

SMART MONDELINGE TAAVAARDIGHEID BEOORDELEN: HET KAN!

Herkent u dit? Leerlingen die elkaar beoordelen en vooruit helpen bij mondelinge taalvaardigheid? Leerlingen die hun voortgang meten aan de hand van valide, heldere en betrouwbare criteria? De docent die de prestatie meet met dezelfde criteria? Kris Buyse en Emma de Rijk hebben de afgelopen periode gewerkt met een beoordelingsmodel dat door de leerling en de docent wordt gebruikt tijdens het oefenen en bij de toetsing van mondelinge taalvaardigheid. In dit artikel wordt het model uitgelegd en een praktijkervaring besproken.

EMMA DE RIJK & KRIS BUYSE

Goed beoordelen van mondelinge taalvaardigheid vinden veel docenten een hele klus. Iedereen weet dat criteria om te beoordelen helder, valide, betrouwbaar en gemakkelijk hanteerbaar behoren te zijn, want dat is prettig voor de docent en scheelt een hoop discussie met leerlingen. Om te beoordelen zijn er veel modellen in omloop en elke docent gebruikt ze. Kris Buyse, hoofd-docent aan de universiteit van Leuven, heeft de afgelopen paar jaar onderzoek gedaan naar het beoordelen van mondelinge taalvaardigheid bij moderne vreemde talen onderwijs. Hij heeft een model ontwikkeld voor het beoordelen van mondelinge taalvaardigheid, dat zowel formatief als summatief bruikbaar is voor de docent én voor de leerling. Emma de Rijk is docent Spaans en heeft dit model het afgelopen jaar op de basisschool en de middelbare school gebruikt voor mondelinge taalvaardigheidstoetsen. In de praktijk ervaart ze dat het imple-

menteren van dit model niet vanzelf gaat, maar oefening baart kunst.

Buyse heeft een beoordelingsmodel ontwikkeld dat de docent moet aanpassen al naargelang de doelstelling van de mondelinge toets. Omdat dit model per toets op maat wordt gemaakt, zou je kunnen spreken van een 'sjabloon'. Leerlingen krijgen het beoordelingsmodel uitgereikt tegelijk met de opdrachten voor de mondelinge toets en gebruiken het om te oefenen (formatieve functie). De docent gebruikt het voor toetsing (summatieve functie). Formatieve en summatieve beoordelingsmodellen kunnen verschillend zijn (Meestringa & Ravesloot, 2013). Met het formatieve model kunnen docent en leerling nagaan waar een leerling zich op een bepaald moment bevindt ten opzichte van het einddoel; dit wordt vaker omschreven met *can-do statements*. Het eindproduct wordt gemeten aan de hand van een aantal criteria die worden omschreven in een beoordelingsformulier (*rubric*). In dit verband is ook de manier waarop

evaluatie door Hattie en Timperley (2007) wordt uitgewerkt interessant: informatie geven waardoor je het einddoel helder krijgt, noemen ze 'feed up'. Informatie ten aanzien van het punt waar de leerling nu staat 'feed back'. Informatie over de volgende stap in de richting van het doel noemen ze 'feed forward'.

Het model van Buyse heeft een driedelige functie. Het is bedoeld om 'dubbel summatief' mee te scoren: globaal (holistisch, synthetisch) en analytisch. Dat wil zeggen dat enerzijds een globaal cijfer 'op gevoel' kan worden gegeven. Anderzijds kunnen punten worden gegeven voor de zes criteria die het model bevat, wat een analytische beoordeling oplevert. In het ideale geval komen deze twee beoordelingen overeen. Is dit niet het geval, dan kunnen de verschillen worden bekeken en kan een definitief cijfer worden vastgesteld. Daarnaast is dit model, dat in leerlingtaal wordt uitgeschreven, geschikt voor *peerevaluatie* (leerlingen die elkaar tussendoor beoordelen, de formatieve functie).

Het beoordelingsmodel uitgelegd

Het model van Buyse (zie kader 1) beoogt SMART bruikbaar te zijn (specifiek, meetbaar, activerend, realistisch en tijdgebonden). De docent maakt zich een heldere voorstelling van de te behalen doelen en omschrijft de zes criteria specifiek in leerlingtaal, vermeldt daarnaast het ERK-streefniveau en de weging per criterium (meetbaar). Dit levert een specifiek, meetbaar, realistisch gereedschap op om te beoordelen. Door dit model in te zetten tijdens het oefenen, werken leerlingen samen actief toe naar het einddoel en evalueren elkaar tussendoor. Een toets is altijd tijdgebonden; de datum van de toets staat vast en het aantal uren dat in lestijd wordt gebruikt voor het oefenen ook. Het beoordelingsmodel is een werkdocument voor alle deelnemers, wat de transparantie, betrouwbaarheid en validiteit van de beoordeling vergroot.

Bij elke toets mondelinge taalvaardigheid hoort een op maat gesneden beoordelingsformulier met zes criteria die in leerlingtaal weergeven wat zal worden



Foto: Anda van Riet

SJABLOON MONDELINGE EVALUATIE

ERK-niveau: _____ **TOTAAL: ___ / 40**
 Naam: _____
 Beoordelaar: _____

| A. OPDRACHT | totaal A: ___ van 15 | B. TAAL | totaal B: ___ van 25 |
|---|----------------------|--|----------------------|
| 1. Vlotheid & hoeveelheid score 1: ___ van 5 Aandachtspunten: Positieve elementen: | | 4. Uitspraak & intonatie totale score 4: ___ van 5 Uitspraak: Intonatie: Positieve elementen: | |
| 2. Communicatieve output (kwantiteit + kwaliteit) gemiddelde score 2: ___ van 5 Activiteit 1: Activiteit 2: Activiteit 3: Attitude / non-verbale communicatie: Positieve elementen: | | 5. Woordenschat totale score 5: ___ van 10 A. Correctheid woorden (algemeen taalgebruik) B. Woordcombinaties (valentie, collocaties, uitdrukkingen) C. Rijkdom, precisie & variatie D. Communicatieve functies Positieve elementen: | |
| 3. Begrip score 3: ___ van 5 Lexicale problemen (concreet/abstract) Pragmatische, discursieve problemen Grammaticale problemen Positieve elementen: | | 6. Grammatica & pragmatisch-discursieve structuren totale score 6: ___ van 10 A. Morfologie B. Syntaxis C. Pragmatisch-discursieve strategieën (register, structuur, cohesie, structuurwoorden, compensatiestrategieën en interactieve strategieën) D. Variatie Positieve elementen: | |
| COMPENSERENDE ELEMENTEN (portfolio, vooruitgang, attitude; range: +1 tot +3) + | | PENALISERENDE ELEMENTEN (portfolio, vooruitgang, attitude; range: -1 tot -3) - | |

Kader 1. Sjabloon

TOELICHTING BIJ HET SJABLOON MONDELINGE EVALUATIE

Dit beoordelingsmodel bestaat uit twee delen: (A) het uitvoeren van de opdrachten en (B) de kwaliteit van de gesproken tekst. Per deel kan een globaal puntenaantal worden gegeven. Ook kan globaal een ERK-niveau worden aangegeven. Daarnaast is het zinvol een sleutel te formuleren: wanneer scoort de leerling voldoende per criterium.

A. Het uitvoeren van de opdrachten wordt kortweg 'de opdracht' genoemd, of 'de taak' en beschrijft de volgende criteria:

- vlotheid en hoeveelheid
 - de hele opdracht of meerdere deelopdrachten worden beoordeeld
 - vlotheid kan worden verfijnd door de duur van pauzes tussen woorden/zinnen te benoemen
 - hoeveelheid kan in een aantal minuten worden aangegeven
- communicatieve output
 - de hele opdracht of meerdere opdrachten worden beoordeeld op het duidelijk overbrengen van de boodschap
 - verfijning is mogelijk door gebruik van vooraf afgesproken strategieën te specificeren
- begrip
 - vragen van de docent of vragen in een tweegesprek worden al dan niet begrepen door de leerling; dit begrip wordt beoordeeld (dit onderdeel kan worden opgevat als 'luistervaardigheid')

B. De gesproken tekst wordt 'taal' genoemd en beschrijft de volgende criteria:

- uitspraak en intonatie
 - uitgesproken woorden en zinnen worden beoordeeld; hierbij kunnen bepaalde klanken of intonatiecriteria worden gespecificeerd
- woordenschat
 - woordenschat wordt beoordeeld per opdracht of in zijn geheel, waarbij specifieke woordgroepen kunnen worden benoemd in subcategorieën, zoals woordcombinaties (collocaties, valenties, uitdrukkingen), rijkdom/precisie/variatie, communicatieve functies; deze toevoegingen kunnen het geheel verzwaren, maar tegelijk uitnodigen om voldoende aandacht te schenken aan deze subcriteria
- grammatica en zinsbouw
 - hier kan met specifieke te beoordelen elementen worden aangevuld, zoals correcte werkwoordsvormen, bepaalde verbindingen van zinnen en wat voor de prestatie verder essentieel wordt gezien op het beoogde niveau.

Bij elk criterium is ruimte voor het invullen van de weging en van aandachtspunten en positieve elementen.

beoordeeld. In het sjabloon van Buyse zijn aanwijzingen gegeven die naargelang de opdrachten kunnen worden vereenvoudigd of nog verder verfijnd in subcriteria. Bij het onderdeel 'Grammatica' kan bijvoorbeeld specifiek worden getoetst op het gebruik van bijvoeglijk naamwoorden of bepaalde verleden tijden. De basis van zes criteria wordt eventueel op de basisschool teruggebracht naar vier, maar vanaf het niveau van de middelbare school kunnen de zes criteria worden gebruikt.

Belangrijk onderdeel van het sjabloon zijn de 'positieve elementen' die bij elk criterium voorkomen. Tijdens het oefenen kunnen *peers* zich hierop richten onder het motto: 'Doe meer van wat goed gaat'. Ook voor de verbeterpunten is ruimte. Het sjabloon is aanpasbaar aan alle soorten taallessen voor algemene of specifieke doeleinden en aan alle niveaus. Het relatieve gewicht van het deel 'Opdracht' en het deel 'Taal' en het gewicht per criterium hangt af van de doelstellingen van de cursus en het niveau; de weging van 'Taal' zou hoger kunnen liggen voor een taalopleiding zoals vertaalwetenschappen of taalkunde dan bij specifiek economische of toeristische doeleinden, waar begrip en het overbrengen van de boodschap (communicatieve output) mogelijk zwaarder wegen. (Voor verdere uitleg en toelichting, zie Buyse, 2012, 2013.)

In een workshop op de Landelijke Studiedag van Levende Talen in november 2013 kwamen docenten bij elkaar om het model te bekijken, te vergelijken met eigen modellen die waren meegenomen en om een uitgewerkt model tijdens de workshop te gebruiken. Docenten vroegen zich in eerste instantie verschillende zaken af: of het niet te veel criteria waren om een schoolexamen mee te beoordelen in de korte tijd die beschikbaar is. Ook vroegen ze zich af of een formatief instrument ook summatief bruikbaar kon zijn.

Het schoolexamen van gymnasium 6

Een beoordelingsmodel inzetten tijdens het oefenen is niet iets wat zonder meer werkt. Leerlingen en docent moeten het in de vingers krijgen door te oefenen. Voor leerlingen van 6 gymnasium werd het model in de praktijk gebruikt bij de voorbereiding en toetsing van het mondeling examen Spaans, als onderdeel van het vak Spaans Elementair (ERK-niveau A2). Het mondeling examen bestond uit twee delen: een presentatie (monoloog) en twee gesprekken (dialoog).

Een van de opdrachten was een gesprek tussen twee leerlingen die foto's beschreven en vragen stelden over elkaars foto's. In groepjes verzamelden leerlingen foto's rond diverse thema's (sport, familie, vakantie, baantjes)

| | | TOTAAL: ___ / 25 | |
|--|----------------------|---|----------------------|
| A. OPDRACHT (ERK-streefniveau A2/B1) | totaal A: ___ van 10 | B. TAAL | totaal B: ___ van 15 |
| 1. Vloeiendheid en hoeveelheid | punten: ___ van 3 | 4. Uitspraak en intonatie | punten: ___ van 3 |
| Compliment: _____ | | Compliment: _____ | |
| Aandachtspunt: _____ | | Aandachtspunt: _____ | |
| heldere boodschap, hier en daar stukt het maar stoort niet | 3 | helder, accent stoort niet | 3 |
| boodschap duidelijk, hier en daar herformuleringen | 2 | functionele pauzes en afwijkingen | 2 |
| veel herformuleringen en niet afgemaakte zinnen | 1 | lange pauzes, accent belemmert begrip | 1 |
| niet te volgen | 0 | niet te volgen | 0 |
| 2. Communicatieve output (kwantiteit en kwaliteit) | punten: ___ van 4 | 5. Woordenschat | punten: ___ van 6 |
| Compliment: _____ | | Compliment: _____ | |
| Aandachtspunt: _____ | | Aandachtspunt: <i>pienso, opino, me gusta</i> | |
| taak 1: tegenwoordige tijd monoloog / dialoog / voorbereid / spontaan taak 2: verleden tijd monoloog / dialoog / voorbereid / spontaan taak 3: hypothese monoloog / dialoog / voorbereid | | | |
| output compleet, samenwerking succesvol | 4 | woordgebruik passen bij het onderwerp | ___ van 1 |
| relevant en coherent met voldoende verbindingen | 3 | combinaties van woorden passen bij het onderwerp | ___ van 2 |
| output stukt regelmatig | 2 | bereik (rijkdom, precisie, variatie) is passend | ___ van 1 |
| gesprek en presentatie sluiten niet aan bij de taak | 1 | communicatieve functionaliteit sluit aan bij interactie | ___ van 2 |
| 3. Begrip | punten: ___ van 4 | 6. Grammatica en structuren | punten: ___ van 6 |
| Compliment: _____ | | Compliment: _____ | |
| Aandachtspunt: _____ | | Aandachtspunt: _____ | |
| beschrijvingen, hoofdpunten en mening zijn compleet begrepen | 4 | vormen kloppen | ___ van 2 |
| vragen en uiteenzettingen zijn begrepen | 3 | samenhang tussen woorden klopt | ___ van 1 |
| begrip van de taken blijkt niet elke keer | 2 | gebruikt verbindingswoorden: es que, por eso, además | ___ van 1 |
| aansluiting op vragen en opdrachten blijkt niet | 1 | gebruikt compensatiestrategie waar nodig | ___ van 2 |
| BONUS: gevuld portfolio, zichtbare voortgang, goede inzet, zichtbaar doorzettingsvermogen | | AFTREKPUNTEN: ontbrekende portfolio-elementen, gebrek aan voortgang, storende houding, afhaken | |

Kader 2. Voorbeeldmodel gebruikt in gymnasium 6 bij Spaans

en oefenden samen met het beschrijven van de foto's. Volgens de expertmethode werden vervolgens de leerlingen opnieuw verdeeld, nu in groepjes van 4 (van elk thema één leerling). Nu vertelde elke leerling over zijn/haar foto's en ze bevroegen elkaar erover, precies zoals tijdens het mondeling ook zou gebeuren. In deze fase werden taken verdeeld: twee observanten beoordeelden het gesprek tussen de beide andere leerlingen.

Bij deze beoordeling gebruikten de leerlingen het beoordelingsmodel (zie kader 2). Ze beschreven positieve punten en aandachtspunten, en scoorden de criteria. Ook scoorden ze specifieke elementen, de subcriteria: het gebruik van *me gusta* ('ik vind het leuk'), *pienso* ('ik denk') en *opino* ('ik meen'). Tijdens het oefenen wisselden de leerlingen regelmatig van rol. Ook werden oefensituaties voor de klas uitgespeeld en klassikaal besproken volgens het model, zodat iedereen het evalueren met de criteria in de vingers kreeg.

Het model hielp te begrenzen tijdens de voorbereidingen, al leek dat in het begin niet zo. Begrenzen welke talige elementen worden beoordeeld was verhelderend voor leerlingen en voor de docent. Het aanbrenge van subcriteria bleek glad ijs, want over de te scoren specifieke grammaticale vormen ontstond discussie; is *se gusta* helemaal fout? Geef je punten als *pienso* viermaal wordt gebruikt en *opino* nooit? In eerste instantie vonden de leerlingen dit onduidelijk. Tijdens het praten over deze punten werd duidelijk wat de norm zou zijn en maakte de irritatie plaats voor interesse. Uiteindelijk bleek het model, ondanks de tijdelijke wrijving, wel te motiveren en een goede bron van informatie over het te behalen einddoel. Het proces van voorbereiding en uitvoering namen de leerlingen zeer serieus. Ze voelden zich uiteindelijk meester van de situatie en gebruikten het beoordelingsmodel om elkaar vooruit te helpen, en dit betekende tijdswinst voor de docent.

Tijdens het schoolexamen was de situatie als volgt: een kwartier voor het mondeling werden twee leerlingen uit de verschillende groepjes een kwartier bij elkaar gezet om vragen over elkaars foto's voor te bereiden. Tijdens het mondeling vervulde elke leerling twee rollen: die van gesprekspartner en die van observant. Aangezien de docent dubbel summatief wilde beoordelen, maar geen tweede docent voorhanden was, werden de mondelingen bovendien opgenomen. De observant noteerde zijn/haar beoordeling van een leerling in het model. Dit telde niet mee, toch namen de leerlingen dit onderdeel serieus. Tijdens de toetsing noteerde de docent alleen het globale cijfer voor de opdracht en voor de taal. Daarna beluisterde zij de opnames voor

de analytische beoordeling, bekeek de scores van de observanten en bepaalde het cijfer. Naderhand, bij het uitdelen van het cijfer, vroegen de leerlingen naar het waarom van het cijfer en dit was een inhoudelijk gesprek over de communicatieve kracht van leerlingen, over werkwoorden en over variatie in taal. Een inspirerend moment.

Conclusies

Mondelinge taalvaardigheid doelgericht oefenen en toetsen binnen de beperkte mogelijkheden kan. Aan de kant van de leerling valt winst te behalen. Leerlingen die een spreek- of gespreksopdracht krijgen en direct daarbij het beoordelingsformulier ontvangen, samen de opdracht voorbereiden en elkaar aan de hand van het model feedback geven, zijn actief betrokken bij hun leerproces. Ze zijn inhoudelijk betrokken bij de uitvoering van de opdracht en accepteren coaching van de docent. Leerlingen krijgen alle criteria onder de knie, enerzijds door te oefenen met spreken en anderzijds door feedback te geven – het resultaat van een transparant, betrouwbaar en valide beoordelingsmodel in gebruik. Doordat leerlingen elkaar tussendoor steeds beoordelen met het model, wordt de leerling meer eigenaar van zijn eigen leerproces, wat motiverend werkt (Dochy & Gijbels, 2010).

Ook aan de kant van de docent kan de inzet van een dergelijk beoordelingsmodel winst opleveren. Tijd voor individuele leerlingen tijdens het oefenen, de wetenschap dat het model werkt en betrouwbaar is en een groep gemotiveerde leerlingen aan het werk hebben, dat is prettig. De roep om onderwijs van deze tijd, waarbij valide, betrouwbare, heldere en gemakkelijk hanteerbare gereedschappen worden ontwikkeld om het leerproces en de leerprestatie van de leerling inzichtelijk te maken, die roep is onmiskenbaar. Hopelijk levert dit model meer docenten winst op. ■

LITERATUUR

- Buyse, K. (2012). Criterios y plantillas para la evaluación del español para fines específicos. In A. van Hoof (Red.), *El español de las profesiones: Artículos seleccionados del IV Congreso Internacional de Español para Fines Específicos* (pp. 186–200). Madrid: Ministerio de Educación, Cultura y Deporte.
- Buyse, K. (2013). 'Come and sit here next to me': Towards a communicative assessment of oral language skills. *Language Learning in Higher Education*, 3(1), 89–107.
- Dochy, F., & Gijbels, D. (2010). Evaluatie. In S. Janssens (Red.), *Leren en onderwijzen* (pp. 121–175). Leuven: Acco.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Meestringa, T., & Ravesloot, C. (2013). Schrijven in de tweede fase. *LTM*, 100(6), 6–11.