

- Merrill, M. D. (2002). First principles of instruction. *Educational Technology, Research and Development*, 50(3), 43–59.
- Oller, J. W., & Jonz, J. (1994). *Cloze and Coherence*. Associated University Press.
- Onderwijsraad & Raad voor Cultuur. (2019). *Lees! Een oproep tot een leesoffensief*. Raad voor Cultuur/Onderwijsraad.
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific studies of Reading*, 18(1), 22–37.
- Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssens, T., Kieft, M., Raedts, M., Steendam, E. van, Toorenaar, A., & Bergh, H. van den. (2008). Observation of peers in learning to write; Practice and Research. *Journal of Writing Research*, 1(1), 53–83.
- Rooijackers, P., Silfhout, G. van, & Bergh, H. van den. (2021). Timmeren met oude spijkers en zonder hout: Waarom leergang Nederlands nu niet echt aan strategie-instructie doen. *Levende Talen Magazine*, 108(1), 18–23.
- Rooijackers, P., Silfhout, G. van, Schuurs, U., & Bergh, H. van den. (Aangeboden). De relatie tussen het vooraf lezen van teksten en het beantwoorden van vragen.
- Rooijackers, P., Silfhout, G. van, Schuurs, U., Mulders, I., & Bergh, H. van den. (2020). Lezen en antwoorden bij teksten met vragen; een cross-sectionele eye-trackstudie onder 52 vwo-leerlingen. *Pedagogische Studiën* 97(4), p. 281–308.
- Rouet, J. F., Vidal Abarca, E., Erboul, A. B., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes*, 31(2), 163–186.
- Salmerón, L., Delgado, P., & Mason, L. (2020). Using eye-movement modelling examples to improve critical reading of multiple webpages on a conflicting topic. *Journal of Computer Assisted Learning*, 36(6), 1038–1051.
- Salmerón, L., & Llorens, A. (2019). Instruction of digital reading strategies based on eye-movements modeling examples. *Journal of Educational Computing Research*, 57(2), 343–359.
- Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology*, 103(4), 877.
- Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, 57(2), 149–174.
- Steensel, R. van, & Houtveen, T. (2020). De vele kanten van leesbegrip; Verslag van een literatuurstudie naar kernelementen van effectief onderwijs in begrijpend lezen. *Levende Talen Tijdschrift*, 21(3), 3–12.
- Wolf, M. (2018). *Reader, come home: The reading brain in a digital world*. Harper.

PATRICK ROOIJACKERS is docent Nederlands aan het Sint-Janslyceum in 's-Hertogenbosch en promovendus (vanwege een Dudoc-Alfabeurs) aan de Universiteit Utrecht. Sinds kort werkt hij als toetsdeskundige bij het Cito. E-mail: p.j.h.rooijackers@uu.nl.

GERDINEKE VAN SILFHOUT is als leerplanontwikkelaar Nederlands en expert toetsing werkzaam bij de afdeling voortgezet onderwijs van SLO. E-mail: g.vansilfhout@slo.nl.

HUUB VAN DEN BERGH is hoogleraar Vakdidactiek Nederlands aan de Universiteit Utrecht. E-mail: h.vandenbergh@uu.nl.

De ontwikkeling van een toets luistervaardigheid Nederlands

URIËL SCHUURS

In het schoolvak Nederlands krijgen de schriftelijke vaardigheden de meeste aandacht en is er relatief weinig gerichte aandacht voor het verbeteren van de spreek- en luistervaardigheid. Toch is er reden wel aandacht aan luistervaardigheid te schenken, al is het maar omdat een groot deel van de kennisoverdracht in hbo en universiteit via het gesproken woord plaatsvindt. Als luistervaardigheid om die reden een prominenter plek zou krijgen in het curriculum van het voortgezet onderwijs, dan is het handig te beschikken over toetsen luistervaardigheid. Dit artikel geeft een beschrijving van de ontwikkeling van en onderzoek naar enkele kijk-luistertoetsen voor de bovenbouw van vwo en havo.

Over welke talige competenties moet een eerstejaars student universiteit of hbo beschikken om probleemloos de studie te kunnen volgen? Op deze vraag is geen eensluidend antwoord mogelijk, maar de kwestie krijgt wel in toenemende mate aandacht (vgl. Van der Westen, 2018). Met name schrijfvaardigheid – met daarbinnen tekststructuur, stijl, register, spelling en grammatica – en leesvaardigheid vormen een bron van zorg bij

eerstejaars studenten; tekenend voor die zorgen is dat een toenemend aantal opleidingen een eigen taaltoets heeft ingevoerd, als middel om taalzwakke instromende studenten te identificeren. Een ontoereikende luistervaardigheid wordt binnen dit kader echter weinig genoemd (vgl. Herelixka & Verhulst, 2014; Bonne et al., 2017). Dat is niet geheel verrassend: anders dan een gebrekkige schrijfvaardigheid valt een ontoereikende luistervaardigheid minder direct op.

Toch is er reden wel aandacht aan luistervaardigheid te schenken, al is het maar omdat een groot deel van de kennisoverdracht in hbo en universiteit via het gesproken woord plaatsvindt. Als luistervaardigheid om die reden een prominenter plek zou krijgen in het curriculum van het daarop voorbereidende voortgezet onderwijs, dan is het handig te beschikken over toetsen luistervaardigheid. Dit artikel geeft een beschrijving van de ontwikkeling van enkele kijk-luistertoetsen die specifiek voor de bovenbouw van vwo en havo zijn ontwikkeld¹, van onderzoek naar de beoordelaarsbetrouwbaarheid van een van die toetsen en van de mogelijkheden om de ontwikkelde toetsen in het onderwijs Nederlands in te zetten.

Het belang van luistervaardigheid in de bovenbouw havo/vwo

In het schoolvak Nederlands krijgen de schriftelijke vaardigheden de meeste aandacht. Meer dan eens is geconstateerd dat er relatief weinig gerichte aandacht is voor het verbeteren van de spreek- en luistervaardigheid. Zo constateren Bonset & Braaksma (2008, p. 137) dat er op school wel wordt ‘gesproken en (vooral) geluisterd door leerlingen (als werkvorm en als communicatiemiddel), maar er wordt weinig expliciete aandacht besteed aan het verbeteren van de mondelinge taalvaardigheid’.

In de bovenbouw havo en vwo worden op veel middelbare scholen de mondelinge vaardigheden voor het schoolvak Nederlands aan de orde gesteld in de vorm van een individuele mondelinge presentatie, maar daarbij wordt doorgaans de spreekvaardigheid – of liever: de presentatievaardigheid – van de individuele leerling beoordeeld, niet de luistervaardigheid van het publiek; wel dient volgens de eindtermen te worden beoordeeld in hoeverre de presenterende leerling adequaat kan reageren op bijdragen van luisteraars of gespreksdeelnemers.

Weliswaar wordt door de Projectgroep Nederlands V.O. (2002) een uitgebreide luisterdidactiek beschreven, maar in de praktijk krijgt luistervaardigheid in de moedertaal weinig aandacht, zowel in het voortgezet onderwijs als in onderzoek (vgl. Bonset & Braaksma 2008, p. 144). Het lijkt wel alsof ervan wordt uitgegaan dat leerlingen weinig moeite hebben met luistervaardigheid (vgl. Berckmoes & Rombouts, 2009; Bonne et al., 2017), of dat ze op den duur vanzelf voldoende luistervaardig worden: zij bevinden zich immers jarenlang dag in dag uit in een natuurlijk *immersion program* waarin voortdurend Nederlands gesproken wordt. Waar het relatief korte stukken spraak betreft, is er ook weinig reden om eraan te twijfelen dat

leerlingen leren de betekenis van mondelinge uitingen te begrijpen, maar het wordt anders wanneer leerlingen moeten luisteren naar een langere verhandeling over een voor hen relatief onbekend en abstract onderwerp om daar de essentie uit te halen, zoals in het academisch onderwijs het geval is.

In de centrale examens Nederlands heeft Luistervaardigheid geen eigen plek. Ook in het schoolexamen (SE) krijgt luisteren weinig aandacht. Meestringa & Ravesloot (2013, p. 14) constateren: ‘Standaard is dat in de schoolexamens eigen schriftelijke toetsen (met gesloten en/ of halfopen vragen), mondelinge prestaties (voordracht, discussie- of debatbijdragen), eigen schrijfopdrachten en het persoonlijk gesprek (over een leesdossier bijvoorbeeld) worden gebruikt.’ Slechts in 18% van de havo-eindexamenklassen en in 10% van de vwo-examenklassen heeft luisteren een plek in het SE (Meestringa & Ravesloot 2013). Ook in de ophanden zijnde onderwijsvernieuwing Curriculum.nu speelt luisteren als zelfstandige vaardigheid geen rol van belang; wel wordt het genoemd als onderdeel van mondelinge vaardigheden.

Daarmee is niet gezegd dat Luisteren vanzelf gaat. Bonset & De Vries (2009, p. 27) beschreven de startcompetenties Luistervaardigheid voor beginnende studenten als volgt: ‘Kan een uitgebreid betoog volgen, ook wanneer het niet duidelijk gestructureerd is en wanneer verbanden slechts worden geïmpliceerd en niet uitdrukkelijk worden benoemd.’ De resultaten op het centraal eindexamen Nederlands onderstrepen jaar na jaar dat veel leerlingen moeite hebben om een schriftelijk betoog goed te volgen – terwijl ze daar uitgebreid op hebben geoefend, de tekst kunnen teruglezen, het woordenboek kunnen raadplegen en het tempo van informatie-aanbieding zelf kunnen bepalen. Omdat luistervaardigheid niet geoefend wordt en de spraak niet terug gehoord kan worden, is er reden om te veronderstellen dat een mon-

deling gepresenteerd betoog voor hen nog lastiger te volgen is. Weirich et al. (2019) wijzen erop dat bij luisterfragmenten van meer dan vijf minuten geheugen- en concentratie-effecten al een rol gaan spelen. Ook Bonne & Vrijders (2015) wijzen op de problemen die studenten ondervinden en bieden enkele instrumenten om de luister- en noteervaardigheid van beginnende studenten te verbeteren. Deygers et al. (2017) signaleren dat de toelatingstoetsen die Vlaamse universiteiten momenteel gebruiken, taaltaken bevatten die voor eerstejaarsstudenten in de praktijk van weinig belang zijn. De respondenten van hun onderzoek geven aan dat noteren, dat wil zeggen luisteren en vervolgens de essentie opschrijven, in de top-drie staat van essentiële studievaardigheden.

Uit de literatuur komt Luisteren naar voren als een complexe vaardigheid die niet per definitie als vanzelf aangeleerd wordt. Bij luisteren spelen kennis van de taal en kennis van de wereld een rol. Met name wanneer luisteraars begripsproblemen ondervinden, wordt bewust gebruik gemaakt van cognitieve en metacognitieve strategieën (Goh, 2002; O’Malley et al., 1989; Vandergrift & Goh, 2012). Voorbeelden van cognitieve strategieën zijn:

- Inferenties maken: informatie uit de tot dusverre begrepen tekst of de directe context gebruiken om een meer samenhangende interpretatie van de tekst mogelijk te maken.
 - Voorspellen: anticiperen op komende informatie.
- Voorbeelden van metacognitieve strategieën zijn:
- Voorbereiden om te luisteren: vooraf analyseren van de intenties van de spreker.
 - Realtime beoordelen van input op relevantie: belangrijke en minder belangrijke informatie scheiden, afhankelijk van de taakvereisten.
 - Begripsevaluatie: evalueren of men de be-

luisterde input goed genoeg heeft begrepen om de taak te voltooien.

Er bestaat nog weinig inzicht in de manier waarop, en op welke momenten een luisteraar deze strategieën inzet. De belangrijkste bevinding tot dusverre is dat scores op luistertoetsen vooral samenhangen met de mate waarin de luisteraar bewust van deze strategieën gebruik maakt (vgl. Rukthong & Brunfaut, 2019). Mogelijk geeft dit richting aan de manier waarop luisteronderwijs en de bijbehorende toetsen kunnen worden ingericht.

Uit een studie van Bonset (2010) komt een ontluisterend beeld naar voren: opleiders uit het wetenschappelijk onderwijs signaleren op het gebied van luistervaardigheid regelmatig een inactieve studiehouding bij studenten. Ze zitten vaak bij (hoor)colleges zonder pen en papier en maken geen aantekeningen. Een student verklaarde desgevraagd dat hij geen idee had hoe hij tegelijkertijd aantekeningen zou moeten maken en luisteren. Tegelijkertijd schatten studenten hun eigen luistervaardigheid in als goed tot zeer goed (4,2 op een vijfpuntschaal). Hun docenten denken daar anders over en vermoeden dat de studenten hun eigen lacunes in taalvaardigheid niet zien, behalve bij correct formuleren en spellen, omdat ze daarmee geconfronteerd worden via speciale taaltoetsen van de universiteit. Blijkens Van Onna & Jansen (2006) is het niet ongebruikelijk dat taalgebruikers het niveau van hun lees-, schrijf- en luistervaardigheid in het Nederlands te hoog inschatten: hun onderzoek onder kantoormedewerkers liet zien dat voor de luistervaardigheid in hun moedertaal (Nederlands) het merendeel op B1- of B2-niveau zat, terwijl de ondervraagden zelf veelal aangaven op B2- of C1-niveau te zitten. Recenter constateren De Wachter & Heeren (2017) dat ruim 25% van de eerstejaars studenten niet goed in staat is notities te maken bij colleges; zij zoeken de

verklaring daarvoor vooral in de simultaneïteit van het luisteren en schrijven.

Al met al lijkt er goede reden om meer systematisch aandacht voor luistervaardigheid in het onderwijs Nederlands te vragen in de vorm van onderwijs en bijbehorende toetsing. Voor havo en vwo kan die aandacht het beste gericht zijn op wat later van de leerlingen wordt geëist als zij doorstromen naar hbo en universiteit. Momenteel is er in het voortgezet onderwijs alleen aandacht voor luistervaardigheid op het vmbo: daar is kijk-luistervaardigheid een verplicht onderdeel van het examen.² Bij dat examenonderdeel krijgen de leerlingen enkele kijk-luisterfragmenten van ongeveer vijf minuten gepresenteerd. Deze fragmenten zijn vaak in tweeën geknipt, zodat nog kortere fragmenten resteren. Tijdens de examenafname mogen de kandidaten de fragmenten net zo vaak terugkijken als ze willen. Ze kunnen de videofragmenten ook pauzeren en desgewenst naar eerdere vragen terugkeren als ze dat willen. Deze vmbo-luistertaken aan de hand van opgeknipte fragmenten en met de mogelijkheid van stopzetten en terugkijken, lijken niet bepaald op wat er in de werkelijkheid gebeurt bij bijvoorbeeld colleges of lessen.

Omdat studenten in vervolgoopleidingen, zowel op hbo als universiteit, te maken krijgen met langdurige en massale hoorcolleges is een goede luistervaardigheid in de bovenbouw van havo en vwo van groot belang. Het is aannemelijk dat training van luistervaardigheid in de bovenbouw van het voortgezet onderwijs van nut kan zijn, en dat toetsing van luistervaardigheid daarbij wenselijk is om de bereikte niveaus in beeld te brengen en het luisteronderwijs eventueel bij te stellen. Wanneer een school ervoor kiest het luisteronderwijs te intensiveren, kan de sectie Nederlands overwegen om luistervaardigheid als apart onderdeel van de schoolexamens op te voeren; ook daarvoor zijn toetsen handig.

Met het oog op wat van de beginnende stu-

dent in dit opzicht wordt verwacht, zijn door Cito enkele voorbeeld-luistertoetsen ontwikkeld waarmee kan worden nagegaan of leerlingen in de bovenbouw havo/vwo in staat zijn de essentie te halen uit een wat langere verhandeling. Hierna worden deze toetsen en de gebruiksmogelijkheden ervan toegelicht.

Toetsvarianten in aanbiedingsvorm en in opdracht

Bij leerlingen in de bovenbouw havo/vwo is de kans groot dat ze na diplomering doorstromen naar academisch vervolgonderwijs. Zoals gezegd wordt aan luistervaardigheid in havo en vwo echter nauwelijks tot geen aandacht besteed en evenmin bestaan er luistertoetsen voor deze doelgroep. Om deze leemte aan toetsinstrumenten te vullen hebben we een aantal kijk-luistertoetsen ontwikkeld met als basis een video-opname die afkomstig is van de Universiteit van Nederland, een website die universitair docenten en hoogleraren een platform biedt om voor een groot publiek een mini-college te geven. Dit leek ons een geschikte basis, omdat luistertoetsen voor de bovenbouw havo/vwo noodzakelijkerwijs een inhoud bevatten die refereert aan een wetenschappelijk vakgebied. We hebben gekozen voor vier videofragmenten waarin een wetenschapper spreekt over een academisch onderwerp uit zijn of haar vakgebied op een manier die vergelijkbaar is met een eerstejaars hoorcollege. Elk van de voordrachten duurt ongeveer 15 minuten; daaraan zijn specifieke verwerkingsopdrachten gekoppeld.

In het begin van het ontwikkeltraject hebben we gezocht naar de beste aanbiedingsvorm van de videofragmenten. Daartoe hebben we vier mogelijke varianten in aanbiedingsvorm van de video onderscheiden, die in toenemende mate van moeilijkheidsgraad voor de leerlingen als volgt kunnen worden omschreven:

1. de video bekijken, eventueel stilzetten en desgewenst terugspoelen, pas daarna de opdrachten maken (vergelijkbaar met wat in vmbo-examens gebeurt);
2. de video maximaal twee keer bekijken, daarna de opdrachten maken;
3. de video bekijken en eventueel stilzetten zonder de mogelijkheid tot terugspoelen;
4. de video één keer in zijn geheel beluisteren, tussentijds aantekeningen maken en achteraf zonder terugkijken de opdrachten maken.

Variant 1 en 2 sluiten het minst aan bij de functionele context van bijvoorbeeld een traditioneel hoorcollege, variant 3 lijkt op een hoorcollege waar studenten kunnen vragen om verduidelijking, en variant 4 lijkt het meest op een massaal bezocht hoorcollege waar doorgaans geen vragen aan de docent worden gesteld. Om validiteitsredenen leek ons op de werkwijze bij variant 3 overigens wel wat af te dingen: bij een massaal hoorcollege kan de individuele student immers ook niet vragen om het college stil te zetten.³ In samenspraak met een resonansgroep van docenten bovenbouw havo/vwo hebben we gekozen voor variant 4.

Voor elk van de vier videofragmenten hebben we vier verschillende toetsformats uitgewerkt.

Versie A

Lees de vragen, kijk naar de video en beantwoord de vragen.

Versie B

Geleide samenvatting: kijk naar de video en maak een geleide samenvatting aan de hand van een schema.

Versie C

Ongeleide samenvatting: kijk naar de video en vat de relevante informatie samen.

Versie D

Formuleer zelf de drie 3 (of 4, of 5) belangrijkste vragen die de docent tijdens een tentamen zou kunnen stellen over dit fragment.

Deze vier toetsformats zijn voorgelegd aan docenten uit de bovenbouw van havo/vwo. Versie A leek de geraadpleegde docenten maar weinig afwijken van wat er gebeurt bij een leestoets – alleen de stimulus is andersoortig van aard. Om die reden vonden ze die minder aantrekkelijk. Versie D leek het meest aan te sluiten bij de praktijk van de eerstejaars student die immers ook weinig ruggensteun krijgt bij het beluisteren van een hoorcollege. Deze versie bracht echter een ander probleem met zich mee: het bleek welhaast onmogelijk om een sluitend beoordelingsvoorschrift op te stellen, de verwachting was dat allerlei antwoorden goed gerekend zouden moeten worden. Versie B en C leek de docenten zowel valide als beoordeelbaar. Het verschil tussen een geleide en een ongeleide samenvatting is, dat bij de laatste de leerling helemaal zelf moet bedenken welke elementen er in de samenvatting aan de orde moeten komen; bij een geleide samenvatting krijgt een leerling in de taakstelling al aanwijzingen over de elementen waaraan in de samenvatting aandacht moet worden geschonken. Bij wijze van voorbeeld is in bijlage 1 de geleide samenvattingsopdracht opgenomen die in het onderzoek gebruikt is, inclusief het bijbehorende schema met inhoudelijke aanwijzingen die de leerlingen kunnen gebruiken en een aanduiding van het maximaal aantal te behalen scorepunten per onderdeel.

Bij de vier videofragmenten hebben we meerdere toetsversies gemaakt. Uit de beschikbare toetsen hebben we uiteindelijk acht varianten geselecteerd voor publicatie.⁴ In tabel 1 staan de karakteristieken van de gekozen toetsversies beschreven.

Bij de keuze van de onderwerpen van de video's is gelet op

- de relatieve onbekendheid van het onderwerp voor bovenbouwleerlingen: de voordracht moest liefst aansluiten op de voorkennis van de leerlingen zonder dat de

inhoud van het college hen al bekend zou zijn.

- de gestructureerdheid van de presentatie: de opbouw van het verhaal moest niet te ingewikkeld zijn.
- de aantrekkelijkheid van de video in algemene zin.

Moeilijkheidsgraad en beoordelaarsbetrouwbaarheid

Om de beoordelaarsbetrouwbaarheid bij dit soort toetsen te achterhalen, is in 2019 één toetsversie gekozen om in de praktijk uit te proberen in een pretest. In de video van

Titel van de video	Tijdsduur video	Soort opdracht voor de leerling	Aantal bevroegde aspecten	Maximum aantal scorepunten
Hoe kun je een liegend individu uit de massa halen?	18:19 min.	A inhoudelijke vragen beantwoorden	7	17
		B geleide samenvatting maken	8	17
Hoe ziet de mens er over duizend jaar uit?	11:55 min.	D inhoudelijke vragen formuleren die een docent zou kunnen stellen over dit fragment	4	8
		C ongeleide samenvatting maken	6	13
Waarom eten we het liefst elke dag patat?	13:26 min.	D inhoudelijke vragen formuleren die een docent zou kunnen stellen over dit fragment	5	10
		C ongeleide samenvatting maken	6	11
Waarom gaat de een sneller vreemd dan de ander?	15:24 min.	A inhoudelijke vragen beantwoorden	9	20
		B geleide samenvatting maken	10	16

Tabel 1. Relevante karakteristieken van acht publiceerbare toetsversies

de hiervoor gekozen toets geeft een promovendus een mondeling verslag over haar promotieonderzoek. Voor wat betreft de aanbestedingsvorm is ervoor gekozen de video één keer in zijn geheel te laten zien en beluisteren; het bijbehorende opdrachtformat betreft de geleide samenvatting aan de hand van een schema. De inhoudelijke elementen die de leerling in de samenvatting terug moet laten komen, sluiten aan bij wat aan eerstejaars studenten wordt geleerd, respectievelijk al bekend verondersteld wordt, over de opzet van wetenschappelijk onderzoek. Het is te verwachten dat dit model toepasbaar is voor alle tot dusverre ontwikkelde toetsen, omdat die allemaal betrekking hebben op voor de leerlingen relatief onbekende onderwerpen die vanuit een academische optiek worden besproken.

Doel van de pretest was enerzijds een indruk te krijgen van de moeilijkheidsgraad van de gekozen toets en anderzijds te achterhalen of docenten het onderling in acceptabele mate eens zijn in hun beoordeling van de antwoorden van de leerlingen.

Afname van de pretest

De pretest is afgenomen bij 205 leerlingen, verspreid over vijf vwo-5-klassen en vier havo-5-klassen van negen scholen. In alle klassen duurde de afname van de toets, inclusief enige uitleg over de bedoeling vooraf en een korte leerling-enquête na afname, in totaal maximaal 1 lesuur. Leerlingen kregen bij de geleide samenvattingsopdracht zoals gezegd ook inhoudelijke aanwijzingen over de deelonderwerpen die ze in de samenvatting aan de orde zouden moeten stellen (zie bijlage 1). Vooraf gaven de docenten een korte uitleg waarin ze benadrukten dat het schema aanwijzingen bevatte over wat er in de samenvatting aan de orde zou moeten komen. Leerlingen mochten tijdens het luisteren aantekeningen maken.

Elk leerlingwerk werd twee keer beoor-

deeld. De eerste keer was de eigen docent de beoordelaar; deze voerde het beoordelingswerk individueel uit op school. De tweede beoordeling werd op een gezamenlijk dagdeel uitgevoerd door andere dan de eigen docenten. Omdat niet alle betrokken docenten op een en dezelfde dag aanwezig konden zijn, is er voor de beoordeling van het werk van enkele klassen een beroep gedaan op vierdejaars LIO's van de Hogeschool Arnhem Nijmegen. Tijdens het tweede, collectieve beoordelingsmoment is er eerst een korte toelichting op de doelen van het project gegeven, daarna is de video vertoond en tot slot werd het leerlingwerk beoordeeld (per klas een unieke beoordelaar).

Resultaten van het onderzoek

De beoordelingen zijn geanalyseerd op twee manieren. In de eerste plaats is er een toets- en itemanalyse uitgevoerd op de data, daarna is Cohens kappa coëfficiënt tussen eerste en tweede beoordelaar berekend voor verschillende groepen (per klas, per schooltype en overall). Bij zeven leerlingen werden een of meer elementen in de samenvatting niet beoordeeld; deze leerlingen zijn van verdere analyse uitgesloten. In de gebruikte toets zijn er tien elementen die in de samenvatting moesten terugkomen; deze inhoudelijke elementen leiden samen tot maximaal 16 scorepunten. De belangrijkste resultaten van de Toets- en Itemanalyse staan weergegeven in tabel 2.

De gemiddelde p-waarde (goedscore) voor de toets als geheel is 0,47. Dit betekent dat de toets gematigd moeilijk is voor de betreffende leerlingen. Uitgespitst naar schooltype zien we een gemiddelde p-waarde van 0,51 voor vwo-leerlingen en een gemiddelde p-waarde van 0,41 voor havo-leerlingen, hetgeen erop duidt dat vwo-leerlingen de toets gemiddeld iets beter maakten dan de havo-leerlingen. Cronbach's alfa van de toets, die uit 10 items bestond, bedraagt 0,54.

Aantal leerlingen:	198	Aantal inhoudelijke aspecten:	10
Minimum test score:	0	Maximum test score:	16
Gemiddelde testscore:	7,46	Standaarddeviatie:	2,44
Gemiddelde p-waarde:	0,466	Standaardmeetfout:	1,66
Gemiddelde RiT:	41,7	Coëfficiënt alfa:	0,54

Tabel 2. Resultaten van de toets- en itemanalyse

Met name twee inhoudelijke elementen van de samenvatting (nummer 4 en 5 in bijlage 1) bleken erg moeilijk, met een goedscore van respectievelijk 0,15 en 0,17. Nadere bestudering leerde dat het correctiemodel bij deze opgaven een onduidelijkheid bevatte: leerlingen werd gevraagd een beschrijving te geven van twee instrumenten die in de video besproken werden, maar ze gaven daar collectief ook een oordeel over. In het correctiemodel was ervan uitgegaan dat leerlingen alleen een beschrijvend antwoord zouden geven en daardoor werden erg veel antwoorden fout gerekend. Als we deze opgaven verwijderen uit de analyse, stijgt de gemiddelde p-waarde tot 0,52. Na de pretest is deze fout in het correctiemodel uiteraard hersteld.

Daarnaast is de beoordelaarsovereenstemming berekend. Hiervoor is gebruik gemaakt van Cohen's kappa coëfficiënt (κ). Deze coëfficiënt houdt rekening met toevalsovereenstemming en wordt om die reden beschouwd als een robuuste maat voor overeenstemming tussen beoordelaars (Heuvelmans & Sanders 1993). Als we afzien van item 5 fluctueert κ tussen 0,55 en 0,79. Dit is te interpreteren als redelijk goed tot goed (Landis & Koch 1977).

De deelnemende leerlingen is ook een enquête voorgelegd om hun mening te vragen

over de kijk-luistertoets. Er werd gevraagd om op een vierpuntsschaal aan te geven hoe leuk en hoe moeilijk ze de toets vinden, hoe zinnig ze dit soort toetsen vinden en of ze genoeg tijd hadden om de toets te maken. Zie voor de resultaten tabel 3.

Uit de antwoorden van de leerlingen kan worden afgeleid dat de meerderheid van de leerlingen het maken van een kijk-luistertoets wel leuk vindt (ongeveer 60% antwoordt 'een beetje leuk' of 'leuk'). Een kleine meerderheid vindt het maken van deze toets relatief gemakkelijk (bij havo 56% en bij vwo 68%). De leerlingen ervaren het maken van een dergelijke toets als 'een beetje zinnig' tot 'zinnig' (bij havo 55%, bij vwo 75%). De leerlingen hadden naar eigen zeggen voldoende tot ruimschoots voldoende tijd voor het maken van de toets (bij havo 90% en bij vwo 95%).

Hieruit mag worden afgeleid dat de invoering van kijk-luistermateriaal tot weinig weerstand bij de leerlingen zal leiden. Gelet op de benodigde tijd is het verleidelijk om langere (of meer) toetsonderdelen te gebruiken: de tijd die de leerling ervoor nodig heeft is minder dan tevoren was ingeschat. Ook is het denkbaar dat een kijk-luistertoets wordt uitgebreid met een andersoortige opdracht, bijvoorbeeld een

schrijftaak die inhoudelijk goed aansluit bij de vertoonde video.

Conclusies en aanwijzingen voor de onderwijspraktijk

Voor leerlingen in de bovenbouw van havo en vwo is een goede luistervaardigheid van groot

belang, maar zij worden zelden geoefend in deze vaardigheid. We hebben vier kijk-luistertoetsen ontwikkeld in verschillende varianten die ingezet kunnen worden door scholen die ervoor kiezen meer aandacht aan luistervaardigheid te geven. Deze toetsen hebben alle een mini-college over een wetenschappelijk onderwerp als basis. Voor een van deze toetsen hebben we exemplarisch de

Vraag 1: Hoe leuk vond je het om deze toets te maken?					
	aantal leerlingen	echt niet leuk	niet zo leuk	beetje leuk	erg leuk
HAVO	111	11%	31%	45%	14%
VWO	78	4%	36%	49%	12%

vraag 2: Hoe moeilijk vond je het om deze toets te maken?					
	aantal leerlingen	heel moeilijk	beetje moeilijk	beetje gemakkelijk	erg gemakkelijk
HAVO	115	1%	43%	42%	14%
VWO	79	0%	32%	53%	15%

vraag 3: Hoe zinnig is het om dit soort toetsen te maken?					
	aantal leerlingen	onzinnig	beetje onzinnig	beetje zinnig	erg zinnig
HAVO	112	13%	31%	40%	15%
VWO	77	8%	17%	49%	26%

vraag 4: Had je genoeg tijd om de toets te maken?					
	aantal leerlingen	veel te weinig tijd	iets te weinig tijd	net genoeg tijd	ruimschoots genoeg tijd
HAVO	112	0%	10%	24%	66%
VWO	77	0%	5%	35%	60%

Tabel 3. Resultaten van de leerling-enquête

moeilijkheidsgraad en de beoordelaarsovereenstemming vastgesteld. Deze toets bleek relatief moeilijk, maar wel betrouwbaar te meten. De beoordelaarsovereenstemming bleek redelijk goed tot goed. Leerlingen vonden het maken van de toets leuk, relatief eenvoudig en zinnig.

Uit de pretest blijkt dat de gepresenteerde kijk-luistertoets aan de moeilijke kant was voor de leerlingen. Wellicht komt dit ook doordat zij een kijk-luistertoets niet gewend zijn en er in het voortgezet onderwijs nauwelijks of niet geoefend wordt met het min of meer zelfstandig verwerken van informatie uit videofragmenten.⁵ Het is bovendien denkbaar dat het nu gebruikte format van de geleide samenvatting leidt tot betere prestaties dan bijvoorbeeld de ongeleide samenvatting, omdat er met behulp van steekwoorden al voor de leerlingen wordt aangeduid welke deelonderwerpen in de samenvatting aan de orde moeten komen.

Uiteraard kent dit onderzoek zijn beperkingen. Hier is gekeken naar het beluisteren van langere spraakfragmenten. Het is denkbaar dat ook andere aspecten van luistervaardigheid aandacht verdienen, bijvoorbeeld die in dialogen en discussies; dit viel buiten de scope van het onderzoek. Verder bevinden de vier door ons gekozen videofragmenten zich alle in de hoek van de psychologie en de biologie; vervolgonderzoek kan duidelijk maken of de resultaten anders zijn als er mini-hoorcolleges over andere onderwerpen worden gebruikt. Tot slot zegt dit pilot-onderzoek weinig over de meettechnische kwaliteit van andersoortige luistertoetsen, bijvoorbeeld met behulp van vragen of een ongeleide samenvatting.

Naar analogie van het oefenen van leesteksten mag verwacht worden dat leerlingen beter presteren op luistertoetsen met langere geluidsfragmenten wanneer ze daar vaker mee oefenen. Er is echter weinig tot geen geschikt oefenmateriaal voorhanden.

Wel is het voor docenten relatief eenvoudig om – analoog aan de nu gepubliceerde toetsen – zelf oefenmateriaal te ontwikkelen. Verder beschrijven Bonne et al. (2017) een oefenprogramma voor studenten hbo en universiteit waarin diverse aspecten van luistervaardigheid aan de orde komen. Er zijn ook suggesties gepubliceerd voor de manier waarop bij leerlingen bewustzijn kan worden gecreëerd over de manier waarop zij luisteren (zie Vandergrift & Tafaghodtari, 2010; Dwars & Weststrate, 2014). Tot slot kan in de klas specifiek geoefend worden met het gebruiken van de in paragraaf 1 genoemde luisterstrategieën, zoals inferenties maken, voorspellen waar het betoog naartoe gaat en de geboden informatie op relevantie beoordelen.

Het ontwikkelde materiaal is niet bedoeld als oefenmateriaal: wanneer leerlingen achteraf de antwoorden te zien krijgen, wordt de waarde ervan als toets uitgehouden. Wel kan het materiaal worden gebruikt als toets in bijvoorbeeld een schoolexamen of als voortgangstoets. En zoals gezegd kunnen docenten naar analogie zelf oefenmateriaal ontwikkelen. Om docenten Nederlands in het voortgezet onderwijs hiermee op weg te helpen, zal een korte instructie hierover gepubliceerd worden op www.cito.nl (zie noot 4). Omdat geleide samenvattingen gemakkelijker lijken dan ongeleide samenvattingen, ligt het voor de hand eerst met geleide samenvattingen te oefenen en daarna pas met ongeleide samenvattingen.

We hebben de ontwikkeling beschreven van enkele instrumenten die het mogelijk maken een andere invulling te geven aan schoolexamens. Met name hebben we onderzocht of het mogelijk is een luistertoets te maken met een moeilijkheidsgraad die geschikt is voor examenklas vwo/havo (niet te moeilijk, niet te eenvoudig) en met een acceptabele betrouwbaarheid. In de geest van curriculum.nu zou een dergelijke luistertoets gecombineerd kunnen worden met een leestaak en/of een

schrijftaak. Momenteel wordt er door Cito een project uitgevoerd dat betrekking heeft op de schoolvakken Nederlands en de moderne vreemde talen, waarbij wordt gezocht naar andersoortige toetsinstrumenten, instrumenten waarbij bijvoorbeeld een beroep wordt gedaan op meerdere taalvaardigheden. De ontwikkeling van de hier beschreven prototypische luistertoetsen maakt onderdeel uit van dit grotere project.

NOTEN

1. Het ontwikkelteam bestond, naast de auteur, uit Karin Heesters en Kirsten van Ingen. Marleen de Jonge en Danaï Psathas voerden een literatuur-review uit naar de positie en het belang van luistervaardigheid in het onderwijs Nederlands, Maarten de Groot leverde psychometrische ondersteuning, Hugo Gitsels gaf waardevolle opmerkingen bij een eerdere versie van dit artikel.
2. Voor BB en KB geldt dat kijk-luisteren een vast onderdeel is van het centraal examen voor scholen die digitaal examen doen; voor de andere vmbo-scholen is het een verplicht onderdeel van het schoolexamen.
3. Dit inzicht is in het coronatijdperk in versnelde mate achterhaald geraakt: steeds vaker neemt de docent het college op, zodat de student dit vanaf huis een of meerdere keren kan bekijken. Veel universitaire colleges zijn al veel langer als podcast beschikbaar; vaak is dit een weergave van de presentatie in de met studenten gevulde zaal. Overigens hielden we voor havo variant 2 (de video tussentijds stilzetten) toch al nadrukkelijk als mogelijkheid open, zeker wanneer voor het eerst aan de hand van soortgelijk materiaal in de klas wordt geoefend met luistervaardigheid.
4. De gekozen varianten zijn inmiddels in downloadbare vorm gepubliceerd, zie <https://www.cito.nl/centrale-toetsen-en-examens/cte-onderzoek>.
5. We spreken van ‘min of meer’ zelfstandig

verwerken, omdat we in dit onderzoek een geleide samenvattingsopdracht hebben gebruikt.

LITERATUUR

- Berckmoes, D., & Rombouts, H. (2009). *Verkennd onderzoek naar knelpunten taalvaardigheid in het hoger onderwijs*. Universiteit Antwerpen.
- Bonne, P., & Vrijders, J. (2015). Ear openers voor docenten en hun studenten: Instrumenten voor luister- en noteervaardigheid. *Levende Talen Magazine*, 102(8), 10–14.
- Bonne, P.J., Vrijders, J., & Casteleyn J. (2017). Leren luisteren naar hoorcolleges. In A. Mottart & S. Vanhooren (red.) *31ste Conferentie Onderwijs Nederlands* (pp.135–138). Skribis.
- Bonset, H. (2010). Nederlands in het voortgezet en hoger onderwijs: Hoe sluit dat aan? *Levende Talen Magazine* 3, 17–21.
- Bonset, H., & Braaksma, M. (2008). *Het schoolvak Nederlands opnieuw onderzocht. Een inventarisatie van onderzoek van 1997 tot en met 2007*. SLO.
- Bonset, H., & Vries, H. de (2009). *Talige startcompetenties Hoger Onderwijs*. SLO.
- Curriculum Nu | Leergebied Nederlands (2019). SLO.
- Deygers, B., Branden, K. van den, & Gorp, K. van (2017). University entrance language tests: A matter of justice. *Language Testing* 35(4), 449–476.
- Dwars, A., & Weststrate, Ch. (2014). *Strategietraining voor Franse Luistervaardigheid in de bovenbouw*. *Levende Talen Magazine* 2, 22–27.
- Goh, C. C. M. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, 30, 185–206.
- Herelixa, Ch., & Verhulst, S. (2014). *Nederlands in het hoger onderwijs. Een verkennende literatuurstudie naar taalvaardigheid en taalbeleid*. Nederlandse Taalunie.

Heuvelmans, A.P.J.M., & Sanders, P.F. (1993). Beoordelaarsovereenstemming. In T.J.H.M. Eggen & P.F. Sanders (red.), *Psychometrie in de praktijk* (pp.443–473). Cito.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174.

Meestringa, T., & Ravesloot, C. (2013). Het schoolexamen Nederlands havo/vwo in kaart. *Levende Talen Tijdschrift*, 14(2), 11–19.

O'Malley, J. M., Chamot, A. U., & Kupper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10(4), 418–437.

Onna, B. van, & Jansen, C. (2006). *Nederland talenland? Over de beheersing van Engels, Duits, Frans en Nederlands in Nederlandse organisaties*. Radboud Universiteit.

Projectgroep Nederlands V.O. (2002). *Nederlands in de tweede fase: Een praktische didactiek*. Coutinho.

Rukthong, A., & Brunfaut, T. (2019). Is anybody listening? The nature of second language listening in integrated listening-to-summarize tasks. *Language Testing* 37(1), 31–53.

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening*. Routledge.

Vandergrift, L., & Tafaghodtari, M. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning* 60(2), 470–497.

Wachter, L. De, & Heeren, J. (2017). Noteervaardigheden in het eerste jaar aan de universiteit. In A. Mottart & S. Vanhooren (red.) *31ste Conferentie Onderwijs Nederlands* (pp.124–126). Skribis.

Weirich, S., Bachinger, A., Trendtel, M., & Krelle, M. (2019). Listening comprehension tests in Germany and Austria: Research report and critical review. *L1 - Educational Studies in Language and Literature* 19, 1–26.

Westen, W. van der (2018). Een kijkje over de schutting. *Wensen vanuit het vervolgon-*

derwijs voor het nieuwe curriculum P.O./V.O. 2018-5-24-Van-der-Westen-W.-Wensen-vervolgonderwijs-voor-het-nieuwe-curriculum-po-en-vo-DEF.pdf (taalbeleid-hogeronderwijs.org)

URIËL SCHUURS werkte achtereenvolgens bij Universiteit Utrecht, Radboud Universiteit Nijmegen en tot voor kort bij het Cito. Op laatstgenoemde werkplek ontwikkelde hij examens Nederlands en voerde hij onderzoek uit op het vlak van taaltoetsing. Momenteel werkt hij bij een bureau voor communicatietraining en -advies.
E-mail: uriel@communicatietrainers.nl.

BIJLAGE 1

**Opdracht bij videofragment:
Vreemdgaan**

Je gaat kijken naar een minicollege dat wordt verzorgd door psycholoog dr. Tila Pronk. Maak van dit college een samenvatting op het uitwerkblad. Maak voor je samenvatting gebruik van onderstaand schema. Zorg ervoor dat je samenvatting tussen de 150 en 200 woorden telt.

SCHEMA

	TYPERING	JOUW UITWERKING
1 pnt	1 Introductie onderzoeksvraag	
1 pnt	2 Belang van onderzoeksvraag	
1 pnt	3 Hypothese van het onderzoek	
1 pnt	4 Eerste onderzoeksinstrument: Zelftest vragenlijst	
1 pnt	5 Tweede onderzoeksinstrument	
2 pnt	6 Hoofdonderzoek	
1 pnt	7 Belangrijkste resultaat van het onderzoek	
3 pnt	8 Werking van de hersenen bij verleidingen	
3 pnt	9 Hoe impulscontrole trainen	
2 pnt	10 Twee belangrijkste conclusies	