

# Toetsing en vernieuwing in het vreemde-talenonderwijs

## De nieuwe examenprogramma's zijn niet het einde

ERIK KWAKERNAAK, AART POWW

De eindexamenprogramma's van de verschillende schooltypen oefenen een belangrijke invloed uit op wat er in de scholen gebeurt. Die programma's zijn daarom ook het belangrijkste centrale sturingsmiddel als het gaat om de vernieuwing van het vreemde-talenonderwijs (vto). Veel andere landen gebruiken bovendien centrale, bindende leerplannen en ook leermiddelencontrole als middelen om de kwaliteit van het onderwijs te bewaken en nodige vernieuwing te bewerkstelligen. Zo niet Nederland. Zelfs de inspectie kan zich uiteindelijk alleen beroepen op de basisdoelen en de eindtermen van de examenprogramma's als officiële norm. Vandaar dat het de moeite loont om de nieuwe examenprogramma's voor havo en vwo kritisch tegen het licht te houden, als een aanzet tot discussie en tot onderzoek over de relatie tussen toetsing en didactiek. Met dit doel werd op 4 juni 1999 in Groningen een conferentie gehouden onder de titel Toetsing en vernieuwing in het vreemde-talenonderwijs. In de statige Senaatskamer van het Groningse Academie-gebouw stonden vanuit hun portretlijsten rijen professoren uit vervlogen tijden wat afwezig over de circa 60 hedendaagse deelnemers heen. Vier sprekers, evenveel als er vaardigheden zijn, waren uitgenodigd om elk een voorzet te geven. Op elke voorzet kwam een reactie

van een tweede spreker, plus een discussie onder leiding van telkens een andere voorzitter (zie het programma in het kader). Het volgende is een vrije weergave van de discussie die in Groningen gevoerd werd. Omwille van de leesbaarheid is gekozen voor de vorm van een opiniërend artikel, niet van een verslag dat de discussie op de voet volgt.

### Uitgangspunt

Is het wel zo dat de eindtoetsing het aangrijpingspunt zou zijn om de praktijk van het vto te vernieuwen? Sommigen zetten vraagtekens bij dit uitgangspunt. Met alle respect voor de kwaliteit van het nu beschikbare toetsmateriaal kun je je afvragen tot welke zinvolle leerlingactiviteiten dit geleid heeft bijvoorbeeld in de basisvorming, of bij lezen en luisteren in de bovenbouw havo/vwo. Ervaringen bij onderwijsontwikkelingsprojecten leren dat het allereerst om de ontwikkeling van programma's gaat en dat toetsingsproblemen pas later aan de orde komen. Bovendien hoeft de afsluitende toetsing niet onmiddellijk terug te slaan op de training. Zo hoeft de gescheiden toetsing van vaardigheden in het eindexamen niet te verhinderen dat ze vóór dat examen geïntegreerd geoefend worden.

Toch leert een korte blik terug in de geschiedenis dat examenprogramma's en

ERIK KWAKERNAAK, AART POWW Toetsing en vernieuwing in het vreemde-talenonderwijs

daardoor mogelijk gemaakte of voorgeschreven toetsvormen wel degelijk de lespraktijk beïnvloed hebben. Tot aan de jaren zeventig had het vertalen een belangrijke plaats in de examenprogramma's. Van 1970 dateert het tot voor kort nog voor havo en vwo geldende programma, waarin stond dat de vier vaardigheden getoetst moesten worden. Vanaf die tijd nam het vertalen in de lespraktijk flink af. Het lbo-mavo-examenprogramma mvt van 1986 was een stevige stimulans voor communicatieve onderbouwmethodes, waardoor bijvoorbeeld het naspelen van alledaagse situaties belangrijker werd dan het navertellen van gelezen teksten ('boekentest') in de tijd daarvóór. Bij de invoering van de basisvorming lieten leermiddelenmakers, docenten en het Cito zich leiden door de kerndoelen, bij de nieuwe Tweede Fase havo/vwo door de eindtermen in het nieuwe examenprogramma. We kennen natuurlijk ook negatieve of tenminste discutabele gevolgen van toetsvormen: we discussiëren nu al meer dan een kwart eeuw over het wijd verbreide gebruik van de cse-meerkeuzetoetsen (centraal schriftelijk eindexamen) van het Cito als oefenvorm. Negatieve en positieve terugslag-effecten van examenprogramma's op het voorafgaande onderwijs moeten zorgvuldig geanalyseerd worden.

### Leesvaardigheid

In die centrale Cito-toetsen voor intensieve leesvaardigheid zijn positieve veranderingen te constateren. Ze gaan nu meer teksten bevatten, die op verschillende manieren gelezen moeten worden: soms alleen selectief, namelijk als er alleen bepaalde informatie uit gehaald moet worden, dan weer meer structurerend, als de hoofdzaken eruit gedestilleerd moeten worden, of gedetailleerd, als het om precies begrip van bepaalde passages gaat. Naast de bekende meerkeuzevragen zullen er nu ook open (kort-antwoord-)vragen worden gesteld.

Naast deze verbeteringen blijft een aantal zaken problematisch. Meerkeuzevragen blijven deel uitmaken van het examen. Een nadeel is en blijft dat de toetsmakers daarvoor hun toevlucht moeten nemen tot inhoudelijk moeilijke en 'ver van het bed liggende' teksten. Anders kunnen ze er geen meerkeuzevragen bij maken die goed discrimineren (dat wil zeggen die door genoeg leerlingen fout worden beantwoord).

Verder zou nodig onderzocht moeten worden wat veel docenten vermoeden: bepaalde leerlingen ontwikkelen door het oefenen met de meerkeuzevragen technieken om deze goed te beantwoorden, en slagen in dat laatste ook zonder dat ze werkelijk hun leesvaardigheid in de vreemde taal verbeteren. Met andere woorden: naarmate een kandidaat meer geoefend heeft met meerkeuzevragen, neemt de kans toe dat niet uitsluitend vaardigheid in het lezen en begrijpen van teksten wordt gemeten, maar ook specifieke vaardigheid in het lezen en beantwoorden van meerkeuzevragen. Dat laatste is niet de bedoeling van het examenprogramma en nauwelijks relevant te noemen voor het leven buiten de school. Grondig onderzoek naar de relatie tussen opdrachtvorm en gemeten vaardigheden is dringend gewenst.

Meer open vraag- of opdrachtvormen, zoals onder meer samenvattingen in telegramstijl of trefwoorden, zijn goed voor de inhoudsvaliditeit van de toetsen en voor de voorafgaande training in de scholen, namelijk door het positieve terugslageffect, wanneer de toetsen en/of de toetsvormen als oefening gebruikt worden. Er is wel een 'maar': ze blijven een probleem voor een betrouwbare beoordeling. Nu de leesvaardigheidstoetsen niet meer machinaal, maar handmatig door de docenten nagekeken gaan worden, bestaat de mogelijkheid om in de loop van de jaren met meer open vraag- en opdrachtvormen te experimenteren.

Kunnen de examens niet ook deelttoetsen voor deelvaardigheden gaan bevatten? Immers, als bijvoorbeeld vaststaat dat het raden van woordbetekenissen uit de context en de (receptieve) woordkennis belangrijk zijn voor leesvaardigheid, kunnen die deelvaardigheden ook als apart onderdeel getoetst worden? Hierover lopen de meningen uiteen. Het Cito houdt zich aan de gewoonte dat in het examen alleen volledige vaardigheden getoetst worden, waarin de deelvaardigheden geïntegreerd zijn. In de voorafgaande training kunnen docenten werken met een leerlingenvolgsysteem waarin ook deelvaardigheden gemeenten worden. Daarmee kan beter en op tijd feedback gegeven worden, zodat de leerling weet aan welke aspecten hij of zij moet werken.

Een probleem is en blijft dat de intensieve leesvaardigheid nog steeds de helft van het cijfer uitmaakt. Is het eigenlijk niet treurig dat de Nederlandse vto-beroepsgroep dit al 30 jaar lang accepteert? Niet alleen de verhouding tussen lees- en de andere vaardigheden ligt daarmee scheef, ook die tussen intensief en extensief lezen. Het ITS-behoefteonderzoek van 20 jaar geleden liet al zien dat er in de praktijk meer extensief dan

intensief wordt gelezen. Vormen van meer extensief lezen kunnen wel op kleine schaal in het centraal examen getoetst worden, maar het blijft daar lastig. Goed is dat het extensief lezen nu eindelijk ook in het havo/vwo-examenprogramma staat, maar het is alleen een handelingsdoel. Het vernieuwingseffect van handelingsdoelen blijft sterk afhankelijk van wat de docenten er in de praktijk mee doen.

Het vto zou er niet in mogen berusten dat de overheid om financiële redenen de luistervaardigheid niet in het centrale examen wil hebben. Zo houdt ze de scheefheid in de gewichtsverdeling tussen de vaardigheden in stand. In elk geval zou bij de leesvaardigheid als deelvak bij Duits en Frans (Duits en Frans 1 in het vwo) luistervaardigheid gevoegd moeten worden, ook in de examinering (school-examen). Uit het veld zijn al talrijke signalen te horen dat lezen alleen te eenzijdig is.

Er bestaat een gemis aan planning op middel-lange en langere termijn. Nieuwe examenprogramma's worden uitgewerkt door ad hoc samengestelde commissies, die in meerdere of mindere mate de wensen en trends van het moment in hun voorstellen verwerken. Weliswaar heeft de overheid bepaald dat de kern-

doelen en de nieuwe examenprogramma's na een aantal jaren geëvalueerd en herzien worden, maar een inhoudelijk langetermijnperspectief is niet te bekennen. Dit is van belang voor de kwaliteit en de ontwikkeling van het vto. Er zou een groep gevormd moeten worden, die als een soort denktank de ontwikkelingen en de continuïteit in de gaten houdt en zo'n langetermijnbeleid voor het vto ontwikkelt.

#### Schrijfvaardigheid

Het is nuttig om bij schrijfvaardigheid een kritische blik te werpen op de 'taalvaardigheden'. Zo noemt het examenprogramma – nogal verwarrend – de taalhandelingen of taalfuncties waarmee gepoogd wordt te specificeren wat de leerling schrijvend moet kunnen, bijvoorbeeld eindterm 29: 'bedanken, zich verontschuldigen, feliciteren en uitnodigen'. Het is een bonte verzameling van hoogst frequente naast zelden voorkomende taalhandelingen. Daarbij komt dat een flink deel ervan gerealiseerd kan worden met behulp van een zeer beperkt aantal uitdrukkingen en frasen, passend op één A4'tje. Andere taalhandelingen, zoals eindterm 30: 'informatie geven en vragen', zijn weer zo veelomvattend dat je er als specificatie weinig tot niets aan hebt.

De examenprogramma's geven niveaoverschillen aan door enige differentiatie in schrijftaken (taalhandelingen en deels tekstsoorten), maar welke kwaliteit het taalgebruik op de verschillende niveaus moet hebben, wordt zo goed als geheel opengelaten. Dat is onbevredigend. Weliswaar is het lastig om vast te leggen welke woordenschat (productief) beheerst moet worden, maar men kan zeer wel bepalen welke grammaticale structuren op de verschillende niveaus correct beheerst moeten worden.

Niet alleen de talige, ook de inhoudelijke eisen zijn onduidelijk. De examenprogramma's stellen wel dat de kandidaten de tekst logisch moeten opbouwen. Maar voor schrij-

ven is, bijvoorbeeld volgens de studiebelastingsuren-adviestabellen van de SLO, bitter weinig tijd. Kun je daarin werkelijk aandacht besteden aan training in het logisch ordenen van de inhoud?

Een schrijftaak kan globaal onderverdeeld worden in het vinden van de inhoud, het ordenen (structuren) van de inhoud en het formuleren. In het vervolgonderwijs hebben ex-leerlingen vaak moeite met het ordenen van de informatie in de tekst en met de structuur en coherentie binnen alinea's. Moet het vto daaraan werken, of is dat alleen een taak voor het moedertaalonderwijs en kun je ervan uitgaan dat de leerlingen het dan vanzelf in de vreemde talen ook goed doen? Stemt het vto zijn schrijfvaardigheidsprogramma af op wat het moedertaalonderwijs doet?

Sinds jaar en dag wordt in het Nederlandse vto overwegend gewerkt met opdrachten met opgegeven, voorgestructureerde inhoudspunten, mede in navolging van de voorbeeldtoetsen van het Cito. De schrijftraining in het vto beperkt zich dus tot de problemen van het formuleren. Maar die zijn voor een deel oplosbaar door het correct hanteren van 'bouwstenen' of frasen, zoals die vanouds in brievenboeken en tegenwoordig in computerprogramma's ter beschikking staan.

Dat leidt tot het volgende probleem: het gebruik van 'naslagmateriaal', zoals dat in het examenprogramma is voorgeschreven. Welk materiaal dat precies moet of mag zijn, wordt weer opengelaten, hetgeen forse verschillen in toetscondities - en in de voorafgaande training - mogelijk maakt. Er lijkt zich een tweedeling af te tekenen, met name over het gebruik van (opzoek)grammatica's en brievenboeken of correspondentie-bouwstenenprogramma's. De vto-'deskundigen' die niet voor de klas staan, vinden dat dat moet, met als rechtvaardiging: het gebruik van alle beschikbare hulpmiddelen moet gestimuleerd, geleerd en mede beoordeeld worden. Veel praktijkmensen in de scholen verwacht-

Programma van de rondetafelconferentie Toetsing en vernieuwing in het vreemdetalenonderwijs gehouden op vrijdag 4 juni 1999 in de Senaatskamer, Academiegebouw, Rijksuniversiteit Groningen

<b>Leesvaardigheid</b>	Eerste spreker: dr Erik Kwakernaak (RUG)* Referent: drs Noud van Zuijlen (Cito) Voorzitter: dr Hilde Hacquebord (RUG)
<b>Schrijfvaardigheid</b>	Eerste spreker: dr Cor Koster (VU) Referent: dr Leijn Melse (Cito) Voorzitter: prof. dr Jan Hulstijn (UvA)
<b>Luistervaardigheid</b>	Eerste spreker: dr Arie Hoeflaak (RUL) Referent: dr John de Jong (Swets & Zeitlinger) Voorzitter: dr Lydius Nienhuis (UU)
<b>Gespreksvaardigheid</b>	Eerste spreker: prof. dr Ludo Beheydt (Louvain-la-Neuve) Referent: drs Han van Toorenburg (SLO) Voorzitter: dr Ingrid Wijgh (VU)
<b>Slotbeschouwing</b>	Dr Aart Pouw (RUG)

Dagvoorzitter: Marjolijn Verspoor (RUG). Initiatief en organisatie: Marjolijn Verspoor (RUG), Aart Pouw (RUG), Erik Kwakernaak (RUG) \*Als eerste spreker over leesvaardigheid zou Pieter Loonen optreden. Hij overleed twee maanden voor de conferentie.

ten negatieve effecten voor de leerhouding van leerlingen. Wat betreft het gebruik van woordenboeken gaan wel steeds meer docenten om.

Net als bij leesvaardigheid is het de vraag of deelvaardigheidstoetsen alleen in het traject vóór het examen of ook in het examen zelf ingezet kunnen worden. Het Cito heeft al jaren geleden deeltoetsen gepubliceerd waarin verschillende deelvaardigheden aan de orde komen, bijvoorbeeld openingen, afsluitingen, opbouw van alinea's, variatie in woordgebruik. Aparte deelvaardigheidstoetsen zijn makkelijker te beoordelen en betrouwbaarder. Maar ook bij geïntegreerde schrijfvaardigheidsopgaven kan de beoordeling betrouwbaarder gemaakt worden: door meer specificatie van de criteria in de examenprogramma's, en door centrale (voorbeeld) toetsen met correctiemodellen.

De vragen zijn talrijk en belangrijk genoeg om het wenselijk te maken dat ook voor schrijfvaardigheid een langetermijnvisie en -beleid ontwikkeld wordt.

#### Luistervaardigheid

Voor luistervaardigheid bestaat nog geen echte didactiek, als gevolg van een gebrek aan wetenschappelijke duidelijkheid over wat luisteren nu eigenlijk precies inhoudt en hoe het werkt. Het lijkt op lezen, maar verschilt daarvan in een aantal wezenlijke opzichten. De lezer kan profiteren van de mogelijkheid om te skimmen, vooruit te lezen en weer terug te gaan, zijn eigen leestempo te bepalen en dat te variëren. Maar de luisteraar zit vast aan lineair luisteren in het tempo dat door de spreker bepaald wordt. Om herhaling of verduidelijking vragen, en ook de band meerdere malen afdraaien kan alleen in bepaalde situaties en in elk geval niet bij het luisteren zoals dat in het examen wordt getoetst. Bij het luisteren in het normale leven is er heel vaak extra visuele informatie; die situatie kan benaderd worden door video, maar het zal nog wel even

duren voordat dat in de examens algemeen ingevoerd wordt. Misschien zal de cd-rom, eerst bij het oefenen, later ook bij het toetsen uitkomst brengen.

De verschillen tussen luisteren en lezen zijn in de examenprogramma's beperkt. Eindterm 4 voor lezen: 'onderdelen van een tekst benoemen en het verband tussen delen van een tekst aangeven' heeft geen pendant bij luisteren. Bij eindterm 5 voor lezen: 'op grond van de tekst conclusies trekken met betrekking tot ...' worden genoemd 'beoogd publiek, taalgebruik en schrijfdoel(en), als ook opvattingen en gevoelens van de auteur'. Bij eindterm 12 voor luisteren blijft dat beperkt tot 'intenties, opvattingen en gevoelens van de spreker'. Daar staat tegenover dat eindterm 13 voor luistervaardigheid: 'op basis van het gehoorde anticiperen op het meest waarschijnlijke vervolg van een gesprek' (wat in de pieptoets gebeurt) juist weer geen pendant heeft bij lezen. Dat is vreemd, omdat het daar ook mogelijk en zinvol is om de leerling te laten anticiperen op het vervolg van een afgebroken tekst.

In de Cito-luistertoetsen is niet zoveel verandering te verwachten als in de leestoetsen. De invoering van open kort-antwoord-vragen is onwaarschijnlijk; het Cito overweegt het gebruik van waar/niet-waar-vragen en matching-opdrachten.

Evenals bij leesvaardigheid ontbreekt het in de examenprogramma's aan duidelijkheid wat de eisen zijn op de verschillende niveaus: de moeilijkheidsgraad van de teksten en die van de vragen. In die zin is het maar gelukkig dat bijna alle docenten de Cito-toetsen gebruiken, al was het alleen maar omdat het moeilijk en tijdrovend is om ze zelf te maken. Daardoor wordt tenminste bij deze vaardigheid de eenheid in het schoolexamen een beetje bewaard. Toch zouden tekstcriteria zoals relevantie, complexiteit en concreetheid van de informatie, spreektempo en dergelijke verder uitgewerkt moeten worden. Net als bij

leesvaardigheid zouden open kortantwoordvragen ingevoerd kunnen worden, met correctiemodellen erbij.

Het eigen karakter van luisteren ten opzichte van het lezen moet meer nadruk krijgen, en er moet nauwkeuriger gespecificeerd worden wat de leerlingen op de verschillende niveaus moeten kunnen. De overheid wil aansluiting zoeken bij het Europees Referentie-kader voor Moderne Talen, waarin verschillende niveaus van taalvaardigheid gedefinieerd worden; de Nederlandse examenprogramma's zullen daar nu aan aangepast gaan worden. Het is een aardige exercitie om de Nederlandse niveaubeschrijvingen naast die van het Referentiekader te leggen. Die laatste zijn getoetst ondermeer aan wat *native speakers* kunnen. Als je goed kijkt blijkt dat onze examenprogramma's enkele eisen stellen waar zelfs *native speakers* moeite mee hebben. Bij luistervaardigheid niveau 4 staat: 'Het spreektempo varieert van normaal tot snel'. Van '*spoken language delivered at fast native speed*' is in het Europees Referentiekader alleen sprake op het hoogste deelniveau 2 (*mastery* - subniveau 1 daaronder heet *effective proficiency*) van niveau C (*proficient user*). Dit niveau C2 is niet voor alle *native speakers* in de volle breedte haalbaar. Bij gespreksvaardigheid worden van de Nederlandse kandidaten van het hoogste tot en met het laagste niveau de taalhandelingen 'om aandacht vragen', 'een klacht uiten' gevraagd en wordt geëist dat de spreker 'zijn taalgebruik enigszins kan aanpassen aan de situatie en zijn gesprekspartner(s)', met dien verstande dat op het hoogste niveau het woordje 'enigszins' weggelaten is. Deze criteria dragen dus niets c.q. vrijwel niets bij aan de onderscheiding van de niveaus. In het Europees Referentie-kader doen ze dat wel, doordat ze daar verdeeld zijn over verschillende niveaus.

Het Europese Referentiekader gaat uit van taalgebruikssituaties; bij analyse daarvan blijkt dat luisteren in ongeveer driekwart

daarvan een rol speelt. Dat wijst in de richting van een vraag die Pieter Loonen op de conferentie aan de orde had willen stellen, als het hem vergund was geweest die mee te maken: is het wel zo vanzelfsprekend als we de laatste 30 jaar aannemen, dat de vier vaardigheden in de toetsing zo strikt gescheiden gehouden worden? In het leven buiten de school, en trouwens ook in de meeste buitenlandse examenprogramma's treden ze vaak gecombineerd op.

Daar is tegen in te brengen dat zo'n scheiding van vaardigheden misschien juist belangrijker is voor de training dan voor de toetsing, of preciezer gezegd: belangrijker voor de diagnostische toetsing dan voor de eindtoetsing, net zoals toetsing van deelvaardigheden in het trainingstraject belangrijker is dan aan het eind. In de training heeft de leerling er belang bij om te weten waar en waarom precies de communicatie met hem minder goed verloopt; dan kan hij er gericht iets aan doen. In de eindtoetsing gaat het er om vast te stellen of het totale communicatieve resultaat goed is.

Ook bij luistervaardigheid blijven meer dan genoeg vragen open waarover de discussie door moet gaan, en niet pas op het moment dat een herziening van de examenprogramma's op de agenda staat.

#### Gespreksvaardigheid

Dat de vaardigheden in het examen gescheiden getoetst moeten worden, heeft een positief terugslag-effect op het onderwijs. Door aan de afzonderlijke vaardigheden een wegingspercentage toe te kennen, worden docenten gedwongen 'gewogen aandacht' aan de verschillende vaardigheden te geven. Wat niet getoetst wordt, wordt niet onderwezen; wat in combinatie met andere onderdelen getoetst wordt – zoals vroeger de spreekvaardigheid met de boekenlijst en/of literatuur –, krijgt makkelijk minder aandacht dan wenselijk is. Juist bij gespreksvaardigheid is dit van het

hoogste belang. Nu deze didactisch lastige vaardigheid expliciet met een bepaald gewicht in het examenprogramma staat, kan niemand haar meer als toevallig nevenproduct van het vto behandelen. Ook beheersingsniveau-eisen moeten per vaardigheid vastgesteld worden. Wie een artikel of boek over biochemie kan lezen, kan niet per se ook deelnemen aan een mondelinge discussie daarover.

De examenprogramma's schieten ook op het punt van de gespreksvaardigheid tekort in precisie. Wat betekenen allerlei abstracte termen nu concreet, bijvoorbeeld 'monoloog', 'uiteenzetting', 'volledigheid', 'uitvoerigheid', 'begrijpelijkheid', 'gepastheid van taalgebruik'? Ook over de toetsvorm wordt nauwelijks iets concreets gezegd; zo wordt niet eens gespecificeerd of het om een geleide toets (zoals bij schrijfvaardigheid gebruikelijk) of om een open vorm moet gaan. In het laatste geval moet de leerling niet alleen formuleren in de vreemde taal, maar ook nog zelf inhouden vinden en ordenen. Hoe ga je dan om met het verschil tussen de ene leerling die zich op de vlakke houdt met wat standaardfrases en een beperkte woordenschat, en de andere die zich vol overgave in een onderwerp stort, in allerlei woord- en andere formuleringsproblemen verzeild raakt en daardoor misschien wel iets interessants zegt, maar helaas slecht formuleert?

Een oplossing kan gevonden worden in een combinatie van open en meer geleide toetsonderdelen. De eerste hebben meer inhoudvaliditeit, omdat spreken in normale communicatie ook vaak vereist dat de spreker tegelijk de inhoud genereert én voor de vreemdtalige formulering zorgt. Dat maakt echter een betrouwbare beoordeling lastiger. Betrouwbaarheid is weer makkelijker te bereiken bij geleide toetsing, omdat het daar eerder mogelijk is om de prestatie van de ene leerling te vergelijken met die van anderen - wat tenslotte de basis van alle toetsing is.

Die combinatie van meer open en meer

geleide toetsonderdelen, ondersteund met goede scoringsadviezen, kunnen voor de docenten verhelderen waar het om gaat, en wat meer is, ook een positieve terugwerking hebben op de training. Daarbij moet wel gewaarschuwd worden voor een overwaardering van geleide toetsing en een mogelijk negatief teruglageffect daarvan. Leerlingen en docenten houden van duidelijkheid in taken en eisen, en dat zou kunnen leiden tot het stampen en reproduceren van frasen en standaardgesprekjes.

De beoordeling van gespreksvaardigheid is en blijft een lastig probleem. Interessant is het recente proefschrift van Jan van Maele uit Louvain-la-Neuve, waarin het ging over het beoordelen van de taalvaardigheid van niet-native sprekers. Beoordelingen door *native speakers* en door *niet-native speakers* werden met elkaar vergeleken. Het bleek dat docenten die geen *native speaker* van de te beoordelen taal waren, anders oordeelden dan de *native speakers*, en dat de beoordelingen van de eerstgenoemde groep bovendien onderling sterk verschilden. Verdere precisering van de aan te leggen criteria is hoogst noodzakelijk.

Het kernprobleem bij gespreksvaardigheid is de organisatie van de training in de klas. Hoe creëer je representatieve situaties met voor de leerlingen relevante thema's? Hoe organiseer je de feedback? Het belangrijkste is in feite niet de eindtoetsing, maar de continue evaluatie en terugkoppeling naar de leerling, zodat die in zijn leerproces gestimuleerd wordt. Daarbij moet de leerling niet afhankelijk gemaakt worden van correctie door de docent, maar zij/hij moet ook leren zelf de kwaliteit van haar/zijn eigen uitingen in de vreemde taal te controleren. Dat is een mogelijke deeloplossing van het grote probleem bij gespreksvaardigheid: hoe krijgt elke leerling de bij zijn vreemdetaalontwikkeling passende feedback, die hij/zij werkelijk kan opnemen en verwerken?

De didactiek van de gespreksvaardigheid staat in feite nog in de kinderschoenen. De meest dringende vragen op dit gebied lijken te zijn: open of geleide opdracht- en toetsvormen; de weging van inhoud en taalvorm bij de toetsing en de consequenties voor de training als de inhoud meegewogen wordt; de organisatie van de gespreksvaardigheidstraining waarbij de leerling genoeg feedback krijgt die zij/hij kan gebruiken in zijn taalleerproces. Meer dan genoeg om een voortzetting van de discussie en een roep om onderzoek te rechtvaardigen.

#### Slotbeschouwing en voornemens

De nieuwe examenprogramma's moderne vreemde talen voor havo en vwo doen een poging om dat wat de leerlingen moeten kunnen, preciezer vast te leggen dan het oude examenprogramma van 1970. Ze leiden de eindtermen af van de vormen waarin natuurlijke communicatie plaatsvindt.

Op de conferentie in Groningen is geweest op de wenselijkheid van toetsen of instrumenten met een diagnostische functie, zodat leerlingen meer helderheid krijgen over het niveau van hun prestaties en over de vraag wat ze aan verbetering kunnen doen. Er wordt een sterk beroep gedaan op de creativiteit van toetsontwikkelaars om nieuwe, valide en betrouwbare toetsvormen te ontwikkelen. Het examenprogramma moet nog op veel punten worden verduidelijkt. Er is weinig gezegd over de inbreng van docenten bij dit alles.

Het zal zaak zijn bij voorstellen tot vernieuwing van de toetsing de uitvoerbaarheid te betrekken. In het onderwijs en zeker in het eindexamen zijn de mogelijkheden beperkt om natuurlijke communicatie in een vreemde taal na te bootsen, terwijl het toch duidelijk het streven van de nieuwe examenprogramma's is om het vto te richten op dat wat de ex-leerlingen buiten de school nodig hebben. De mogelijkheden om in dat kader valide en betrouwbare toetsing te realiseren, zijn beperkt.

Als belangrijke discussiepunten kwamen op de conferentie naar voren:

- het ontbreken van een beleids- en ontwikkelingsperspectief voor het Nederlandse vto op middellange en langere termijn;
- de vraag of de strenge scheiding van vaardigheden en het taboe op aparte deelvaardigheidstoetsen in de examens gerechtvaardigd zijn;
- de open- en onduidelijkheid van de examenprogramma's over toetsvormen en belangrijke toetscondities in de schoolexamens zoals het 'naslagmateriaal' bij schrijfvaardigheid;
- de gebrekkige specificatie van taalhandelingen ('taalvaardigheden'), beoordelingscriteria en beheersingsniveaus;
- de weging van intensief lezen ten opzichte van extensief lezen en de overige vaardigheden;
- de problematische validiteit van meerkeuzevragen als toetsvorm als ze ook als oefenmiddel gebruikt worden;
- de onduidelijke onderscheiding van inhoudelijke en talig-formele aspecten bij productieve taken (schrijf- en gespreksvaardigheid) in de toetsing en de consequenties daarvan voor de voorafgaande training.

Duidelijk is dat er ook behoefte bestaat aan onderzoek: naar de ervaringen van docenten en leerlingen met de nieuwe examenprogramma's en de gebruikte toetsvormen, naar de effecten ervan op leermiddelen, op docenten en op leerlingen, naar nieuwe toetsvormen en beoordelingsprocedures, naar mogelijkheden om de examenprogramma's en – vooral – de vto-praktijk te verbeteren.

De deelnemers hebben afgesproken de discussie voort te zetten op een volgende conferentie in Amsterdam of Utrecht. De denktank is in Groningen in beweging gezet en rolt naar het midden van het land. Zal hij Zoetermeer en Den Haag ooit bereiken?