

Nederlandse examens en de Europese schaal van taalvaardigheid

GÉ STOKS EN JOHN DE JONG

*Elders in dit nummer zijn de achtergronden van het Europees Referentiekader en de Europese schaal van taalvaardigheid uitvoerig beschreven. In deze bijdrage zullen we eerst kort ingaan op het belang van de afstemming van de Nederlandse examens op die Europese schaal van taalvaardigheid. Daarna zal het project Excef worden beschreven, een door SLO en Citogroep uitgevoerd onderzoek naar de mogelijkheden om de Nederlandse examens af te stemmen op de Europese schaal van taalvaardigheid.**

Belang van de Europese schaal van taalvaardigheid

Als een Nederlandse leerling in het buitenland wil laten zien hoe goed hij zijn talen beheerst, kan hij niet volstaan met het noemen van het cijfer op zijn examenlijst. Een zeven voor Engels op de havo zegt iemand in het buitenland weinig. Ook wanneer het cijfer zou worden uitgesplitst naar vaardigheden blijft het probleem bestaan dat een cijfer voor mensen die het onderwijssysteem in Nederland niet kennen geen duidelijke indicatie is voor het bereikte niveau van taalvaardigheid. Daarom is ook de Europese schaal van taalvaardigheid van het Europees

Referentiekader ontwikkeld. Wanneer in Europa alle examens voor de moderne vreemde talen aan de Europese schaal worden gerelateerd ontstaat er veel meer duidelijkheid en transparantie over de niveaus die behaald worden. In verschillende landen van Europa is men inmiddels hiermee bezig en ook de instanties aangesloten bij ALTE (Association of Language Testers in Europe) hebben hun examens (Cambridge, Goethe, DELF/DALF, Cervantes, etc.) aan de Europese schaal gerelateerd.

Over de manier waarop kwalificaties aan de schaal kunnen worden gerelateerd bestaat nog onduidelijkheid. Soms wordt dat op een intuïtieve manier gedaan. Men vraagt een aantal docenten (of leerlingen/studenten) welk niveau zij denken dat een gemiddelde leerling heeft behaald op de Europese schaal en als er onder de bevraagde docenten voldoende overeenstemming bestaat, stelt men dat dat het behaalde niveau is. Omdat het Ministerie van Onderwijs, Cultuur en Wetenschappen van mening was dat de inschatting van de Nederlandse examens op een zorgvuldige manier moest plaatsvinden is dit probleem nader verkend in het zogenaamde Excef-project. Dit project is door SLO en Citogroep uitgevoerd. In het navolgende worden zowel de gevolgde procedure om tot een

inschatting te komen, als de uitkomsten van een verkennend onderzoek gepresenteerd. Het gehele onderzoek kan gratis aangevraagd worden bij de SLO (Van Hest, De Jong en Stoks 2001).

Het bepalen van de niveaus

Het probleem van de vaststelling van de niveaus is in het Excef project vanuit twee invalshoeken onderzocht.

1. De SLO heeft de examenprogramma's voor vmbo en havo/vwo geanalyseerd om te achterhalen welk niveau met deze examenprogramma's nagestreefd wordt. Daartoe zijn de formuleringen van de eindtermen vergeleken met die van de niveaudecriptoren in het Europees Referentiekader.
2. De Citogroep heeft voor een beperkt aantal talen en een beperkt aantal niveaus een onderzoek uitgevoerd, waarin docenten Frans, Duits en Engels de examens voor lezen, luisteren en spreken geclassificeerd hebben in termen van het Europees Referentiekader. Met andere woorden, zij moesten proberen aan te geven op welk niveau van het Referentiekader de opgaven in de examens zich bevonden.

Ten behoeve van het project zijn de niveaubeschrijvingen in het Nederlands vertaald. Verder zijn de zes niveaus van de Europese schaal van taalvaardigheid nogmaals onderverdeeld zodat er in totaal 12 niveaus kunnen worden onderscheiden. Dit is gedaan omdat binnen veel schoolsystemen, ook in het buitenland, behoefte bestaat aan fijnere niveauonderscheidingen, bijvoorbeeld tussen vmbo en havo of tussen havo en vwo. De hoogste niveaus (C1 en C2) worden binnen een schoolstelsel vrijwel niet bereikt, dus moeten de prestaties van 95% van alle leerlingen in het voortgezet onderwijs binnen de niveaus A1.1 tot en met B2.2 gepositioneerd worden.

Analyse van examenprogramma's

Voor de analyse zijn de examenprogramma's van vmbo en havo/vwo geanalyseerd. De descriptoren van het Referentiekader zijn vergeleken met de formuleringen in de examenprogramma's. Daarbij bleek, zoals wel was te verwachten, dat de systematiek die gebruikt is bij de examenprogramma's niet altijd overeenkomt met die van het Referentiekader.

Op de eerste plaats zijn er binnen ieder van de vaardigheden deelaspecten die in het examenprogramma (en ook in de examens) niet aan bod komen. Dit duidt direct al op een probleem bij de analyse van de examenprogramma's in termen van het Europees Referentiekader: de beperkte dekking van het totale domein maakt het niet goed mogelijk om zonder meer aan een examenprogramma het label van een niveau uit het Referentiekader te koppelen.

Verder worden in het Europees Referentiekader bij leesvaardigheid bijvoorbeeld de niveaus voornamelijk gedefinieerd in termen van het lezen van tekstsoorten en de mate waarin deze worden begrepen, terwijl de eindtermen in de examenprogramma's spreken van handelingen op basis van gelezen teksten zoals relevante informatie aangeven, hoofdgedachte aangeven, betekenis van belangrijke elementen aangeven, vergelijken en conclusies trekken. Al deze zaken zijn op zich wel te interpreteren als lezen en begrijpen, maar de systematiek is van een andere aard. In het Europees Referentiekader wordt het niveau onderscheid bij de receptieve vaardigheden beschreven in termen van de kenmerken van tekstsoorten, tekstonderwerpen en moeilijkheidsgraad van de teksten. Daarnaast worden aparte niveaubeschrijvingen gegeven van mogelijke taken naar aanleiding van de teksten. In de Nederlandse eindexamenprogramma's worden geen niveauonderscheidingen gehanteerd met betrekking tot de taken.

Het Europees Referentiekader bevat meer

dan 20 schalen die betrekking hebben op gespreksvaardigheid of aspecten daarvan. Bovendien wordt in het Europees Referentiekader systematisch onderscheid gemaakt tussen mondelinge interactie en mondelinge productie. Qua intentie kan men stellen dat in de examenprogramma's met name sprake is van mondelinge interactie (gespreksvaardigheid). Bij de Cito spreektoetsen is er echter zelden sprake van echte interactie, maar van tamelijk korte antwoorden op stimuli. Dat zou dus betekenen dat er eerder sprake is van mondelinge productie. In de algehele analyse worden alle schalen (dus zowel die voor mondelinge productie als voor mondelinge interactie) op hun relevantie bekeken en wordt aangegeven welke schalen/categorieën voor het formuleren van eindtermen voor gespreksvaardigheid van belang kunnen zijn (zie Staatsen & Stoks 2001). Voor de vaststelling van de beoogde niveaus op havo en vwo is de toelichting op het examenprogramma (de zogenoemde niveaumeter) die ten grondslag ligt aan de niveaubeschrijvingen op havo en vwo geanalyseerd en is bekeken welke niveaus hier geambieerd worden.

Uit de analyses blijkt dat de niveaus op het vmbo zich voor wat de receptieve vaardigheden betreft op de niveaus A2.2 en B1.2 bevinden, met uitzondering van het laagste niveau voor Frans (A2.1). Van de productieve vaardigheden is alleen de gespreksvaardigheid onderzocht. Die bevindt zich op de niveaus A2.1 en A2.2.

Zoals was te verwachten bevinden de niveaus voor havo en vwo zich voor wat de receptieve vaardigheden betreft op het niveau B2.1 (havo) en B2.2 (vwo). Voor gespreksvaardigheid is de spreiding iets groter en varieert van B1.1 voor Frans havo tot B2.2 voor Engels en Duits vwo.

Analyse van de examens en examenopgaven

De Citogroep heeft in het kader van dit project een beperkt onderzoek verricht naar de feitelijke

niveaus van de examens. Tien mvt-docenten uit het Nederlandse voortgezet onderwijs hebben de examens voor lezen en de eindtoetsen voor luisteren en spreken geïnterpreteerd in termen van het Europees Referentiekader. Zij moesten proberen om als groep zo veel mogelijk tot een gelijke, taalonafhankelijke, interpretatie van de niveau-omschrijvingen te komen. Per vaardigheid kregen de docenten (tekst)fragmenten of leerlingprestaties aangeboden voor elk van de vier examenniveaus (mavo-C, -D, havo, vwo) en voor elk van de drie talen (Engels, Duits, Frans).

Voor luisteren en lezen beoordeelden de docenten eerst de teksten of tekstgedeelten zonder de opgaven. Daarna beoordeelden zij deze in combinatie met de bijbehorende opgaven. Voor spreken werd de taalproductie van leerlingen beoordeeld: bandopnamen van leerlingen bij de uitvoering van een aantal spreekopdrachten uit spreektoetsen havo en vwo van de Citogroep.

De docenten werden geïnstrueerd om al het materiaal individueel te beoordelen en alleen met de andere leden van hun taalgroep te overleggen in geval van algemene interpretatieproblemen. Zij mochten dus niet overleggen over niveau-inschattingen van specifieke (tekst)fragmenten en opgaven.

Bij de classificatie werd docenten verzocht de (tekst)fragmenten met bijbehorende opgaven en de leerlingprestaties zoveel mogelijk in te delen in de subniveaus (A2.1, A2.2, B1.1, B1.2, etc.), waarbij A2.1 bijvoorbeeld een niveau representeert dat ongeveer halverwege A1 en A2 ligt, terwijl A2.2 overeenkomt met volledige beheersing op niveau A2. Waar zij geen subniveaus konden onderscheiden gaven zij een classificatie in hoofdniveaus (A1, A2, B1, etc.) aan.

Docenten stemden vaak niet overeen in hun classificaties en vertoonden onderling soms grote afwijkingen. Classificaties waarover docenten in meerderheid overeenstemden werden opgenomen in een databestand

en nader geanalyseerd. De classificaties bij lezen en luisteren kwamen over het algemeen op een lager subniveau uit (bijvoorbeeld B1.1 in plaats van B1.2) wanneer de docenten rekening hielden met de bij de tekst(fragmenten) behorende opgaven.

De opgaven voor een bepaald examen bleken te zijn verspreid over een aantal niveaus van het Referentiekader. Maar wat is dan het niveau van de kandidaten? In de praktijk moeten leerlingen bij de examens Lezen en Luisteren zestig tot vijftig procent van de opgaven goed beantwoorden om een voldoende te krijgen. Wanneer we dit gegeven vertalen naar de spreiding over de niveaus van het Referentiekader, dan zou een kandidaat van een bepaalde opleiding minimaal opgaven moeten aankunnen die liggen op onge-

veer tweederde van het hoogste niveau. Voor Duits havo Lezen opgaven op niveau B2.1 en voor Engels vwo Lezen C1.1. Voor spreken kunnen de niveau-eisen direct worden afgeleid uit de leerlingresultaten: aangezien alleen leerlingen met minimaal een 5,5 in het onderzoek waren opgenomen, geven de laagst op de Referentiekaderschaal geclassificeerde leerlingen de cesuurscore aan.

Conclusie

Uit de onderzoeken van SLO en Citogroep blijkt dat het niveau dat volgens de examenprogramma's wordt nagestreefd niet altijd gelijk is aan het feitelijk bij de eindexamens behaalde niveau (zie tabel).

Vaardigheid	Taal	Schooltype	Schatting		
			Examenprogramma	Eindexamen	
Lezen	Engels	vwo	B2.2	<	C1.1
		havo	B2.1	<	B2.2
	Duits	havo	B2.1	=	B2.1
		havo	B2.1	>	B1.2
Luisteren	Engels	vwo	B2.2	<	C1.1
		vwo	B2.2	=	B2.2
	Duits	havo	B2.1	=	B2.1
		mavo-d	—*	—	B1.1
		mavo-c	B1.1	>	A2.2
	Frans	vwo	B2.2	>	B1.2
Spreken	Engels	vwo	B2.1	>	B1.2
		havo	B1.2	=	B1.2
	Duits	vwo	B2.1	>	A2.2
		havo	B1.2	>	A2.2
	Frans	vwo	B1.2	>	A1.1
		havo	B1.1	—	Geen data

*Het nieuwe examenprogramma voor het vmbo bevat in het verrijksdeel geen luistervaardigheid en kan dus niet met het oude d-niveau worden vergeleken.

Uit de vergelijking blijkt dat er verschillen tussen de talen zijn, die niet in de examenprogramma's tot uitdrukking komen. Vanwege de wenselijke transparantie verdient het aanbeveling met de bestaande verschillen tussen de talen rekening te houden en dat onderscheid in het examenprogramma tot uitdrukking te brengen. Het examenniveau voor Engels is het hoogst, Duits ligt ertussenin en dat voor Frans is het laagst. Voor Duits komen behalve bij Spreken de examenprogramma's en de examens redelijk goed overeen in termen van het Europees Referentiekader. Bij Engels komen de examens behalve bij Spreken hoger uit dan de examenprogramma's, bij Frans juist lager. Voor Spreken geldt dat bij de examens (behalve bij havo Engels) een soms beduidend lager niveau wordt gerealiseerd dan in de examenprogramma's wordt verondersteld.

In het onderzoek worden enkele aanbevelingen gedaan, onder andere om bij de herziening van de examenprogramma's formuleringen en niveaus te kiezen die een rechtstreekse relatie hebben met de categorieën van het Europees Referentiekader. Dit zou niet alleen voor vmbo en havo/vwo moeten gebeuren, maar voor alle taalvaardigheidsprogramma's in het hele Nederlandse onderwijs. Daarbij is het wel van belang dat het nu door de Citogroep verrichte onderzoek grootschaliger wordt opgezet dan binnen de kaders van het Excef-project mogelijk was.

* Het eindrapport van het Excef-project met als titel *Nederlandse taalkwalificaties in Europees verband, Nederlandse examenniveaus voor Frans, Duits en Engels en hun plaats op de Europese schaal van taalvaardigheid* kan besteld worden bij de SLO, t.a.v. Angela Ligtenberg, tel. 053-4840285 of per e-mail <a.ligtenberg@slo.nl>

Voor belangstellenden zijn beide achtergrondstudies (Staatsen & Stoks 2001) en Van Hest, Beltman en Kleintjes 2001) bij SLO en Citogroep op te vragen.

LITERATUUR

Council of Europe (2001). *Common European Framework for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Hest, E. van, H. Beltman & F. Kleintjes (2001). *Koppeling MVT-examens aan het Europees Referentiekader*. Arnhem: Citogroep.

Hest, E. van, John H.A.L. de Jong & G. Stoks (2001). *Nederlandse taalkwalificaties in Europees verband, Nederlandse examenniveaus voor Frans, Duits en Engels en hun plaats op de Europese schaal van taalvaardigheid*. SLO: Enschede.

Staatsen, F. & G. Stoks (2001). *Analyse van de examenprogramma's voor vmbo en havo/vwo ten behoeve van de afstemming op de Europese schaal van taalvaardigheid*. SLO: Enschede.