

- tasks. *Studies in Educational Evaluation*, 70, 1–13.
- Rijt, J. van, Wijnands, A., & Coppen, P.-A. (2019b). Dutch teachers' beliefs on linguistic concepts and reflective judgement in grammar teaching. *L1-Educational Studies in Language and Literature*, 19, 1–28.
- Rijt, J. van, Wijnands, A., & Coppen, P.-A. (2020). How secondary school students may benefit from linguistic metaconcepts to reason about L1 grammatical problems. *Language and Education*, 34(3), 231–248.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107–120.
- Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and Integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research*, 18(4), 453–473.
- Svalberg, A. (2007). Language awareness and language learning. *Language Teaching*, 40(4), 287–308.
- Tellier, A., & Roehr-Brackin, K. (2017). Raising children's metalinguistic awareness to enhance classroom second language learning. In M. del Pilar García Mayo (Ed.) *Learning foreign languages in primary school: Research insights*, 115 (pp. 22–48). *Multilingual Matters*.
- Wijnands, A., Rijt, J. van, & Coppen, P.-A. (2021). Learning to think about language step by step: A pedagogical template for the development of cognitive and reflective thinking skills in L1 grammar education. *Language Awareness*, 30(4), 317–335.

GIJS LEENDERS is promovendus aan de Universiteit Utrecht en eerstegraadsdocent Nederlands op het Amadeus Lyceum in Vleuten. Binnen het programma van Dudoc-Alfa werkt hij onder begeleiding van Rick de Graaff en Marjo van Koppen aan zijn promotietraject getiteld: *Van taalvergelijking naar taalbewustzijn: Een contrastiefgrammatische leerlijn voor de vakken Nederlands, Engels en Duits*. Hij onderzoekt hierin hoe het grammaticaonderwijs op middelbare scholen kan bijdragen aan het bevorderen van bewuste taalvaardigheid. E-mail: <G.p.m.leenders@uu.nl>

RICK DE GRAAFF is hoogleraar Didactiek van de Vreemde Talen en Tweektalig Onderwijs aan de Universiteit Utrecht. Zijn onderwijs en onderzoek is gericht op de didactiek en leeropbrengsten van vreemdetalenonderwijs in het basis-, voortgezet en hoger onderwijs. Hij doet praktijkonderzoek in samenwerking met talendocenten en vakdidactici om het leren en onderwijzen van talen betekenisvoller te maken, met speciale belangstelling voor tweetalig en vroeg vreemdetalenonderwijs evenals de rol van meertaligheid. E-mail: <R.degraaff@uu.nl>

MARJO VAN KOPPEN is hoogleraar Variatielinguïstiek van het Nederlands aan de Universiteit Utrecht en senior onderzoeker Variatielinguïstiek aan het Meertens Instituut te Amsterdam. Haar onderwijs en onderzoek is gericht op de taalkunde van het Nederlands. Ze doet enerzijds onderzoek naar variatie in mensentaal en anderzijds naar wat de communicatie tussen dieren, in het bijzonder koeien en katten, ons vertelt over de cognitieve en sociale taalvaardigheden van dieren- en mensentaal. E-mail: <J.m.vankoppen@uu.nl>

VERSCHENEN PROEFSCHRIFT

Instructie in tekststructuur als brug naar beter tekstbegrip?

ROEL VAN STEENSEL

Bogaerds-Hazenberg, S. T. M. (2023). *Text structure instruction in Dutch primary education: Building bridges between research and practice*. Universiteit Utrecht. 306 blz. Isbn 978 94 6483 085 9.

In december 2023 verschenen de uitkomsten van het PISA-onderzoek, waaruit bleek dat het niveau van begrijpend lezen van Nederlandse vijftienjarigen (opnieuw) aanzienlijk is afgenomen (Meelissen et al., 2023). Ongeveer een derde van de leerlingen loopt zelfs het risico om laaggeletterd het onderwijs te verlaten. Uit het PIRLS-onderzoek, dat ongeveer een half jaar daarvoor verscheen, bleek dat er ook onder leerlingen in groep 6 sprake is van een daling in de leesvaardigheid (Swart et al., 2023). In het licht van uitkomsten zoals die van PISA en PIRLS is het belangrijk om te kijken naar recente inzichten over wat onderwijs in begrijpend lezen meer of minder effectief maakt. In juni 2023 promoveerde Suzanne Bogaerds-Hazenberg aan de Universiteit Utrecht op een proefschrift over tekststructuuronderwijs in het basisonderwijs. Hieronder beschrijf ik de centrale aannamen, opzet en voornaamste uitkomsten van haar promotieonderzoek en ga ik in op de vraag wat de conclusies van haar onderzoek betekenen voor de onderwijspraktijk.

Twee centrale aannamen

Bogaerds-Hazenberg startte haar promotieonderzoek vanuit twee aannamen. De

eerste is dat tekstbegrip gebaat is bij de combinatie van kennis over leesstrategieën en kennis over tekststructuren. Kennis over leesstrategieën omvat declaratieve kennis (kennis over welke leesstrategieën er zijn), procedurele kennis (kennis over hoe die strategieën uit te voeren) en conditionele kennis (kennis over wanneer en waarom het zinvol is leesstrategieën in te zetten). Voor effectief gebruik van leesstrategieën, stelt Bogaerds-Hazenberg, is conditionele kennis essentieel. Leesstrategieën moet je toepassen in situaties waarin ze zinvol zijn: strategisch lezen impliceert flexibel en selectief gebruik van strategieën. Juist daarvoor is de combinatie met tekststructuuronderwijs handig: het maakt flexibel en selectief strategiegebruik mogelijk, doordat lezers hun aanpak van een tekst kunnen laten sturen door kennis over hoe die tekst georganiseerd is. Bijvoorbeeld: als je weet dat je een beschouwing voor je hebt en je weet wat een beschouwing inhoudt, dan weet je dat je om tot begrip van die tekst te komen, moet anticiperen op een afweging van argumenten. Met andere woorden, op basis van je kennis over tekststructuren kun je voorspellen (leesstrategie) welk type informatie je tegen zult komen en waar je dus je aandacht op moet richten.

De tweede aanname waarmee Bogaerds-Hazenberg haar promotieonderzoek begon, was dat de effectiviteit van gecombineerd