Beroepsvereniging van docenten
Nederlands als Tweede Taal

VAKWERK 5

Achtergronden van de NT2-lespraktijk

LEZINGEN CONFERENTIE BVNT2 IN BLANKENBERGE 2009

Vakwerk 5

Achtergronden van de NT2-lespraktijk

Simon Verhallen (red.)

Vakwerk 5

Achtergronden van de NT2-lespraktijk

Lezingen Conferentie BVNT2 in Blankenberge
juni 2008

BV NT2 – Amsterdam – 2009

Dit is een uitgave van de Beroepsvereniging van docenten NT2. Inlichtingen over de vereniging zijn verkrijgbaar bij de secretaris van de BV NT2 of op de website (bvnt2.org).

Postadres:

BVNT2
Carola van der Voort
Frans Halsplantsoen 32
1399 EW Muiderberg

Mailadres:

Bestuur@bvnt2.org

Omslagontwerp: Marjolein Overdevest
Tekstopmaak: Simon Verhallen
Druk: VU Reprografie Amsterdam

Copyright: Artikelen of fragmenten van artikelen mogen worden overgenomen, mits bedoeld voor educatieve doeleinden en mits de bron duidelijk wordt vermeld. In alle andere gevallen is schriftelijke toestemming van het bestuur van de Beroepsvereniging van docenten NT2 noodzakelijk.

Voorwoord

"Voor de nieuwsgierige, voor de kritische en de onderzoekende NT2-docent." Dat zou een passende titel voor deze bundel artikelen zijn, maar we houden ons graag aan goede gewoonten. En dus heet deze bundel, zoals Bart Bossers de eerste noemde: Vakwerk.

Traditiegetrouw zijn de inleiders van de grote BNVT2-conferentie, ditmaal gehouden in Blankenberge in Vlaanderen, gevraagd hun bijdragen aan de lezingenreeks aldaar ook in een tekst om te zetten ten behoeve van de bundel Vakwerk. Enkele inleiders hadden de inhoud van hun bijdrage al in een artikel elders gepubliceerd of toegezegd. Die vindt u dus niet hier. Daarnaast is nu een artikel opgenomen dat bij de afronding van Vakwerk 4 niet kon worden meegenomen, maar dat we graag hier plaatsen.

Waarom is deze bundel speciaal voor de nieuwsgierige, etc., NT2-docent?
Er wordt de gelegenheid geboden om ook eens kritisch naar taaltoetsing te kijken; is de steeds terugkerende roep om meer taaltoetsen nu wel echt in het voordeel van de docenten en de leerlingen/cursisten? Er worden ook hulpmiddelen geboden om eens heel anders naar de schrijfproducten van de NT2-leerders te kijken en naar de uitspraakproblemen, en daarbij worden nieuwe didactische modellen gepresenteerd: voor wie het anders wil aanpakken. Ook op het gebied van tekstbegrip en leesvaardigheid worden ons enkele prikkelende beschouwingen voorgehouden en vervolgens is er een werkmodel voor de waarlijke echt onderzoekende geesten onder de NT2-docenten. Kortom, een bundel met teksten die, ingevolge de doelstellingen van de BVNT2, de NT2-docent deskundiger en vakkundiger kunnen maken.

Sible Andringa, van de UvA en lid Staatsexamencomissie NT2, geeft een algemene en kritische beschouwing over taaltoetsing en NT2-toetsen: ethische kwesties moeten geregeld aan de orde gesteld worden. Men mag niet klakkeloos achter elke toetsbeweging aanrennen.
De bijdrage van Folkert Kuiken aan deze bundel is het directe gevolg van de masterclass die hij op de conferentie in Blankenberge heeft gehouden. De deelnemers aan de masterclass waren zeer enthousiast en ook erg verbaasd over het feit dat het opzetten van eigen actieonderzoek veel gemakkelijker was gebleken dan ze hadden verondersteld. De deelnemers die we achteraf spraken zeiden dat ze hoopten dat deze aanpak in het NT2-onderwijs meer navolging zou kunnen krijgen. Om die reden hebben we Folkert Kuiken gevraagd de

inhoud van de masterclass in een artikel om te zetten. Mogelijk dat het ook een richtsnoer en een hulpmiddel kan zijn bij activiteiten van opleiders en aanbieders van inburgeringscursussen wanneer de verschillende Leerlastonderzoeken vanwege de minister van WWI (Wonen, Wijken en Integratie) zullen worden opgezet in Nederland.
Op de conferentie in Blankenberge hield hoogleraar Huub van den Berg van de Universitiet van Utrecht een lezing, die door taaldocenten en NT2-docenten die in het vmbo en mbo werken, heel interessant werd gevonden. Het onderzoek van Van den Berg en van zijn collega's – Jentine Land promoveerde onlangs op het onderzoek dat hier beschreven is - laat zien dat zwakke lezers wel eens veel meer achterop raken door goed bedoelde versimpelingen in teksten die zij op school en in de beroepsopleiding te lezen krijgen. Vooral docenten die in hun teams betrokken worden bij het schrijven van leermateriaal – en dat komt nogal eens voor – kunnen hun voordeel doen met de onderzoeksuitkomsten die door het trio Land, Sanders en Van den Berg worden gepresenteerd.
Jan Strybol uit Gent breekt een lans voor een andere aanpak bij het feedback geven op schrijfproducten van cursisten. Als de docent iets minder de rode pen hanteert en zichzelf af en toe laat vervangen door anderen, dan ontstaan er nieuwe mogelijkheden. Jan presenteert een systematische aanpak; zijn voorbeelden zijn zeer aansprekend.
In deze bundel is als vijfde bijdrage een artikel opgenomen van Catia Cucchiarini en Ambra Neri uit Nijmegen over het verbeteren van uitspraakfouten. Op grond van kwalitatieve en kwantitative analyses van een drietal databases komen zij tot een inventaris van veel voorkomende uitspraakfouten en ze stellen voor dat de NT2-docent zich vooral daar op concentreert.

Simon Verhallen, maart 2009

INHOUD

Sible Andringa
Over de zuiverheid van toetsen. 9

Catia Cucchiarini & Ambra Neri
Correctieve feedback op uitspraak: welke fouten maken NT2-leerders? 24

Folkert Kuiken
Actieonderzoek in het NT2-onderwijs: de docent in actie. 35

Jentine Land, Huub van den Bergh & Ted Sanders
Zwakke teksten voor 'zwakke' lezers. 53

Jan Strijbol
Een rollenmodel voor feedback 61

Over de Beroepsvereniging van docenten NT2 75

OVER DE ZUIVERHEID VAN TOETSEN

SIBLE ANDRINGA

Universiteit van Amsterdam, ACLC

Inleiding

De roep om taalvaardigheidstoetsen wordt steeds groter. We willen graag precies weten wat een leerling al wel en nog niet kan; we willen zeker weten dat nieuwkomers het Nederlands voldoende beheersen; we willen weten of pabo-studenten rekenen en taal voldoende beheersen, en we willen zelfs zeker weten dat onze kleuters goed beslagen ten ijs komen als ze beginnen aan het basisonderwijs. Als je deze dingen wilt weten, dan moet je toetsen, want meten is weten. Deze roep om toetsen wordt vaak gevoed door zorgen over de kwaliteit van het onderwijs. Er is de laatste jaren veel te doen geweest over het onderwijs, dat volgens allerlei onderzoeken achteruit holt, veroorzaakt door de vloed aan vernieuwingen die zijn doorgevoerd.

Toetsen wordt in deze discussie vaak geopperd als een middel om dit kwaad te stoppen. Recente voorbeelden zijn het rapport van de Commissie Meijerink over het reken- en taalonderwijs in Nederland ("Over de drempels met taal en rekenen," 2008) en het rapport van de Commissie Dijsselbloem over de onderwijsvernieuwingen van de jaren negentig ("Parlementair onderzoek onderwijsvernieuwingen," 2008). Beide rapporten bevelen het instellen van toetsen aan om de kwaliteit van het onderwijs te monitoren en te verbeteren. Toetsen zijn ook ontdekt als politiek instrument om controle uit te oefenen op de inburgering van nieuwe Nederlanders; taalvaardigheidstoetsen spelen daarin een prominente rol.

Deze tekst wil duidelijk maken dat het gebruik van toetsen niet zonder gevolgen is. Ik beschrijf welke impact toetsen hebben op het onderwijs, maar ook wat toetsen betekenen voor de leerling, omdat dit perspectief nogal eens buiten beschouwing wordt gelaten. Ook komen de technische beperkingen van taalvaardigheidstoetsen aan bod. Centraal staat de vraag of de toetsen waar de commissies Meijerink en Dijsselbloem om vragen wel zullen leiden tot de onderwijsverbeteringen die worden beoogd.

Waarom toetsen?

Een toets probeert zo objectief mogelijk vast te stellen wat iemand weet of kan; toetsconstructeurs doen erg hun best om toetsen te maken die dat kunnen. Voor de toetsende instantie – dat kan de overheid zijn of de school of een docent – is het bijna vanzelfsprekend dat een toets objectief uitdrukt wat iemand kan of weet. Het gaat hen niet in eerste instantie om de score zelf, maar om de effecten die toetsresultaten met zich meebrengen. Globaal gezien kun je twee motieven onderscheiden voor het gebruik van toetsen.

Formatieve toetsen moeten goed studiegedrag bewerkstelligen en er dus voor zorgen dat cursisten of leerlingen datgene kunnen of weten wat van ze gevraagd wordt; ze kunnen ook gebruikt om het programma aan te passen. Formatieve toetsen zijn ook bedoeld om beslissingen te nemen over het lesprogramma.

Summatieve toetsen worden gebruikt om aan het eind van een cursus of programma goede en door iedereen geaccepteerde beslissingen over de toetsnemers te nemen. Deze toetsen worden dus vaak gebruikt voor selectiedoeleinden. Heel vaak verenigen toetsen beide functies in zich.

De meeste toetsen zijn formatieve toetsen. De repetities en overhoringen die veelvuldig worden gebruikt op het voortgezet onderwijs zijn voorbeelden van formatieve toetsen. Een toets bepaalt een standaard of norm waaraan een toetsnemer moet voldoen. Dat dwingt cursisten tot het investeren van tijd en energie in dat wat getoetst wordt en moedigt ze aan tot het vergroten van hun kennis of het verbeteren van hun vaardigheden. De extrinsieke motivatie die toetsen bewerkstelligen is voor veel docenten de belangrijkste reden voor het inzetten van toetsen. Formatieve toetsen dienen ook vaak een summatief doel. Het cijfer op de toets is vaak één van de vele cijfers op basis waarvan wordt besloten of een cursist de cursus met voldoende gevolg heeft afgelegd. Sommige, wellicht wat cynisch aangelegde, docenten vinden zelfs dat een toets het enige middel is dat nog tot gewenst studie- en leergedrag kan motiveren.

Summatieve toetsen zijn toetsen waarbij over het algemeen veel op het spel staat. Het niet halen van een summatieve toets heeft over het algemeen vervelende gevolgen. Deuren die toegang geven tot een nieuw leerjaar of tot een bepaalde opleiding gaan dicht. Het kan ook zo zijn dat een onvoldoende resultaat leidt tot het opnieuw moeten doen van de cursus of tot plaatsing in bijspijkerklassen. De toetsnemer zal een summatieve toets dus ervaren als een onvermijdelijke horde die genomen moet worden, en dat leidt vaker wel dan niet tot gevoelens van stress en angst bij de toetsnemers. Dit wordt nog versterkt door het feit dat je vaak maar één kans hebt

De aantrekkelijkheid van toetsen

Toetsen kunnen hun formatieve en summatieve functies zo effectief vervullen doordat toetsing eigenlijk altijd plaats vindt in een hiërarchische context. Een toets heeft meestal consequenties voor de toetsnemer, omdat de toetsende instantie de macht heeft om beslissingen te nemen op basis van het toetsresultaat. In het geval van de docent en de leerling behoeft dat geen uitleg. Maar ook overheden maken graag gebruik van toetsen om invloed uit te oefenen op scholen, bijvoorbeeld om de kwaliteit van het onderwijs te waarborgen of te verbeteren. Een mooie illustratie van dit gebruik van toetsen vormt het rapport van de commissie Meijerink ("Over de drempels met taal en rekenen," 2008). De aanleiding van dit rapport is bezorgdheid over de kwaliteit van het onderwijs, vooral op het gebied van rekenen en taal. Zo stelt het rapport: "Jongelui die het mbo en het hoger onderwijs binnenstromen, blijken toch wel erg veel moeite te hebben met lezen, spellen en rekenen." (p. 5) Het rapport adviseert continue aandacht voor taal en rekenen in de vorm van leerlijnen die beginnen in het basisonderwijs en doorlopen in het voortgezet onderwijs. De commissie heeft daartoe in niveaus, van jaar tot jaar, beschreven wat een leerling minstens moet kunnen. De niveaus lijken op die van het Europees referentiekader. Om ervoor te zorgen dat er daadwerkelijk voldoende aandacht zal worden besteed aan rekenen en taal adviseert de commissie het gebruik van, het liefst verplicht gestelde, toetsen: "toetsing garandeert aandacht voor rekenen en taal". De opgeleverde niveaubeschrijvingen vormen de toetsdoelen per onderwijsjaar en onderwijsniveau. Het advies is in grote lijnen overgenomen door de overheid in de hoop dat het zal leiden tot een verbetering van de taal- en rekenvaardigheid.

Een prettige bijkomstigheid van toetsen is dat ze efficiënt en goedkoop zijn, goedkoper dan de meeste andere opties om de onderwijskwaliteit te bevorderen, zoals bijvoorbeeld investeren in docentcompetenties. Verder geeft een toets controle over de cesuur. Als je vindt dat een te groot deel slaagt of zakt voor de toets, dan kan eenvoudig de norm worden aangepast. Dat dit niet slechts een theoretische mogelijkheid is, bewijst de discussie die momenteel plaats vindt rondom de zak/slaag grens van het Basisexamen Inburgering, de Toets Gesproken Nederlands die via de computer wordt afgenomen bij sommige buitenlanders die zich in Nederland willen vestigen. In de politiek is veel steun voor het verleggen van de grens van A1min naar A1, een mogelijkheid die momenteel wordt onderzocht.

In de samenleving genieten toetsen enorm veel vertrouwen. Ze geven het gevoel dat hetgeen dat getoetst wordt, serieus wordt genomen door de overheid of het onderwijsveld. Verder worden toetsen gezien als eerlijk en

democratisch: iedereen heeft gelijke kansen zichzelf te bewijzen. Toetsen hebben een aura van wetenschappelijkheid en objectiviteit en ze worden vervaardigd door instanties met aanzien (Shohamy, 2001). Dat is waarschijnlijk waarom toetsuitslagen zelden worden betwist en alom geaccepteerd. Het is als leek niet makkelijk om te begrijpen hoe een toetsscore tot stand komt; de verantwoording is gegoten in wetenschappelijk jargon dat moeilijk is te begrijpen zonder specialistische kennis. Dat maakt het voor een leek bijna onmogelijk een toetsresultaat ter discussie te stellen. Toch zijn er soms best goede redenen om vraagtekens te zetten bij de scores van taalvaardigheidstoetsen, een punt dat verderop uitgebreid besproken zal worden. Duidelijk is dat er via toetsing heel wat geduw en getrek plaats vindt tussen de partijen in het onderwijsveld. Of dat bezwaarlijk is hangt niet alleen af van de kwaliteit van de toets, maar ook van de impact die toetsen hebben op die partijen. Daarover gaat de volgende paragraaf.

De impact van toetsen

Uit het bovenstaande blijkt al dat toetsen een bepaalde impact hebben op hun omgeving. Het effect dat een toets heeft op zijn omgeving wordt *terugslag*, in het Engels washback, genoemd. Messick (1996) definieert washback als de mate waarin het invoeren en gebruik van een toets docenten en taalleerders aanmoedigt tot activiteiten die ze anders niet zouden doen en die taalverwerving stimuleren of hinderen (p. 241). Washback kan dus positief of negatief zijn. Hughes (1993; in Wall, 2000) onderscheidt drie vormen van washback. Een toets heeft allereerst een bepaalde impact op de deelnemers, dat zijn de toetsnemers, maar ook op docenten, materiaalontwikkelaars, beleidsmakers, etc. Via de deelnemers heeft een toets ook gevolgen voor de onderwijsprocessen. Er wordt bijvoorbeeld tijd gemaakt om leerlingen voor te bereiden op een toets, en toetsresultaten zorgen soms voor bijstelling van het onderwijsprogramma. Tenslotte heeft een toets ook effect op het product, dat wil zeggen, op datgene wat geleerd wordt en hoe dat geleerd wordt. In deze paragraaf ligt de nadruk op de impact van toetsen op het onderwijs en de leerling. De maatschappelijke en politieke impact van toetsing wordt verder buiten beschouwing gelaten.

Impact op de docent en de onderwijsprocessen

Onderzoek naar washback heeft aangetoond dat toetsen negatieve effecten op het onderwijs kunnen hebben. Toetsen kunnen bijvoorbeeld gevolgen hebben voor de activiteiten die plaatsvinden in de klas. Het invoeren van een toets kan ertoe leiden dat docenten meer aandacht gaan besteden aan zaken die leerlingen helpen het goed te doen op de toets. Het is absoluut niet uitgesloten dat ze daarbij waardevolle activiteiten als het geven van feedback en het creëren van interactie links

laten liggen ten koste van activiteiten die toegespitst zijn op de toets. Shohamy (2001) onderzocht welke activiteiten er plaatsvonden in de klas vlak voor en vlak na een nationale toets Arabisch in Israël, een toets die geen gevolgen had voor de toetsnemers (iets wat overigens niet werd geloofd). Voorafgaand aan de toets werd er veel herhaald, kreeg men veel huiswerk op, lag de nadruk op gedecontextualiseerd leren en moest er veel uit het hoofd geleerd worden, werd de moedertaal als taal van instructie gebruikt en werd er veel overhoord. Na de toets werd er nieuwe stof behandeld, veelal in de doeltaal, werd er minder huiswerk opgegeven, werden er realistische en communicatieve contexten gecreëerd om het Arabisch te oefenen, en werd er weinig overhoord. Meestal werkt het invoeren van een toets beperkend. Er is dan veel aandacht en tijd voor de dingen die getoetst worden en te weinig ruimte voor dingen die niet voorkomen in de toets.

De effecten van toetsen zijn niet noodzakelijkerwijs negatief. Negatieve washback wordt vaak gezien als het gevolg van een imperfecte toets, zodat je door de toets te verbeteren de negatieve gevolgen ervan kunt verminderen (Bailey, 1996). Daar zit zeker een kern van waarheid in. Een taalvaardigheidstoets die authentieke en natuurlijke vormen van taalgebruik toetst in een realistische context leidt tot ander, en waarschijnlijk meer gewenste vormen van voorbereiding en leergedrag dan een toets die taalvaardigheid op een meer gedecontextualiseerde manier meet (Messick, 1996). Een valide toets leidt bij instellingen, docenten en leerlingen tot gedrag en programma-aanpassingen dat het leren ten goede komt. Docenten zullen de getoetste stof grondiger behandelen en ervoor zorgen dat alle stof behandeld wordt. Met andere woorden, een meer valide toets leidt tot positieve washback, is de gedachte. Wat het uiteindelijke effect is van toetsen op de onderwijsprocessen is moeilijk te zeggen. Dat het onderwijs zich voegt naar de toets is duidelijk, maar in welke mate dat gebeurt is van veel factoren afhankelijk. Van belang zijn bijvoorbeeld de beslissingen die genomen worden op basis van de toets. Een toets die vaststelt of je een bepaald hoofdstuk opnieuw moet bestuderen heeft minder impact dan een toets die bepaalt of je door mag naar een nieuw studiejaar. Een andere factor is de docent. Toetsresultaten maken vergelijkingen tussen klassen mogelijk. Sommige docenten voelen daar meer druk van uitgaan dan andere docenten en zullen meer aandacht besteden aan zaken die leerlingen helpen het goed te doen op de toets.

Impact op de leerling

Het is moeilijk te onderzoeken welk effect een toets heeft op de leerling. Onderzoek hiernaar is gedaan op basis van vragenlijsten en interviews, en daarom enigszins anekdotisch van aard. Duidelijk is in ieder geval dat toetsen aansporen tot leren. Toetsen, en dan vooral de

gevolgen die het niet halen van de toets heeft, vormen de extrinsieke motivatie die de meeste leerlingen nodig hebben om de tijd en energie te investeren die nodig is om vooruit te gaan. Ook hier weer is het belang van een valide toets duidelijk: als dat wat getoetst wordt goed overeenstemt met hetgeen een leerling geacht wordt te kunnen, dan zal er een positief effect uitgaan van de toets. Maar het is erg lastig vast te stellen of een toets daadwerkelijk die positieve impact uitoefent. Vaak wordt wel aangetoond dat leerlingen vooruitgaan, maar het middel waarmee dat wordt aangetoond is de toets zelf. Dat laat ruimte voor de mogelijkheid dat de leerlingen vooral vaardig zijn in het maken van de toets, en niet zozeer in de vaardigheid die de toets beoogt te meten.

Toetsen hebben dus invloed op wat en hoe een leerling leert, en kunnen het leren versnellen. Positief aan toetsing voor de leerling is verder het de mogelijkheid biedt vooroordelen weg te nemen die er zouden kunnen leven bij docent en klasgenoten. Via de toets kan de leerling laten zien wat hij of zij werkelijk kan. Een toets is in principe objectief (het scoren ervan overigens niet noodzakelijkerwijs). Een leerling kan zichzelf dus bewijzen. Verder kunnen toetsen leerlingen helpen door specifieke problemen aan het licht te brengen. Het kan een opluchting zijn te ontdekken dat iemand dyslexie heeft, waardoor tegenvallende schoolresultaten verklaarbaar worden.

Zo goed als afwezig in de washback-literatuur is het effect van toetsing op de psyche van de leerling. Dat is empirisch niet vast te stellen zonder ethisch een bedenkelijke weg in te slaan. Je kunt een groep niet in tweeën delen en willekeurig onvoldoendes uit gaan delen. Ook hier ben je aangewezen op het bevragen van de ervaringen die mensen met toetsen hebben. Shohamy (2001) laat testnemers aan het woord. Zij rapporteren bijvoorbeeld dat een toets aan de wieg stond van of juist in de weg stond voor een bepaalde carrière; dat ze zich door toetsen gestigmatiseerd voelden als mislukking; dat ze heel veel stress ervoeren; dat toetsen oneerlijk zijn, of nutteloos. In hoeverre dit allemaal klopt is onduidelijk; je kunt nooit vaststellen hoe iemands carrière zou zijn verlopen als de toets wel of juist niet was gehaald. Niettemin is het waarschijnlijk dat deze dingen gebeuren. Iedereen krijgt te maken met toetsing, en iedereen heeft de gevolgen van toetsen wel eens ondervonden.

Het probleem van toetsen is dat er voor elke leerling die bovengemiddeld scoort er ook iemand is die benedengemiddeld scoort. Wat gebeurt er met iemand die keer op keer moet constateren dat hij of zij minder 'goed' is dan de rest? Bijna iedereen heeft wel eens geprobeerd iets te leren waarin hij of zij niet goed bleek. Mensen die dat bij zichzelf constateren geven meestal op. Waarom zou je nog energie steken in iets waarin je niet goed bent gebleken of waar je geen talent voor hebt? Onvoldoendes zouden hetzelfde effect kunnen hebben op iemands

leergedrag. Een leerling die op school zit is meestal niet in een positie om te stoppen met school, maar kan wel stoppen met leren. Bovendien brengt het halen van een onvoldoende zonder te leren misschien wel minder gezichtsverlies met zich mee dan hard je best doen en toch een onvoldoende halen. Leerlingen die wel door (moeten) zetten, ervaren toetsen waarschijnlijk als uitermate stressvol, en onvoldoendes als uitermate pijnlijk. Dat is met name het geval bij toetsen waar veel belangen op het spel staan. Zakken voor het examen is niet alleen pijnlijk, maar kan ook vervelende gevolgen hebben als een verblijfsvergunning of een baan op het spel staat. De hoeveelheid stress die dan om de hoek komt kijken is bijzonder groot, wat de prestaties weer negatief kan beïnvloeden.

Toetsen spelen waarschijnlijk ook een rol bij de perceptie die de docent van de leerling heeft. Al genoemd is dat toetsen vooroordelen die docenten zouden kunnen hebben op basis van sociaal economische achtergrond kunnen wegnemen. Maar toetsen kunnen ook een negatieve rol spelen. Leerlingen kunnen vereenzelvigd worden met hun toetsresultaten. Als iemand een goed toetsresultaat heeft behaald, dan is dat meestal omdat die persoon ook daadwerkelijk vaardig is of veel weet. Een lage score kan duiden op achterblijvende kennis of vaardigheden. Echter als een toetsscore tot stand komt onder flinke stress of zonder dat er voldoende concentratie was bij het maken van de toets, dan zal de toetsscore juist dat reflecteren, niet de iemands werkelijke kennis of vaardigheid. In zulke gevallen is het heel makkelijk om te miskennen wat iemand werkelijk in zijn mars heeft.

Het moge duidelijk zijn dat toetsen de grondslag vormen voor belangrijke, soms toekomstbepalende beslissingen. Bovendien beïnvloeden ze onderwijsprocessen, soms ten goede, maar niet noodzakelijkerwijs, en ze kunnen flinke persoonlijke gevolgen hebben. De vraag die zich aandient en die centraal staat in de volgende paragraaf is: Hoe goed zijn onze taalvaardigheidstoetsen eigenlijk? Zijn de beslissingen die we nemen op basis van taalvaardigheidstoetsen wel verantwoord.

De technische beperkingen van taalvaardigheidstoetsing

Het gebrek aan theorieën van taalvaardigheid

Taalvaardigheid bestaat en dus kun je het toetsen. Maar makkelijk is dat niet. Een van de problemen is dat het schort aan een goede theorie waarop toetsen gebaseerd kunnen worden (Hulstijn, In press). Als gevolg daarvan is het moeilijk om een schaal te maken waarlangs taalvaardigheid gemeten kan worden. Vergelijk het met het meten van snelheid in kilometers per uur. De algemeen geaccepteerde theorie is dat je snelheid meet aan de hand van afstand en tijd. Dus als je wilt weten

hoe snel iets of iemand was, dan dien je te weten hoeveel kilometer er werd afgelegd, en hoeveel tijd dat kostte. De dimensies aan de hand waarvan je snelheid bepaalt zijn geaccepteerd en goed meetbaar. Het probleem bij taalvaardigheid is dat we niet precies weten welke dimensies ten grondslag liggen aan het construct. Er zijn veel nogal uiteenlopende theorieën van taalvaardigheid, die eigenlijk geen van allen voldoende empirisch zijn onderbouwd om te dienen als basis voor een taalvaardigheidstoets (Hulstijn, In press). Het is dan ook niet verwonderlijk dat de vele taalvaardigheidstoetsen die er zijn weliswaar op elkaar lijken, maar nooit precies hetzelfde zijn. De verschillen tussen de toetsen en toetsvormen reflecteren verschillen in hoe men over taalvaardigheid denkt.

Een hieraan gerelateerd probleem is dat uit de theorieën van taalvaardigheid die er zijn geen duidelijke meetschaal voortkomt, met als gevolg dat de generaliseerbaarheid van het toetsresultaat in het geding komt. Een vergelijking met een natuurkundig verschijnsel maakt ook dit probleem inzichtelijk Als je hebt vastgesteld dat iets 31 meter lang is, dan is dat hoe dan ook meer dan 30 meter, en je weet ook hoeveel meer. Dit meetresultaat heeft een duidelijke betekenis. Het is echter niet eenvoudig om zinvolle uitspraken te doen over iemands taalvaardigheid op basis van een score van bijvoorbeeld 30 van de 40 vragen goed. Want hoe taalvaardig ben je als je 30 scoorde, en hoeveel taalvaardiger is iemand met score 31 ten opzichte van iemand met een score van 30? Deze vragen zijn moeilijk te beantwoorden, doordat de toetsschalen die worden gebruikt (in dit voorbeeld een schaal van 0 tot 40) betekenisloos zijn; ze zijn niet verankerd in een theorie die een toetsscore kan koppelen aan een taalvaardigheidsniveau. Ze komen voort uit het gekozen toetsdesign, niet uit het te meten construct.

Een gevolg hiervan is dat het moeilijk is om inhoudelijke uitspraken te doen over iemands taalvaardigheid. Uitspraken over de prestaties van individuele leerlingen worden daarom vaak gedaan door ze te vergelijken met de prestaties van een grote groep. Dit wordt normgerichte interpretatie genoemd. Als de norm bijvoorbeeld op 40 procent wordt gelegd, dan moet de leerling ervoor zorgen dat zijn/haar score bij de beste 60 procent hoort. Interpretatie op basis van een norm is nog steeds inhoudsloos. Daarom is domeingerichte interpretatie te prefereren. Dit wil zeggen dat er een inhoudelijk criterium wordt gesteld waaraan de cursist met voldoen. Tegenwoordig wordt het Europees Referentiekader (*Common European Framework of Reference for Languages: Learning, teaching, assessment*, 2001) hier vaak voor gebruikt. Een taalleerder moet dan bijvoorbeeld bewijzen dat hij of zij in staat is om 'met enige omhaal van woorden een gesprek te voeren over alledaagse dingen' (vrij naar het Europees referentiekader). Het probleem van het gebruik van

criteria is dat het een subjectief oordeel vereist om vast te stellen of aan het gestelde doel is voldaan.

Betrouwbaarheid, validiteit en efficiëntie

Om zinvolle uitspraken te doen over iemands taalvaardigheid maken toetsconstructeurs niet alleen gebruik van normen en criteria, ze stellen ook hoge eisen aan de *betrouwbaarheid* en de *validiteit* van hun toetsen. Of een toets betrouwbaar is, is vast te stellen aan de hand van bepaalde statistische procedures en is uit te drukken in een maat voor betrouwbaarheid, de betrouwbaarheidscoëfficiënt. Er wordt allereerst gekeken naar hoe de items van een toets zich gedragen. Elk item moet informatie bijdragen over de te meten vaardigheid: items die dat niet doen vallen af. Een andere veelgebruikte procedure voor het betrouwbaar stellen van een toets is de zogenaamde *split-half* methode. De toetsitems worden willekeurig onderverdeeld in twee subtoetsen, waarna de correlatie tussen de twee subtoetsen wordt berekend. Een hoge correlatie is gewenst, want dat betekent dat de twee subtoetsen de kandidaten min of meer hetzelfde ordenen. Zo zijn er meer methodes voor handen om de betrouwbaarheid van een toets vast te stellen. Betrouwbaarheidsbepaling is eigenlijk niets anders dan een statistische procedure waarmee je controleert of de items van je toets zich netjes gedragen. Door de items van je toets in een statistisch keurslijf onder te brengen, kun je met een behoorlijke mate van zekerheid verschillen in vaardigheid aantonen tussen verschillende toetsnemers.

Een betrouwbare toets geeft helaas niet de garantie dat je een zelfde ordening zou vinden als je 40 totaal andere toetsitems zou gebruiken. De toetsscore is afhankelijk van de items die zijn geselecteerd voor opname in de toets. Een score op woordenschat van 30 goed op 40 items zegt vooral dat je van de 40 *bevraagde* woorden er 30 wist. Het is onduidelijk hoe deze score zich verhoudt tot een score op 40 andere items, en belangrijker, of er dan eenzelfde rangorde van testnemers te voorschijn zou komen. De hoop van de toetsconstructeur is dat het resultaat generaliseerbaar is; als er steeds willekeurig 40 andere woorden geselecteerd worden voor de toets, dan zou dat steeds hetzelfde scorepatroon op moeten leveren, zodat je in ieder geval uitspraken kunt doen over wie sterk is en wie zwak. Een hoge betrouwbaarheid geeft een indicatie dat dat zo is, maar kijkt feitelijk niet buiten de toets.

Validiteit is een ander belangrijk middel om de kwaliteit van een toets te waarborgen. Een toets is valide als hij daadwerkelijk datgene meet wat je beoogt te meten. Valideren is in de hoofdzaak geen statistische procedure, maar het is een evaluatieproces waarbij de toetsconstructeur controleert, bij voorbeeld, of een luistertoets ook echt luistervaardigheid meet. Bij dit evaluatieproces kun je verschillende vormen van validiteit betrekken. Ik noem er twee.

Inhoudsvaliditeit vereist dat de toets alle aspecten van de te meten taalvaardigheid goed afdekt, maar niet meer dan dat. Taalvaardigheidtoetsen en inhoudsvaliditeit staan per definitie op gespannen voet met elkaar. Het voorbeeld van de woordenschattoets uit de vorige alinea illustreert dit punt. Je kunt niet alle woorden toetsen, en dus moet je een hele kleine steekproef trekken. *Ecologische validiteit* is ook een belangrijk validiteitsaspect. De C-test is een gatentekst-procedure waarbij om de zoveel woorden de tweede helft van een woord moet worden aangevuld door de toetsmaker. C-toetsen blijken behoorlijk voorspellend te zijn voor zowel lees-, luister-, schrijf- en spreekvaardigheid in de tweede taal (Klein-Braley, 1997). Maar het inzetten van een C-test voor het meten van iemands schrijfvaardigheid zou je een schending van de ecologische validiteit kunnen noemen, want de procedure staat te ver af van het te meten construct. Validiteit komt onder druk te staan op het moment dat efficiëntie een rol gaat spelen. Paradoxaal genoeg is dat vaak het geval bij toetsen waar grote belangen op het spel staan, omdat er dan vaak hele grote groepen getoetst moeten worden (Shohamy, 2001), zoals bij het centraal schriftelijk aan het eind van het voortgezet onderwijs of bij de Cito eindtoets. Vanwege de grote aantallen moesten er wellicht concessies gedaan worden aan de toetsvorm, zodat het toetsen niet te veel tijd kost. Om die reden kan men besluiten tot makkelijk scoorbare toetsvormen zoals meerkeuze, en tot een relatief klein aantal vragen (Marchant, 2004). Deze concessies zijn nodig om ervoor te zorgen dat de grote aantallen toetsen nog verwerkt kunnen worden. Ze zouden echter wel eens duur kunnen zijn in termen van impact.

Taalvaardigheidsmetingen zijn niet per definitie slecht of onnauwkeurig, maar de perfecte toets bestaat niet. Toetsen zijn nooit honderd procent valide en betrouwbaar. Er wordt bijvoorbeeld veel gebruik gemaakt van meerkeuze, ook al zijn de meeste onderzoekers het er over eens dat de taak zelf – het kiezen uit meerdere opties – een extra dimensie toevoegt die niet direct van doen heeft met de te meten vaardigheid. Veel mensen ervaren dit ook zo en oefenen daarom de taak. De consequentie hiervan is dat taalvaardigheidstoetsen wel taalvaardigheid meten, maar noodzakelijkerwijs enigszins grof zijn, een beetje alsof je iemands lengte wilt bepalen met een meetlint dat alleen halve meters kent. Het verdient daarom aanbeveling om toetsresultaten met enig wantrouwen te benaderen, helemaal als er belangrijke beslissingen op het spel staan.

Conclusies en aanbevelingen

Docenten zijn vaak degenen die toetsen afnemen, en soms degenen die besluiten tot het inzetten van toetsen in hun onderwijs. Dat betekent

dat zij een bepaalde verantwoordelijkheid dragen. Zij dienen toetsen kritisch te benaderen. Belangrijk is bijvoorbeeld dat docenten nadenken over de washback die uitgaat van de toetsen die ze gebruiken. De toets is heel bepalend voor wat leerlingen leren, voor waar zij energie in zullen steken, dus is het van belang dat de toets een goede afspiegeling is van datgene wat de leerling geacht wordt te weten of te kunnen. De belangrijkste aanbeveling is daarom dat de toetstaken zoveel mogelijk authentiek zijn (Bailey, 1996; Messick, 1996; Shohamy, 2001) en zoveel mogelijk worden geplaatst in een betekenisvolle en realistische context. De taken die bijvoorbeeld worden gebruikt in een spreekvaardigheidtoets moeten zo goed mogelijk lijken op de spreektaken die taalleerders in het dagelijkse leven moeten uitvoeren. Vormen van toetsing die de nadruk leggen op uit het hoofd leren, op het invullen van gaten, het kiezen uit meerdere opties doen dat typisch niet, en het is dus maar zeer de vraag of de voorbereidingen die leerlingen treffen voor het maken van dergelijke toetsen leiden tot het soort kennen of kunnen dat beoogd wordt. Bovendien kunnen ondermaatse toetsen funest zijn voor iemands motivatie. Van Avermaet (Van Avermaet, 2003) illustreert dit als hij beschrijft hoe zijn zoon de lust tot leren vergaat vanwege een onredelijke toets. Nadenken over de washback van hun toetsen is overigens zeker niet alleen een taak van docenten, maar vooral ook van toetsontwikkelaars. ALTE, de Association of Langauge Testers in Europe, heeft een *Code of Practice* ontwikkeld als richtlijn voor taaltoetsontwikkelaars ("The ALTE Code of Practice," 1994). Brown (2000) doorzocht de literatuur op zoek naar toetsconstructie-principes die positieve washback bewerkstelligen, wat een bruikbare lijst aanbevelingen heeft opgeleverd voor toetsconstructeurs. Die lijsten kunnen ook heel bruikbaar zijn voor docenten om toetsen op hun merites te beoordelen (zie links in de literatuurlijst).

Een ander belangrijk advies is dat de docent een toetsbeleid ontwikkelt. Hij moeten bedenken welke beslissingen absoluut genomen moeten worden en of toetsen nodig zijn om die beslissingen te nemen. Er zijn vaak minder belastende alternatieven voor toetsen. Portfolio's, zelfreflecties en observaties zijn bijvoorbeeld uitstekende alternatieven voor toetsen die een formatief doel dienen. Het voordeel van dat soort vormen van evaluatie is dat de leerling veel meer controle heeft. Een toets is iets wat een leerling overkomt; een portfolio is volledig zijn/haar eigen verantwoordelijkheid. Als er besloten wordt tot het gebruik van een toets, bedenk dan of de gekozen toets de te nemen beslissing wel rechtvaardigt. Een goed uitgangspunt voor een toetsbeleid is dat het aantal toetsen wordt geminimaliseerd. Toetsing lijkt misschien onschuldig als er geen belangrijke beslissingen worden genomen op basis van de scores, maar het is de vraag of dat wel zo is. Waarschijnlijk is dat niet de perceptie van de leerling, en het betekent bovendien voor een aantal leerlingen dat ze

een onvoldoende zullen moeten incasseren. Toets niet als het enige doel is informatie te verzamelen ten behoeve van het dossier van de leerling, en vraag verantwoording als u van hogerhand gevraagd wordt de leerlingen te toetsen.

Ook belangrijk is dat de docent de behaalde toetsscores bekijkt in het licht van wat hij of zij weet van de leerling of cursist. Vooral als iemand slecht gescoord heeft is het van belang uit te zoeken hoe dat komt. Het is nogmaals niet zo waarschijnlijk dat een goede score het resultaat is van toeval, maar als een leerling een slecht resultaat behaalt, is het maar zeer de vraag of dat komt doordat leerlingen echt zwak zijn in de gemeten vaardigheid; een lage score kan ook veroorzaakt worden door een gebrek aan concentratie, faalangst, of doordat de getoetste iets anders aan zijn hoofd had. In zulke gevallen is de ruis in de toetsscore erg groot, en zegt het toetsresultaat nagenoeg niet zo veel over iemands vaardigheid.

Tot slot

Deze bijdrage moet niet worden gelezen als een pleidooi tot het afschaffen van toetsen. Als ze goed gebruikt worden, dan kunnen toetsen een uitermate krachtig middel zijn om positieve veranderingen in het onderwijs te bewerkstelligen. Bovendien kunnen toetsen aanjagend werken Maar dat stelt wel hoge eisen aan de kwaliteit van de toets en de competenties van docenten om toetsen verstandig te gebruiken. Didactici lijken zich onvoldoende bewust van de impact van toetsen. Opvallend bijvoorbeeld is dat in didactische handboeken het onderwerp toetsing en toetsconstructie soms afwezig is, terwijl toetsen een belangrijke rol spelen in het het vormgeven van onderwijsprogramma's. Dat doet geen recht aan de feitelijke situatie dat toetsen een sterk aanjagend en sturend effect hebben, en dat het succes van een taalleerprogramma ook sterk afhangt van de wijze waarop wordt geëvalueerd. Als toetsing wel wordt besproken, dan mist meestal het onderwerp washback. Ook dat is zorgelijk, omdat de beslissingen die op basis van toetsen worden gemaakt grote, zichtbare en onzichtbare, gevolgen kunnen hebben voor de toetsnemer. Terughoudendheid en voorzichtigheid is geboden, helemaal gezien het feit dat taalvaardigheidstoetsen hun beperkingen kennen. Ook al is er veel tijd en aandacht besteed aan de betrouwbaarheid en de validiteit van een taalvaardigheidstoets, ze zijn nooit perfect, domweg omdat we de kennis nog niet hebben om perfecte toetsen te maken.

In dit licht moeten ook de aanbeveling van de Commissies Meijerink en Dijsselbloem worden bekeken. Beide rapporten zijn sympathiek in hun oprechte bezorgdheid voor de beschreven problematiek, maar het is zeer de vraag of de voorgestelde oplossing – in beide gevallen het instellen

van (verplichte) toetsen – het gewenste effect zal hebben. Met name dit soort *high stakes* toetsen kunnen leiden tot ingrijpende aanpassingen van onderwijsprogramma's en lesactiviteiten. Er wordt niet veel onderzoek gedaan naar de effecten die toetsen hebben op de toetsnemers en op de onderwijs- en leerprocessen; het precieze effect van het instellen van deze toetsen valt dan ook niet te voorspellen, maar twee waarschuwingen zijn wel te geven: er zullen onvoldoendes vallen en het onderwijs zal zich naar de toets voegen.

Literatuur

The ALTE Code of Practice. (1994). Retrieved juni 2008, 2008, from http://www.alte.org/quality_assurance/index.php

Bailey, K. M. (1996). A review of the Washback concept in language testing. *Language Testing, 13*, 257-279.

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. White Plains, NY: Pearson Education.

Brown, J. D. (2000). University entrance examinations: Strategies for creating positive washback on English language teaching in Japan. *JALT Testing & Evaluation SIG Newsletter* Retrieved 07-2008, 2008, from http://www.jalt.org/test/bro_5.htm

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

Hughes, A. (1993). Backwash and TOEFL 2000. (Unpublished manuscript).

Hulstijn, J. H. (In press). Language proficiency: Construct and measurement. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* Amsterdam: Benjamins.

Klein-Braley. (1997). C-tests in the context of reduced redundancy testing: an appraisal. *Language Testing, 14*(1), 47-84.

Marchant, G. J. (2004). What is at stake with high stakes testing? A discussion of issues and research. *Ohio Journal of Science, 104*(2), 2-7.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*, 241-254.

Over de drempels met taal en rekenen. (2008). Enschede: Expertgroep doorlopende leerlijnen taal en rekenen.

Parlementair onderzoek onderwijsvernieuwingen. (2008). Den Haag: Commissie Parlementair Onderzoek Onderwijsverniewingen.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow: Pearson Education.

Van Avermaet, P. (2003). Ethiek in taaltoetsing. In T. Koole & J. Nortier (Eds.), *Artikelen van de vierde sociolinguistische conferentie* (pp. 455-467). Delft: Eburon.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning: can this be predicted or controlled? *System, 28*, 499-509.

CORRECTIEVE FEEDBACK OP UITSPRAAK; WELKE FOUTEN MAKEN NT2-LEERDERS.

CATIA CUCCHIARINI & AMBRA NERI

Radboud Universiteit, Taalwetenschap: Centre for Speech and Language Technology

Inleiding

De manier waarop we praten is ons visitekaartje. Uitspraak speelt daarbij een grote rol: een slechte uitspraak kan een negatieve indruk maken tijdens een sollicitatiegesprek en tot vooroordelen leiden over de inhoudelijke competenties van een sollicitant, ook al gebruikt hij/zij verder correcte terminologie en grammatica. Over het algemeen is het moeilijk voor volwassen leerders om de klanken van een tweede of vreemde taal correct te leren uitspreken en voor een grote meerderheid van hen is het onhaalbaar om te leren spreken als een moedertaalspreker. Meestal is dat ook niet per se nodig om te kunnen functioneren in de maatschappij: als iemand zich makkelijk verstaanbaar kan maken voor een luisteraar, zodat er doeltreffend en mogelijk ook nog plezierige communicatie kan plaatsvinden, dan is dat vaak voldoende. Bij de meeste uitspraakcursussen ligt tegenwoordig de nadruk dan ook op verstaanbaarheid, maar ook dit doel kan voor vele volwassen NT2-leerders moeilijk te bereiken zijn.

Als we bijvoorbeeld naar klanken kijken, één van de aspecten van uitspraak, dan zien we dat een groot probleem bij de verwerving van klanken in een tweede taal de invloed is vanuit de moedertaal van de spreker, zowel bij de waarneming van klanken (bijvoorbeeld sommige NT2-leerders horen geen verschil tussen de 'a' van 'man' en de 'aa' van 'maan' omdat zij dat contrast niet kennen in hun moedertaal) als bij de productie van klanken (zo wordt 'bul', bijvoorbeeld, door vele leerders uitgesproken als 'boel'). Moeilijkheden bij de verwerving van klanken in een tweede taal kunnen ook gerelateerd zijn aan de specifieke kenmerken van bepaalde klanken in de tweede taal, zoals bijvoorbeeld bij de Nederlandse 'g/ch', een klank die relatief weinig voorkomt in andere talen en die voor vele NT2-leerders een struikelblok blijkt te zijn.

NT2-docenten die deze leerders proberen te helpen, worstelen vaak met een gebrek aan tijd die besteed kan worden aan uitspraaktraining. Bovendien werken NT2-docenten meestal met groepen cursisten met verschillende taalachtergronden die uiteenlopende problemen hebben.

Aanvullende zelf-studie is niet altijd de ideale oplossing: leerders kunnen vaak moeilijk hun eigen uitspraak beoordelen en horen hun eigen fouten helemaal niet. Individuele begeleiding door een vakkundig docent zou waarschijnlijk meer helpen, maar daarvoor is er meestal, helaas, geen tijd of geld.

Om de uitspraak toch te kunnen verbeteren zou men kunnen kiezen voor een efficiënte aanpak waarbij de training wordt gericht op fouten die het meest frequent, het meest storend voor de verstaanbaarheid, en het meest hardnekkig zijn. Maar wat zijn deze fouten voor het Nederlands als tweede taal?

In het kader van de ontwikkeling van een geautomatiseerde uitspraaktraining op basis van automatische spraakherkenning, zijn we op zoek gegaan naar informatie over uitspraakfouten van NT2-leerders in de literatuur. Dit soort informatie is namelijk cruciaal voor het maken van automatische uitspraaktrainingsystemen, omdat die nauwkeurig moeten worden voorgeprogrammeerd: ze moeten de juiste fouten kunnen opsporen en daar feedback op kunnen geven.
Tegelijkertijd is informatie over de meest frequente, storende en hardnekkige NT2-uitspraakfouten ook belangrijk voor docenten die optimale uitspraaktraining willen bieden aan groepen NT2-leerders met verschillende taalachtergronden.
Uit ons literatuuronderzoek bleek snel dat systematische informatie over veel gemaakte uitspraakfouten van NT2-leerders echter schaars is. We besloten dus om zelf infomatie te verzamelen op basis van een aantal databases met spraak van NT2-leerders die we ter beschikking hadden. In dit artikel doen we verslag van onze bevindingen en proberen we een inventaris van uitspraakfouten op te stellen die als richtlijn zou kunnen dienen voor NT2-docenten.

De achtergrond

Veel van de informatie over de problemen die anderstaligen ervaren bij het verwerven van de uitspraak van het Nederlands is afkomstig van spraaktherapeuten, van NT2-docenten en van wetenschappers die om verschillende redenen het Nederlands van anderstaligen hebben onderzocht (Aan de Wiel et al. 1991, Coenen 1991, Tupker 1995, Doeleman 1998). Het beeld dat uit de beschikbare literatuur onstaat is dat er over het algemeen meer problemen zijn met klinkers dan met medeklinkers. Dit kan te maken hebben met het feit dat het Nederlands relatief veel klinkers heeft vergeleken met andere talen, 13 éénklanken (monoftongen), 3 tweeklanken (diftongen) en een paar klinkers in leenwoorden uit andere talen. Onder de medeklinkers (consonanten) lijkt vooral de g/ch (zoals in geef en lachen) nogal problematisch te zijn: een klank die relatief zeldzaam is in andere talen (Maddieson 1984).

Veel van deze studies naar uitspraakproblemen in het Nederlands zijn echter niet systematisch van aard. Systematische studies zijn bijvoorbeeld wel uitgevoerd voor leerders van Turkse en Japanse afkomst (Van Boeschoten 1989, De Graaf 1986), maar uitgebreid onderzoek met leerders van verschillende modertalen is vooralsnog niet voorhanden. Daarom besloten wij om zelf die informatie te verzamelen aan de hand van verschillende analyses van verschillende spraakdatabanken, zoals in de volgende sectie zal worden uitgelegd.

Aanpak

Om informatie te vergaren over de aard en de frequentie van uitspraakfouten bij leerders van het Nederlands met verschillende moedertalen maakten wij gebruik van drie verschillende spraakdatabases die wij ter beschikking hadden. De resultaten van de drie analyses combineerden wij vervolgens om een inventaris van fouten op te stellen die naar onze mening prioriteit zouden moeten krijgen in NT2-uitspraak-training. Hieronder zullen wij achtereenvolgens de databases beschrijven, de motivatie om die te gebruiken voor ons onderzoek en de resultaten die op basis daarvan verkregen zijn.

Spraakdatabase 1: NT2, voorgelezen spraak

Allereerst keken we naar de opnames van voorgelezen spraak van 54 NT2-leerders met verschillende moedertalen en verschillende vaardigheidsniveaus (zie tabel 1) die verzameld waren in het kader van eerder onderzoek naar de mogelijkheden voor het automatisch scoren van uitspraakkenmerken (Cucchiarini et al, 2000). De sprekers hadden gemiddeld al drie jaar in Nederland gewoond en hadden het Nederlands formeel geleerd tijdens minimaal enkele maanden. De meeste sprekers bleken in feite gevorderden of half-gevorderden te zijn. We konden aannemen dat de uitspraakfouten in dit spraakmateriaal geen beginnersfouten waren en dus vooral representatief zullen zijn voor hardnekkige fouten.

Het materiaal bestond uit drie sets van fonetisch rijke zinnen. Dit zijn zinnen waarin alle klanken van het Nederlands voorkomen en waarbij de meest gangabare klanken vaker vertegenwoordigd zijn. Dit materiaal leek bijzonder geschikt voor dit doel. Ten eerste omdat het makkelijker is om spraak te beoordelen als alle sprekers hetzelfde voorlezen. Ten tweede ook omdat vergelijkingen makkelijker zijn als het materiaal gebalanceerd is en alle klanken aanwezig zijn.

Tabel 1: Verdeling van de sprekers op basis van moedertaal en taalvaardigheidsniveau (laag tv, gemiddeld tv en hoog tv).

Moedertaal	Laag	Gemiddeld	Hoog	Totaal
Arabisch	1	3	2	6
Chinees/Japans	1	1	1	3
Turks	1	2	0	3
Italiaans	0	2	4	6
Spaans	1	2	2	5
Engels	2	2	2	6
Duits	1	3	2	6
Frans	1	3	1	5
Zweeds/Noors	1	3	0	4
Pools/Russisch	0	2	2	4
Servo-Kroatisch	0	2	4	6
Totaal	9	25	20	54

Spraakdatabase 2: NT2, semi-spontane spraak

Uit onderzoek is gebleken dat de fouten die worden gemaakt in voorgelezen spraak vaak verschillen van de fouten die in spontane spraak voorkomen. Klanken worden in voorgelezen spraak meestal nauwkeuriger gearticuleerd omdat de spreker alleen bezig is met het uitspreken van een tekst en niet ook met het formuleren van een gedachte (Dickerson 1975, Lin 2001, Cucchiarini et al. 2002). Voorgelezen spraak lijkt ook sterker beïnvloed te worden door de schrijfwijze (Young-Scholten 1997). Daarom keken we ook naar fragmenten van semi-spontane spraak die verzameld was in het kader van een pre-test van de Profieltoets Inburgering (Cucchiarini et al., 2002). Deze spraak bestaat uit de antwoorden die 43 kandidaten met verschillende moedertalen hebben gegeven op vragen die ze van te voren niet kenden en kan daarom worden beschouwd als niet voorbereide, semi-spontane spraak.

Vanwege de context waarin dit spraakmateriaal werd verzameld was het niet mogelijk om gegevens te verkrijgen over de sprekers, behalve hun vaardigheidsniveau (23 sprekers waren beginners en 20 half-gevorderden). Maar juist vanwege die context is dit spraakmateriaal een handige bron van informatie over fouten die in meer realistische omstandigheden kunnen voorkomen. Bovendien werden de sprekers voor deze analyse door werknemers van de CITO-groep zodanig geselecteerd dat ze een zo representatief mogelijke steekproef vormden qua geslacht, moedertaal en vaardigheidsniveau.

Omdat bekend is dat bij spontane productie van spraak mensen vaak moeilijke woorden en/of klanken proberen te vermijden uit angst om fouten te maken (Piske et al. 2001), hebben wij ook gecontroleerd of onze spraakdatabase compleet was. De bedoeling van onze analyse was immers juist de moeilijke Nederlandse klanken te identificeren door fouten op te sporen. Als die moeilijke klanken zouden ontbreken juist omdat ze moeilijk zijn, zou de analyse een verkeerd beeld opleveren. Daarom vergeleken we onze database met drie verschillende andere databases (zie Neri et al. 2006). Uit de vergelijking bleek dat onze database compleet was en we besloten die verder te gebruiken voor de analyse van uitspraakfouten.

Spraakdatabase 3: NVT, voorgelezen spraak

Ten derde keken we naar spraak die was voorgelezen door 20 Italiaanse leerders van Nederlands als vreemde taal. Omdat deze leerders netjes verdeeld waren in 5 groepen op basis van de hoeveelheid training die zij hadden gekregen in het Nederlands, konden wij inzicht krijgen in welke fouten hardnekkig zijn en problematisch blijven voor leerders ook nadat ze jaren onderwijs Nederlands hebben gevolgd.

De 20 Italiaanse sprekers waren studenten tussen 19 en 25 jaar die Nederlands studeerden op de opleiding voor Vertalers en Tolken van de Universiteit van Triëst (Italië). De keuze voor deze groep leerders was vooral pragmatisch van aard. Gegeven onze contacten met de opleiding wisten we dat het zou lukken om een redelijk representatieve steekproef samen te stellen. Alle proefpersonen waren moedertaalsprekers van het Italiaans die pas op de universiteit begonnen waren met Nederlands te leren. Zij hadden minstens 150 uur les Nederlands gevolgd, ze waren sterk gemotiveerd om Nederlands te leren en ze hadden alle aangegeven het belangrijk te vinden om hun uitspraak te verbeteren, al vonden zij grammaticale correctheid belangrijker.

De procedure: annotatie en analyse

De 1329 spraakfragmenten (810 uit database 1, 319 uit database 2 en 200 uit database 3) werden geanalyseerd om uiteindelijk een selectie te kunnen maken van uitspraakfouten die ernstig genoeg zijn om er aandacht aan te besteden in de les, cq in ons automatisch uitspraaktrainingsprogramma.
De eerste stap in dit proces is de annotatie: alle spreekfragmenten werden beluisterd en van kenmerken voorzien. Bij zowel de annotatie, als bij de analyse en de selectie van de fouten werden de volgende criteria gehanteerd:

1. Opvallend
2. Storend voor de communicatie
3. Frequent
4. Persistent
5. Gemeenschappelijk voor meerdere moedertalen

De spraakfragmenten werden eerst geannoteerd op uitspraakfouten die opvallend en als storend werden ervaren. Vervolgens werden de fouten geanalyseerd in kwantitatief en kwalitatief opzicht om uiteindelijk fouten te selecteren die frequent, persistent en gemeenschappelijk zijn voor meerdere moedertalen. De focus op fouten die opvallend en storend zijn sluit aan bij recente benaderingen van het uitspraakonderwijs. De nadruk op frequente fouten heeft te maken met het feit dat die een groot effect hebben op de algemene indruk van iemands uitspraak. De reden om te kiezen voor fouten die gemeenschappelijk zijn voor meerdere moedertalen is gerelateerd aan het feit dat in NT2-klassen in Nederland cursisten met verschillende moedertalen bij elkaar zitten. Ten slotte, de keuze om vooral persistente fouten te selecteren is gebaseerd op de constatering dat het niet zoveel zin heeft om aandacht te besteden aan het corrigeren van fouten die toch uit zichzelf verdwijnen.

Allereerst werd het spraakmateriaal auditief geanalyseerd door een expert die annotaties maakte van opvallende fouten. De betrouwbaarheid van de expert werd op verschillende manieren getoetst om te voorkomen dat de annotaties te subjectief zouden zijn. Uit de tests bleek dat de annotaties van de expert grotendeels overeenkwamen met die van vijf andere experts aan wie ook gevraagd was een onderdeel van het spraakmateriaal te beoordelen.

Vervolgens werden de geannoteerde fouten zowel kwantitatief als kwalitatief geanalyseerd om frequentiepatronen te kunnen ontdekken en om de aard, de mogelijke oorzaak en het potentiële effect van een fout te kunnen bepalen. Dit resulteerde in een inventaris van uitspraakfouten waar volgens ons als eerste aandacht aan zou moeten worden besteed.

De Resultaten

Spraakdatabase 1: NT2, voorgelezen spraak

Uit de analyse van spraakdatabase 1 blijkt allereerst dat klinkers voor iedereen meer problematisch zijn dan medeklinkers, zowel in relatieve als in absolute zin. Klinkers vormden slechts 40% van de (doel)klanken in deze database en werden in 8% van de gevallen weggelaten of fout uitgesproken. Medeklinkers, die 60% vormden van de (doel)klanken, werden in 2% van de gevallen foutief uitgesproken.

Wij selecteerden 8 klinkers en 5 medeklinkers die het vaakst foutief werden uitgesproken in het hele spraakmateriaal en keken naar de meest frequente realisaties van die fouten om een algemeen beeld te krijgen. Vervolgens gingen wij op dezelfde manier te werk voor de verschillende groepen van moedertalen, om te begrijpen of sprekers met verschillende taalachtergronden afwijkende foutpatronen vertoonden ten opzichte van het algemeen beeld. Zoals duidelijk blijkt uit tabel 2 (klinkers) en tabel 3 (medeklinkers), gelden veel foutpatronen voor sprekers van verschillende moedertalen.

Tabel 2: Voorkomen van meest frequente klinkerfouten (overall trend) bij de verschillende moedertalen. Kruisjes geven aan welke fouten *niet* voorkomen voor sprekers van een bepaalde moedertaal

	Doelklinker							
	(stomme e)	/a/	/ui/	/u/	/aa/	/eu/	/ij/	/uu/
	Gerealiseerd als							
Moedertaal	Weggelaten /e/, /ee/	/aa/	/au/ /oi/	/oe/ /uu/	/a/	/uu/ /oo//oe/	/ee-ie/ /ai//aai/	/oe/
Arabisch	✓	✓	✓	✓	✓	✓	✓	✓
Chinees/Japans	✓	✓	✓	✓	✓	✓	✓	☒
Turks	✓	✓	✓	✓	✓	✓	✓	✓
Italiaans	✓	☒	✓	✓	✓	✓	☒	✓
Spaans	✓	✓	✓	✓	✓	✓	✓	✓
Engels	✓	✓	✓	✓	✓	✓	✓	✓
Duits	✓	✓	✓	✓	✓	✓	✓	✓
Frans	✓	✓	✓	✓	✓	☒	✓	☒
Zweeds/Noors	✓	✓	☒	✓	✓	☒	✓	☒
Servo-Kroatisch	✓	✓	✓	✓	✓	✓	✓	✓
Russisch/Pools	✓	✓	✓	✓	✓	✓	☒	✓

Tabel 3: Voorkomen van meest frequente medeklinkerfouten (overall trend) bij de verschillende moedertalen.

Doelklank	/r/	/t/	/g/	/w/	/h/
Gerealiseerd als	Weggelaten /l/, /a/	Weggelaten /d/	(goal) /h/,/k/	/oe/ /f/	Weggelaten /g/
Arabisch	✓	✓	✓	✓	☒
Chinees/Japans	✓	✓	✓	✓	✓
Turks	☒	✓	✓	✓	☒
Italiaans	✓	✓	✓	☒	✓
Spaans	☒	✓	✓	✓	✓
Engels	☒	✓	✓	✓	☒
Duits	✓	✓	✓	✓	☒
Frans	✓	✓	✓	✓	✓
Zweeds/Noors	✓	✓	☒	☒	☒
Servo-Kroatisch	✓	✓	✓	✓	✓
Russisch/Pools	✓	✓	✓	No	✓

Naast het weglaten en verwisselen van klanken, keken wij ook naar de invoegingen omdat deze problemen kunnen signaleren met specifieke klanken of contexten. De meest frequente invoeging bleek de *sjwa (*de stomme–e), en dit vond meestal plaats aan het einde van woorden en lettergrepen na een medeklinker, of in moeilijk te articuleren clusters van medeklinkers, zoals dat in *afspraak,* dat bijvoorbeed gerealiseerd wordt als (affesepraak). Hoewel sjwa-invoeging ook in het Nederlands voorkomt (bijvoorbeeld in het woord 'werk'; werruk), was dit verschijnsel niet correct toegepast in de geannoteerde gevallen. Ook werd de /h/ vaak incorrect ingevoegd aan het begin van woorden. Na de klank /ng/ zoals in het woord *lang*, troffen we een nog sterker patroon van invoegingen van /g/ , /ch/ en /k/; deze invoegingen kwamen bij de meeste sprekers voor.

Vervolgens hebben we nog eens opnieuw gekeken naar de fouten die we hadden gevonden. Het doel van deze tweede analyse was om fouten uit te filteren of anders te rangschikken op basis van andere belangrijke factoren naast de frequentie: waren alle gevonden fouten echt representatief voor fouten die gemaakt zouden worden in spontane spraak, en welke zouden de uitspraak meer bemoeilijken? Sommige fouten konden bijvoorbeeld een gevolg zijn van het type spraakmateriaal

dat was gebruikt en zouden misschien minder frequent zijn in spontane spraak. Bijvoorbeeld, de sjwa werd het vaakst foutief voorgelezen. Een derde van de weglating van deze klank vond plaats na de klank /zj/ in Franse leenwoorden als *garage* (die gerealiseerd werd als garaazj). Deze weglating zou eigenlijk helemaal correct zijn in het Frans en in sommige andere talen waarin dit leenwoord voorkomt. Bovendien is de sjwa aanwezig in vele talen en kan makkelijk geproduceerd worden door de stemorganen in een 'rust houding' te houden, zoals we bijvoorbeeld doen als we aan het denken zijn en 'uh' zeggen. Waarschijnlijk had de fout dus puur te maken met het feit dat de sprekers de Nederlandse uitspraak van deze woorden (nog) niet kenden en dat ze mogelijk de uitspraak toepasten die zij wel kenden uit hun moedertaal of uit het Frans. Zo'n fout zou dan makkelijk verdwijnen zodra de spreker meer bekend wordt met de Nederlandse uitspraak van dit type woorden. De sjwa werd ook bijna 40% van de keren vervangen door /e/ en /ee/ als het weergegeven was door de letter *e*. De context van die foutieve verwisseling duidt aan dat het waarschijnlijk om interferentie van het woordbeeld gaat. Ook dit probleem zou mogelijk minder vaak voorkomen in spontane spraak die niet geproduceerd wordt op basis van een geschreven tekst. De klinkers /a/, /aa/, /u/ en /uu/ werden vaak met andere Nederlandse klanken verwisseld. Dit zou juist problematisch kunnen zijn voor de communicatie omdat er misverstanden kunnen ontstaan. Om deze redenen lijkt het ons verstandiger om dit laatste type fouten prioriteit te geven boven sjwa-fouten in uitspraaktraining.

Onder de medeklinkers, vonden we veel foutieve uitsptraak van de /r/. Vele van deze verwisselingen werden door Chinese, Japanse en Duitse sprekers gemaakt. Het is bekend dat deze klank veel verschillende realisaties kent in het Nederlands, afhankelijk van de positie in het woord en als gevolg van regionale variatie, zoals bij de tongpunt -r, de Wassenaarse -r en de huig -r (Vieregge and Broeders 1993, Booij 1999, Van de Velde en van Hout 2001). Omdat blijkbaar zoveel verschillende varianten mogelijk zijn, leek het ons niet prioritair om aandacht te gaan besteden aan kleine nuances in de uitspraak van de /r/.

De klank /t/ werd door verschillende sprekers vaak weggelaten aan het einde van woorden en vooral in consonantenclusters als in *belooft* en *gelokt*. Hoewel t-weglating vrij normaal is in sommige Nederlandse dialecten (Goeman 1999) en in snelle spraak (Booij 1999), kunnen deze fouten ook het gevolg zijn van (foutieve) assimilatie door daaropvolgende klanken/woorden. Verder is het ook denkbaar dat voor sommige leerders het moeilijk is om de klank /t/ aan het einde van clusters van medeklinkers uit te spreken, omdat dit verder weinig voorkomt in andere talen (Tarone 1980, Lin 2001). In ieder geval kan een t-weglating aan het einde van een woord de communicatie zwaar bemoeilijken aangezien /t/ een morfologische functie heeft in Nederlandse werkwoorden: *beloof* betekent

immers iets anders dan *belooft*. Daarom lijkt het belangrijk om de uitspraak van deze klank wél te trainen, maar voornamelijk aan het einde van woorden of binnen consonantenclusters. De klank /g/ of /ch/ is relatief zeldzaam in andere talen (Maddieson 1984). Het is daarom denkbaar dat aandacht voor deze specifieke klank leerders zou kunnen helpen om makkelijker te communiceren en om 'meer Nederlands' te klinken. De uitspraak van de /w/ op een wijze die we kennen van Surinaamse en Brabantse sprekers is weer een geheel ander geval (Rietveld and Van Heuven 1997, Booij 1999). Zo'n uitspraak mag dan wel ervaren worden als afwijkend van de standaard uitspraak, maar zal moeilijk de communicatie belemmeren en hoeft daarom geen prioriteit te krijgen.

Om vergelijkbare redenen kozen we ervoor om de invoeging van de stomme –e, de sjwa, (meestal na medeklinkers, aan het einde van een lettergreep/woord) in onze lijst van belangrijke fouten op te nemen. Door sjwa-invoeging kan een nieuwe lettergreep ontstaan en dit kan woordherkenning bij luisteraars zwaar bemoeilijken (Cutler 1984). Invoeging van /g/, /ch/ en /k/ na de /ng/ in woorden als *bewegingen* (die als bewegingen, bewegingken en bewegingchen uitgesproken werden) kreeg daarentegen een lagere prioriteit omdat het alleen frequent was voor sommige groepen van sprekers (Chinezen, Japanners, Engelstaligen, Zweden en Noren hadden namelijk geen moeite met de klank, Polen Russen en Duitsers in mindere mate dan de andere sprekers).

De inventaris van ernstige fouten die overbleef na de kwalitatieve analyse bestond uit de volgende klanken:
- voor de klinkers de eenklanken /a/, /aa/, /u/, /uu/, /eu/ en /stomme-e/ en de tweeklanken /ui/ en /ij/,
- voor de medeklinkers /g/ en /t/ (waarbij vooral gekeken zou moeten worden naar de context waarin de fout is gemaaky) en de /r/ en de /ng/ (die slechts voor een beperkt aantal leerders problematisch bleken).

Spraakdatabase 2: NT2, semi-spontane spraak

De analyse van database 2 leverde een lijst van opvallende, frequente en hardnekkige fouten die gemaakt worden door sprekers van verschillende moedertalen in spontane spraak. Ook in dit geval bleken klinkers meer problematisch te zijn dan medeklinkers. Klinkers vormden 39% van de doelklanken en werden in 4% van de gevallen verkeerd uitgesproken, tegen 1% bij de medeklinkers. Hoewel we globaal naar alle fouten keken van beginners en gevorderden samen, vonden we een opmerkelijk verschil tussen de twee groepen in klinkers/medeklinkers foutpatronen: beginners maakten bijna net zoveel fouten bij klinkers (53%) als bij medeklinkers (47%), terwijl die verhouding bij gevorderden op 81%/19%

lag. Dit doet vermoeden dat veel fouten bij medeklinkers vanzelf verdwijnen naarmate de spreker vordert in het Nederlands, terwijl klinkerfouten hardnekkig zijn.

De meeste fouten werden gemaakt bij de eenklanken /a/, /aa/, /u/, /uu/, /o/, /e/ en /i/ en bij de tweeklanken /ui/ en /ij/. De fouten lijken veel op die van spraakdatabase 1. Op basis van het type verwisselingen die we vonden, lijkt het verstandig om de training vooral te richten op contrasten van verwisselbare klanken, zoals /a/ en /aa/, en tussen /u/ en /uu/ omdat zo'n training mogelijk twee problemen in een keer kan oplossen. Foutieve realisaties van de sjwa komen bovendien heel zelden voor in deze database, hetgeen onze hypothese lijkt te bevestigen dat deze klank niet zo moeilijk is om uit te spreken en dus geen prioriteit hoeft te krijgen in uitspraaktraining.

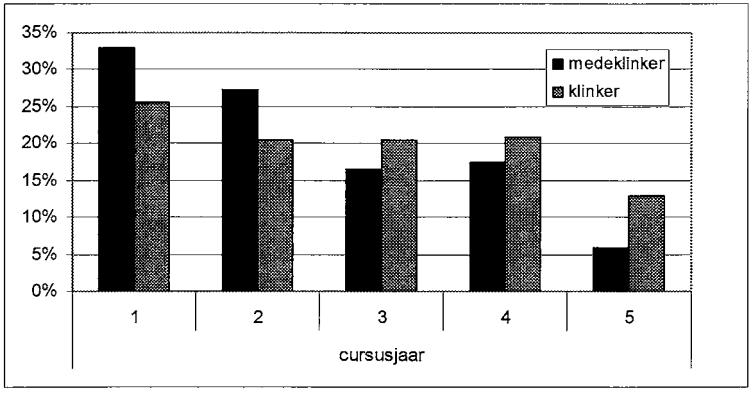
Bij de medeklinkers werd /g/ het vaakst foutief uitgesproken. Deze klank zou aandacht moeten krijgen en de uitspraak daarvan zou geoefend moeten worden in vergelijking met de /h/ of in clusters zoals in /sg/, omdat die vaak verward worden.

De kenmerken van de sprekers in deze studie en de aard van het materiaal zijn voor ons redenen om aan te nemen dat dit de fouten zijn die overblijven na jaren onderwijs in het Nederlands, maar uiteraard zou deze hypothese onderbouwd moeten worden met data uit experimenten waarin tijd en instructie meer systematisch zijn gevarieerd.

Spraakdatabase 3: NVT, voorgelezen spraak

Het doel van deze laatste studie was om te kijken welke fouten gemaakt worden in een NVT-context (Nederlands als Vreemde Taal) en of die fouten anders zijn vanwege minder blootstelling aan onderwijs Nederlands. We vonden dat ook in dit geval meer fouten bij klinkers worden gemaakt dan bij medeklinkers. Ook in dit geval lijken de meeste fouten die gemaakt werden bij de sjwa veroorzaakt te zijn door de schrijfwijze en niet zozeer door echte problemen met de uitspraak. Wat wel vaak gebeurde is dat een klinker vervangen werd door een andere klinker die erop lijkt. Het kan dus een goede strategie zijn om die klinkers samen aan te bieden om de aandacht te vestigen op de onderlinge verschillen, zoals /a/ - /aa/, /u/ - /uu, oe/, / ee/ - /e/ en /i/ - /ie/ en voor de tweeklanken /ui/ - /au/.

Met betrekking tot de relatie tussen type fouten en onderwijs Nederlands vonden wij dat fouten bij klinkers vaak lang problematisch blijven, ondanks blootstelling aan de taal (zie Fig. 1).



Figuur 1: Percentages klinker- en medeklinkerfouten per cursusjaar (1=eerstejaars; 5=afgestudeerden) relatief tot het totaal aantal fonemen van hetzelfde type.

Ondanks de verschillen tussen de Italiaanse en de Nederlandse medeklinkers vonden wij dat medeklinkerfouten bijna volledig verdwijnen bij de meest gevorderde groep (zie Fig. 1). Slechts drie medeklinkers, /h/, /d/ en /g/ werden fout uitgesproken door deze cursisten. Ook invoegingen van klanken waren zeldzaam. Het lijkt erop dat een les over de uitspraak van de /ng/ al heel veel zou kunnen helpen om invoegingen te voorkomen. Het vermijden van sjwa-invoeging, daarentegen, zou moeilijker kunnen blijken voor veel cursisten, ook wanneer aandacht wordt besteed aan prosodie. Het is namelijk bekend dat het aanpassen van prosodie over het algemeen problematischer is (Hirst and Di Cristo 1998).

Conclusies

Op basis van de analyses van de drie databases die hierboven zijn gepresenteerd kunnen we een algemeen beeld krijgen van de meest problematische uitspraakfouten van NT2-leerders met verschillende moedertalen. De drie analyses vertonen enkele kleine verschillen maar toch vooral veel overeenkomsten. Een klein verschil tussen voorgelezen spraak en meer spontane spraak heeft betrekking op de fouten met de sjwa. Zoals te verwachten was, lijken veel van deze fouten een gevolg te zijn van de schrijfwijze. Als mensen niet voorlezen, maar spontaan praten, zijn die fouten ook minder frequent. Hetzelfde lijkt van toepassing te zijn op de invoeging van de /k/ en de /ch/ na de /ng/, want deze komen niet voor in meer spontane spraak. Bij de medeklinkers blijken verder de /g of ch/ en de /h/ problematisch te zijn en die worden ook vaak met elkaar verward.

Verder zien we overal terugkomen dat klinkers moeilijker zijn dan medeklinkers en dat die ook langer problematisch blijven. De klinkers /a/ en /aa/ worden vaak met elkaar verward en het zou daarom misschien goed zijn om die samen contrastief te oefenen. Op vergelijkbare wijze zou men de klanken /u/ en /uu/ samen kunnen oefenen, die ook vaak als /oe/ worden gerealiseerd en de tweeklank /ui/ die vaak met /au/ wordt verward.

Hoewel er waarschijnlijk meer data uit meer moedertalen nodig zijn om een compleet beeld te krijgen, denken wij dat deze lijst al handvatten kan bieden aan docenten voor het verbeteren van uitspraakfouten bij NT2-leerders in gemengde klassen. Uiteraard is deze inventaris puur indicatief. De keuze van de fouten kan namelijk aangepast worden aan de specifieke situatie, bijvoorbeeld als een groep duidelijk oververtgenwoordigd is waarbij sommige fouten duidelijk meer frequent zijn, zoals bijvoorbeeld de /r/ bij Japanse cursisten. Verder bevat deze inventaris geen informatie over prosodie, terwijl we weten dat dat ook heel belangrijk is voor de verstaanbaarheid.

Ten slotte, de pedagogische waarde van een dergelijk inventaris is uiteraard gerelateerd aan de methode waarin het gebruikt wordt. De lijst fouten die op basis van deze analyses werd samengesteld, werd gebruikt in het programma Dutch-CAPT om automatisch feedback te geven op 11 uitspraakfouten op basis van automatische spraakherkenning. Het onderzoek dat met dit systeem is uitgevoerd heeft laten zien dat deze manier van correctieve feedback geven wel degelijk kan helpen om de uitspraak van problematische klanken te verbeteren. Bij de groep proefpersonen die automatische correctieve feedback kreeg werden significant minder fouten gemaakt bij die 11 klanken dan bij proefpersonen die die feedback niet kregen (Neri et al., 2008).

Literatuur

Aan de Wiel, Margreet C., Gertrude A. van den Brink, and Saskia G.E. Struijk van Bergen (1991) Diagnostiek van uitspraakproblemen van tweede-taalverwervers van het Nederlands. *Levende Talen* 466: 507-511.

Booij, Geert. (1999) *The Phonology of Dutch.* Oxford: Clarendon Press.

Coenen, Josée (1991) *Uitgesproken Nederlands.* Groningen: Wolters-Noordhoff.

Cucchiarini, Catia, Helmer Strik and Louis Boves (2000) Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology. *Journal of the Acoustical Society of America* 107: 989-999.

Cucchiarini, Catia, Helmer Strik and Louis Boves (2002) Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America* 111: 2862-2873.

Cutler, Anne (1984) Stress and accent in language production and understanding. In *Intonation, Accent and Rhythm. Studies in Discourse Phonology,* Dafydd Gibbon and Helmut Richter (eds.), 77-90. Berlin/New York: de Gruyter.

De Graaf, Tjeerd (1986) De uitspraak van het Nederlands door buitenlanders. *Logopedie en Foniatrie* 58: 343-347.

Dickerson, Lonna J. (1975) The learner's interlanguage as a system of variable rules. *TESOL Quarterly* 9: 401-407

Doeleman, Rianne (1998) *Native Reactions to Nonnative Speech.* Tilburg: Tilburg University Press.

Goeman, Ton (1999) *T-deletie in Nederlandse Dialecten. Kwantitatieve analyse van structurele, ruimtelijke en temporele variatie.* Doctoral dissertation, The Hague: HAG

Gussenhoven, Carlos (1999) Dutch. In *Handbook of the International Phonetic Association, Part II Illustrations of the IPA*, Cambridge: Cambridge University Press, 74-77.

Hirst, Daniel and Albert Di Cristo (1998) A survey of intonation systems. In *Intonation Systems: A Survey of Twenty Languages,* Daniel Hirst and Albert Di Cristo (eds.), 1-44. Cambridge: Cambridge University Press.

Lin, Yuh-Huey (2001) Syllable simplification strategies-A stylistic perspective. *Language Learning* 51: 681-718.

Maddieson, Ian (1984) *Patterns of Sounds.* Cambridge: Cambridge University Press.

Neri, Ambra, Catia Cucchiarini and Helmer Strik (2006) Selecting segmental errors in non-native Dutch for optimal pronunciation training. *Internationl Review of Applied Linguistics* 44: 357–404.

Neri, Ambra, Catia Cucchiarini and Helmer Strik (2008) The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch *ReCALL* 20(2): 225-243.

Piske, Thorsten, Ian R. A. MacKay and James E. Flege (2001) Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191-215.

Rietveld, Toni and Vincent J. Van Heuven (1997) *Algemene Fonetiek.* Bussum: Coutinho.

Tarone, Elaine (1980) Some influences on the syllable structure of interlanguage phonology. *International Review of Applied Linguistics* 4:143-63.

Tupker, Yvonne (1995) *Uitspraak Onderwijzen*. Manuscript. Tilburg: Katholieke Universiteit Brabant.

Van Boeschoten, Joost A. (1989) *Verstaanbaarheid van klanken in het Nederlands gesproken door Turken*. Dordrecht: ICG Printing.

Van de Velde, Hans and Roeland van Hout (2001) (eds.) *'r-atics. Sociolinguistic, phonetic and phonological characteristics of* [illegible], Brussels: Etudes et Travaux, ILVP-ULB.

Vieregge, Wilhelm H. and Ton Broeders (1993). Intra- en interspeaker variation of /r/ in Dutch. In *Proceedings of Eurospeech '93*, Berlin, 267-270.

Young-Scholten, Martha (1997) Second-language syllable simplification: Deviant development or deviant input? In *New Sounds 97: Proceedings of the Third Symposium on the acquisition of Second-Language Speech,* James Leather and Allan James (eds.), 351-360. Klagenfurt: University of Klagenfurt.

ACTIEONDERZOEK IN HET NT-2ONDERWIJS; DE DOCENT IN ACTIE.

FOLKERT KUIKEN

Universiteit van Amsterdam,
Nederlandse taalkunde / Duale master Nederlands als tweede taal

Inleiding

Goede docenten reflecteren voortdurend op hun eigen lespraktijk. Voorafgaand aan de les bedenken ze wat ze met hun cursisten willen bereiken en na afloop stellen ze vast of het gestelde doel is behaald. Een dergelijke reflecterende houding is van groot belang in het onderwijs, getuige ook de literatuur daarover, zoals in Wallace's (1991) reflectieve benadering of Litjens' (1990) pleidooi voor een reflecterende docent.

Maar wie zichzelf en het eigen onderwijs ter discussie stelt, kan nog meer bereiken door middel van actieonderzoek.[1] Met actieonderzoek verbeteren docenten hun *eigen* lespraktijk op basis van *door henzelf* uitgevoerd onderzoek. Een andere belangrijke component is de *systematiek*: de voor het onderzoek benodigde gegevens worden systematisch verzameld, zodat op basis daarvan gerichte actie kan worden ondernomen om *de praktijk te veranderen*.

In deze bijdrage licht ik het belang van dit soort onderzoek nader toe. Vervolgens ga ik dieper in op de kenmerken van actieonderzoek en op de verschillende fasen van het onderzoek. Dat betreft in de eerste plaats het formuleren van een juiste vraagstelling. Ten tweede de verschillende mogelijkheden die er zijn om gegevens over het gekozen probleem te verzamelen. In de derde plaats manieren om de verzamelde gegevens te ordenen, met elkaar te vergelijken en op grond daarvan conclusies te trekken die tot veranderingen in de lespraktijk leiden. En ten vierde de evaluatie en rapportage van het uitgevoerde onderzoek. Tot slot bespreken we hoe een afgerond onderzoek weer de aanzet tot nieuw onderzoek kan vormen en hoe actieonderzoek gekoppeld kan worden aan andere processen van onderwijsvernieuwing.

Waarom actieonderzoek?

[1] In plaats van actieonderzoek wordt ook wel de term ontwikkelingsonderzoek gehanteerd (Hacquebord 2003).

Men kan zich afvragen wat het nut is van actieonderzoek. Er wordt immers al zoveel onderzoek naar tweedetaalverwerving en tweedetaalonderwijs verricht en moet de docent zich nu ook nog zonodig de rol van onderzoeker aanmeten? Veel docenten zullen echter de ervaring hebben dat – hoe belangrijk wetenschappelijk onderzoek ook mag zijn – de resultaten van theoretisch onderzoek naar tweedetaalverwerving niet zomaar naar de lespraktijk vertaald kunnen worden, terwijl dat ook bij toegepast onderwijsonderzoek lang niet altijd het geval is. Het karakteristieke van actieonderzoek is nu juist dat het zich richt op vragen uit de lespraktijk waar docenten zelf mee worstelen. En wie kent die praktijk beter dan zijzelf, dus waarom zouden docenten zelf dan ook niet de aangewezen personen zijn om de kwesties waar zij in hun praktijk tegenaan lopen aan een nader onderzoek te onderwerpen?

Mocht het verrichten van actieonderzoek door docenten zelf hiermee nog niet voldoende beargumenteerd zijn, dan verschaft Burns (1999) docenten nog een aantal aanvullende motieven voor het doen van actieonderzoek. In de eerste plaats is actieonderzoek, zoals boven vermeld, een krachtig reflectiemiddel, in de zin dat het een kritisch perspectief biedt op de eigen praktijk. In de tweede plaats leidt het systematisch observeren van de NT2-lessen tot een beter inzicht in de dynamiek van de klas, wat veranderingen in het curriculum teweeg kan brengen met als uiteindelijk gevolg een verbetering van de leerprestaties.

Naast deze meer algemene motieven geven docenten ook zelf aan baat te hebben bij actieonderzoek. Voor veel docenten betekent het doen van actieonderzoek een (nog) grotere betrokkenheid bij het onderwijs. Voor anderen betekent het een uitbreiding van de eigen professionele ontwikkeling, aansluitend bij een vaak al bestaande wens om de onderzoeksvaardigheden te vergroten. Het doen van actieonderzoek vormt dan als het ware een natuurlijke uitbreiding van de onderwijspraktijk, waardoor ook het zelfbewustzijn van de docenten toeneemt.

Veel Angelsaksische landen (Australië, Engeland, Amerika en Canada) kennen een langere traditie van actieonderzoek dan Nederland. Maar de laatste tijd neemt de belangstelling hiervoor ook in het talenonderwijs in Nederland toe, onder meer dankzij publicaties van Hacquebord (2003) en Ponte (2002a, b). Opvallend genoeg lijkt actieonderzoek vooral in het voortgezet onderwijs ingang te vinden (Galema 2003, van der Hagen 2005, Hofstetter, Schaareman & de Vries 2005, Kroon 2007, Kuiken 2006, van der Mee 2003). In de volwasseneneducatie is daarvan (vooralsnog?) minder sprake. Verder worden steeds meer docenten tijdens hun opleiding met actieonderzoek in aanraking gebracht (zie SpeZiaal 2008). Hopelijk laten deze (aspirant)docenten zich niet ontmoedigen door de obstakels die ze in de praktijk soms moeten overwinnen bij het opzetten van actieonderzoek als ze eenmaal zelf voor de klas staan, zodat het bij

die ene ervaring blijft. Om dit soort redenen pleit Burns (1999) dan ook voor het opzetten en uitvoeren van gezamenlijk actieonderzoek, omdat het nu eenmaal eenvoudiger is om gemeenschappelijke problemen samen te delen en te lijf te gaan. Samen sta je sterker en kun je een breder draagvlak creëren om zonodig zaken te veranderen. En met meer mensen is het simpelweg gemakkelijker om meer evidentie aan te dragen om aan te tonen dat er iets moet veranderen.

Wat is actieonderzoek?

Bij actieonderzoek staat, zoals gezegd, de onderwijspraktijk centraal. De focus van het onderzoek ligt op de natuurlijke (les)situatie: concrete, praktische kwesties uit de praktijk die voor het onderwijs c.q. de les van onmiddellijk belang zijn. Het is onderzoek vóór de docent, dat tegelijkertijd ook dóór de docent wordt uitgevoerd. De docent is bij actieonderzoek zowel regisseur als acteur (Hacquebord 2003).

In de literatuur wordt actieonderzoek aan de hand van verschillende kenmerken getypeerd. Een eerste serie kenmerken betreft de gerichtheid op de concrete lespraktijk. Karakteriseringen als *contextgebonden, kleinschalig* en *lokaal* duiden alle op onderzoek dat zich richt op problemen binnen een specifieke (les)situatie. Een tweede serie kenmerken heeft te maken met de verandering en verbetering van de onderwijspraktijk die het actieonderzoek zich ten doel stelt. Hierbij gaat het om aanduidingen als *evaluerend* en *reflecterend*. In de derde plaats valt bij actieonderzoek vaak het woord *participerend*. Deze term slaat op het belang van samenwerking bij het opzetten en uitvoeren van actieonderzoek zoals dat in de vorige paragraaf werd genoemd. Ten vierde wordt actieonderzoek vaak gekenschetst als *interpreterend*. Daarmee wordt aangegeven dat bij actieonderzoek niet zozeer uitgebreide kwantitatieve, statistische procedures worden ingezet zoals dat bij grootschalig onderzoek vaak het geval is, maar dat veeleer gebruik gemaakt wordt van meer etnografische, kwalitatieve methoden van onderzoek. Verderop, bij de analyse van onderzoeksgegevens, komen we op dit punt terug.

Belangrijk bij actieonderzoek is, zoals in de inleiding werd gesteld, de systematische manier waarop het onderzoek wordt uitgevoerd en gepland. In wat volgt gaan we in op de verschillende onderdelen van actieonderzoek en de volgorde waarin deze moeten worden doorlopen: planning, actie en reflectie. Tijdens de *planningsfase* gaat het om het verkennen van het onderzoeksterrein, het afbakenen van een onderwerp, waarbij een onderzoeksvraag wordt geformuleerd – met daaraan gekoppeld eventueel een hypothese – en er wordt een planning gemaakt. De *actie* bestaat uit het verzamelen van de voor het onderzoek benodigde gegevens en de analyse daarvan, het trekken van een conclusie op basis

van de gevonden resultaten en het vaststellen en invoeren van mogelijke veranderingen en verbeteringen. In de *reflectiefase* wordt verslag uitgebracht over de bevindingen en wordt het onderzoek geëvalueerd. Voor een gedetailleerde beschrijving van deze onderdelen, inclusief planningskaarten, schematische overzichten en observatieformulieren, zie Ponte (2000b). We beginnen nu met het kiezen van een onderzoeksobject en de formulering van een vraagstelling.

Hoe begin je?

Met het bedenken van een geschikt onderwerp en een goede onderzoeksvraag hebben we meteen een van de lastigste onderdelen van actieonderzoek te pakken. Onderzoekers willen vaak te veel, waardoor het onderzoeksobject te groot wordt en niet meer te overzien is. En net als bij ander onderzoek geldt ook voor actieonderzoek dat bij een goede onderzoeksvraag de helft van het antwoord al in zicht is. Maar daarover straks meer. Want voordat het zover is moeten er soms de nodige hindernissen uit de weg worden geruimd voordat het onderzoek überhaupt kan starten. We gaan daarom eerst in op enkele van die belemmeringen en op wat eraan gedaan kan worden om dat soort barrières te slechten.

De obstakels die docenten kunnen tegenkomen op de weg naar actieonderzoek zijn van velerlei aard. Ze kunnen zowel de docenten van buitenaf bedreigen als door henzelf worden opgeworpen. In de laatste categorie betreft het zaken als de tijd die het docenten ontbeert om dit soort onderzoek te doen; de hoeveelheid extra werk die ze zichzelf hiermee op de hals halen; onzekerheid over de eigen onderzoeksvaardigheden; het gebrek aan waardering ondanks het extra werk en de extra tijd die ze erin stoppen: ze krijgen er niet meer salaris door en, in tegenstelling tot universitaire docenten, krijgen ze er geen academische waardering voor. Daar komen de door de buitenwereld opgeworpen barrières dan nog bij: weinig steun van de omgeving die dit soort activiteiten als een bedreiging kan opvatten; collega's die onderzoek doen als iets perifeers of negatiefs zien; een directie die bang is dat negatieve aspecten van het onderwijs naar buiten komen; algeheel scepticisme over het nut van (actie)onderzoek.

Uit ervaring blijkt dat er verschillende strategieën zijn om dit soort hindernissen te overwinnen. In de eerste plaats is het aan te raden om klein te beginnen, om een duidelijke, realistische planning te maken en om daarvoor tijd te claimen bij de directie. Een tweede strategie is om contact te leggen met mensen die hetzelfde (willen) doen, om een groepje op te richten van mensen die in actieonderzoek geïnteresseerd zijn, om elkaars lessen te observeren en te fungeren als zogeheten kritische vrienden. Later kan dit uitgroeien tot een netwerk(je) of het organiseren

van een (incidentele) workshop. In alle gevallen is het belangrijk om een goede organisatie bij te houden: het maken van notulen, het vaststellen van duidelijke afspraken, het bijhouden van een logboek, enzovoort, zodat de gezette stappen traceerbaar zijn en daarvan verantwoording kan worden afgelegd.

Als deze hobbels eenmaal zijn genomen kan het actieonderzoek van start gaan. Zoals gezegd geldt daarbij: begin klein, kies een overzichtelijk onderwerp en formuleer een duidelijke vraagstelling. Meestal werkt het het best om te beginnen met een concreet probleem. Om een dergelijk probleem te definiëren kan gebruik worden gemaakt van de volgende vragen:

- Ik wil graag meer weten van ...
- Het lukt mijn leerders niet om ... Wat kan ik daaraan doen?
- Ik wil graag de manier veranderen waarop mijn leerders ...
- Ik zou in mijn klas meer aandacht willen besteden aan ... Hoe kan ik dat doen?
- Ik zou met mijn collega het volgende wel eens willen uitproberen: ... Wat zou er gebeuren als we dat deden?
- Waarom doen sommige leerders in mijn klas ... en andere ... Hoe kom ik erachter wat hier aan de hand is?

De onderwerpen en vragen die hieruit voortkomen, kunnen over van alles gaan: hoe kan ik de zelfstandigheid van mijn leerders stimuleren, hoe kan ik ervoor zorgen dat mijn leerders buiten de les de nodige input in het Nederlands krijgen, hoe ga ik om met een continue in- en uitstroom van cursisten, hoe breng ik meer differentiatie in mijn les aan, hoe bewerkstellig ik dat de wat stillere leerders meer hun mond open doen, enzovoort, enzovoort. Bedenk daarbij dat er ook onderwerpen zijn waar, zeker op korte termijn, niets aan te doen is. Ga daar gerust aan voorbij.

Overigens is er niets op tegen om ruim de tijd te nemen voor een goede afbakening van het onderzoeksonderwerp en voor het verwoorden van de onderzoeksvra(a)g(en). Hoe specifieker en concreter deze zijn, hoe beter. Galema (2003) beschrijft hoe een docent installatietechniek in het vmbo constateert dat zijn leerlingen het theorieboek niet lezen waardoor zij slechte cijfers halen. Na een gesprek hierover met de leerlingen en een aantal kritische vrienden komt hij erachter dat de leerlingen volledig afknappen op de theorie en in plaats daarvan meer praktijkgericht bezig willen, meer met elkaar willen samenwerken en meer contact willen met klanten en opdrachtgevers. Dit leidt ertoe dat de leerkracht zijn aanvankelijke vraag 'hoe krijg ik mijn leerlingen aan het lezen?' ombuigt naar de specifiekere en realistischer vraagstelling 'hoe kan ik de theorie integreren met de praktijk, zodat de leerlingen minder hoeven te lezen en meer plezier krijgen in de lessen?'

Figuur 1 bevat een schematisch overzicht van de tien stappen die in de planningsfase van het onderzoek kunnen worden doorlopen.

Planning van actieonderzoek in tien stappen

1. Overleg met collega's van de onderwijsinstelling die geïnteresseerd zijn in actieonderzoek.
2. Bespreek de knelpunten en hindernissen en hoe deze overwonnen kunnen worden.
3. Ga na waar weinig of niets aan te doen valt en wat wel mogelijk is.
4. Maak met collega's een lijst van punten die het begin van een onderzoek zouden kunnen zijn.
5. Verken deze gebieden. Observeer de eigen klas een week lang op (een van) die punten en maak de balans op. Welke verdere informatie is nodig en waarop kan gefocust worden?
6. Kom weer met collega's bij elkaar en vergelijk de uitkomsten van de observaties.
7. Focus op de gemeenschappelijke punten die het begin van een onderzoek zouden kunnen vormen.
8. Formuleer zo mogelijk specifiekere vragen.
9. Maak een voorlopig actieplan en bepaal:
   - wie bij het onderzoek betrokken zal zijn;
   - wie er verder geïnformeerd moet worden;
   - welke strategieën met elkaar afgesproken moeten worden;
   - welke bronnen er nodig zijn, zowel materieel als immaterieel.
10. Stel een planning op en neem daarin de volgende informatie mee:
   - de totale duur van het project;
   - een realistisch aantal bijeenkomsten;
   - de tijd die in ieder geval nodig is om tussen twee bijeenkomsten gegevens te kunnen verzamelen;
   - de tijd die nodig is voor rapportage en terugkoppeling.

Figuur 1: Planning van actieonderzoek in tien stappen

Is er eenmaal een onderwerp vastgesteld met een daaraan gekoppelde vraagstelling, dan is het handig om een beschrijving van het onderzoeksproject te maken. In het algemeen bevat een dergelijke projectbeschrijving de volgende elementen:

- een beschrijving van het onderwerp in algemene zin;
- een beschrijving van de lesonderdelen die met het onderwerp te maken hebben en daarmee aan verandering onderhevig zijn;
- een plan hoe dit aangepakt gaat worden;
- een overzicht van alle betrokkenen (onderzoekspartners, supervisoren, collega's, cursisten, etc.);
- een overzicht van de benodigde praktische zaken (materialen, apparatuur, etc.);

- een planningsschema;
- een overzicht van dataverzamelingstechnieken;
- een beschrijving van de manier waarop de verzamelde gegevens verwerkt en geanalyseerd worden;
- een aanduiding hoe er over de resultaten van het onderzoek gerapporteerd gaat worden.

In de volgende paragrafen zullen we achtereenvolgens ingaan op de laatste drie onderdelen van de projectbeschrijving: het verzamelen van de gegevens, data-analyse en rapportage.

Hoe verzamel je de benodigde gegevens?

Het spreekt vanzelf dat het soort gegevens en de manier waarop deze verzameld worden afhangen van de onderzoeksvraag. Omdat het bij actieonderzoek gaat om concrete problemen uit de alledaagse lespraktijk, ligt het voor de hand om voor de dataverzameling praktisch uitvoerbare en haalbare instrumenten in te zetten. Gezien de aard van het onderzoek gaat het daarbij vooral om kwalitatieve, exploratieve technieken. Deze technieken worden vaak onderverdeeld in observationele en niet-observationele technieken.

Observationele technieken worden ingezet om een beeld te krijgen van het (taal)gedrag van een of meerdere leerders. In het algemeen is het bij observatie handig om rekening te houden met de volgende aandachtspunten:
- Bepaal *wat* het doel van de observatie is (interactie, betrokkenheid van de leerders, feedback, etc.); probeer niet alles te doen.
- Bepaal de locatie *waar* geobserveerd gaat worden (de klas, de kantine, etc.).
- Bepaal *wie* er geobserveerd wordt (een of meerdere leerders, de docent, etc.).
- Bepaal of het geobserveerde gedrag meteen ('on time') wordt vastgelegd (bijvoorbeeld met behulp van een observatieschema); doe het anders zo spoedig mogelijk daarna.
- Wees bij het observeren en het scoren van de observaties zo objectief mogelijk.
- Zorg ervoor dat de observatie een representatief beeld vormt van wat de geobserveerden normaal gesproken doen.

Bij het maken van observaties kan van verschillende technieken gebruik worden gemaakt. We noemen er hier een aantal:
- Aantekeningen: de onderzoeker maakt notities van wat relevant is voor het onderwerp van onderzoek.

- Dagboek: de onderzoeker legt dagelijks vast wat van belang is voor het gekozen onderwerp.
- Observatieschema: het geobserveerde gedrag wordt geregistreerd met behulp van een bestaand, aangepast dan wel zelf ontworpen observatiesysteem. Voorbeelden van (bestaande) observatieschema's, die in originele dan wel aangepaste vorm gebruikt kunnen worden zijn het, Engelstalige, *COLT* (Spada & Fröhlich 1995) en de *Kijkwijzer* (van Eerde e.a. 2007) en de *Observatieschema's Leersucces* (Alons, e.a. 2007). Ook in Ponte (2000b) zijn verschillende observatielijsten opgenomen.
- Audio en/of video-opnames: het voordeel van audio- en video-opnames is uiteraard dat ze meerdere keren kunnen worden beluisterd en bekeken. Daarin schuilt echter ook het gevaar, want hoe vaker en preciezer dit gebeurt, hoe tijdrovender deze procedure is. Een tussenweg is vaak om uit het opgenomen materiaal fragmenten te selecteren die representatief zijn voor het totaal.
- Foto's: het gebruik van foto's is een wat onderbelichte techniek in actieonderzoek. Toch is uit foto's veel relevante informatie te vinden over bijvoorbeeld de houding van de leerders (onderuitgezakt of betrokken), het bordgebruik door de docent, de indeling van de klas, etc. En dankzij de digitale camera zijn dit soort gegevens tegenwoordig in een mum van tijd beschikbaar.
- Kaarten: in plaats van foto's kan de indeling van de klas ook met pen en papier in kaart worden gebracht (wie kruipt altijd weg in een hoekje, richt de docent zich tot alle leerders, etc.). Ook sociogrammen (bijvoorbeeld wie heeft tijdens de les contact met wie en hoe vaak) kunnen gemakkelijk op papier worden gezet.

Naast het maken van observaties zijn er verschillende niet-observationele technieken met behulp waarvan gegevens verzameld kunnen worden om de onderzoeksvraag te beantwoorden:
- Interviews: deze kunnen individueel dan wel in groepsverband bij de betrokkenen worden afgenomen. Afhankelijk van de mate van sturing kan de onderzoeker kiezen voor een meer of minder gestructureerd interview, waarbij in het eerste geval alle vragen van tevoren vastliggen, terwijl de interviewer in het tweede geval de vragen mede zal laten afhangen van de verkregen antwoorden.
- Vragenlijsten: het in laten vullen van vragenlijsten is minder tijdrovend dan het afnemen van interviews; daar staat tegenover dat de interviewer bij een onduidelijk antwoord de mogelijkheid heeft om door te vragen. Om die reden worden interviews en vragenlijsten vaak met elkaar gecombineerd. Uiteraard zijn de antwoorden die binnenkomen afhankelijk van de kwaliteit van de vragen en is het altijd raadzaam om een vragenlijst eerst in kleine kring uit te proberen alvorens deze in

groten getale te verspreiden. Ook is het belangrijk om een afweging te maken tussen open en gesloten vragen en vragen die met behulp van een vijf- of zevenpuntsschaal zijn te beantwoorden. De laatste twee categorieën zijn gemakkelijk te scoren, maar zullen geen andere antwoorden opleveren dan die al door de vragensteller zijn bedacht, wat bij open vragen niet het geval is.

- Biografie: leerders kan gevraagd worden om een (taalleer)biografie van zichzelf op te stellen. Aan de hand daarvan kan de onderzoeker bijvoorbeeld nagaan welke gemeenschappelijke thema's daaruit voortkomen en wat (gedeelde) positieve of negatieve ervaringen zijn.
- Documenten: hierbij valt te denken aan het verzamelen van portfolio's, lesplannen, lesmateriaal, testuitslagen, etc. Als alternatief kan leerders gevraagd worden een brief te schrijven, bijvoorbeeld een brief waarin ze aan een beginnende taalleerder advies moeten uitbrengen over het leren van Nederlands als tweede taal. Op die manier laten ze vaak meer van zichzelf zien dan bij het invullen van een 'droog' formulier. Ook het maken van een tekening waarmee ze hun leerproces weergeven kan een dergelijke functie vervullen. Is er op die tekening bijvoorbeeld een plaats ingeruimd voor de docent en waar staat deze? Schijnt de zon er of pakken zich boven de horizon grauwe wolken samen?
- Metafoor: vraag leerders waar het leren van Nederlands als tweede taal mee te vergelijken is. Is het een spannende reis, waarbij steeds nieuwe, interessante plekken worden aangedaan? Is het een nachtmerrie, waaruit de leerder zo snel mogelijk hoopt te ontwaken? Is het een nuttige verplichting, zoiets als dienstplicht, waar je aanvankelijk erg tegen opziet, maar die je later toch niet had willen missen? Het voordeel van dit soort technieken is dat leerders er veel van zichzelf in kwijt kunnen, zonder dat ze het gevoel hebben gedwongen te worden om het achterste puntje van hun tong te laten zien.

Uiteraard is het ook mogelijk om observaties te combineren met een of meer van de bovenstaande niet-observationele technieken.

Hoe analyseer je de verzamelde gegevens?

Bij de analyse van de verzamelde gegevens gaat het om twee zaken: in de eerste plaats om een beschrijving van *wat* de data laten zien; in de tweede plaats om een verklaring *waarom* de data het verkregen beeld opleveren. Om tot een goede beschrijving van de verzamelde gegevens te komen moeten deze worden *geordend*, *gecodeerd* en met elkaar *vergeleken*. Daarna vindt de *interpretatie* van de data plaats om tot een verklaring van de resultaten te komen: wat betekenen deze resultaten, hoe zijn we tot deze resultaten gekomen en waarom zijn dit de resultaten?

We zullen dit illustreren aan de hand van een voorbeeld. Stel docenten willen weten op welke manier zij feedback geven op vragen van hun leerders. Ze hebben daartoe video-opnames van hun lessen gemaakt. Van deze opnames hebben ze het tweede kwartier van elke les uitgeschreven. Dit zijn in feite de ruwe gegevens waarover ze beschikken. De verschillende manieren van feedback worden nu eerst geordend en gecodeerd aan de hand van een observatieschema (met categorieën als: geen feedback, minimale respons, vraag om verduidelijking, begripscontrole, herhaling, het goede model aanreiken, uitbreiding, expliciete afkeuring, etc.). Daarna worden de gegevens met elkaar vergeleken: welke vorm van feedback komt het meest voor, krijgt de ene leerder een ander soort feedback dan de ander, zijn er verschillen tussen docenten onderling, etc. Vervolgens wordt er gezocht naar een verklaring van de gevonden resultaten: hoe komt het dat een docent op de ene leerder zo anders reageert dan op de ander, waarom maakt de ene docent veel vaker gebruik van expliciete afkeuring dan de ander, wat is het effect daarvan bij de leerders, enzovoort? Op basis hiervan kunnen docenten zich bezinnen op de manier waarop zij in het vervolg het beste op reacties van hun leerders kunnen reageren.

In het algemeen moet onderzoek aan twee belangrijke eisen voldoen: betrouwbaarheid en validiteit. Met de eerste eis wordt bedoeld dat de verzamelde gegevens een betrouwbaar beeld moeten geven van het gedrag van degenen die worden onderzocht. Een intelligentietest die op maandag een score van 120 oplevert en op dinsdag een score van 100 is niet betrouwbaar te noemen. Bij validiteit gaat het erom of onderzoekers inderdaad datgene meten wat ze willen meten. Wie de luistervaardigheid van NT2-leerders wil toetsen door hen naar de fileberichten te laten luisteren moet zich afvragen of op deze manier niet eerder het geheugen en de geografische kennis van de leerders getoetst worden dan hun luistervaardigheid. Bij psychologische toetsen (zoals de intelligentietest) en bij medische testen worden zeer hoge eisen ten aanzien van betrouwbaarheid en validiteit gesteld. En dat is maar goed ook. Bij actieonderzoek kan meestal niet aan die zeer strenge eisen worden voldaan, wat voor sommigen reden van kritiek is. De vraag is echter of bij actieonderzoek wel dezelfde strenge eisen in acht moeten worden genomen, wat uiteraard niet wegneemt dat bij elk onderzoek een zo hoog mogelijke kwaliteit moet worden nagestreefd. Om ook bij actieonderzoek toch een zo hoog mogelijke betrouwbaarheid en validiteit te garanderen kan één observatie door één onderzoeker bij één groep leerders een te beperkt beeld opleveren. Daarom is het veiliger om het onderzoeksobject van verschillende kanten en op verschillende manieren te benaderen met betrekking tot:

- de methode: verzamel data op verschillende manieren;

- het tijdstip: verzamel data op verschillende momenten;
- de plaats: verzamel data bij verschillende groepen, onder verschillende omstandigheden;
- de onderzoeker: laat de data verzamelen door verschillende onderzoekers;

de theorie, het analysekader: analyseer de data vanuit verschillende perspectieven.

Hoe rapporteer je over het onderzoek?

Als actieonderzoek gedefinieerd kan worden als onderzoek voor en door docenten en als het onderzoek zich vooral richt op een concreet probleem uit de eigen lespraktijk met als doel de eigen leerkrachtvaardigheden te vergroten en/of de prestaties van de leerders te verbeteren, dan is het de vraag of rapportage van het onderzoek wel zin heeft, omdat de docenten het toch vooral voor zichzelf en de eigen leerders doen. Toch is het om verschillende redenen verstandig om verslag van het onderzoek uit te brengen. In de eerste plaats als verantwoording naar directe collega's, directie, cursisten en andere betrokkenen toe om te laten zien hoe de onderzoekstijd is besteed, wat het onderzoek heeft ingehouden en wat dat heeft opgeleverd. In de tweede plaats naar docent-onderzoekers in den lande toe. Misschien niet eens zozeer vanwege de resultaten, als wel door de manier waarop het onderzoek is aangepakt, waar anderen bij hun eigen onderzoekgerichte benadering van het onderwijs wellicht hun voordeel mee kunnen doen. Verder hoeft er bij rapportage van onderzoek niet meteen aan een lijvig onderzoeksrapport gedacht te worden, maar zijn er ook allerlei alternatieve rapportages mogelijk. Hieronder zetten we een aantal mogelijkheden op een rij, zowel wat betreft schriftelijke als mondelinge verslaglegging.

Schriftelijke rapportage:
- Verslag: traditionele rapportage, waarbij de tekst voor een breed publiek is bestemd (externe rapportage) of de auteurs vooral leden van de eigen instelling (intern rapport) als doelgroep voor ogen hebben (bijvoorbeeld Hofstetter, Schaareman & de Vries 2005).
- Artikel in een tijdschrift: bijvoorbeeld in een Nederlandstalig tijdschrift als *Les, Levende Talen Tijdschrift* of *Toegepaste Taalwetenschap in Artikelen*. Hier gaat het dus om externe rapportage (bijvoorbeeld Kroon 2007).
- Update in een nieuwsbrief, personeelsbulletin, etc.: voorbeelden van interne rapportage, waarbij niet perse gewacht hoeft te worden op een afgerond eindverslag, omdat dit soort media zich ook bij uitstek lenen voor tussentijdse rapportage om te laten zien wat de stand van zaken van het onderzoek is.

- Portret: een meer alternatieve vorm van verslaglegging, waarbij een beeld wordt geschetst van een saillante (groep) leerder(s) die representatief is voor het gehele onderzoek (bijvoorbeeld Kuiken 2006). Overigens geven sommige tijdschriften soms ook de voorkeur aan een interview met een onderzoeker boven een uitgewerkt onderzoeksverslag (bijvoorbeeld Galema 2003, van der Hagen 2005, van der Mee 2003).

Mondelinge rapportage
- Presentatie: de klassieke lezing, individueel of *en groupe* (met het hele onderzoeksteam), al dan niet interactief, voor een select publiek (intern, eigen collega's) dan wel openbaar (tijdens een studiedag of conferentie, bij voorbeeld van de BVNT2).
- Video, foto's of poster: als alternatief voor het traditionele praatje kan een video, fotoserie of poster worden samengesteld, omdat beelden vaak meer tot de verbeelding spreken dan woorden. Uiteraard kunnen de twee (presentatie en video/foto's; poster met praatje) heel goed met elkaar gecombineerd worden.

Een klassiek onderzoeksverslag bevat meestal de volgende onderdelen: 1. Inleiding; 2. Probleemstelling; 3. Manier van aanpak; 4. Resultaten; 5. Conclusies. In de inleiding wordt op de achtergrond van het onderzoek ingegaan: waarom is het uitgevoerd en wat kan de lezer in het verslag verwachten? In de tweede sectie worden de probleemstelling en de onderzoeksvra(a)g(en) geformuleerd: wat is het probleem, wat was daar al over bekend (op basis van literatuur en eigen ervaringen of die van anderen) en welke vraag wordt beantwoord? Bij de manier van aanpak wordt uitgelegd hoe de beantwoording van de onderzoeksvraag is aangepakt, wie er aan het onderzoek meededen en hoe de gegevens zijn verzameld. In de sectie over de resultaten wordt gerapporteerd wat de analyse van de gegevens heeft opgeleverd. Tot slot wordt samengevat wat er nu bekend is over het probleem, wat op basis van de resultaten anders moet en wat er eventueel verder nog onderzocht moet worden, kortom conclusies, discussie en aanbevelingen. Maar zoals gezegd, het is ook zeker mogelijk om, afhankelijk van het publiek, de resultaten van actieonderzoek op een andere, alternatieve manier te presenteren.

Tot slot

Het verrichten van eigen onderzoek is een spannende bezigheid. Voor veel docenten die aan actieonderzoek zijn begonnen smaakt het naar meer. Vooral als het enthousiaste reacties oplevert bij collega's en cursisten. En dat is dikwijls het geval. Ook hier geldt maar al te vaak dat als er één schaap over de dam is, er spoedig meer volgen. Als collega-

docenten die eerst angstvallig de kat uit de boom hebben gekeken zien hoe lonend onderzoek kan zijn, zijn ze vaak bereid om daar een volgende keer ook aan mee te doen. Dat kan het moment zijn om een docentennetwerk ('taalwerkplaats' of welke andere benaming men daar ook aan wil geven) in het leven te roepen. Als dat eenmaal loopt kunnen er volgende stappen worden gezet: het uitvoeren van onderzoek opnemen in een deskundigheidsbevorderingstraject, contact leggen met mogelijke onderzoekspartners (van zusterinstellingen, hbo of universiteit), of het koppelen van actieonderzoek aan andere onderwijsvernieuwingen. Maar of actieonderzoek nu op kleine of grote schaal wordt verricht, het draagt in alle gevallen op z'n minst bij tot een betere reflectie op de eigen onderwijspraktijk en in de meeste gevallen tot vergroting van de eigen leerkrachtvaardigheden en tot verbetering van de cursistenprestaties.

Literatuur

Burns, A. (1999) *Collaborative action research for English language teachers*. Cambridge: Cambridge University Press.

Eerde, D. van, H. Hacquebord, M. Hajer, M. Pulles, C. Raymakers (2007) *Kijkwijzer voor taalgericht vakonderwijs*. Enschede: SLO/Platform Taalgericht Vakonderwijs.

Farrell, T.S.C. (2007) *Reflective language teaching. From research to practice*. London: Continuum.

Galema, K. (2003) Je eigen praktijk onderzoeken. Actieonderzoek op het Montessori-College Oost in Amsterdam. *Les* 124, 4-7.

Hacquebord, H. (2003) Ontwikkelingsonderzoek in het talenonderwijs. Oude traditie, veelbelovend concept? *Levende Talen Tijdschrift* 4, 2, 18-23.

Hagen, M. van der (2005) Kritische vrienden nodig bij actieonderzoek. *12-18* januari 2005, 28-29.

Hofstetter, T., Schaareman, E. & Vries, S. de (2005) *Vak en Taal. Actieonderzoek H4A 2004/2005.* Interne publicatie. Groningen: Reitdiep College/Kamerlingh Onnes.

Kroon, H. (2007) Actieonderzoek maakt veranderingen duurzaam. *12-18* juni 2007, 30-33.

Kuiken, F. (2006) *Tot op het bot gemotiveerd. Vier portretten van Amsterdamse vo-scholen bij de ontwikkeling en invoering van taalbeleid.* Amsterdam: Dienst Maatschappelijke Ontwikkeling, Gemeente Amsterdam.

Litjens, P. (1990) *Nederlands als tweede taal. Een handboek voor de volwasseneneducatie*. Amersfoort/Leuven: Acco.

Mee, G. van der (2003) 'Het gaat altijd over je eigen probleem.' Actieonderzoek moet lespraktijk veranderen. *Het Onderwijsblad* 21, 12-15.

Ponte, P. (2002a) *Actieonderzoek door docenten. Uitvoering en begeleiding in theorie en praktijk.* Dissertatie. Leuven/Apeldoorn: Garant.
Ponte, P. (2002b) *Onderwijs van eigen makelij. Procesboek actieonderzoek in scholen en opleidingen.* Soest: Nelissen.
Spada, N. & Fröhlich, M. (1995) *The Communicative Orientation of Language Teaching Observation Scheme (COLT).* Sydney: NCELTR Publications.
SpeZiaal (2008) Tips voor de praktijk. Professionalisering moet je doen! *SpeZiaal* 3, 2, 21-24.
Wallace, M.J. (1991) *Training foreign language teachers. A reflective approach.* Cambridge: Cambridge University Press.
Wallace, M.J. (1998) *Action research for language teachers.* Cambridge: Cambridge University Press.

ZWAKKE TEKSTEN VOOR "ZWAKKE LEZERS"

Jentine Land, Huub van den Bergh & Ted Sanders

Universiteit van Utrecht, UIL-OTS / Universiteit van Amsterdam, ILO

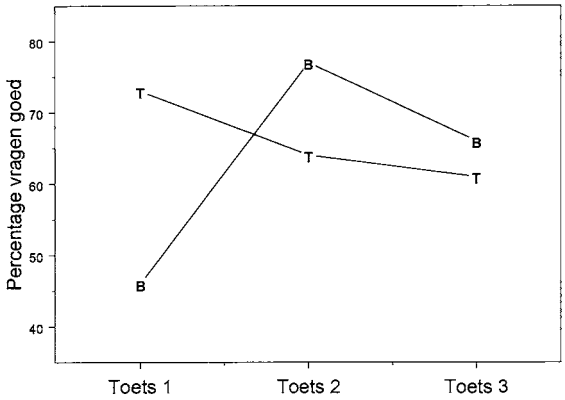
Inleiding

Leerlingen uit het VMBO, met name de leerlingen uit de basisberoeps- en of kaderberoepsgerichte leerweg zijn over het algemeen slechte lezers. Zij lezen weinig in hun vrije tijd, vinden lezen niet leuk en zijn vaak zwakkere lezers. Maar op school moeten ze lezen, en de maatschappij eist als het ware dat ze ook goed leren lezen. We kunnen dan ook de vraag stellen of het hen helpt als teksten gemakkelijk te begrijpen en leuk om te lezen zijn. De vraag is hoe zien teksten eruit die gemakkelijk te begrijpen en leuk om te lezen zijn. Daar gaan we in deze bijdrage verder op in. Eerst nog even de kwestie van de zwakke lezer.

Wanneer zijn lezers zwakke lezers?

Een uitspraak als VMBO-ers zijn zwakke lezers, is gemakkelijk gedaan en het lijkt wel alsof iedereen het daarmee eens is. Toch moet een dergelijke uitspraak meteen genuanceerd worden.

In een onderzoek zijn bij leerlingen uit de theoretische leerweg (voorheen mavo) en leerlingen uit een basisberoepsgerichte leerweg (voorheen lbo) drie tekstbegriptoetsen afgenomen (Van den Bergh, 2006). Volgens de gangbare gedachtegang kunnen leerlingen in de theoretische leerweg beter lezen, en behalen ze hogere tekstbegripscores. Voor de eerste tekstbegriptoets in dit onderzoek klopt dat (zie Figuur 1), daar halen de 'theoretici' betere prestaties dan de 'beroepsgerichten'. Maar op de tweede en derde toets is het verschil tussen beide groepen leerlingen omgedraaid. Daar halen juist de leerlingen in de beroepsgerichte leerweg (gemiddeld) hogere scores tekstbegrip (zie Figuur 1). Is dat niet vreemd? Nou, eigenlijk niet.



Figuur 1. Gemiddeld aantal goed beantwoorde vragen op drie tekstbegriptoetsen voor leerlingen uit de theoretische (T) en basisberoepsgerichte (B) leerweg.

De resultaten uit dit onderzoek laten zich goed verklaren. Twee belangrijke factoren spelen een rol. Ten eerste zijn de leerlingen uit de basisberoepsgerichte leerweg 'koks-in-opleiding'. Ten tweede zijn de drie tekstbegriptoetsen geheel verschillend van aard. De tweede tekst gaat over het bereiden van een gerecht en de derde toets is een recept, terwijl de eerste toets een algemene tekstbegriptoets is. Dit onderzoek laat dus zien dat er nogal wat afhangt van de tekst waarover vragen gesteld worden; niemand zal er gek van opkijken als koks (in opleiding) erg goed zijn in het lezen van recepten; tekstbegrip wordt beïnvloed door de achtergrondkennis van de lezer.

Deze kanttekening, over de relatie tussen tekstbegrip en achtergrondkennis (qua werkdomein en qua tekstsoort) van de lezer, neemt echter niet weg dat bepaalde kenmerken van teksten van invloed kunnen zijn op het begrip van die tekst; de hoofdgedachte in een slecht gestructureerde tekst is bijvoorbeeld minder makkelijk te achterhalen dan in een goed gestructureerde tekst. In dit artikel willen we verder ingaan op twee kenmerken van teksten: structuurkenmerken en stijlkenmerken.

Structuurkenmerken van teksten: Fragmentatie versus Coherentie

Over de eisen waaraan een tekst moet voldoen, waarmee we wellicht de leerlingen in het VMBO tot lezen kunnen verleiden, bestaan verschillende

ideeën. Zo stond op een website van een educatieve uitgever in 2003 de volgende tekst: 'Sla een willekeurige pagina op en het zal u direct opvallen: in *[leergang xxx]* staan geen lange zinnen of moeilijke woorden. Het taalgebruik is voor 100% afgestemd op de VMBO-leerling. Zelfs de meest complexe eindtermen van de overheid zijn in *[leergang xxx]* omgezet in klare taal'. Klaarblijkelijk is het mogelijk, althans volgens deze uitgever, om een tekst zo te formuleren, en zo weer te geven, dat iedereen deze begrijpt.

Er zijn verschillende theoretische noties om te bepalen of een tekst gemakkelijk en/of begrijpelijk is. De meest bekende is zonder enige twijfel het gebruik van zogenaamde leesbaarheidsformules. Aan de hand van tellingen van eenvoudige tekstkenmerken, zoals zinslengte, aantal lettergrepen per woord, etc., wordt dan een oordeel gegeven over de moeilijkheid van een tekst. Ook worden teksten vaak herschreven, opdat zij makkelijker leesbaar worden. Hiervoor geven zulke leesbaarheidsformules duidelijke handvatten. Zoals: Maak de zinnen korter. De teksten worden dan gemakkelijker. De teksten zijn beter te begrijpen. En: Vermijd lange woorden. De teksten worden dan gemakkelijker. Dus komen er teksten met korte zinnen en makkelijke woorden. In schoolboeken voor het VMBO komen we zulke teksten vaak tegen. Het verband tussen zinnen wordt in dergelijke teksten vaak in het midden gelaten (zie Figuur 2).

Steden worden belangrijker
In het begin waren de heren nog de baas in de steden.
Stedelingen moesten allerlei werk doen.
Dit was voor de kooplieden en ambachtlieden lastig.
Ze vroegen aan hun heer stadsrechten.
Een stad met stadsrechten mocht zichzelf besturen.
...

Figuur 2. Een fragment uit een geschiedenisboek voor het VMBO.

Als resultaat krijgen we staccato-teksten, waarin elke zin op een nieuwe regel begint. Inbedding is taboe, en bijzinnen komen niet voor. We noemen dit *fragmentatie*. Het idee dat hieraan ten grondslag ligt is dat kortere zinnen minder belastend zijn voor het werkgeheugen, en dat lezen daardoor minder inspannend is en het dus leuker is om te lezen. Kortom, fragmentatie helpt. Dergelijke teksten komen regelmatig voor in leermethoden, al verschilt de mate van fragmentatie zowel van vak tot vak als van leerweg tot leerweg. Zo komen bij de zaakvakken en economie vaker gefragmenteerde teksten voor dan bij wiskunde, en komen

gefragmenteerde teksten vaker voor in methodes voor de basisberoepsgerichte leerweg (zie Land, 2008).

Een geheel ander uitgangspunt is een tekst waarin de *coherentie* centraal staat; in een coherente tekst wordt de samenhang tussen de zinnen juist benadrukt. Deze moet makkelijk af te leiden zijn; coherentiesignalen (*omdat, daardoor, ten eerste ...*) lijken het leesproces te versnellen en lijken te leiden tot een beter begrip van de tekst, althans in laboratorium situaties. Door de heldere structuur van de tekst en de hulp van talige signalen moet het voor de lezer makkelijk zijn om een tekst te begrijpen. De lezer heeft dan minder moeite om een mentale representatie te maken van de tekst in kwestie. Dit leidt tot een beter begrip van de stof die in de tekst aan bod komt. Dit gemak zal dus door de lezer gewaardeerd moeten worden.

In het voorafgaande zijn twee tegengestelde intuïties over de begrijpelijkheid van teksten besproken. Maar, let wel, voor beide uitgangspunten is zowel theoretische als empirische evidentie voor handen. Natuurlijk roept dit om een experiment, waarin nagegaan kan worden welke van de twee hypothesen juist is. Hiervoor moeten we van een tekst twee versies maken: één die past bij de gefragmenteerde hypothese, en één die past bij de coherentiehypothese (zie Figuur 3).

Gefragmenteerde tekst	Coherente tekst
Slavernij in Amerika	**Slavernij in Amerika**
In de 19e eeuw waren er veel plantages in de zuidelijke staten van Amerika.	In de 19e eeuw waren er veel plantages in de zuidelijke staten van Amerika, maar er waren te weinig arbeiders. Daarom lieten de Amerikanen slaven uit Afrika voor hen werken. De slaven wilden graag naar Amerika komen, want zo konden ze geld verdienen voor hun familie. Toen ze eenmaal in het zuiden van Amerika werkten, was het echter niet zo fijn als ze dachten. De slaven hadden namelijk helemaal geen rechten. Daardoor kregen ze geen geld voor het werk. Ook konden de bazen hen ongestraft mishandelen (...)
Er waren te weinig arbeiders.	
De Amerikanen lieten slaven uit Afrika voor hen werken.	
De slaven wilden graag naar Amerika komen.	
Ze konden geld verdienen voor hun familie.	
Het werk in het zuiden van Amerika was niet zo fijn als ze dachten.	
De slaven hadden helemaal geen rechten.	
Ze kregen geen geld voor het werk.	
De bazen konden hen ongestraft mishandelen (...)	

Figuur 3. Een gefragmenteerde en een coherente versie van één tekst (overgenomen uit Land, 2008).

Als de fragmentatiehypothese juist is, zullen leerlingen gefragmenteerde teksten (zie Figuur 3) beter begrijpen en hoger waarderen. Als de coherentiehypothese juist is geven leerlingen de voorkeur aan de coherente tekst en begrijpen zij juist deze beter.

Bij een dergelijk experiment spelen tenminste twee problemen waar rekening mee gehouden moet worden. Ten eerste, als aangetoond kan worden dat er bij één tekst een verschil is tussen de beide versies, zegt dat nog niks over (fragmentatie versus coherentie bij) een andere tekst. Er moet aangetoond worden dat bij een groot aantal teksten een systematisch verschil is tussen de beide versies. Vandaar dat er in ons onderzoek voor acht verschillende teksten een gefragmenteerde èn een coherente versie geconstrueerd zijn.

Ten tweede, hoe wordt gemeten of een tekst begrepen is? Zijn de standaard meerkeuzevragen echt de beste manier om tekstbegrip te meten? In dit onderzoek zijn drie manieren om tekstbegrip te meten gecombineerd: meerkeuzevragen, tijdbalkvragen en schemavragen. Bij tijdbalkvragen moeten de leerlingen gegeven tekstelementen in de juiste volgorde zetten. En, bij schemavragen moeten de leerlingen de verbanden uit de tekst reproduceren of woorden uit de tekst op de meest logische manier rangschikken (zie Kamalski, Sanders, Lentz & Van den Bergh, 2005; Land 2008).

De beide versies van de acht teksten zijn voorgelegd aan 582 leerlingen van alle niveaus in het VMBO (LWOO, basis, kader en theoretische leerweg). Naast het tekstbegrip is ook hun waardering van de teksten bevraagd. De belangrijkste resultaten kunnen als volgt samengevat worden:

1. Coherentie werkt beter dan Fragmentatie (± 20% meer vragen correct);
2. Coherentie werkt beter bij zowel leerlingen in de beroepsgerichte leerweg als leerlingen in de theoretische leerweg;
3. Het verschil tussen beide versies is kleiner bij leerlingen uit de theoretische leerlingen dan bij basisberoepsgerichte leerlingen. Coherentie is vooral nuttig voor lagere opleidingsniveaus;
4. Er is een hoofdeffect van niveau: leerlingen uit de theoretische leerweg behalen hogere tekstbegripscores;
5. Er is geen effect op waardering van de tekst; gefragmenteerde en coherente versies van een tekst worden in gelijke mate gewaardeerd.

Op grond van deze resultaten kan geconcludeerd worden dat kenmerken die de structuur van een tekst bevorderen een positieve invloed hebben op het tekstbegrip van VMBO-leerlingen; optimaliseren van de coherentie

van teksten blijkt meer invloed te hebben dan het minimaliseren van de cognitieve belasting door fragmentatie. Het is dan op z'n zachtst gezegd opmerkelijk dat de leerboeken, en met name die voor de lagere niveaus, aan fragmentatie meer belang gehecht wordt. Een mismatch!

Er moet echter wel (tenminste) één kanttekening gemaakt worden: de teksten in dit onderzoek zijn wel erg duidelijk gemanipuleerd. Het is goed ook na te gaan of een minder extreme manipulatie even veel invloed heeft op het begrip van de tekst.

Stijlkenmerken: Identificatie versus distantie

In d Teksten onderscheiden zich natuurlijk niet alleen qua mate van coherentie. Een andere dimensie waarop teksten vaak van elkaar verschillen is de mate waarin een lezer zich met de tekst kan identificeren. Met sommige teksten kan een lezer zich heel gemakkelijk identificeren, terwijl dit bij andere teksten veel moeilijker is. Kijk maar eens naar het voorbeeld in Figuur 4. Deze beide teksten komen uit een geschiedenisleergang voor hetzelfde jaar (voor dezelfde leerweg) van het VMBO.

Een 'identificerende' tekst	Een 'distantiërende' tekst
Boris ziet een poster van Stalin. In 1950 is Stalin aan de macht. Hij is de leider van Rusland. Stalin is een strenge heerser. Het volk heeft hem niet gekozen. Hij greep zelf de macht. Stalin is voor het communisme. Wie daar tegen is, pakt hij op. Die mensen worden vermoord. Ook vrienden van Boris zijn vermoord. Dat durft hij niet te zeggen. Anders pakken ze hem ook op.	De communisten sluiten na de revolutie in 1917 onmiddellijk vrede met Duitsland. Daarna willen zij hun andere plannen uitvoeren (bron 3). Overal in het land ontstaan sovjets. Een sovjet is een raad van boeren of arbeiders. Zo'n sovjet beslist over het bestuur. Maar lang niet alle Russen zijn het eens met de communisten. Een felle burgeroorlog breekt uit tussen voor en tegenstanders. Tijdens deze strijd wordt de tsaar vermoord. In 1921 is de burgeroorlog voorbij. De communisten hebben gewonnen. Rusland heet voortaan de Sovjet-Unie.

Figuur 4. Voorbeelden van teksten uit verschillende geschiedenis leergangen (overgenomen uit Land, 2008).

Beide typen teksten, aan de ene kant de meer verhalende en identificerende teksten, en aan de andere kant de meer distantiërende teksten, komen in de leerboeken voor het VMBO voor. Het eerste type tekst wordt gekenmerkt doordat er een personage in voorkomt (bij voorbeeld: *ik vind luizen geen probleem*). Een flink deel van de tekst is in de onvoltooid tegenwoordige tijd weergegeven, en er wordt vaak in de directe rede 'gesproken'. Bij meer distantiërende teksten is er een vertellersinstantie (bij voorbeeld: *Koning Lodewijk XIV vond luizen geen probleem*). De tekst is vaak in de onvoltooid verleden tijd, er komen geen emoties naar voren, en er worden relatief veel jaartallen genoemd.

Net zoals bij de structuurkenmerken zijn er bij deze stijlkenmerken twee rivaliserende hypothesen. Volgens de eerste hypothese versterkt identificatie de betrokkenheid van de lezer. En, omdat de afstand tussen lezer en schrijver kleiner wordt, zijn redeneringen en keuzes makkelijker te begrijpen. Volgens de distantiërende hypothese leidt identificatie alleen maar af. Hierdoor is er te veel aandacht voor bijzaken, wat de tekst alleen maar moeilijker maakt.

Ook voor de toetsing van deze beide hypothesen is een experiment opgezet waarin twee versies van een gemanipuleerde tekst vergeleken worden wat betreft begrijpelijkheid en waardering (zie Figuur 5).

Een 'distantiërende' tekst	Een 'identificerende' tekst
In de 18de eeuw werkten veel kinderen in een grote fabriek. De kinderen vonden twee dingen verschrikkelijk: de stank en het lawaai. De stank werd veroorzaakt door de wol waarmee de kinderen moesten werken. En de oude machines zorgden voor het lawaai. De kinderen moesten 14 uur per dag werken. Er werd streng gecontroleerd of ze wel hard genoeg werkten. Want anders zouden de machines stil staan en zou er te weinig verdiend worden.	Stefan is 10 jaar en hij werkt in een grote fabriek in de 18de eeuw. Hij staat achter de weefmachine. Hij vindt twee dingen verschrikkelijk: de stank en het lawaai. 'Waarom stinkt het hier zo?' vraagt hij. 'Dat komt door de wol waarmee we werken,' legt zijn oude, kromme buurman vriendelijk uit. Stefan wil nog meer weten. 'Waarom is er dan zoveel lawaai?' 'Omdat de machines heel oud zijn,' hoort Stefan een man met lang haar zeggen. 'Hoeveel uur moet jij werken?' '14 uur,' wil Stefan antwoorden, maar ...

Figuur 5. Een voorbeeld van een 'distantiërende' en een 'identificerende' tekst (overgenomen uit Land, 2008).

Teksten zoals die in Figuur 5 zijn voorgelegd aan 561 leerlingen uit de tweede klas van het VMBO. Alle leerwegen waren hierbij vertegenwoordigd. Tekstbegrip en tekstwaardering van de leerlingen is op identieke wijze gemeten als in het experiment naar de effecten van structuurkenmerken.

Of de manipulatie effectief geweest is, moet onder andere blijken uit de antwoorden van de leerlingen op de waarderingsvragen. Gelukkig bleken de leerlingen de teksten in de 'identificerende' conditie iets meer te waarderen, en iets spannender te vinden dan de (vergelijkbare) teksten in de 'distantiërende' conditie.

Uit de statistische analyse blijkt dat leerlingen distantiërende teksten iets beter begrijpen dan identificerende teksten. Hierbij moet wel opgemerkt worden dat het verschil tussen beide condities afhankelijk is van de wijze waarop het tekstbegrip gemeten wordt; bij meerkeuzevragen en tijdbalkvragen is wel, maar bij de schemavragen is geen significant verschil tussen distantiërende en identificerende teksten (zie Land (2008) voor een nadere analyse van effectieve en minder effectieve identificerende/distantiërende kenmerken).

We kunnen dan ook concluderen dat identificerende teksten (iets) meer waardering van de leerlingen krijgen, maar minder goed begrepen worden.

Conclusies

Samenvattend kan gesteld worden dat er zowel qua structuurkenmerken als qua stijlkenmerken tegenstrijdige intuïties bestaan. De consequentie hiervan is dat niet alle teksten, ook die in de leergangen, voor vmbo-leerlingen duidelijk zijn. Belangrijk is dat vmbo-ers serieus genomen worden, en dat (juist) zij goede teksten verdienen; teksten die qua structuur en qua stijl geoptimaliseerd zijn. En dat zal niet alleen opgaan voor vmbo-ers, maar ook voor mbo-leerlingen. Duidelijk is in elk geval dat de complexiteit niet wordt verminderd door alleen maar korte zinnen met korte woorden te gebruiken.

De inzet van dergelijk onderzoek moet zijn: goede teksten voor vmbo-ers. Te vaak is er nu sprake van een mismatch tussen tekstkenmerken en de mate waarin leerlingen daarvan profiteren; zwakke teksten worden geproduceerd voor zwakke lezers. Dat bij dergelijke teksten gekapitaliseerd moet worden op dat wat leerlingen al weten en/of kunnen staat als een paal boven water.

Eén en ander neemt niet weg dat we in wetenschappelijke zin slechts het topje van de spreekwoordelijke ijsberg zichtbaar hebben gemaakt. We hebben er bij voorbeeld nog geen idee van in hoeverre structuur- en stijlverschillen zich al tijdens het lezen van een tekst een rol

spelen. En, in hoeverre verschillen in het leesproces bepalend zijn voor verschillen in tekstbegrip. Onderzoek op dit terrein is dringend gewenst.

Literatuur

Bergh, H. van den (2006). *Zeker weten door zuiver meten?* Oratie Universiteit Utrecht.

Land, J.F.H. (2009). *Zwakke lezers, sterke teksten.* Effecten van tekst- en lezerskenmerken op het tekstbegrip en de tekstwaardering van vmbo-leerlingen. Stichting Lezen Reeks 13. Eburon, Delft.

Kamalski, J., Sanders, T., Lentz, L. & Bergh, H. van den (2005). *Hoe kun je het beste meten of een leerling een tekst begrijpt?* Een vergelijkend onderzoek naar vier methoden. *Levende Talen Tijdschrift*, *4*, 3-9.

EEN ROLLENMODEL VOOR PROCESFEEDBACK

Over wie wat kan doen en vooral wanneer

JAN STRYBOL

PCVO Het Perspectief Gent

Inleiding

Feedback is in vele lessituaties een soort ritueel geworden met vaste ingrediënten. In dit artikel stellen we dat ritueel in vraag door feedback te herbekijken vanuit drie eenvoudige vragen: wat, wanneer en wie. We beperken ons hier vooral tot feedback op schrijven maar veel van de beschouwingen gelden eigenlijk ook voor spreken.

Wat: Er wordt in onze klassen voornamelijk aandacht besteed aan vormcorrectie. Reeds jaren zijn de meningen verdeeld over de zin hiervan. Vooral bij feedback op schrijven gaat daarom veel aandacht naar het verbeteren van de methode van feedback zelf (b.v. correctiemodellen met symbolen). Dat is een stap vooruit, maar de focus blijft op het aantonen van de fouten (m.a.w. op het besef). Er zijn echter veel andere factoren die meespelen in de leerevolutie van de cursist (b.v. motivatie, attitude, ontkenning, angst) en daardoor is de efficiëntie zeker nog niet optimaal. Inhoudelijke feedback of motiverende woorden kunnen in bepaalde situaties veel meer teweegbrengen dan de gangbare vormcorrectie.

Wanneer: Feedback is vooral gericht op het eindproduct, terwijl procesfeedback veel meer zin heeft. Met proces bedoelen we twee zaken:

- het taalleerproces: iedere leerder zit in een fase van het leerproces en de feedback moet daaraan aangepast zijn
- het productieproces (schrijven en spreken): het is veel logischer om tijdens de productiefase zelf feedback te geven in plaats van achteraf

Wie: Feedback wordt meestal enkel door de docent zelf gedaan, maar in feite kan je het ook vanuit een rollenmodel bekijken. Deze rollen kunnen ingevuld worden door de docent, de cursist zelf, medecursisten, moedertaalsprekers, woordenboeken, de computer, etc. Elk van deze invullingen heeft zijn beperkingn maar ook zijn sterktes. Zaak is daarin

een juiste keuze te maken en zo te komen tot een feedbackmethode die veel rijker is dan de eenrichtingsfeedback vanuit de docent.

Wat?

De uitgangspositie van dit artikel is dat feedback vandaag een soort aanvaard ritueel is geworden. Wat het schrijven betreft, komt het nog vaak hierop neer dat de cursist zijn taak afwerkt en het dan aan de docent geeft. Die voert vervolgens zijn deel van het contract uit, namelijk correctie van de fouten met in het beste geval advies over de tekstopbouw of de duidelijkheid en tenslotte een waardeoordeel (b.v. een cijfer). De cursist krijgt het document terug en zal vaststellen wat er fout is gegaan plus eventueel de fouten verbeteren (omdat dat zo gevraagd wordt). Beide partijen zijn 'tevreden' want ieder heeft zijn taak van het contract volbracht. Maar de vraag is eigenlijk: heeft de cursist voldoende geleerd daaruit? Is het een tevredenheid omdat het ritueel is afgehandeld, of om wat de feedback heeft aangeleerd? Met andere woorden: is dit wel een goed contract?

In de ESL-literatuur (Engels als tweede taal) is daar al jarenlang een polemiek over met talloze studies die onderzoeken of/wanneer/hoe het nu wel of niet zin heeft de leerders expliciet te wijzen op fouten. Truscott (1) stelde onomwonden dat deze vorm van feedback geen enkele zin heeft. Meer nog: dat het een negatieve werking heeft en de docent zijn energie beter aan andere dingen kan besteden. Het verwerven van de juiste vormen volgt een eigen pad (dat we niet kunnen sturen als docent) en het belangrijkste is voldoende input (lezen en luisteren) te geven gecombineerd met herhaalde output (spreken en schrijven).

Daar kunnen we aan toevoegen dat vormfeedback bij de cursist een overdreven focus teweegbrengt op de vormuiting (en dit nog vaak op het verkeerde moment). Dat kan negatieve gevolgen hebben op de motivatie en de spreek/schrijfangst vergroten. Vaak is het dus gewoon niet de best mogelijke feedback van dat moment: inhoudelijke feedback of motiverende woorden kunnen bijvoorbeeld veel meer teweegbrengen.

Twee cursisten van mij, Bebe en Jovanka, hadden allebei het gevoel dat ze zich in het Nederlands niet even mooi konden uitdrukken als in hun eigen taal. De teksten die ze schreven (zie de Bijlage bij dit artikel) zijn echter wel heel mooi en op dat moment is het veel belangrijker om hen daarop te wijzen en hen dus te stimuleren om zich nog meer uit te drukken in het Nederlands. Dat ze daarbij fouten maken is normaal, maar dat is nu niet aan de orde. Vormcorrecties zouden ook afbreuk doen aan de stimulerende boodschap. Als zij hier in de eerste plaats te horen

krijgen dat ze echt iets waardevols hebben gemaakt, zal dat op de lange termijn veel meer effect hebben. Bovendien is het zelfs goed mogelijk dat je daardoor een intrinsieke motivatie opwekt bij de cursisten om hun tekst alsnog foutloos te maken (want de tekst heeft aan waarde gewonnen).

Een andere illustratie is de anekdote van 'de Amerikaanse manager' die aan een van zijn teams vraagt om een rapport te maken. Tot drie keer toe neemt hij het verslag mee naar huis en zegt de volgende dag: 'Is dat alles wat jullie kunnen'? Als het team uiteindelijk zegt dat ze echt niet beter kunnen, antwoordt hij: 'Ok, dan ga ik het eens lezen'. Ook onze cursisten kunnen vaak beter presteren dan ze eigenlijk doen, maar dan moeten we durven uit het rituele contract te stappen.

Daar zit allemaal wel wat in, maar het blijft voor docenten en cursisten toch moeilijk om de expliciete feedback te laten gaan. Dana Ferris (2) schreef later een antwoord op Truscott en zij voert zowat de beweging aan die ervoor ijvert wel expliciete feedback te geven omdat het werkt onder bepaalde voorwaarden en met inachtneming van verschillende factoren die de goede werking van de feedback verhinderen of juist helpen.

Precies omdat er veel factoren kunnen meespelen is het moeilijk om bewijzen te vinden voor deze visies. Onderzoeksresultaten worden dan ook vaak betwist omdat bijvoorbeeld
- enkel een bepaald type leerders werd onderzocht
- niet de juiste manier van feedback werd gehanteerd
- de langetermijnefficiëntie niet in kaart wordt gebracht enz.

De titels van onderstaande artikels spreken voor zichzelf:
(1) Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning.*
(2) Ferris, D. R. (1999a). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing.*
Waarop Truscott dan weer inging met: Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing.*
En ook daarop kan je commentaar van Ferris lezen in Ferris, D.R. (2003). *Response to Student Writing. Implications for Second Language Students.* LEA, London.

Er zijn natuurlijk al verschillende pogingen geweest om vormfeedback te verbeteren. Heel wat docenten werken nu met een codesysteem, waarbij een symbool aangeeft welk type fout de cursist gemaakt heeft. De

bedoeling is dat de cursist (geholpen door het symbool) zelf op zoek gaat naar de verbetering. Er moet met andere woorden een inspanning worden gedaan, wat de kans op leren vergroot.

Tegen deze methode zijn er echter ook een aantal bezwaren: cursisten zijn geen taalkundigen, de werkdruk van de docent verhoogt, het is niet altijd duidelijk om welk type fout het precies gaat, maar vooral: de focus van de feedback blijft op de vormelementen in de tekst en die feedback is eenvormig. We tonen namelijk welke fouten en tekortkomingen er zijn, maar is er ook informatie over het belang ervan? En de haalbaarheid om die (op het niveau van de cursist) te remediëren? Soms weten cursisten zelf reeds dat ze een bepaalde fout maken (m.a.w. het besef is er al). Heeft het dan zin om hen opnieuw die fout te tonen? Of moeten we een stap verder gaan? Bij dat soort fouten is het veel belangrijker om te werken aan de motivatie van de cursist: hem te tonen waarom die fout geremedieerd moet worden. Of aan de andere kant dat hij daar nu nog niet teveel energie hoeft in te steken.

Die eenvormigheid is vreemd als je bekijkt hoeveel factoren er meespelen (niveau, attitude, motivatie, sterktes en zwaktes van de leerder, doel van de oefening, moment van de feedback, etc.). De vraag is dus waarom vele docenten enkel op deze manier werken. Er moeten toch ook andere contracten mogelijk zijn die beter passen bij verschillende situaties. Met andere woorden: zou er niet meer maatwerk kunnen zijn in feedback? Een belangrijke vraag waar we verder in dit artikel nog op terugkomen.

Wanneer?

Feedback is vaak vooral gericht op het eindproduct, terwijl procesfeedback veel meer zin heeft. Met proces bedoelen we beide zaken, namelijk het *taalleerproces* (iedere leerder zit in een fase van het leerproces en de feedback moet daaraan aangepast zijn) en ook het *productieproces* zelf (het is veel logischer om tijdens de productiefase zelf feedback te geven in plaats van achteraf).

Het leerproces

Elke NT2-leerder zit in een bepaalde fase van het leerproces van de taal. Dit traject verloopt heel divers voor verschillende leerders, afhankelijk van een resem factoren waaronder leerstijl, interesse, moedertaalinterferentie, hoeveelheid input en output enz. Kortom, in een welbepaalde klas heeft geen enkele cursist dezelfde verworvenheden en tekortkomingen.

Daar komt nog bovenop dat de beheersing van een aantal grammaticale elementen van de taal (wellicht) een eigen volgorde en schema volgt waar noch de docent, noch de cursist vat op heeft. Met andere woorden: soms heeft feedback op een grammaticale tekortkoming helemaal geen zin omdat de tijd nog niet rijp is.

Als docent hebben we echter een dwingende attitude tegenover fouten. Als we onze rol als docent t.o.v. een cursist vergelijken met de rol van een ouder t.o.v. een kind zien we heel verschillende gevoelens met betrekking tot de fout, de verantwoordelijkheid en het geloof in de verbetering.

Mijn zoontje Bram denkt bijvoorbeeld dat het werkwoord 'hebben' eigenlijk 'kebben' is. Hij zegt dus ik keb, jij kebt en wij kebben. Als ouder vind je zulke fouten schattig en ook helemaal niet erg. Je hebt er geen problemen mee te geloven dat het allemaal wel in orde komt (behalve bij extreme gevallen natuurlijk). Ouders voelen zich maar voor een stuk zelf verantwoordelijk om de fout op te lossen: de kinderen leren dat wel op de een of andere manier als de tijd rijp is. Het verdwijnen van de fout vinden ouders zelfs jammer want hun kleintje is groot geworden.

Mijn cursist Sahar slaagt er maar niet in de constructie van de bijzin onder de knie te krijgen. Leerkrachten hebben bij zulke fouten van hun cursisten een heel andere reactie: ze storen zich aan de fouten, geloven niet dat het in orde komt als ze er zelf niks aan doen, voelen zich 100% verantwoordelijk om het direct op te lossen en zijn heel tevreden als de fout verdwijnt.

Hierbij kan men terecht bedenken dat de vergelijking niet helemaal opgaat en dat het bij NT2-leerders inderdaad niet altijd vanzelf goed komt. We moeten dus als docent wel degelijk iets doen, maar we moeten afleren altijd als doel te hebben de tekortkoming (bijvoorbeeld een grammaticale fout, ondermaatse kennis van woordenschat) onmiddellijk zelf in orde te brengen want het werkt zo niet. Het is beter de volgende evolutielijn bij het leren voor ogen te houden:

tekortkoming	besef	motivatie	remediëring	vooruitgang

Als een leerder van de eerste niveaus een tekortkoming vertoont die behoort tot de zeer frequente basiselementen van de taal, is dat vrij normaal. Maar als een leerder van een hoog niveau diezelfde fout maakt, moet daar een andere soort feedback op komen. In het eerste geval gaat het vooral om informatie die het besef aanbrengt, terwijl in het tweede geval de motivatie om te remediëren heel sterk moet worden geprikkeld. Feedback moet dus een bepaalde evolutie kennen en niet noodzakelijk

gericht zijn op de onmiddellijke oplossing van het probleem. Als we in niveau A2 de cursisten het besef bijbrengen dat het voltooid deelwoord in het Nederlands complex is en dus aandacht (en studie) zal vragen, dan hebben we op dat niveau onze job gedaan. Als we verwachten dat de cursist daardoor het probleem ineens foutloos beheerst, zijn we daarin niet billijk. Andersom is het op een hoog niveau niet genoeg om de cursist te tonen dat hij het voltooid deelwoord fout heeft geschreven: de feedback moet zijn dat dit niet langer aanvaardbaar is.

Een ander voorbeeld heeft te maken met vermijdingsstrategieën: een cursist kan heel eenvoudige woorden en zinnen gebruiken om iets uit te leggen en daardoor vermijden dat hij fouten maakt (of hapert bij het spreken). De tekst van Andi (zie Bijlage bij dit artikel) illustreert dit punt. Voor een beginnende taalleerder is dat een prima strategie en mag dat gerust aangemoedigd worden. Maar bij een halfgevorderde moet de feedback eigenlijk gericht zijn op het besef dat het hoog tijd wordt om die strategie te ontgroeien. Omgekeerd is het belangrijk om cursisten die complexere zinnen of moeilijker woordgebruik durven te hanteren te motiveren om dit verder te zetten en hen dus niet zomaar op de noodzakelijk gemaakte fouten te wijzen.

Als we rekening houden met deze lineaire evolutie maken we al een stap voorwaarts, maar zoals we in het begin aangaven, hebben cursisten vooral een individuele leerlijn. De ene cursist heeft een bepaald aspect van de taal vrij snel onder de knie terwijl een andere daar heel lang mee blijft sukkelen. De ene beseft zijn fout, de andere niet. Ook motivatie, attitude, ontkenning, fossilisatie en dergelijke spelen een grote rol. Idealiter weten we als docent ongeveer wat de moeilijkheden zijn van onze cursisten, zodat we hen aangepaste feedback op maat kunnen geven.

Maar is dat wel mogelijk, kun je je afvragen? Docenten zijn geen supermensen die elke situatie en elke cursist direct goed inschatten en zo komen tot de goede beslissing in verband met feedback. Daarom is het erg belangrijk om over feedback te praten met je cursisten en hen uit te leggen wat je doet of wil doen. Een belangrijke bron van informatie over besef, motivatie en het leerproces van de cursist is die cursist zelf. Dat alles past bovendien heel goed in het idee van het sociaal-constructivisme en het autonome leren: in plaats van gewoon feedback te geven, toon je aan de cursist welke mogelijkheden er zijn tot feedback en wat de voor-en nadelen zijn. De cursist voelt zich betrokken want hij tekent met jou een nieuw contract waar hij zelf in gelooft en dat is een heel andere insteek dan het rituele contract waar we mee begonnen zijn.

Zo zijn we ook gekomen tot een deel van het antwoord op de vraag of feedback niet wat meer op maat kan bestaan. We beschikken over een veel groter arsenaal aan feedbacksignalen dan het aantonen van vormfouten: aanmoediging, waarschuwing, confrontatie, appreciatie, belang aanduiden, ontkenning tegengaan enzovoort. Als we die op het juiste moment toepassen en bovendien rekening houden met billijkheid, haalbaarheid en het scheppen van een 'veilig kader', maakt onze feedback veel meer kans om een goede voedingsbron te zijn voor echte motivatie en remediëring. Hoe dit precies moet, is stof voor een artikel op zich en zal ook wel blijven afhangen van het vingertoppengevoel van elke docent met zijn cursisten.

Het productieproces

Als een cursist iets schrijft, zit hij niet alleen in een bepaalde fase van zijn leerproces, maar het schrijven zelf is eveneens een proces op zich. Eigenlijk is het de logica zelf: als je feedback (en een waardeoordeel) geeft op een afgewerkt product, dan heeft het niet zoveel zin om dat nog aan te passen. Veel beter is het om tijdens de productiefase zelf reeds feedback te geven zodat de cursist tijdig kan bijsturen. De cursist is ook meer gemotiveerd om dit te doen want hij moet sowieso het product nog afwerken dus zal hij zeker rekening houden met de feedback en die incorporeren. Bovendien heeft hij er nog niet zoveel tijd en moeite in gestoken en zal dus meer geneigd zijn om zaken te herbeginnen.

Waarom gebeurt het dan zo weinig? Wellicht omdat men niet altijd inziet dat schrijven (en spreken) eigenlijk processen zijn en dat men dus kan werken in verschillende stadia:

de conceptuele fase	het omzetten in taal	de revisie	het eindproduct

Bij korte taaluitingen gebeurt dit proces weliswaar heel snel en worden zelfs stadia overgeslagen. Maar het voordeel is dan weer dat eventuele productfeedback daarop eveneens snel, kort en overzichtelijk is en dat er minder weerstand zal zijn van de cursist om te herbeginnen aangezien er ook nog niet zoveel moeite werd gedaan dan in langere schrijfstukken en gesprekken.

Zeker bij langere taaluitingen is het wel degelijk mogelijk om het proces in stadia te laten uitvoeren en feedback te voorzien tussen de stadia door. Die feedback kan bovendien gekoppeld worden aan een rollenmodel:

Procesmodel schrijven:

fase 1 outline	**fase 2** kladversie	**fase 3** eindversie

Bijhorend rollenmodel:

rol 1 geïnteresseerde lezer	**rol 2** corrector	**rol 3** evaluator

Feedback:

feedback 1 inhoud	**feedback 2** vorm	**feedback 3** waardeoordeel

Waarom deze procesindeling zo weinig gebruikt wordt, heeft wellicht ook te maken met de tijd die zo'n procesaanpak vraagt en daar hebben docenten meestal geen overschot van. Dat argument is ergens wel aannemelijk, maar zeker geen reden om nooit aan procesfeedback te doen. Bovendien is het zo dat de feedback niet noodzakelijk door de docent zelf moet gebeuren. De rollen hierboven kunnen door verschillende actoren ingevuld worden, waaronder de cursisten zelf. Zeker dat laatste is interessant, want dan wordt feedback een deel van het leerproces en speelt het tijdsargument dus niet zoveel rol meer.

Wie?

De docent zelf kan in principe alle rollen vervullen en dat wordt vandaag ook vaak verwacht omdat hij die autoriteit heeft en omdat er dan controle is over de feedback. Toch zijn daar nadelen aan verbonden:

- De werkdruk wordt groter en dus is er minder tijd over om extra aandacht te geven.
- De cursisten blijven afhankelijk van het oordeel en de sturing van de docent.
- De docent zelf is een individu en heeft dus ook een beperkte perceptie.

De autoriteit van de docent wordt trouwens niet altijd zomaar aangenomen. Iedereen zal wel al eens de opmerking gekregen hebben dat die fouten toch niet zo belangrijk zijn en dat andere

moedertaalsprekers daar geen probleem mee hebben. (M.a.w. de docent is een autoriteit voor schooltaal, maar daarbuiten gelden andere regels).

Kunnen we ons dan niet de vraag stellen of de rollen van het model niet door andere actoren kunnen worden ingevuld? Wat de evaluator betreft, is dat niet zo eenvoudig: de docent zal steeds de grootste rol daarin spelen. Maar waarom zouden peers niet geïnteresseerde lezers kunnen zijn die met evenveel autoriteit als de docent kunnen zeggen of een outline interessant is, of alles duidelijk is en of er voldoende informatie of argumenten aanwezig zijn?

Voor de rol van corrector moet er autoriteit zijn. Moedertaalsprekers uit de omgeving van de cursist voldoen daaraan maar hebben het 'nadeel' dat ze een sociaal-affectieve band hebben met de cursist: zij zijn niet per definitie geïnteresseerd in hoe de cursist iets schrijft of zegt maar in wat die cursist te vertellen heeft. Daarom horen we vaak van cursisten dat hun nederlandstalige vrienden hen nooit verbeteren als ze praten. Toch is ook hier een mouw aan te passen door bijvoorbeeld een taalbuddy aan te stellen, of door die feedback enkel op bepaalde momenten te laten gebeuren.

Peers kunnen wel gemotiveerd worden om aan elkaar feedback te geven omdat ze een gemeenschappelijk doel hebben: de taal leren. Zij hebben dan weer een autoriteitsprobleem, maar ook daar bestaan oplossingen voor. Als niet één maar verschillende peers feedback geven op een schrijfproduct, verhoogt de autoriteit. Als ze bovendien niet zeggen dat er iets fout is, maar gewoon aangeven dat zij dat denken , valt het argument van de autoriteit weg. De cursist krijgt feedback en moet dan zelf beslissen wat ermee te gaan doen. Het indienen van de kladversie (met opmerkingen van de peers) samen met de eindversie geeft bovendien de mogelijkheid dat de docent achteraf nog als autoriteit optreedt. Het kan juist interessant zijn met de cursist te overleggen waarom hij een bepaalde opmerking heeft verbeterd of juist genegeerd. Zulke vormen van peer feedback en conferencing openen dus een debat en dat is een interessanter leermoment dan het vaststellen van de door de docent gegeven vormcorrecties.

Ook dingen kunnen als corrector optreden: woordenboeken en spellingcorrectors zijn daar een voorbeeld van. Als in het einddproduct van een schrijfopdracht nog spellingfouten staan terwijl afgesproken was een spellingcorrector/woordenboek te gebruiken, is een eenvoudige vormcorrectie van de docent niet de juiste feedback. In dat geval kan de docent bijvoorbeeld weigeren om zijn deel van het feedbackcontract te vervullen zolang de (opzoekbare) spellingfouten nog aanwezig zijn. De

mogelijkheden van de computer worden op dat gebied trouwens steeds groter. In de nabije toekomst zal het bijvoorbeeld mogelijk worden om een tekst te scannen op het type woordgebruik: behoren de woorden van de tekst tot de meest frequente in de taal of zijn er ook (voldoende) woorden van een iets moeilijker niveau.

Tot slot

Dit zijn slechts een paar voorbeelden van de manier waarop de rollen in het model kunnen ingevuld worden. Als afluiter wil ik nog een voorbeeld geven van de doorvoering van dit model op feedback bij spreken. Correctie bij spreken is een hekel punt omdat het de motivatie van de spreker aantast en omdat er schroom bestaat bij de docent op dat gebied. Als een cursist iets interessants vertelt, is het moeilijk hem te zeggen dat hij eigenlijk een inversie had moeten gebruiken. In mijn eigen klassen heb ik deze rol daarom laten invullen door de 'vervelende robot'. Dat is de docent zelf die 'verandert' in een robot die geprogrammeerd is om fouten te analyseren. Een robot heeft geen gevoelens en is niet geïnteresseerd in de inhoud van wat er verteld wordt. Dit is een duidelijke afspraak met de cursist die bereid is de confrontatie met deze robot aan te gaan. Er is geen schroom meer en de vormcorrecties kunnen vrijuit gegeven worden. Dat is zeker het geval als de cursist niet iets persoonlijks vertelt, maar bijvoorbeeld een verhaal herhaalt. De reacties van de cursisten zijn zeer positief: eindelijk iemand (en gelukkig een machine) die hen onomwonden zegt wat er nog verbeterd kan worden.

Literatuur

Ferris, D. R. (1999a). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing.*

Ferris, D.R. (2003). *Response to Student Writing. Implications for Second Language Students.* LEA, London.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning.*

Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing.*

Bijlage: Voorbeeldteksten van cursisten

Voorbeeldtekst: Bebe

"Congo is een grote land met 12 provincies. Elke provincie heeft een hoofdstad en veel dorpen, voor dat, hebben we twee sorte van vrouwen en manen.
In de dorp, leven mensen door de velde, kleine handel en door uittrekking van diamant, goud en andere erts in somige dorpen, hier wonen mensen meestal samen en het leven is beetje moeilijk.
Maar in de hoofdsteden, zijn mensen allemaal aktief.
Vanaf het oorloog, vrouwen en manen moeten flexibel en creatief zijnvoor iets krijgen om te leven, want de fonctionnaris hebben geen salaris.
Dus, de vrouwen moeten moedig zijn. Ze verkopen bijna als om kinderen te zorgen ".

Voorbeeldtekst: Jovanka

"In Los Angeles zijn veel planten verwoesten omdat de lucht zo vuil is. Als movie stars en surfers adem de lucht in, mogelijkheid worden ze uitgestorven. Er zijn twee verschillende vogels soorten in Los Angeles. Een mooie vogel en een vogel die met plastic churgerie mooi wordt. Sommige mensen in Los Angeles zijn milieuvriendelijk doordat ze televisieprogrammas recycleren. Duurde autos als Mercedes Benz zijn in Los Angeles wegwerpproducten."

Voorbeeldtekst: Andi

"Vrienden en familie voor Indonesië zijn heel belangrijk. Maar in België zijn de oude mensen zoals oma en opa meer onafhankelijk.

In Indonesië zijn de oude mensen niet afhankelijk. Ze moeten hun kinderen hebben om voor hen te zorgen.
In Indonesië zijn mannen de baas. Als man en vrouw getrouwd zijn. Man is het hoofd van de familie en meestal blijf de vrouw thuis om voor de kinderen te zorgen. En de man gaat uit werken. Vrouwen in Indonesië zijn meer afhankelijk van hun man. Bijvoorbeeld: ze hebben ook verschillende taken. Veel mannen van België helpen ook in het huishouden. In België mannen en vrouwen zijn gelijkwaardig. Maar in Indonesië niet.
De vrouwen in België zijn meer vrij om te doen wat ze zelf willen zonder dat iemand een opmerking maakt. Bijvoorbeeld: met hun kledij en roken op straat. In mijn land zijn ze meestal niet zo vrij."

Over de Beroepsvereniging van Docenten NT2

Sinds de oprichting op 28 januari 2003 houdt de BV NT2 zich bezig met informatie-uitwisseling, belangenbehartiging en deskundigheids-bevordering ten behoeve van docenten Nederlands als Tweede taal in Vlaanderen en in Nederland.

Het bestuur bestaat uit:

Annemieke Zevenhuijzen, voorzitter
Carola van der Voort, secretaris
Simon Verhallen, penningmeester
Douwe Brouwer, bestuurslid
Marjolijn Kaak, bestuurslid
Jan Strijbol, bestuurslid
Maritsa de Meester, bestuurslid

Informatie over de activiteiten van de BV NT2 is te vinden op de eigen website van de vereniging: www.bvnt2.org. De website wordt ook door leden gebruikt als virtuele ontmoetingsplaats, overlegruimte en discussieplatform. De leden worden regelmatig per elektronische Nieuwsbrief, e-mail en per post op de hoogte gehouden van geplande activiteiten van de vereniging. Daarnaast publiceert de BV NT2 regelmatig in tijdschriften zoals *LES, Profiel* en *Alfa-nieuws.*

Er is een aantal werkgroepen gevormd rond de volgende thema's:

- NT2 en Logopedie
- Externe Contacten
- Staatsexamen en Inburgeringsexamen
- Alfabetisering
- Conferenties
- Zelfstandig Ondernemen en NT2

De BV NT2 is een groeiende, actieve vereniging. Diverse leden fungeren als contactpersoon door binnen hun werkkring de informatie van de vereniging verder te verspreiden en collega's enthousiast te maken voor de activiteiten van de BV NT2. Aanmelding als lid en/of contactpersoon is mogelijk via onze website.

Eerder verschenen in deze reeks

In 2004 is verschenen: een bundel met lezingen gehouden op de eerste grote NT2-conferentie georganiseerd door de BV NT2. Deze conferentie vond plaats te Heino in 2003. De volgende artikelen zijn erin opgenomen:

Folkert Kuiken : *Focus on Form in het onderwijs Nederlands als tweede taal*
Ellie Liemberg: *Het Staatsexamen NT2 en de Europese taalniveaus*
Jan H. Hulstijn: *Oefenen in woord-voor-woord verstaan als vorm van impliciet leren*
Ruud Stumpel: *Het Pocketwoordenboek Nederlands als tweede taal*
Bart Bossers: *Zoveel lezers, zoveel wensen...Leesvaardigheidstraining in de NT2-klas*

In 2005 verscheen Vakwerk 2, met lezingen gehouden op de tweede grote NT2-conferentie georganiseerd door de BV NT2 (Vierhouten 2004). De bundel bevat de volgende artikelen:

Kris van den Branden: *Vijf vragen over taakgericht taalonderwijs*
Jeanne Kurvers: *Schrijven voor beginners*
Peter Schoenaerts: *Drama in het NT2-onderwijs*
Peter Jordens: *Inversie bestaat niet*
Bart Bossers: *Vormen van grammatica-onderwijs en hun effect op de taalverwerving*
Annemie Decavele & Leen Verrote: *Typisch?! (Inter)culturele communicatie in de NT2-klas*

Vakwerk 3 verscheen in 2006. Erin opgenomen zijn de lezingen gehouden op de NT2-conferentie in Hoeven-Bovendonk, 2006. De bundel bevat de volgende artikelen:

Wilfried Decoo: *Op de golven van talenmethodiek: een blik vanuit het kraaiennest.*
Jan-Arjen Mondria: *Mythen over vocabulaireverwerving.*
Ineke van de Craats: *Obstakels op de NT2-snelweg.*
Alied Blom: *Geen grammatica in het NT2-onderwijs, maar wat dan?*
Folkert Kuiken: *Focus on form als alternatief voor grammaticaonderwijs.*
Bart Bossers: *NT2 en het nieuwe leren.*
Chris van Veen: *De nadruk op uitspraak. Theoretische overwegingen en praktische voorbeelden.*

Vakwerk 4 verscheen in 2006. Erin opgenomen zijn de lezingen gehouden op de NT2-studiemiddag in Amsterdam over Uitspraak en Verstavaardigheid in juni 2007. Deze studiemiddag heeft een unieke verzameling bijdragen over uitspraakonderwijs NT2 opgeleverd. De bundel bevat de volgende artikelen:

Chris van Veen: *Uitspraak: onderzoek en onderwijs*
Ineke van de Craats: *Spraak en alfabetisering*
Laurent Rasier: *Uitspraakvaardigheid en de NT2-/NVT-leerder. Van theorie en onderzoek naar onderwijspraktijk*
Hinke van Kampen & Ruud Stumpel: *De nadruk op uitspraak. Theoretische overwegingen en praktische voorbeelden.*
Chrissy Hosea: *'Wat hoor ik nou?' Waarom NT2-leerders verstavaardigheids-problemen hebben.*

Yolande Timman: *De kleinste taalelementen in het grote CP-verband. De drieweg-aanpak van uitspraak.*
Annemarie Groot & Margreet van Wijk: *Net even anders: luisteren gericht op verstavaardigheid*
Margreet Verboog: *Verstaanbaar spreken: wat helpt en wat niet?*
Vincent van Heuven: *Uitspraak van het Nederlands door Turkse immigranten*
Karolien Thio & Simon Verhallen: *Lessen Uitspraak en Melodie opnieuw onder de aandacht.*
Bart Bossers: *Luister goed! Verstavaardigheid als oorzaak van uitspraakproblemen.*

De bundels Vakwerk zijn, zolang de voorraad strekt, verkrijgbaar bij het secretariaat van de BV NT2.