

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/76466>

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

Taalvaardigheidstoets Aanvang Lager onderwijs

Achtergronden

Steunpunt Nederlands als Tweede Taal
Blijde Inkomststraat 7
B-3000 Leuven

Catia Cucchiarini
Koen Jaspert

© Steunpunt NT2, Leuven 1995

ISBN 90-75511-02-7

Niets uit deze uitgave mag zonder voorafgaande schriftelijke toestemming van het Steunpunt Nederlands als Tweede Taal en van beide auteurs worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm, computer-software of op welke wijze dan ook.

Voorwoord

Voor u ligt de Taaltoets Aanvang Lager onderwijs. Aanleiding tot het ontwikkelen van deze toets werd gevormd door de situatie van allochtone kleuters in het Vlaamse onderwijs. Van deze groep kleuters is geweten dat velen onder hen beginnen aan het lager onderwijs met een serieuze taalvaardigheidsachterstand. Die taalvaardigheidsachterstand wordt door heel veel mensen uit de onderwijspraktijk aangewezen als de voornaamste oorzaak van het zwakke presteren van deze allochtone leerlingen in de eerste, zo cruciale fase van het lager onderwijs. Hoe groot deze achterstand dan wel was, en hoe ermee omgegaan kon worden, was veel minder duidelijk. Aan deze leemte in noodzakelijke kennis van de leerlingssituatie proberen wij met deze toets tegemoet te komen.

De toets gaat uit van de visie dat taalvaardigheid niet een toepassing inhoudt van een groot aantal los te verwerven elementen en regels, maar een geïntegreerde vorm van talig handelen veronderstelt. Dit betekent dat deze toets er niet op gericht is de taalvaardigheidsachterstand toe te gaan schrijven aan een tekort op een of ander onderdeel van taal, wat dan in remediëring extra aandacht dient te krijgen. De TAL probeert een algemeen beeld te schetsen van de mate waarin een leerling over de taalvaardigheid beschikt die noodzakelijk is om in het eerste leerjaar te functioneren.

Wie onvoldoende taalvaardig blijkt, zal de kans moeten krijgen die taalvaardigheid te verhogen door beter ondersteund te functioneren in een omgeving die rijk is aan kansen om taal te verwerven. Dat de TAL dit soort benadering van taalvaardigheidsverwerving veronderstelt, geeft meteen aan dat de toets aansluit bij wat in Vlaanderen onder de analytische benadering van taalonderwijs wordt verstaan: taalvaardigheid wordt al doende opgebouwd, en niet

via het leren kennen en vervolgens toepassen van allerlei onderdelen van het taalsysteem.

De TAL werd ontwikkeld door de Werkgroep Taaltoetsen. Deze werkgroep werd opgericht in het kader van de werking rondom het Onderwijsvoorrangsbeleid in de Vlaamse scholen. In deze werkgroep waren medewerkers actief van PMS-centra verbonden met alle Vlaamse onderwijsnetten. Een lijst met namen van de mensen die van de Werkgroep deel hebben uitgemaakt, vindt men hieronder. De Vlaamse Onderwijsraad zorgde voor de coördinatie van deze werkgroep. De wetenschappelijke ondersteuning van de werking werd verzorgd door het Steunpunt NT2, dat ook verantwoordelijk was voor het betrouwbaarheids- en validiteitsonderzoek dat in dit Achtergrondenboek wordt besproken. Deze tekst werd geschreven door Catia Cucchiaroni en Koen Jaspaert. De Werkgroep fungeerde hierbij als een soort resonansgroep: zij voorzagen eerdere versies van de tekst steeds van kritische commentaar. De lay-out en vormgeving van deze uitgave waren in handen van Yvonne Van den Lemmer.

Tenslotte moet zeer zeker de grote inbreng genoemd worden van de PMS-Centra bij het tot stand komen van deze toets. Met name de PMS-Centra die OVB-scholen begeleiden, werden steeds bereid gevonden om toetsversies af te nemen, de resultaten te verwerken op een dergelijke wijze dat de Werkgroep ze kon analyseren, en feedback te geven over hun ervaringen tijdens de toetsafname. Zonder die bereidheid van medewerkers van om en bij de vijftig PMS-Centra van de verschillende onderwijsnetten, was deze toets er nooit gekomen. Wij hopen dan ook ten zeerste dat hij voor hen een bruikbaar instrument blijkt, en dat het eindprodukt in hun ogen de vele inspanningen wettigt.

Deelnemers aan de Werkgroep Taaltoetsen:

Natacha Aleksic, Piet Bracke, Edgard Cocquet, Lieve Corneillie, Betty Cranshoff, Jos Cré, Catia Cucchiarini, Viviane Eraly, Kris Geerinck, Luc Heyerick, Koen Jaspert, Walter Magez, Kristine Olieslagers, Els Palmaers, Anne Reynaerts, Ronny Schoofs, Paty Thuys, Inge Uyt den Houwen, Hilde Van Acker, Saskia Van Berkel en Rita Van den Bosch.

Inhoudstafel

1. Inleiding	1
2. Op zoek naar een geschikte toets	4
2.1 Toetscategorieën	5
2.2 Criteria waaraan de toets zou moeten voldoen	10
2.3 Welke toetsen voldoen aan de gestelde eisen?	12
3. Het ontwikkelen van een nieuwe toets: de TAL	13
3.1 Onderzoek naar taalvaardigheidseisen	13
3.2 Toetsconstructie	15
3.3 Eerste proefonderzoek	18
3.3.1 Proefpersonen en procedure	19
3.3.2 Resultaten	20
3.4 Tweede proefonderzoek	21
3.4.1 Proefpersonen en procedure	21
3.4.2 Resultaten	21
3.5 Definitieve samenstelling van de TAL	24
4. Betrouwbaarheid en validiteit	25
4.1 Betrouwbaarheid	25
4.2 Validiteit	27
5. De TAL als taalvaardigheidstoets	29
5.1 Betrouwbaarheid	29
5.1.1 Proefpersonen en procedure	29
5.1.2 Resultaten	30
5.2 Validiteit	32

6. De TAL als signaal voor potentiële problemen	34
6.1 Drempelbepaling	34
6.2 Validiteit	36
7. Gebruik van de TAL	45
7.1 Algemene opmerkingen	45
7.2 Concretisering van het gebruik	47
8. Conclusies	51
Referenties	52

Taalvaardigheidstoets Aanvang Lager onderwijs

1. Inleiding

Al geruime tijd wordt in de literatuur de nadruk gelegd op het belang van schoolse taalvaardigheden voor het slagen in het onderwijs. Hoewel taalvaardigheid geen garantie biedt voor schoolsucces - er zijn nu eenmaal veel andere factoren die daarbij een rol spelen - is een goede beheersing van de schooltaal een belangrijke voorwaarde om op school te slagen. Taal is immers het communicatiemiddel tussen de leerkracht en de leerling (Van Bon 1982; Jaspaert 1991).

Dat schoolse taalvaardigheden aanzienlijk verschillen van algemene taalvaardigheden is intussen een geaccepteerd feit (Cummins 1979, 1984; Hulstijn 1984; De Haan 1985, 1987; Litjens 1989). De twee soorten taalvaardigheid worden vaak aangeduid met de acronyemen die Cummins (1980) heeft geïntroduceerd, namelijk BICS (Basic Interpersonal Communicative Skills) en CALP (Cognitive-Academic Language Proficiency). BICS refereren naar dagelijkse taalvaardigheden, terwijl CALP voor het meer cognitief belastend, minder contextrijk taalgebruik staat dat typerend is voor schoolse situaties. In eerste instantie werden deze twee typen van taalvaardigheid als een dichotomie gepresenteerd. Later (Cummins 1984) worden BICS en CALP voorgesteld als mogelijke posities langs twee continua. Het eerste continuüm staat voor de hoeveelheid contextuele ondersteuning en het tweede voor de mate van cognitieve inspanning die vereist wordt. Elk voorbeeld van communicatief gedrag kan in principe langs deze twee continua geplaatst worden. Belangrijk hierbij is echter dat met de taalgebruiker rekening wordt gehouden. Volgens Cummins (1984: 15) worden taaltaken niet als zodanig getypeerd als cognitief-veeleisend of contextrijk, maar met betrekking tot bepaalde taalgebruikers. Het is dus best mogelijk dat

een en dezelfde taak voor de ene taalgebruiker cognitief veeleisend of contextarm is en voor de andere niet. Dit maakt het ook mogelijk om BICS en CALP in een ontwikkelingsperspectief te plaatsen. Bijvoorbeeld, zoals Cummins (1984: 13) zegt: 'As mastery is developed, specific linguistic tasks and skills travel from the bottom towards the top of the vertical continuum' (het continuüm dat gaat van cognitief weinig eisende tot cognitief veeleisende taken, C.C., K.J.).

Het belang van de taalgebruiker voor het BICS-CALP onderscheid wordt vaak over het hoofd gezien, met het resultaat dat BICS en CALP als statisch worden opgevat: sommige taken horen nu eenmaal bij de BICS en sommige bij de CALP. In aansluiting hierbij wordt ook vaak gesuggereerd dat op school een verschuiving zou optreden van het gebruik van BICS in de onderbouw van het basisonderwijs naar een gebruik van CALP in de bovenbouw (Litjens 1990: 46). Volgens deze voorstelling van zaken zouden kinderen pas in de laatste jaren van het basisonderwijs moeilijkheden ondervinden, wanneer van BICS naar CALP wordt overgegaan. Als men echter BICS en CALP dynamisch opvat en vanuit die optiek het taalgebruik thuis en op school vergelijkt, dan is het duidelijk dat heel wat kinderen al vanaf het begin van hun schoolcarrière met cognitief ingewikkelde, contextarme taal geconfronteerd worden. Dat is met name het geval bij kinderen voor wie de taal die op school gesproken wordt niet de moedertaal is (Cummins 1984: 15).

Gezien het belang van taalvaardigheid voor schoolsucces, is het essentieel dat op school aandacht besteed wordt aan het ontwikkelen van de nodige taalvaardigheden. Om dit mogelijk te maken, moet de school zicht hebben op de positie van de verschillende leerlingen ten aanzien van de taalvaardigheidseisen die in het onderwijs worden gesteld. Hiervoor zijn specifieke instrumenten nodig, zoals bijvoorbeeld toetsen waarmee men op cruciale momenten in de schoolloopbaan kan vaststellen of leerlingen over voldoende

taalvaardigheid beschikken.

Het ontwikkelen van toetsen die hiervoor bedoeld zijn, is een van de doelstellingen van de Werkgroep Taaltoetsen, waarin vertegenwoordigers van verschillende PMS-Centra samenwerken met medewerkers van de Cel-Migranten van de Vlor en van het Steunpunt NT2 van de K.U. Leuven. Het eerste doel dat deze werkgroep zich heeft gesteld, is het ontwikkelen van een taalvaardigheidstoets voor de overgang kleuteronderwijs-lager onderwijs, de zogenaamde Taalvaardigheidstoets Aanvang Lager onderwijs (TAL). Over uitgangspunten, constructie en normering van deze toets wordt in onderhavige tekst uitvoerig gerapporteerd. Vooraleer wij overgaan tot het bespreken van de TAL, lijkt het ons echter zinvol om even uit te leggen waarom het nodig was om een dergelijke toets te ontwikkelen. Met andere woorden, men zou zich kunnen afvragen waarom we niet voor hetzelfde doel een al bestaande toets gebruikt hebben.

In paragraaf 2.1 bespreken we sommige van de toetscategorieën die in de literatuur onderscheiden worden. We beperken ons tot die categorieën die nodig zijn om duidelijk te maken naar welke toets we op zoek waren. In 2.2 gaan we in op de criteria waaraan deze toets zou moeten voldoen. In 2.3 wordt gekeken welke van de bestaande toetsen aan deze eisen voldoen. Aangezien geen van de toetsen geschikt bleek te zijn, is besloten om een nieuwe toets te ontwikkelen, de TAL. Hierover wordt in paragraaf 3 gerapporteerd. Meer specifiek, paragraaf 3.1 gaat over het onderzoek naar de taalvaardigheidseisen in het eerste leerjaar, dat als uitgangspunt is gebruikt voor het ontwikkelen van de TAL. In paragraaf 3.2 wordt de toetsconstructie beschreven en wordt duidelijk gemaakt hoe de TAL zich tot al bestaande toetsen verhoudt. In 3.3 komt het eerste proefonderzoek aan de orde. We beschrijven eerst de proefpersonen en de procedure (3.3.1) en daarna de resultaten (3.3.2). In paragraaf 3.4 wordt over het tweede proefonderzoek gerapporteerd. Ook in dit

geval worden eerst de steekproef en de procedure besproken (3.4.1) en vervolgens de resultaten (3.4.2). Paragraaf 3.5 gaat over de definitieve samenstelling van de TAL. In paragraaf 4 gaan we in op de begrippen betrouwbaarheid en validiteit. Paragraaf 5 gaat over de TAL als taalvaardigheidstoets. In 5.1 wordt over het betrouwbaarheidsonderzoek gerapporteerd. Ook in dit geval beschrijven we eerst de proefpersonen en de procedure (5.1.1) en daarna de resultaten (5.1.2). Vervolgens wordt in paragraaf 5.2 de validiteit van de TAL als taalvaardigheidstoets behandeld. Paragraaf 6 is gewijd aan de functie van de TAL voor het signaleren van potentiële problemen. In 6.1 komt de drempelbepaling aan de orde en in 6.2 de validiteit van de TAL als signaal voor potentiële problemen. Paragraaf 7 gaat over de manier waarop de TAL gebruikt zou moeten worden. In paragraaf 8 worden enkele conclusies getrokken.

2. Op zoek naar een geschikte toets

Zoals in de inleiding al is uitgelegd, was er behoefte aan een toets die gebruikt kon worden om vast te stellen of kinderen bij de aanvang van het lager onderwijs over genoeg taalvaardigheid beschikken om daadwerkelijk van dat onderwijs te kunnen profiteren. Een dergelijke toets zou uiteraard aan bepaalde voorwaarden moeten voldoen. Alvorens we ingaan op de criteria die gehanteerd zijn bij het zoeken naar een geschikte toets, bespreken we kort enkele toetscategorieën die in de vakliteratuur worden onderscheiden.

Bij sommige van de indelingen die hieronder genoemd worden gaat het niet zozeer om een tegenstelling, maar eerder om een continuüm, waarvan de extremen twee tegenovergestelde eigenschappen van toetsen representeren. Aangezien een volledige presentatie van toetstypologieën ons te ver zou leiden, zullen we ons hier beperken tot het bespreken van die toetscategorieën die in onze

specifieke situatie relevant werden gevonden (voor meer informatie, zie Henning 1987; Baker 1989; Hughes 1989; Davies 1990).

2.1 Toetscategorieën

Een eerste indeling van toetsen heeft als uitgangspunt het doel waarvoor ze moeten dienen. Op basis hiervan kan men geschiktheidstoetsen van vorderingentoetsen onderscheiden.

Bij een geschiktheidstoets gaat het om het selecteren van kandidaten voor een bepaald doel. Het doel kan verschillend zijn, bijvoorbeeld het volgen van een opleiding of het krijgen van een baan. Bij de selectie kan men uitgaan van wat de kandidaten moeten kunnen en alleen degenen kiezen die dat kunnen, of men kan beslissen om de beste kandidaten uit de groep te nemen.

Een vorderingentoets dient om vast te stellen of iemand geleerd heeft wat in een cursus behandeld is. Het zal duidelijk zijn dat een vorderingentest afhankelijk is van een bepaald curriculum, terwijl geschiktheidstests over het algemeen curriculum-onafhankelijk zijn.

Een tweede indeling van toetsen kan worden gemaakt op basis van het type informatie dat ze opleveren. Bij sommige toetsen gaat het vooral om het bepalen van de positie die een kandidaat inneemt ten opzichte van andere kandidaten. De informatie die uit de toets komt, wordt uitgedrukt in percentielscores. De resultaten van de toets zeggen nauwelijks iets over wat de kandidaat kan. Het interpreteren van de toetsresultaten is pas mogelijk nadat de toets onderworpen is aan een vrij grootschalig onderzoek dat het mogelijk maakt om normen te bepalen voor de relevante populatie (normeringsonderzoek). In de Engelstalige literatuur worden deze toetsen aangeduid als 'norm-referenced' en de meting wordt ook wel eens relatief genoemd in tegenstelling tot de absolute meting die bij 'criterion-referenced' toetsen gebruikt wordt (Hudson & Lynch 1984). In het Nederlands zou men kunnen spreken van

‘normgerelateerde’ respectievelijk ‘criteriumgerelateerde’ toetsen.

Bij normgerelateerde toetsen is het heel belangrijk om goed te differentiëren tussen toetsdeelnemers. Deze toetsen zijn ook vaak bedoeld om de sterkste of de zwakste kandidaten uit een groep te selecteren. Door de verschillen tussen de deelnemers zo duidelijk mogelijk te maken, wordt de kans groter dat men een goed gefundeerde selectie maakt. Daarom streeft men in deze toetsen naar een maximale spreiding van de scores.

Bij criteriumgerelateerde toetsen is het daarentegen niet noodzakelijk om een normeringsonderzoek uit te voeren om de toetsresultaten te kunnen interpreteren. Men kan al conclusies trekken op basis van de score van een kandidaat. De score wordt namelijk gerelateerd aan het criterium, dat wil zeggen aan de parameter die men met de toets wil meten (Hudson & Lynch 1984). De score op een criteriumgerelateerde toets moet een beschrijving kunnen geven van wat de kandidaat kan of van wat hij van een bepaalde stof beheerst. Een goede definitie van criteriumgerelateerde toetsen wordt door Popham (1978: 93) gegeven: ‘A criterion-referenced test is used to ascertain an individual’s status with respect to a well-defined behavioral domain’.

Het is duidelijk dat in dit soort toetsen het verkrijgen van een grote spreiding van de scores niet het eerste doel is. De items worden dan ook niet gekozen op basis van hun discriminatievermogen, maar op basis van hun relevantie ten opzichte van het criterium dat in de toets centraal staat. De drempel die men moet halen om te kunnen slagen wordt in dit type toetsen bepaald op basis van een analyse van de minimumvoorwaarden waaraan men zou moeten voldoen.

Zowel criterium- als normgerelateerde toetsen hebben hun voor- en nadelen. In zekere zin zijn de uitkomsten van een criteriumgerelateerde toets makkelijker te interpreteren dan die van een normgerelateerde toets, omdat de doelstellingen concreet en zinvol zijn. Als iemand slaagt dan weet men ongeveer wat hij/zij kan. Bij

normgerelateerde toetsen is dat niet altijd het geval. Anderzijds is het zo dat juist omdat er bij criteriumgerelateerde toetsen in principe geen vergelijking met andere kandidaten vereist is, het soms moeilijker kan zijn om de toetsuitslag te interpreteren. Met andere woorden, men weet niet altijd precies wat het betekent om de drempel te bereiken. Verder is het duidelijk dat de keuze tussen de twee soorten toetsen voor een groot deel zal afhangen van het doel waarvoor getoetst wordt (Carrol 1982: 101). Voor een uitgebreidere discussie van voor- en nadelen van deze en de later te bespreken toetscategorieën verwijzen we naar de relevante literatuur (Popham 1978; Henning 1987; Baker 1989; Hughes 1989; Davies 1990).

Een derde onderscheid kan worden gemaakt tussen discrete en integratieve toetsen. Met discrete toetsen kan men afzonderlijke elementen toetsen, bijvoorbeeld specifieke grammaticale structuren zoals de verleden tijd van werkwoorden of het meervoud van zelfstandige naamwoorden. Integratieve toetsen, zoals een interview of een dictee, geven daarentegen een globaler beeld van iemands taalvaardigheid. Het is belangrijk om op te merken dat het discreet of integratief karakter van de toets niet alleen bepaald wordt door de handeling die wordt uitgevoerd, maar ook door de manier van scoren. Bijvoorbeeld, een in principe integratieve toets zoals een interview kan discreet worden gemaakt als de beoordelaar alleen op het gebruik van de verleden tijd let.

Discrete toetsen waren vooral populair onder de voorstanders van de psychometrische benadering (Lado 1961). Deze auteurs gingen ervan uit dat taal uit verschillende elementen bestaat, die in isolatie geleerd en getest kunnen worden. Het onderscheid in de vier deelvaardigheden lezen, luisteren, schrijven en spreken stamt uit deze traditie.

In de jaren '70 kreeg de psychometrische benadering zware kritiek van voorstanders van de Unitary Competence Hypothesis (Oller 1979). Volgens deze onderzoekers vormt taalvaardigheid een groot

geheel dat niet in afzonderlijke deelvaardigheden kan worden opgedeeld. Waar ze zich vooral tegen verzetten, was het toetsen van aparte elementen in afwezigheid van context. Hun standpunt was dat context juist heel belangrijk is in normale communicatieve situaties, soms nog belangrijker dan kennis van de afzonderlijke elementen.

Het type toetsen dat de aanhangers van de UCH voorstelden, zijn integratieve toetsen zoals cloze en editingtests. Een cloze test bestaat uit een tekst waarin enkele woorden zijn verwijderd. De kandidaat moet dan de ontbrekende woorden toevoegen om de tekst te vervolledigen. Een editingtest, daarentegen, wordt verkregen door twee teksten met elkaar te combineren. Men neemt de eerste tekst als uitgangspunt en in die tekst worden op een willekeurige manier woorden uit de tweede tekst toegevoegd. De taak van de kandidaat is dan de tekst te herstellen door de overtollige woorden te schrappen. Het is duidelijk dat met deze toetsen meerdere aspecten van taalvaardigheid tegelijkertijd worden aangesproken.

Cloze en editingtests bleken echter ook hun voor- en nadelen te hebben, zoals alle andere toetsen. Bovendien vond men dat de uitgangspunten van Oller en zijn medewerkers niet echt goed gefundeerd waren (Vollmer 1983). De unidimensionaliteit van taalvaardigheid bleek namelijk onhoudbaar te zijn, zoals Oller zelf moest toegeven (Oller 1983).

Ondanks deze ontwikkelingen, moet men erkennen dat Oller en de zijnen een belangrijke bijdrage hebben geleverd aan het onderzoek naar taalvaardigheid. Hun grootste verdienste is dat ze gewezen hebben op de tekortkomingen van discrete toetsen, een techniek die niemand tot op dat moment in twijfel had getrokken. Met hun kritiek hebben ze duidelijk gemaakt dat taalvaardigheid meer is dan de som van de afzonderlijke elementen en dat kennis van de elementen geenszins garandeert dat men ook in staat is die elementen te gebruiken in concrete situaties (zie ook Morrow 1979).

Verder kunnen toetsen worden onderscheiden langs het

continuüm direct/indirect. Er is sprake van directe toetsing wanneer de kandidaat gevraagd wordt om te doen wat uiteindelijk gemeten moet worden. Bijvoorbeeld, iemand laten spreken als het gaat om het meten van uitspraak. Bij indirecte toetsing, daarentegen, gaat het om het meten van kennis of vaardigheden waarvan men weet of vermoedt, soms op basis van een theorie, dat ze goede indicatoren zijn voor de criteriumprestatie. Wat gedaan wordt in de toetssituatie is niet iets dat men onder normale communicatieve omstandigheden zou moeten doen. Om het voorbeeld van uitspraaktoetsing nog eens te gebruiken, men zou uitspraak kunnen toetsen door de kandidaat te vragen of paren van woorden met elkaar rijmen of niet (Lado 1961).

Het belang van directe toetsen werd vooral onderstreept door de voorstanders van communicatieve toetsen (zie Morrow 1979; Carrol 1982). Volgens deze onderzoekers zijn toetsen zoals cloze en editingtests ook niet geschikt voor het meten van communicatieve vaardigheid omdat ze te indirect van aard zijn. Bijvoorbeeld, ze vonden dat een cloze toets niet veel kan zeggen over iemands vermogen om te communiceren, omdat het toevoegen van ontbrekende woorden niet iets is dat men onder normale omstandigheden zou doen. Enerzijds zijn directe toetsen inderdaad te prefereren boven indirecte toetsen omdat ze makkelijker te construeren zijn en vanwege de grotere overeenkomst tussen toetssituatie en werkelijkheid. Anderzijds zijn resultaten van directe toetsen minder generaliseerbaar dan die van indirecte toetsen, omdat ze meer beperkt zijn tot de specifieke situatie of vaardigheid die getoetst is.

Een vijfde dimensie waarlangs toetsen onderscheiden kunnen worden, betreft het meer of minder subjectieve karakter van de meting. Het verschil tussen objectieve en subjectieve toetsen heeft inderdaad niet zozeer betrekking op de inhoud van de toets, maar op de manier van scoren. Meerkeuzevragen zijn een typisch voorbeeld

van objectieve toetsing. De prestatie van de kandidaat zal door elke beoordelaar op dezelfde manier worden gescoord. Interviews en opstellen zijn daarentegen typische subjectieve toetsen. Er is namelijk een grote kans dat de manier waarop ze gescoord worden van beoordelaar tot beoordelaar zal verschillen. Omdat objectieve scoring altijd te prefereren is, probeert men in dit soort gevallen enige objectiviteit te verkrijgen door de beoordelaars eerst gezamenlijk te trainen of door van te voren overeengekomen beoordelingscriteria te gebruiken.

2.2 Criteria waaraan de toets zou moeten voldoen

De belangrijkste vereiste voor een toets die moet uitmaken of kinderen voldoende taalvaardig zijn voor het eerste leerjaar is dat hij uitgaat van de taalvaardigheidseisen die het onderwijs aan kinderen in het eerste leerjaar stelt. Het ontwikkelen van een dergelijke toets vergt in de eerste plaats een grondige analyse van het typisch schoolse taalgebruik en in de tweede plaats een selectie van die elementen van de schooltaal die noodzakelijk zijn om van het onderwijs in het eerste leerjaar te kunnen profiteren.

Een tweede belangrijk criterium bij de keuze van de toets is de visie op taal leren die aan de toets ten grondslag ligt. Tegenwoordig zijn er heel veel taaltoetsen op de markt, maar ze gaan niet allemaal van dezelfde visie uit. Zo zijn er toetsen die sterk gebaseerd zijn op het idee dat kennis van de elementen en de grammaticale regels van een taal fundamenteel is voor het leren van die taal. Dit was duidelijk niet ons uitgangspunt.

De aanpak waarvoor men hier geopteerd heeft, sluit goed aan bij de visie op taalonderwijs die door het Steunpunt NT2 gepropageerd wordt. Volgens deze visie hebben de behoeften van leerlingen in verband met het Nederlands vooral betrekking op taalvaardigheid (zie ook Jaspaert 1991). Het begrip taalvaardigheid verwijst naar het

vermogen om een taal te gebruiken in concrete communicatieve situaties. In deze zin onderscheidt taalvaardigheid zich van een andere, verwante notie, namelijk taalkennis. Deze refereert naar de kennis die iemand heeft van de kenmerken en regels van een taal.

Eén visie op taalleren gaat uit van het idee dat taalkennis in zekere zin een voorwaarde zou zijn voor taalvaardigheid. Met andere woorden, dat kennis van grammaticale regels absoluut noodzakelijk is om een taal te leren spreken en verstaan (voor voorbeelden, zie McLaughlin 1985 en Carter & McCarthy 1989). Onderzoek naar taalverwerving heeft echter uitgewezen dat kennis van grammaticale regels niet altijd leidt tot betere beheersing van een taal voor communicatieve doeleinden of, anders gezegd, 'it is the *use* of language and not just the possession of the tools of *usage* that ultimately decides the user's effectiveness in handling communicative situations' (Carrol 1982: 11; voor meer informatie over het verschil tussen taalvaardigheid en taalkennis, zie ook Ellis 1990: 174-198).

In een onderwijssituatie is het duidelijk dat leerlingen het Nederlands in de eerste plaats nodig hebben om zich verstaanbaar te maken en om te kunnen begrijpen wat ze lezen en wat tegen hen gezegd wordt. Vooral dit laatste, receptieve aspect blijkt cruciaal te zijn in het onderwijs. Productieve vaardigheden zijn zeker ook belangrijk, maar het is duidelijk dat receptieve vaardigheden eerder aangesproken worden. Voordat een leerling gebruik maakt van zijn productieve vaardigheden om te laten zien dat hij iets geleerd heeft, moet hij een beroep doen op zijn receptieve vaardigheden om de boodschap van de leerkracht te kunnen begrijpen (Jaspaert 1991: 3-4). In aansluiting hierbij heeft men zich bij het zoeken naar een geschikte toets vooral gericht op toetsen met een sterk accent op receptieve vaardigheden. Gezien de nadruk op taalvaardigheid werd bovendien de voorkeur gegeven aan integratieve toetsen boven discrete toetsen.

De laatste eis waaraan de toets zou moeten voldoen is van praktische aard. Om organisatorische problemen te voorkomen en ook om tijd te besparen, werd gekozen voor een toets die niet individueel moet worden afgenomen.

2.3 Welke toetsen voldoen aan de gestelde eisen?

Aan de hand van bovengenoemde criteria is voor een aantal bestaande toetsen nagegaan of ze geschikt waren voor het doel dat we voor ogen hadden. Na een analyse van deze toetsen moest helaas geconstateerd worden dat er geen enkele was die aan de gestelde eisen voldeed. Het is echter niet onze bedoeling om hier alle toetsen te gaan bespreken die mogelijk in aanmerking kwamen. We zullen ons daarom beperken tot het aangeven van de belangrijkste redenen waarom de toetsen niet geschikt bleken en tot het noemen van een paar van deze toetsen als voorbeeld.

Ten eerste bleek dat bijna alle toetsen die voorhanden zijn niet de taalvaardigheidseisen in het onderwijs als uitgangspunt hebben. Er zijn bijvoorbeeld wel toetsen waarmee nagegaan kan worden of kinderen aan het begin van het basisonderwijs bepaalde belangrijke begrippen beheersen, zoals de Begrippentoets (Verhoeven 1982) van het Cito. Bij deze toets is het echter niet helemaal duidelijk hoe de getoetste begrippen geselecteerd zijn. Met andere woorden, er wordt niet precies uitgelegd of deze begrippen geselecteerd zijn op basis van een analyse van de eisen die gesteld worden bij de aanvang van lezen en rekenonderwijs. Verder is het ook niet duidelijk of deze toets zomaar gebruikt kan worden in de Vlaamse situatie.

Een andere toets die in overweging werd genomen is de Taaltoets Allochtone Kinderen. In verband met deze toets moet echter opgemerkt worden dat hij zich niet richt naar schoolse taalvaardigheden, maar juist naar 'basale taalvaardigheden (BICS)' (Verhoeven & Vermeer 1989: 34). Verder leek de TAK ook niet

geschikt omdat deze toets individueel moet worden afgenomen, wat veel tijd in beslag neemt. Bovendien geldt ook voor deze toets dat hij niet genormeerd is voor Vlaamse leerlingen. Wel is het zo dat Brusselse normen bestaan voor deze toets, maar die gelden niet voor de Vlaamse situatie in het algemeen.

De twee bovengenoemde toetsen leken ook niet geschikt omdat ze eerder discreet van aard zijn, terwijl we juist op zoek waren naar een meer integratieve toets. Hetzelfde geldt ook voor een andere toetsbatterij die nogal vaak gebruikt wordt, de Taaltests Voor Kinderen (TVK, Van Bon 1982). Aangezien er geen toets kon worden gevonden die aan de eisen voldeed, werd er uiteindelijk besloten om over te gaan tot het ontwikkelen van een nieuwe toets die geschikt zou zijn voor het beoogde doel. In de volgende paragraaf bespreken we de verschillende fasen in de constructie van deze toets.

3. Het ontwikkelen van een nieuwe toets: de TAL

3.1 Onderzoek naar taalvaardigheidseisen

Conform de criteria die in 2.2 beschreven zijn, is er geopteerd voor een toets van de receptieve schoolse taalvaardigheid waarover kinderen zouden moeten beschikken om in het eerste leerjaar mee te komen. Daarvoor was het nodig om een representatieve steekproef van die schoolse taal te verkrijgen. Met het oog hierop is er ten eerste een onderzoek uitgevoerd naar de taalvaardigheidseisen die in het eerste leerjaar worden gesteld.

In het kader van dit onderzoek werd een beschrijving gemaakt van de taal waarmee kinderen in schoolboeken geconfronteerd worden in het eerste leerjaar en van het mondelinge taalgebruik van 16 leerkrachten uit het eerste leerjaar van 16 verschillende scholen. De scholen werden geselecteerd op basis van hun geografische

ligging, het net waartoe ze behoren en het percentage allochtonen (voor verdere informatie hieromtrent, zie Jaspaert & Schrooten 1995). De leerkrachten werd gedurende een halve klastag een microfoontje opgespeld, zodat alles kon worden opgenomen wat ze tegen de kinderen zeiden. Deze gegevens werden achteraf getranscribeerd en geanalyseerd.

Deze analyse leverde een woordfrequentielijst op, die ongeveer 185.000 tokens bevat. Uit deze lijst zijn volgens bepaalde criteria (voor meer details hierover, zie Jaspaert & Schrooten 1995) ongeveer 1.400 woorden geselecteerd, die typisch geacht kunnen worden voor de taal die in het eerste leerjaar gebruikt wordt. Deze 1.400 woorden zijn ingedeeld in de volgende drie categorieën:

1. woorden met frequentie van voorkomen hoger dan 1000
2. woorden met frequentie van voorkomen tussen 1000 en 50
3. woorden met frequentie van voorkomen tussen 50 en 10

Vervolgens heeft elke deelnemer aan de Werkgroep 50 woorden uit elke categorie geselecteerd, op basis van de volgende criteria:

- woordsoort:

- a. substantieven
- b. werkwoorden
- c. voegwoorden
- d. adjectieven
- e. bijwoorden
- f. voornaamwoorden
- g. voorzetsels

- de conceptuele relevantie van het begrip in (één van) de volgende drie sferen:

- a. eigen lichaam
- b. de directe omgeving
- c. de wereld daarbuiten

- de mogelijkheid om het begrip samen met een context weer te geven in een tekening

De verschillende keuzes van de deelnemers zijn daarna besproken en na overleg is een eerste steekproef van 150 woorden samengesteld, die uiteindelijk in een proefonderzoek gebruikt is.

3.2 Toetsconstructie

Na een uitvoerige analyse van de voor- en nadelen van verschillende toetsformaten, is er gekozen voor een combinatie van voorgelezen teksten en tekeningen. Voor elk van de geselecteerde woorden werd geprobeerd een item te construeren voor de toets. Daarbij was het niet de bedoeling dat het woord zelf het item was, maar dat het gebruikt werd als aanzet om een natuurlijke taalcontext te creëren waarin dat woord zou kunnen voorkomen. Op die wijze is ernaar gestreefd stukjes taalaanbod te verzamelen die vergeleken kunnen worden met het doorsnee taalaanbod in het eerste leerjaar. Door aan elk stukje taalaanbod een vrij simpele opdracht te verbinden, kunnen we nagaan of een leerling in staat is het stukje taalaanbod te verwerken. De leerlingen kunnen hun antwoord op elk item kenbaar maken door een kruisje te zetten op een gedeelte van een tekening of op één van de plaatjes die bij de tekst horen (voor meer details, zie paragraaf 3.5).

We kunnen nu proberen deze keuze langs de bovengenoemde vijf dimensies te plaatsen. Aangezien de toets bedoeld is om vast te stellen of kinderen over voldoende taalvaardigheid beschikken om in het eerste leerjaar mee te kunnen, is het duidelijk dat het om een geschiktheidstoets gaat: men wil namelijk weten of kinderen in talig opzicht 'geschikt' zijn voor het eerste leerjaar.

Ten tweede is er het onderscheid tussen normgerelateerde en criteriumgerelateerde toetsen. Aan de ene kant is de TAL een

criteriumgerelateerde toets omdat de bedoeling is vast te stellen of kinderen bepaalde taaltaken kunnen uitvoeren waarvan men denkt dat ze belangrijk zijn om in het eerste leerjaar mee te kunnen. Om welke taken het gaat, is bepaald op basis van een analyse van de taalvaardigheid die van kinderen in het eerste leerjaar wordt verwacht. De doelstellingen zijn dus concreet en zinvol.

Aan de andere kant is de TAL aan een vrij grootschalig onderzoek onderworpen (meer dan 3000 leerlingen) dat het mogelijk zal maken de scores van veel kinderen met elkaar te vergelijken, zoals bij normgerelateerde toetsen het geval is. Op basis van dit onderzoek kon bijvoorbeeld al een begin worden gemaakt met het bepalen van drempels, zoals beschreven in paragraaf 6.1. De waarde van deze drempels zal verder aan de hand van het validiteitsonderzoek (zie paragraaf 6.2) kunnen worden getoetst. Met andere woorden, uit dit onderzoek zal blijken wat een bepaalde taalvaardigheidsdrempel inhoudt voor het functioneren in het eerste leerjaar.

Op basis van bovengenoemde kenmerken, lijkt de hier voorgestelde toets een goed compromis tussen criterium- en normgerelateerde toetsen. De TAL heeft namelijk wel de voordelen van criteriumgerelateerde toetsen, maar niet de nadelen ervan. Een dergelijke combinatie wordt soms ook in de literatuur aangeraden (Carrol 1982: 10).

Ook voor wat betreft het onderscheid direct/indirect is het niet makkelijk om de TAL uitsluitend aan één van deze twee categorieën toe te wijzen. Bovendien hebben we al eerder opgemerkt dat het bij vele van deze classificaties niet om een dichotomie gaat, maar om een continuüm. Dit betekent dat een toets niet per se direct of indirect hoeft te zijn, maar ergens tussen in kan liggen. Dit blijkt het geval te zijn voor de TAL. Aan de ene kant is het zo dat de kinderen gevraagd wordt iets te doen dat ze ook onder normale omstandigheden op school moeten doen, namelijk een boodschap

begrijpen. Aan de andere kant is de manier waarop de kinderen hun antwoord kenbaar kunnen maken, namelijk een (deel van de) tekening aankruisen, niet iets dat hun onder normale omstandigheden gevraagd wordt. Vanwege dit laatste kenmerk zou de TAL eerder als een indirecte toets beschouwd kunnen worden.

Verder is er eerder sprake van integratief dan van discreet toetsen. De te toetsen woorden worden niet in isolatie aangeboden, maar zijn ingebed in een context. Het gaat dus niet om het herkennen van afzonderlijke elementen, maar om het begrijpen van verbanden en verhoudingen tussen de elementen en hun context. Omdat kinderen altijd één plaatje uit meerdere plaatjes moeten kiezen, lijkt dit op een meerkeuzevragentoets. Het gaat dus om een objectieve vorm van toetsing, aangezien geen subjectieve mening van beoordelaars wordt gevraagd.

Een laatste punt dat genoemd moet worden in verband met de TAL betreft de mogelijke partijdigheid van deze toets. Er is sprake van bias of partijdigheid van een toets als de te meten parameter systematisch overschat of onderschat wordt als functie van het behoren tot een bepaalde groep. Dit kan het geval zijn wanneer het oplossen van de toetsopgaven andere kennis vereist dan die welke met die specifieke toets gemeten wordt en wanneer niet alle groepen over die extra kennis in vergelijkbare mate beschikken (Uiterwijk 1994: 6-7).

De partijdigheid van een toets kan verschillende oorzaken hebben. Er is bijvoorbeeld sprake van culturele bias wanneer voor het beantwoorden van de toetsvragen kennis van een bepaalde cultuur cruciaal is, terwijl die kennis niet de te meten parameter is. Aangezien het de bedoeling is om de TAL af te nemen bij alle kinderen in het eerste leerjaar, lijkt het nu legitiem om zich af te vragen of deze toets ook geschikt is voor allochtone kinderen. Met andere woorden, men zou kunnen nagaan of de TAL niet cultureel gebiased is.

De kwestie rondom de partijdigheid van toetsen heeft de laatste jaren veel aandacht gekregen, juist als gevolg van de aanwezigheid van grote contingenten allochtone kinderen in het onderwijs (Extra & Verhoeven 1985, Driessen 1990, Klatter-Folmer & Vallen 1995). Men is zich namelijk gaan afvragen of de toetsen die tot op dat moment in het onderwijs gehanteerd werden ook wel geschikt waren om kinderen te toetsen die een heel andere culturele achtergrond hebben. Dit is natuurlijk een heel belangrijk aspect waarmee altijd rekening moet worden gehouden bij het kiezen van een toets. Afhankelijk van het doel van een toets kunnen de effecten van culturele bias meer of minder sterk zijn.

Ook in verband met de TAL is het belangrijk om te kijken naar het doel van de toets. Zoals al eerder opgemerkt werd, is de TAL bedoeld om vast te stellen of kinderen in talig opzicht voldoende uitgerust zijn om in het eerste leerjaar mee te kunnen. Het is dus logisch dat deze toets een zo goed mogelijke representatie van de taalvaardigheidseisen in het eerste leerjaar moet geven. Hoe groter de overeenkomst tussen de twee, hoe beter de signaalfunctie van de toets. Op basis van de manier waarop de toets geconstrueerd is, mogen we aannemen dat de TAL inderdaad een goede weergave is van de taalvaardigheidseisen in het eerste leerjaar. Als nu zou blijken dat in die eisen een zekere mate van culturele bias besloten ligt, dan is het logisch dat die bias ook in de TAL zit. Bias wordt alleen een probleem als de bias in de TAL anders zou zijn dan die in de taalvaardigheidseisen in het onderwijs, maar op basis van de constructie van de TAL mogen we aannemen dat dat niet het geval is.

3.3 Eerste proefonderzoek

Op basis van de 150 geselecteerde woorden is een voorlopige versie van de toets samengesteld. De woorden zijn verwerkt in een aantal

teksten die door de proefleider worden voorgelezen. Uit deze teksten zijn 38 toetsitems gedestilleerd. Elke tekst werd voorzien van een tekening of van een aantal plaatjes die gebundeld waren in een boekje. Op basis daarvan konden de kinderen hun antwoord kenbaar maken door een kruisje te zetten op één van de elementen in de tekening of op één van de plaatjes.

Om een eerste indruk te krijgen van hoe de toets functioneert, is in het najaar 1992 deze voorlopige versie aan een proefonderzoek onderworpen. De bedoeling van deze proefafname was vooral inzicht te krijgen in de hanteerbaarheid van deze toetsvorm, de keuze van de taken, de duidelijkheid van de instructies en het al dan niet functioneren van de gekozen procedure.

3.3.1 Proefpersonen en procedure

Voor het proefonderzoek is een quotasteekproef van kinderen uit OVB-scholen gebruikt. De kinderen behoorden tot verschillende etnische groepen, namelijk Vlamingen (14), Turken (22), Marokkanen (38) en Italianen (10). Vijf kinderen hadden een andere herkomst dan de hiervoor genoemde. Het onderzoek werd gespreid over twee afnamemomenten, van 41 kinderen is de toets aan het eind van de kleuterschool afgenomen en van 48 aan het begin van het eerste leerjaar. Van elke leerling werden naam, klas, school, leeftijd, etnische herkomst en moedertaal genoteerd.

Met het oog op het verzamelen van zoveel mogelijk observatiemateriaal, werd de omvang van de proefgroep per afname beperkt tot vijf kinderen. Vóór de echte toetsafname werden enkele oefenitems aangeboden om de kinderen vertrouwd te maken met de uit te voeren taak. De toetsen werden door leden van de werkgroep afgenomen.

Zoals boven vermeld, kregen de kinderen een boekje voorgelegd waarin een aantal tekeningen stond. Bij elke tekening werd vooraf een verhaaltje voorgelezen. Op deze manier werd een context

aangebracht. Iedere tekst werd maar één keer door de proefleider voorgelezen, zonder het te testen begrip te benadrukken. De kinderen werd vervolgens gevraagd om één of meer elementen van de tekening aan te kruisen. Hoewel er geen echte tijdsdruk was, was het wel de bedoeling dat alle kinderen de bladzijden van het boekje samen zouden omslaan. Om dit mogelijk te maken, wachtte de proefleider tien seconden na elke voorgelezen tekst.

3.3.2 Resultaten

Uit de resultaten van het vooronderzoek bleek het gemiddelde van de gehele toets vrij hoog te liggen, 28,2 op 38. De behaalde scores variëren tussen 3 en 38. De betrouwbaarheid van de toets is hoog, Cronbachs alpha (zie ook 4.1) bedraagt 0,93. Ook de item-betrouwbaarheid blijkt hoog te zijn, met uitzondering van een enkel item dat uiteindelijk verwijderd werd.

De gemiddelde scores voor de twee subgroepen 3de kleuterklas en 1ste leerjaar bedragen 26,2 (SD = 9) respectievelijk 29,9 (SD = 7,6). De scores voor de verschillende nationaliteiten blijken significant te verschillen. De Turkse groep behaalt de laagste score (GEM = 19,1; SD = 9,3) en de Vlaamse de hoogste (GEM = 34,6; SD = 4,1).

Op basis van deze gegevens is het echter niet verantwoord om conclusies te trekken over het prestatieniveau van de verschillende subgroepen. Daarvoor is het aantal proefpersonen te klein. Wel heeft het proefonderzoek belangrijke informatie opgeleverd over het functioneren van de verschillende items en van de toets in zijn geheel. Sommige items, zoals bijvoorbeeld 'taart' en 'verrassing', bleken onderling verwisselbaar te zijn. Met andere woorden, veel kinderen gingen de taart aankruisen als er gevraagd werd om de verrassing aan te kruisen. Verder bleek dat sommige tekeningen niet duidelijk genoeg waren of dat de voorwerpen in de tekeningen niet herkenbaar waren. Op basis van de resultaten van het proefonderzoek werd besloten de toets enigszins aan te passen. De

herziene versie van de toets is vervolgens voor een tweede proefonderzoek gebruikt.

3.4 Tweede proefonderzoek

In het voorjaar 1993 werd er een tweede proefonderzoek uitgevoerd om na te gaan of de toets in zijn herziene vorm beter functioneerde dan de vorige versie. De herziene versie bevatte 52 items plus 8 items die dubbel getoetst werden. Het ging eigenlijk om dezelfde woorden die in twee verschillende contexten werden aangeboden. Het idee hierbij was om na te gaan of het scoren op deze items verschillend kan zijn als de context verandert terwijl het te toetsen begrip hetzelfde blijft. Verder was deze tweede proefafname bedoeld om inzicht te krijgen in de interne consistentie van de toets, de moeilijkheidsgraad en het prestatieniveau van de verschillende subgroepen.

3.4.1 Proefpersonen en procedure

De steekproef voor het tweede vooronderzoek bestond uit 324 leerlingen uit twee verschillende schooljaren (153 uit midden derde kleuterklas en 171 uit midden eerste leerjaar) en verschillende herkomstgroepen (106 Vlamingen, 81 Turken, 82 Marokkanen, 20 Italianen en 35 overigen). De toetsen werden afgenomen door PMS-medewerkers uit een tiental PMS-Centra van de verschillende netten. Ook in dit geval werden van elke leerling naam, klas, school, leeftijd, etnische herkomst en moedertaal genoteerd. Bij deze proefafname werd dezelfde procedure gebruikt als bij de eerste.

3.4.2 Resultaten

De herziene versie van de toets met 60 items bleek ook ruim betrouwbaar, Cronbachs alpha bedraagt 0,92. In Tabel 1 zijn de gemiddelde scores en standaarddeviaties voor de verschillende

etnische groepen weergegeven.

Tabel 1: Gemiddelden (GEM) voor de verschillende herkomstgroepen, standaarddeviatie (SD) en aantal proefpersonen (N)

Herkomst	GEM	SD	N
Vlamingen	54,5	5,6	106
Turken	44,8	11,4	81
Marokkanen	46,5	6,9	82
Italianen	46,3	11,0	20
Overigen	50,4	6,1	35

Als de gegevens voor de verschillende etnische groepen uitgesplitst worden naar leerjaar, dan blijkt het verschil tussen 3de kleuterklas en 1ste leerjaar het grootst te zijn bij de Turkse leerlingen. In Tabel 2 zijn deze resultaten weergegeven.

Tabel 2: Gemiddelden (GEM) opgesplitst naar herkomst en klas, standaarddeviatie (SD) en aantal proefpersonen (N)

Klas	Herkomst	GEM	SD	N
3de kleuterklas	Vlamingen	53,0	5,2	55
	Turken	37,5	12,5	32
	Marokkanen	43,1	6,7	37
	Italianen	41,5	12,3	11
	Overigen	48,6	5,5	18
1ste leerjaar	Vlamingen	56,1	5,5	51
	Turken	49,6	7,6	49
	Marokkanen	49,2	5,9	45
	Italianen	52,3	5,1	9
	Overigen	52,3	6,3	17

Zoals boven vermeld, werden er 8 van de 60 items dubbel getoetst. Om precies te zijn, acht begrippen werden aangeboden in twee verschillende contexten. Deze woorden zijn: *huilen*, *knie*, *straat*, *winkel*, *hond*, *hangen*, *onder* en *wolk*. Hoewel het kernwoord hetzelfde is in beide gevallen, zou men kunnen stellen dat het om verschillende items gaat omdat de context anders is. Dit wordt bevestigd door de resultaten. Sommige leerlingen blijken inderdaad anders te scoren afhankelijk van de context waarin het begrip wordt aangeboden. Als voorbeeld geven we in Tabel 3 de kruistabel voor het eerste van de acht begrippen, *huilen*. Het feit dat de resultaten verschillen afhankelijk van de context waarin de begrippen worden aangeboden geeft extra steun aan het idee dat deze toets niet de kennis van afzonderlijke woorden meet, maar het vermogen om woorden in hun natuurlijke context te begrijpen.

Tabel 3: Kruistabel van scores op item *huilen* in twee verschillende contexten

context 1	context 2		
	0	1	
0	14	51	65
1	16	243	259
	30	294	324

Zoals te verwachten, blijkt de context vooral voor minder taalvaardige leerlingen een belangrijke rol te spelen. Aangezien het niveau van taalvaardigheid varieert voor de onderzochte etnische groepen, blijkt het ook dat deze groepen aanzienlijk van elkaar verschillen in de manier waarop ze met de dubbel getoetste items omgaan. Vooral de allochtone groepen scoren verschillend, afhankelijk van de context waarin het te toetsen begrip wordt

aangeboden. Van de 106 Vlamingen die in de steekproef opgenomen zijn, scoren er maar twee verschillend op één item. Van de 81 Turken en 82 Marokkanen zijn er 20 respectievelijk 40 die verschillend scoren op een of meerdere items. Met andere woorden, het correct interpreteren van een bepaald begrip lijkt voor meer taalvaardige leerlingen (Vlamingen) minder afhankelijk te zijn van de context waarin het begrip wordt aangeboden dan voor minder taalvaardige leerlingen (allochtonen).

3.5 Definitieve samenstelling van de TAL

Aangezien de dubbele items alleen maar ter controle waren opgenomen, zijn ze uiteindelijk niet opgenomen in de toets. Verder werden er twee items uitgesloten (*mond* en *eigenlijk*) die een lage item/total correlatie hadden (0,17). In zijn definitieve vorm bevat de toets 23 scenario's en 50 items. Deze zijn in de volgende drie subtests ingedeeld.

Bij subtest 1, die 10 items bevat (items 1.1 tot en met 1.10), worden telkens één of twee zinnestjes voorgelezen waarin één of meer van de te toetsen woorden zijn verwerkt. Bij elk item krijgen de kinderen een blad met vier verschillende plaatjes. Na het voorlezen wordt er door de proefleider een vraag gesteld waarop de kinderen kunnen antwoorden door één van de plaatjes aan te kruisen.

Subtest 2 bevat 31 items (items 2.1.1 tot en met 2.10.5). Bij één tekening horen meerdere items. De proefleider leest een verhaaltje voor en vraagt de kinderen telkens om iets op de tekening aan te kruisen. De aan te kruisen delen op de tekeningen kunnen voorwerpen zijn of weergaven van bepaalde handelingen die overeenkomen met de te toetsen begrippen.

Ten slotte, bij subtest 3 (9 items) krijgen de kinderen weer een kort verhaaltje te horen dat bij een tekening past. Pas nadat de proefleider het verhaaltje heeft voorgelezen mogen de kinderen hun

potlood pakken en een kruisje zetten op die delen van de tekening waar ze vragen over krijgen.

Deze definitieve vorm van de TAL is uiteindelijk van een groot aantal leerlingen afgenomen in het kader van het betrouwbaarheids-onderzoek. Verdere gegevens over steekproef en procedure zijn te vinden in paragraaf 5.1.1.

4. Betrouwbaarheid en validiteit

Vooraleer een toets zoals de nieuw ontwikkelde TAL ruim ingezet kan worden, dient eerst de vraag beantwoord te worden of de toets wel betrouwbaar en valide is. Betrouwbaarheid en validiteit zijn twee belangrijke criteria waaraan elk meetinstrument, en dus ook een taalvaardigheidstoets zoals de TAL, zou moeten voldoen. Beide begrippen worden hieronder kort toegelicht.

4.1 Betrouwbaarheid

Als we een taalvaardigheidstoets bij een proefpersoon afnemen, dan willen we natuurlijk dat de behaalde score zo goed mogelijk het taalvaardigheidsniveau van de betreffende persoon aangeeft en zo min mogelijk beïnvloed wordt door andere factoren, zoals het tijdstip van afname, de temperatuur, het humeur van de proefpersoon enzovoort. Van een meetinstrument wensen wij dat bij een hypothetische herhaling van de meting hetzelfde resultaat verkregen wordt. Met andere woorden, we willen dat het instrument consistent meet.

Toch weet men dat bij een tweede afname van dezelfde toets het bijna onmogelijk is om dezelfde score te verkrijgen. Dit heeft te maken met het feit dat allerlei externe factoren een invloed kunnen hebben op de toetsscore. Als gevolg hiervan zal elke meting een zekere foutmarge bevatten. In de meettheorie is dit een geaccepteerd

feit.

Men gaat ervan uit dat elke score uit twee componenten bestaat: de zogenaamde 'true score' of 'betrouwbare score' (Drenth & Sijtsma 1990: 99) en de meetfout. Aangezien elke meting een foutcomponent bevat, kan de 'betrouwbare score' nooit verkregen worden. Wel kan men deze score proberen te benaderen door een groot aantal keren dezelfde meting uit te voeren. Het gemiddelde berekend over alle verrichte metingen is dan een benadering van de 'betrouwbare score'. De veronderstelling hierbij is dat de externe invloeden elkaar tegenwerken.

Terwijl er soorten metingen zijn die zonder problemen een groot aantal keren verricht kunnen worden, denk bijvoorbeeld aan het bepalen van de lengte van een plankje met een meetlat, is dat voor veel tests niet mogelijk. Het afnemen van de test zelf heeft al een effect op de proefpersoon, zodat hij/zij bij de tweede meting niet in alle opzichten dezelfde zal zijn. Hij/zij zal bijvoorbeeld iets geleerd hebben van de toets.

Omdat het in vele situaties in de onderzoekspraktijk niet mogelijk is om de 'betrouwbare score' te benaderen, zijn er speciale technieken ontwikkeld om de grootte van de meetfout, en dus de nauwkeurigheid van een meetinstrument, te schatten. De mate van nauwkeurigheid of consistentie van een meetinstrument noemt men de *betrouwbaarheid*: hoe kleiner de meetfout hoe hoger de betrouwbaarheid. Hoe hoger de betrouwbaarheid van een meting, hoe groter de kans is dat de verkregen score een correcte weergave is van de getoetste parameter en niet een door toeval bereikt resultaat.

Over het algemeen wordt betrouwbaarheid bepaald volgens vier verschillende methodes, de parallelvorm- en de test-hertest-methode, die beide uitgaan van herhaalde metingen, en de splitsings- en de interne-consistentie-methode, die op een enkele meting gebaseerd zijn. In het onderzoek dat hier gerapporteerd wordt, is de interne-consistentie-methode gehanteerd, waarvan Cronbachs alfa de

bekendste coëfficiënt is. Deze methode is efficiënt omdat men, net als bij de splitsingsmethode, met één meting kan volstaan. Het voordeel boven de splitsingsmethode is echter dat op deze manier alle problemen met het verkrijgen van parallelle testhelften voorkomen kunnen worden, omdat aangetoond is 'dat alfa gelijk is aan de gemiddelde splitsingsbetrouwbaarheid, berekend over alle mogelijke splitsingen van een test in twee helften' (Drenth & Sijtsma 1990: 112).

Deze korte toelichting van het begrip betrouwbaarheid en andere kenmerken was bedoeld om meer duidelijkheid te geven over de relevantie van deze notie voor de toetspraktijk, met het oog op het bespreken van het betrouwbaarheidsonderzoek in paragraaf 5.1 (voor verdere details, zie Drenth & Sijtsma 1990). Voordat we overgaan tot dit aspect van de toetsevaluatie, is het zinvol om in de volgende paragraaf nog even stil te staan bij een tweede belangrijk begrip in de psychometrie, namelijk validiteit.

4.2 Validiteit

Het feit dat een toets betrouwbaar is, is natuurlijk een heel wenselijk resultaat, maar het is nog niet voldoende om te weten of een toets wel geschikt is voor een bepaalde doelstelling. Van een toets willen we namelijk niet alleen dat hij een bepaald verschijnsel consistent meet, maar ook dat hij dat verschijnsel meet dat we met die specifieke toets van plan zijn om te meten. Een betrouwbare taalvaardigheidstoets die geen taalvaardigheid blijkt te meten, maar het kunnen omgaan met meerkeuzevragen, zou natuurlijk niet geschikt zijn voor onze doelstelling.

De mate waarin een toets daadwerkelijk meet wat hij pretendeert te meten noemt men de *validiteit* van de toets. Het bepalen van de validiteit van een toets is minstens even belangrijk, zo niet belangrijker, als het bepalen van de betrouwbaarheid. Strikt genomen

is validiteit belangrijker omdat validiteit ook betrouwbaarheid impliceert. Men zegt inderdaad dat een toets valide is als hij consistent meet wat men wil meten. Validiteit impliceert dus betrouwbaarheid maar niet andersom of, anders gezegd, betrouwbaarheid is een noodzakelijke maar niet voldoende voorwaarde voor validiteit.

In de psychometrie onderscheidt men verschillende soorten validiteit, die voor een deel gerelateerd zijn aan het doel van de toets. We zullen ons hier beperken tot het soort validiteit die relevant is voor een toets als de TAL, namelijk inhoudelijke en predictieve validiteit (voor verdere informatie, zie Cureton 1951 en Drenth & Sijsma 1990)

Inhoudelijke validiteit heeft te maken met de vraag of de inhoud van de toets representatief en compleet genoeg is om te meten wat gemeten moet worden. Bij curriculumafhankelijke toetsen kan inhoudelijke validiteit bereikt worden als er een goede selectie gemaakt wordt van wat in het curriculum aan bod komt. Bij geschiktheidstoetsen, daarentegen, is de toets inhoudelijk valide als hij een representatieve steekproef bevat van de basisvaardigheden die nodig zijn voor wat men wil meten. Het is belangrijk om op te merken dat inhoudelijke validiteit niet empirisch vastgesteld kan worden: er is nu eenmaal geen coëfficiënt van inhoudelijke validiteit (Henning 1987: 94). Het bepalen daarvan berust op het oordeel van deskundigen. Op deze manier is ook de inhoudelijke validiteit van de TAL bepaald, zoals in paragraaf 6.1 duidelijk zal worden.

Onder predictieve validiteit verstaat men de mate waarin voorspellingen gedaan op basis van de toetsprestaties, bevestigd worden door observaties op een later tijdstip. Het bepalen van predictieve validiteit is belangrijk voor toetsen die pretenderen iets te zeggen over toekomstige prestaties. Typische voorbeelden van dit soort toetsen zijn de toelatingsexamens voor bepaalde opleidingen. Het doel van zulke examens is uitspraken te doen over de

hypothetische toekomstige prestaties van de kandidaten als ze toegelaten zouden worden tot die opleiding. Men gaat ervan uit dat iemand die te laag scoort op het toelatingsexamen niet in staat zal zijn de opleiding met succes te volgen.

De voorspellende waarde (predictieve validiteit) van dergelijke uitspraken, en dus van het examen, kan nagegaan worden door de examenuitslagen te vergelijken met gegevens over de prestaties van de kandidaten gedurende de opleiding. Een hoge positieve correlatie tussen de twee soorten data zal dan betekenen dat het examen een goede voorspeller is van succes in die bepaalde opleiding.

5. De TAL als taalvaardigheidstoets

5.1 Betrouwbaarheid

5.1.1 Proefpersonen en procedure

De steekproef voor het betrouwbaarheidsonderzoek bestond in totaal uit 3018 kinderen uit een 100-tal scholen verspreid over heel Vlaanderen (met uitzondering van West-Vlaanderen). Bij 2564 kinderen werd de TAL tijdens de laatste maand van de kleuterschool of de eerste maand van het eerste leerjaar afgenomen. Verder werden er ook 418 kinderen aan het begin van de derde kleuterklas getest. De steekproef bevatte zowel autochtone als allochtone leerlingen. De allochtone groep bestond uit kinderen van Turkse, Marokkaanse, Italiaanse en Joodse herkomst, plus een kleine groep kinderen met diverse etnische herkomst. De precieze onderverdeling van de steekproef naar klas en etnische herkomst is af te lezen in Tabel 4.

In de periode juni/september 1993 werden de toetsen van bovengenoemde leerlingen afgenomen door PMS-medewerkers uit een 50-tal PMS-Centra van de verschillende netten. Net als in het proefonderzoek kreeg elk kind een boekje waarin de verschillende

prenten gebundeld waren. Op elk boekje werden datum van onderzoek, klas, school, naam leerling, alfabetisch nummer, geboortedatum, leeftijd, etnische herkomst en omgangstaal moeder-kind genoteerd. Vóór de eigenlijke toets werden enkele instructie-items doorgenomen om de kinderen duidelijk te maken hoe ze moesten antwoorden. Elke proefleider tekende een boom, een zon en een boot op het bord en deed voor hoe de kinderen een kruisje moesten zetten op het juiste antwoord. Vervolgens konden de kinderen zelf in hun eigen boekjes oefenen. Er werd uitleg gegeven totdat ieder kind begrepen had wat de bedoeling was.

Bij de toetsafname werd door de proefleider bij elke tekening uit het boekje een context voorgelezen op een natuurlijke verteltoon. De kinderen werd daarna verzocht om een kruisje te zetten op één van de plaatjes bij een item of op één van de elementen van de tekening die bij het item hoorde.

Elke tekst werd maar één keer aangeboden en elke opdracht werd maar één keer gegeven. Na elke tekst wachtte de proefleider tien seconden om de kinderen de mogelijkheid te geven de bladzijde samen om te slaan. Er werd gewerkt met groepjes van maximaal vijf kinderen per proefleider.

5.1.2 Resultaten

De TAL blijkt in hoge mate betrouwbaar te zijn (Cronbachs alfa=0,94). Ook blijkt dat elk item tot die betrouwbaarheid bijdraagt, zodat de toets in zijn geheel als betrouwbaar instrument kan worden geaccepteerd. Deze betrouwbaarheid geldt ook voor alle subgroepen die in de steekproef onderscheiden kunnen worden, zowel naar leeftijd bij afname als naar etnische achtergrond. Alleen voor de groep Joodse leerlingen ligt de betrouwbaarheid van de toets beduidend lager (Cronbachs alfa=0,86).

De algemene gemiddelde score op de toets bedraagt 40,1. Er kunnen grote verschillen geconstateerd worden tussen de

verschillende leeftijdsgroepen en de groepen van verschillende etnische herkomst (zie Tabel 4).

Tabel 4: Resultaten op TAL per etnische groep per schoolniveau

	GEM	SD	N
Autochtonen	45,3	4,4	1256
begin derde kleuter	41,2	6,4	183
einde derde kleuter	45,5	4,0	460
begin eerste leerjaar	46,4	3,0	608
overige	47,0	2,0	5
Turkse herkomst	32,1	11,0	559
begin derde kleuter	25,5	12,2	66
einde derde kleuter	31,2	10,9	215
begin eerste leerjaar	34,4	10,2	278
Marokkaanse herkomst	37,5	9,2	586
begin derde kleuter	28,0	11,5	82
einde derde kleuter	36,7	9,6	152
begin eerste leerjaar	39,9	6,5	322
overige	41,4	8,0	30
Italiaanse herkomst	41,1	7,4	139
begin derde kleuter	35,5	9,1	35
einde derde kleuter	42,4	4,8	24
begin eerste leerjaar	43,2	5,9	80
Joodse herkomst	41,7	6,0	188
begin eerste leerjaar	41,0	6,0	188
Andere etnische herkomst	37,0	10,7	290
begin derde kleuter	28,0	13,5	52
einde derde kleuter	37,1	9,9	71
begin eerste leerjaar	39,8	8,2	166
overige	35,0	0,0	1

5.2 Validiteit

Met betrekking tot de TAL kunnen twee validiteitsvragen gesteld worden. Enerzijds claimt de TAL dat hij de taalvaardigheid meet van leerlingen bij het begin van het lager onderwijs. Men kan zich afvragen of dit wel zo is: wat wordt precies bedoeld met die taalvaardigheid, en hoe weten we dat de TAL die vaardigheid inderdaad meet (inhoudsvaliditeit). Anderzijds willen we op basis van de TAL te weten komen of leerlingen het taalvaardigheidsniveau hebben bereikt dat vereist wordt om het eerste leerjaar met profijt te volgen. Het gaat dus om het signaleren van potentiële problemen, om ervoor te zorgen dat de juiste maatregelen genomen worden ter voorkoming van die problemen.

Met het oog hierop is het natuurlijk heel belangrijk om te weten of er een samenhang bestaat tussen de TAL-scores en de mate van succes in het eerste leerjaar. De TAL heeft enkel een signaalfunctie met betrekking tot de schoolprestaties van leerlingen in het eerste leerjaar, als aangetoond kan worden dat de vaardigheid die erdoor gemeten wordt, een voorwaarde is om goed te presteren op school. In dat geval is het duidelijk dat leerlingen met onvoldoende taalvaardigheid (een te lage TAL-score) grote kans hebben om in het eerste leerjaar zwak of zelfs onvoldoende te presteren. Dit maakt het mogelijk om op tijd maatregelen te nemen om ervoor te zorgen dat hun taalvaardigheid groter wordt en dat ze daardoor een grotere kans maken om het eerste leerjaar met succes af te ronden.

Om te weten of de TAL een goede indicator is, moeten we dus weten of TAL-scores samenhangen met prestaties in het eerste leerjaar. Dit is een vraag naar predictieve validiteit: op basis van metingen die op een bepaald moment zijn verricht, worden uitspraken gedaan over prestaties op een later tijdstip. Dus naast inhoudelijke validiteit is ook predictieve validiteit relevant in verband met de TAL. Het feit dat predictieve validiteit belangrijk is, betekent

echter niet dat de TAL onafwendbaar toont dat leerlingen met een lage TAL-score verloren zijn. Integendeel, de TAL is juist ontworpen om mogelijke problemen snel op het spoor te komen om vervolgens daar iets aan te doen.

Voordat de TAL deze signaalfunctie kan vervullen, is het in ieder geval belangrijk om aan te tonen dat de TAL taalvaardigheid meet bij het begin van de lagere school, en dat die taalvaardigheid nodig is om het eerste leerjaar met succes door te komen. Het eerste punt wordt hieronder behandeld, terwijl het tweede in paragraaf 6 aan bod zal komen.

Dat de TAL meet in welke mate een leerling over de taalvaardigheid beschikt die van hem verwacht wordt in het eerste leerjaar, valt af te leiden uit de manier waarop de TAL is geconstrueerd (zie ook 3.1 en 3.2). Op basis van een onderzoek naar de taalvaardigheidseisen die in het eerste leerjaar worden gesteld, zijn stukjes taalaanbod verzameld waarvan aangenomen mag worden dat ze vergelijkbaar zijn met het taalaanbod in het eerste leerjaar. Deze stukjes taalaanbod fungeren als aanleiding voor het uitvoeren van bepaalde taaltaken. De mate waarin leerlingen in staat zijn de taken uit te voeren bepaalt de toetsprestatie. We kunnen ervan uitgaan dat de toetsprestatie van een leerling een goede indicator is voor de mate waarin hij in staat zal zijn om het taalaanbod in het eerste leerjaar te verwerken: de taak, het verwerken van taalaanbod, is vergelijkbaar en de elementen die in dat taalaanbod zitten, zijn relevant in het eerste leerjaar.

6. De TAL als signaal voor potentiële problemen

6.1 Drempelbepaling

Voor de bepaling van drempels met betrekking tot deze toets kan men verschillende richtingen uit. Men kan proberen deze drempels te bepalen op basis van toetsinterne criteria, of men kan die bepalen door gebruik te maken van zogenaamde ankers, informatie over de leerlingen die uit andere bronnen komt dan deze toets.

Met gebruikmaking van toetsinterne criteria kan men verschillende redeneringen volgen. Een eerste gaat ervan uit dat het eerste leerjaar gemodelleerd is naar de gemiddelde autochtone leerling. Dat wil zeggen dat wie goed wil kunnen functioneren in dat eerste leerjaar minimaal over de taalvaardigheid moet beschikken die typisch is voor die gemiddelde autochtone leerling. Uit Tabel 4 valt af te leiden dat die vaardigheid overeenkomt met een toetsscore van 46 (= gemiddelde score van leerlingen getoetst op einde derde kleuterklas en begin eerste leerjaar). Met andere woorden, wie geen 90% scoort op de toets, is minder vaardig in het Nederlands dan in een eerste leerjaar wordt verwacht. Dit resultaat komt vrij goed overeen met de toetsconstructie, aangezien we juist die vaardigheid hebben willen toetsen die in het eerste leerjaar wordt verwacht. Wie over die verwachte vaardigheid beschikt, wordt verondersteld de toets in principe perfect te kunnen maken. Een score van 90% als drempel beantwoordt dus aan die verwachtingen.

Men kan echter ook stellen dat een score van 90% als drempel wel vrij hoog is, en er een heleboel kinderen zullen zijn die lager scores en toch behoorlijk zullen functioneren in het eerste leerjaar. De reden hiervoor kan tweevoudig zijn. Naarmate de drempel die wordt gehanteerd dichterbij het absolute maximum van de toets ligt, wordt de kans groter dat storende factoren de toewijzing van een leerling tot een bepaalde groep (boven of onder de drempel) mee gaan

bepalen. Elke verstrooidheid kan tot gevolg hebben dat een redelijk taalvaardige leerling de drempel niet haalt. Ten tweede kan men zich afvragen of het wel waar is dat een leerling over een gemiddelde 'autochtone' taalvaardigheid moet beschikken om in het eerste leerjaar te kunnen functioneren. Men kan aanvoeren dat de aandacht voor zorgbreedte in de school ervoor kan zorgen dat ook leerlingen met een lagere taalvaardigheid nog vrij behoorlijk kunnen functioneren. Uitsluitel in deze discussie kan enkel verkregen worden door af te stappen van de toetsinterne drempelbepaling en na te gaan hoe de validiteit van een aantal drempels op een andere wijze kan worden bepaald.

Bij gebrek aan andere bruikbare taalvaardigheidstoetsen, zijn we als het ware verplicht die validiteit te gaan bepalen door middel van een follow-up onderzoek. In dat onderzoek volgen we een aantal leerlingen van wie we eerder de taalvaardigheid hebben vastgesteld in het eerste leerjaar. Dit onderzoek moet duidelijk maken wat het verband is tussen een bepaalde drempel qua taalvaardigheid enerzijds en het functioneren in het eerste leerjaar anderzijds. Dergelijk onderzoek werd in samenwerking met de PMS-Centra van de drie onderwijsnetten opgezet in het schooljaar 1993-1994. Hierover wordt in paragraaf 6.2 gerapporteerd.

Voorlopig kan men echter de redenering naar de te hoge drempel van 90% ter harte nemen door zich wat voorzichtiger op te stellen. Een mogelijke redenering is dat men niet al te ver onder dat gemiddelde niveau mag vallen. In dat verband zijn weerom verschillende berekeningswijzen mogelijk. Wij hebben ervoor geopteerd te stellen dat men minstens een niveau moet halen dat overeenkomt met het gemiddelde niveau van de autochtone leerling min een standaardafwijking voor die leerling. In dit geval komen we dan uit bij een minimumscore van 42, of de vaststelling dat men niet onder de 85% mag scoren.

Welke van de twee normen men wil gebruiken, hangt voorlopig af van wat men wil doen. Wil men ervoor zorgen dat men absoluut zeker alle leerlingen op het spoor komt die moeilijkheden kunnen krijgen op basis van hun taalvaardigheidsniveau, en neemt men daarbij op de koop toe dat men ook een groep potentiële probleemgevallen selecteert die er in werkelijkheid geen zullen blijken te zijn, dan moet men de 90% drempel hanteren. Wil men daarentegen enkel een groep leerlingen selecteren die met grote waarschijnlijkheid problemen zullen kennen, dan kan men best de lagere drempel gebruiken.

Het validiteitsonderzoek moet in deze zaak meer duidelijkheid brengen. Niet alleen moet daaruit meer inzicht voortkomen in het belang van de verschillende drempels, ook de wijze waarop die drempels zich verhouden tot de verschillende groepen die kunnen worden onderscheiden, zal duidelijk gemaakt moeten worden. Het is immers niet zonder meer gegeven dat elk van die drempels precies hetzelfde betekent voor Turkse als voor Marokkaanse leerlingen, voor kinderen getoetst op het einde van de derde kleuterklas als voor kinderen getoetst bij het begin van het eerste leerjaar.

In elk geval geeft het uitgevoerde betrouwbaarheidsonderzoek ons al voldoende grond om te stellen dat we met een goede taalvaardigheidsmeting te maken hebben. Uitgaande van het principe dat kinderen die over de verwachte vaardigheid beschikken in theorie in staat moeten zijn om 100% te scoren, blijkt ook dat grote groepen leerlingen dit niveau niet halen.

6.2 Validiteit

De TAL claimt de taalvaardigheid te meten van leerlingen bij het begin van het lager onderwijs. Daarbij gaat het niet om om het even welke vaardigheid. De TAL wil de vaardigheid meten waarvan aangenomen kan worden dat men ze nodig heeft om in het eerste

leerjaar te functioneren. Met het oog hierop werd in de loop van het schooljaar 1993-1994 een follow-up onderzoek uitgevoerd.

Op basis van de resultaten op de TAL uit het betrouwbaarheids-onderzoek werd een quotasteekproef van 203 leerlingen uit het eerste leerjaar getrokken. Bij de selectie van de leerlingen werd rekening gehouden met hun TAL-score, hun etnische herkomst, de geografische ligging van hun school, het net waartoe de school behoort en de TAL-scores van de andere leerlingen in hun klas (voor verdere details omtrent de selectieprocedure, zie Van Berkel 1995). Op deze wijze werden 35 scholen geselecteerd waarvan telkens een klas in de loop van het eerste leerjaar gevolgd zou worden. Naast de geselecteerde 203 leerlingen werden ook hun klasgenoten gevolgd. Dit resulteerde in een steekproef van in totaal 526 leerlingen.

Van deze leerlingen werden in januari, maart en juni toetsen afgenomen. In januari ging het enkel om een toets technisch lezen, in maart werden er behalve een toets technisch lezen, ook een spellingtoets en twee rekentoetsen afgenomen. In juni, ten slotte, werden aan de toetsbatterij ook nog twee toetsen begrijpend lezen toegevoegd. Het gaat om toetsen uit het Leerlingvolgsysteem van het Cito (voor verdere informatie hieromtrent, zie Van Berkel 1995). De toetsen werden afgenomen door PMS-medewerkers uit 28 PMS-Centra uit de verschillende netten. Naast de toetsgegevens werd informatie verzameld over de frequentie waarmee elke leerling de kleuterschool had bezocht, de inschatting van de leerling door de leerkracht van het eerste leerjaar, en het onderwijsprogramma dat de leerling in het eerste leerjaar te verwerken kreeg.

Uit het onderzoek bleek dat de toetsprestaties van een leerling op elk meetmoment een sterke onderlinge samenhang vertoonden. Er was dus duidelijk veeleer sprake van een soort algemene leerprestatie, dan van verschillende prestaties op taal, rekenen, enzovoort. We hebben dan ook besloten om door middel van factoranalyse een maat voor die algemene leerprestatie te bepalen. Op het voorspellen van

deze algemene maat voor leerprestatie werd het onderzoek verder afgestemd. In wat volgt zal naar deze maat verwezen worden met PRESTATIE. De resultaten van de principale-componentenanalyse zijn in Tabel 5 weergegeven.

Tabel 5: Resultaten van de principale-componentenanalyse

Variabele	Communaliteit	Factor	Eigenvalue	% Var
Leeskaart 3	0,48754	1	3,77961	63,0
Spelling	0,67247			
Rekenen	0,65860			
Sommen	0,53732			
Verwijsrelaties	0,68681			
Betekenisrelaties	0,73688			

Hoewel er drie toetsen voor technisch lezen zijn afgenomen op het derde meetmoment (Leeskaart 1, 2 en 3), is uiteindelijk alleen Leeskaart 3 opgenomen in de principale-componentenanalyse. Dit om te voorkomen dat technisch lezen een te groot gewicht zou krijgen. Bovendien bleken de drie leeskaarten hoog met elkaar te correleren (gemiddeld 0,93).

Een eerste belangrijk gegeven in verband met de relatie tussen taalvaardigheid en leerprestatie is de correlatie tussen de TAL-score en PRESTATIE. Deze correlatie bedraagt 0,46. Verder onderzoek heeft echter uitgewezen dat voor TAL-scores boven 30 Pearson r 0,53 bedraagt. Voor TAL-scores onder 30 blijken niet zozeer de afstanden tussen leerlingen, maar vooral hun rangorde sterk samen te hangen met PRESTATIE. Dit blijkt onder andere uit het feit dat de correlatie hoger blijkt te zijn wanneer de TAL-scores omgecodeerd worden naar een ordinale variabele met drie categorieën (0,50) of wanneer de correlatiecoëfficiënt voor ordinale variabelen, Spearman ρ , berekend wordt (0,55).

De waarde van de correlatie tussen TAL-scores en PRESTATIE geeft aan dat er sprake is van een duidelijke samenhang tussen de taalvaardigheid zoals die met de TAL gemeten wordt en de leerprestatie in het eerste leerjaar. Als we deze correlatie vergelijken met die van andere toetsen zoals, bijvoorbeeld, de leesrijpheidstesten die gebruikt worden om leesmoeilijkheden te voorkomen, blijkt dat in die gevallen men van redelijke predictieve validiteit spreekt bij een correlatie van $r = 0,37$ en $r = 0,40$ (Van Dongen & Van Leent 1981: 24).

Met het aantonen van de samenhang is echter nog niet duidelijk geworden welke verwachtingen met betrekking tot leerprestaties met welke TAL-scores kunnen worden verbonden. Als de TAL wordt afgenomen op het einde van de derde kleuterklas of het begin van het eerste leerjaar, welke score moet een leerling dan halen om aan te nemen dat hij taalvaardig genoeg is om het eerste leerjaar met succes te volgen? Onderstaande tabel kan ons helpen om op die vraag een antwoord te vinden.

Tabel 6: Decielen voor de PRESTATIE-variabele, gemiddelde TAL-scores (GEM TAL), standaarddeviatie (SD) en aantal proefpersonen (N)

deciel	GEM TAL	SD	N
	41,60	7,58	421
1,00	35,20	10,75	44
2,00	36,84	8,64	38
3,00	39,15	7,31	46
4,00	41,26	4,54	42
5,00	41,80	7,55	41
6,00	41,95	6,27	40
7,00	43,09	7,11	43
8,00	44,29	5,04	42
9,00	45,64	4,31	44
10,00	46,76	2,19	41

Tabel 6 geeft de gemiddelde TAL-score weer voor verschillende decielgroepen op de variabele PRESTATIE. Leerlingen die behoren tot de groep van 10% zwakst presterende leerlingen op het einde van het eerste leerjaar, scoorden op de TAL gemiddeld 35,2. Leerlingen die behoorden tot de groep tussen 10% en 20% scoorden op de TAL gemiddeld 36,8 enzovoort.

De vraag die nu beantwoord moet worden, is waar de grens getrokken moet worden tussen de voldoende en onvoldoende prestatie in het eerste leerjaar. Wij hebben bij het vaststellen van die drempel in grote lijnen twee redeneringen gevolgd. We zijn uitgegaan van de wetenschap dat we voor dit onderzoek vooral leerlingen hebben geselecteerd van wie we verwachten dat ze mogelijke zwakke presteerders zijn. Dit betekent dat de groep leerlingen die aan het validiteitsonderzoek heeft deelgenomen, zeker zwakker uitvalt dan de doorsnee leerling in het eerste leerjaar. Vandaar dat het zeker niet overdreven is te stellen dat leerlingen die deel uitmaken van de zwakste 50% procent van de groep, beschouwd moeten worden als zwakke presteerders. Uit bovenstaande tabel valt af te leiden dat die groep in termen van zijn TAL-score getypeerd kan worden als een groep die lager scoort dan 42. Die score van 42 kan dan ook beschouwd worden als een soort ondergrens op de TAL: wie lager scoort, loopt zeer grote kans om in het eerste leerjaar serieuze moeilijkheden te krijgen.

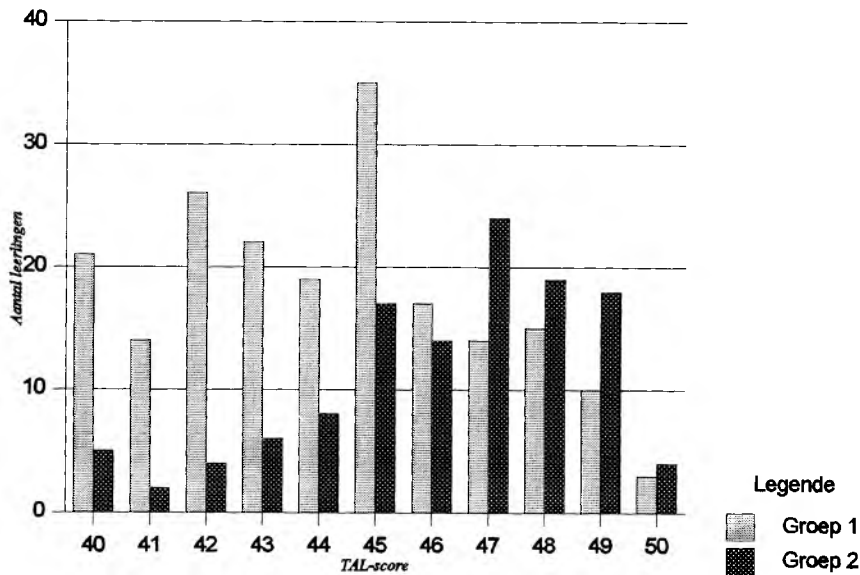
Men kan echter ook redeneren dat het behoren tot de beste 50% van de groep met betrekking tot de leerprestaties in het eerste leerjaar, een te laag criterium is. In de groep die aan het onderzoek heeft deelgenomen, zijn proportioneel heel veel migrantenleerlingen opgenomen. Men zou kunnen stellen dat we nu net aan dit onderzoek begonnen zijn omdat we weten dat deze leerlingen in zeer groten getale zwak presteren, en dat we daarom bij het bepalen van wat een voldoende prestatie in het eerste leerjaar is, beter uit kunnen

gaan van de resultaten van de Vlaamse leerlingen uit het onderzoek.

Van deze Vlaamse leerlingen kan dan weer gesteld worden dat ze in vergelijking met de gemiddelde schoolpopulatie een zwakke groep vormen. Weerom lijkt het dan niet overdreven om pas te spreken van een voldoende prestatie in het eerste leerjaar als leerlingen even goed of beter scoren dan de gemiddelde Vlaamse leerling uit het onderzoek. Die gemiddelde Vlaamse leerling situeert zich op deciel 7 van de gehele groep. Met verwijzing naar Tabel 6 betekent dit dat hij 8 of hoger scoort op de PRESTATIE-variabele.

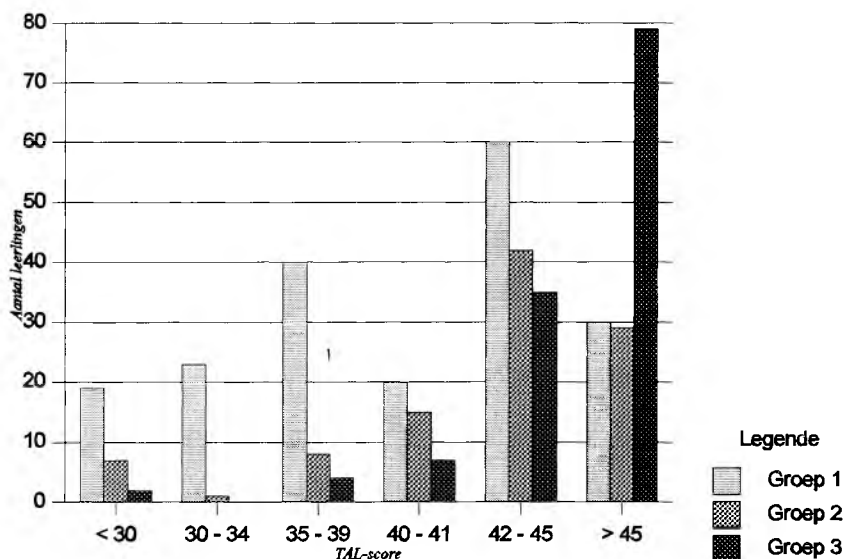
In Tabel 6 kan men ook vaststellen dat die leerling gemiddeld 45 of meer scoort op de TAL. We zouden dus ook kunnen stellen dat alle leerlingen die op de TAL onder de 45 scoren, een meer dan gemiddelde kans maken om zwak te presteren in het eerste leerjaar. Als we echter het halen van de drempel 'score van de gemiddelde Vlaamse leerling' systematisch in verband brengen met de TAL-score (zie Figuur 1), dan blijkt dat leerlingen met een TAL-score van 45 niet vaker de gestelde drempel halen dan leerlingen met een score van 43 of 44, maar dat dit wel geldt voor leerlingen met een score van 46 of meer. Deze vaststelling brengt er ons toe om de tweede drempel niet te leggen op 45, maar op 46. Het blijkt dus dat de drempels die uit het validiteitsonderzoek naar voren komen, overeenkomen met de drempels die op basis van toetsinterne criteria konden worden bepaald.

Analoog aan het zoeken naar een striktere drempel dan die van 42, kunnen we ook op zoek gaan naar een lagere TAL-drempel. Van leerlingen met een TAL-score lager dan 42, hebben we geconstateerd dat de kans zeer groot is dat zij zwak presteren in het eerste leerjaar. Van leerlingen met een score tussen 42 en 46 hebben we gesteld dat zij een significant grotere kans dan gemiddeld lopen om tot die groep van zwakke presteerders te horen.



Figuur 1: Indeling van de leerlingen op basis van hun TAL-score en hun PRESTATIE-score (groep 1: leerlingen die lager scoren dan de gemiddelde Vlaamse leerling; groep 2: leerlingen die hoger scoren dan de gemiddelde Vlaamse leerling)

Is er nu ook niet een ondergroep te bepalen voor wie zo goed als vast staat op basis van hun aanvangstaalvaardigheid, dat ze in het eerste leerjaar zwak zullen presteren? Uit Figuur 2 kan worden afgeleid dat dit niet erg zinvol is. We zien dat verschillen in TAL-scores lager dan 40 weinig informatie bieden over de mogelijkheden van leerlingen om op een voldoende niveau te presteren in het eerste leerjaar. Van de groep met TAL-score lager dan 30 zijn er procentueel evenveel die een van de twee criteria halen als van de groep met een TAL-score tussen 35 en 39. De groep leerlingen met een TAL-score van 40 of 41 scoort wel lichtjes beter, maar de echte toename van voldoende presteerders komt er maar vanaf een TAL-score van 42.



Figuur 2: Indeling van de leerlingen op basis van hun TAL-score en hun PRESTATIE-score (groep 1: leerlingen die lager scoren dan de gemiddelde leerling; groep 2: leerlingen die lager scoren dan de gemiddelde Vlaamse leerling; groep 3: leerlingen die hoger scoren dan de gemiddelde Vlaamse leerling)

We zouden dus weliswaar een onderdrempel kunnen hanteren van 40, maar het gebruik van die drempel verfijnt eigenlijk niet de mogelijkheden om de TAL als voorspellende toets te gebruiken: uit de cijfers blijkt immers duidelijk dat men zich zorgen moet maken voor het gros van de leerlingen die lager scoren dan 42, en dat er binnen die groep een aantal leerlingen is die er toch in slagen op een voldoende wijze te presteren. Aangezien die laatste groep zowel leerlingen betreft met een zeer lage als met een lage TAL-score, heeft het niet veel zin om hier nog een derde TAL-drempel in te voeren.

Samengevat kan worden gesteld dat leerlingen die op de TAL lager scoren dan 42, over onvoldoende taalvaardigheid beschikken om in

een eerste leerjaar normaal te functioneren, zodat de kans heel erg groot is dat ze dat eerste leerjaar afsluiten met tegenvallende leerprestaties. Sommige van deze leerlingen vinden blijkbaar een manier om, ondanks hun gebrek aan taalvaardigheid, toch redelijk tot goed te presteren. Deze mogelijkheid om te compenseren, treffen we aan zowel bij leerlingen met zeer lage TAL-scores, als bij leerlingen die maar net onder de drempel scoren.

Van leerlingen die boven die drempel van 42 scoren, zien we dat leerlingen die maar net boven de drempel scoren een grotere kans op zwak presteren vertonen dan leerlingen die ruim boven die drempel scoren. Een TAL-score onder 46 geeft aan dat een leerling tot de 'risicogroep' behoort. Leerlingen die boven 46 op de TAL scoren blijken over het algemeen wel redelijk goed te presteren in het eerste leerjaar. Sommige leerlingen uit deze groep blijven onder de verwachte prestaties. We nemen aan dat dit in de meeste gevallen een gevolg is van andere dan taalvaardigheidsfactoren.

De conclusies die in bovenstaande tekst getrokken worden met betrekking tot de voorspellende waarde van de TAL, kunnen nog op verschillende wijzen verfijnd worden. Zo kan verder worden onderzocht wie nu de leerlingen zijn met een zwakke TAL-score die toch redelijk presteren in het eerste leerjaar. Kan op basis van de onderzoeksgegevens een idee gegeven worden van hoe het komt dat zij ondanks hun zwakke taalvaardigheid niet uit de boot lijken te vallen? Een vergelijkbare lijn van onderzoek kan opgezet worden naar de leerlingen die ondanks een goede TAL-score, toch zwak presteren. Welke factoren lijken voor dit zwak presteren verantwoordelijk te zijn? Ook kan het gegeven 'goed presteren' in het eerste leerjaar verder worden uitgespit. We hebben wel geconstateerd dat er sprake is van zoiets als algemene schoolprestaties, maar dit hoeft niet te betekenen dat er daarnaast ook niet interessante patronen aan te geven zijn voor elke vorm van

schoolprestatie afzonderlijk. Ook beschikken we over gegevens om na te gaan hoe dat zwak presteren zich in het eerste leerjaar ontwikkelt. Is het iets wat pas op het einde van het eerste leerjaar duidelijk wordt, of is het al van bij het begin te merken? Het zou ons te ver leiden om vragen van deze aard hier te behandelen. De geïnteresseerde lezer verwijzen we daarom naar Van Berkel (1995).

7. Gebruik van de TAL

7.1 Algemene opmerkingen

De vraag die hier nog beantwoord dient te worden, is hoe de geconstateerde samenhang tussen TAL-score en leerprestatie in het eerste leerjaar gebruikt kan worden om het onderwijs efficiënter te maken. In dit verband is het belangrijk erop te wijzen dat het feit dat de TAL predictief valide bleek, niet geïnterpreteerd moet worden als een onontkoombare vaststelling, als een aanwijzing van het feit dat leerlingen met een lage TAL-score weinig of niets van het leren in het eerste leerjaar terecht zullen brengen. De geconstateerde samenhang is veel meer een signaal dat die leerlingen grote problemen zullen krijgen, **tenzij men iets onderneemt om die problemen te ondervangen**. Een lage TAL-score moet gezien worden als een uitdaging aan het adres van het onderwijs om er voor te zorgen dat de voorspelling m.b.t. zwakke leerprestaties die eraan verbonden wordt, niet uitkomt. De TAL levert niet een maat voor het succes of het falen waarmee een leerling de kleuterschool afsluit, dan wel een beschrijving van de kenmerken van de leerling waarmee het lager onderwijs zal moeten zien te werken. In deze zin moet men het gebruik van de TAL zien in het licht van wat onder zorgverbreding in het onderwijs verstaan wordt.

Zorgverbreding is er op gericht een onderwijs uit te bouwen dat

elke leerling optimale kansen biedt om tot leren te komen. Hiertoe is het belangrijk de beginsituatie van elke leerling met betrekking tot een aantal relevante factoren in kaart te brengen en het onderwijs zo te organiseren dat met die beginsituatie rekening wordt gehouden. Vaststellen dat een leerling over te weinig taalvaardigheid beschikt om het onderwijs in het eerste leerjaar met vrucht te volgen, moet worden gezien als het vaststellen van een degelijk relevant kenmerk van de beginsituatie. De actie die op deze vaststelling moet volgen, is niet het afschrijven van de leerling als onbekwaam om het onderwijs in dat eerste leerjaar te volgen, maar wel het plannen en organiseren van het onderwijs in het eerste leerjaar op een wijze dat die leerling er wel tot leren kan komen.

Hierbij is het ook belangrijk er aan te herinneren dat de TAL aansluit bij wat de analytische visie op taalonderwijs wordt genoemd. Deze visie houdt in dat men taalvaardigheid niet ziet als iets wat uiteengelegd kan worden in verschillende onderdelen die elk onafhankelijk van elkaar opgebouwd, geëvalueerd en geremedieerd kunnen worden. Als een leerling een achterstand in taalvaardigheid vertoont, moet men in deze visie niet proberen uit te zoeken hoe die achterstand vertaald dient te worden in achterstanden op verschillende onderdelen, om dan vervolgens in remediëring het onderdeel dat de grootste achterstand vertoont, aan te pakken. In de analytische visie gaat men ervan uit dat een dergelijke achterstand een globale werking vereist, waarbij men de leerling kansen biedt om zijn taalvaardigheid in zijn geheel te verbeteren. Dergelijke kansen kan men onvoldoende bieden als men die achterstand uitsluitend aan wil pakken via remediëring. Wij gaan er dan ook van uit dat de planning van het onderwijs op basis van de TAL in de eerste plaats betrekking moet hebben op wat er structureel in de klas gebeurt, en slechts in tweede instantie aanwijzingen voor remediëring biedt. In de zin dat de TAL taalvaardigheid niet opdeelt in verschillende componenten die onafhankelijk van elkaar worden afgetast, en hij

geen aanwijzingen biedt naar remediëring van taalvaardigheids-onderdelen bij individuele leerlingen, wijkt hij af van de klassieke diagnostische toetsen. Het is belangrijk dit in de gaten te houden wanneer het concrete gebruik van de TAL ter sprake komt.

7.2 Concretisering van het gebruik

Als we bij het gebruik van de TAL de bovenstaande algemene principes in het achterhoofd houden, kunnen we twee richtingen uit. We kunnen ons baseren op de gemiddelde TAL-score van de klas, of we kunnen de score van elke individuele leerling als vertrekpunt nemen. Aangezien dat laatste het best gezien kan worden als een mogelijke aanvulling bij het eerste, bespreken we eerst de samenhang tussen de gemiddelde TAL-score van een klas en de manier waarop met die klas wordt omgegaan.

Een eerste vaststelling is dat de vaardigheid om de schooltaal te gebruiken, een belangrijke voorwaarde voor schoolsucces in het eerste leerjaar is gebleken. Eigenlijk staat deze vaststelling in schril contrast tot de invulling van het vak 'Taal' in dat eerste leerjaar. De traditie wil dat het vak 'Taal' in het eerste leerjaar nagenoeg volledig gericht wordt op het leren lezen en schrijven. Andere, meer op algemene taalvaardigheid gerichte activiteiten, komen nauwelijks aan bod. In het licht van de samenhang tussen die vaardigheid en schoolprestaties ligt het voor de hand om de ontwikkeling van die vaardigheid in alle omstandigheden te ondersteunen. Ongeacht de TAL-score van de gemiddelde leerling in de klas, kan men stellen dat aandacht voor wat leerlingen met taal moeten of kunnen doen, meer dan nuttig is in elk eerste leerjaar.

Het in de planning opnemen van aandacht voor taalvaardigheids-opbouw wordt des te noodzakelijker naarmate het algemene

taalvaardigheidsniveau van de klas lager uitvalt. Voor een klas die een gemiddelde TAL-score lager dan 42 haalt, heeft het weinig zin om het onderwijs van meet af aan toe te spitsen op het aanleren van elementaire lees- en rekenvaardigheden. De kans is zeer groot dat het gros van de kinderen in de klas niet in staat is ten volle te begrijpen wat tegen hen gezegd wordt, laat staan dat ze in staat zouden zijn om op deze terreinen tot normale leerprestaties te komen.

Het gezond verstand schrijft hier voor dat deze leerlingen in de eerste plaats steun moeten krijgen bij de opbouw van de noodzakelijke taalvaardigheid. In dit geval achten wij het noodzakelijk dat de leerlingen bij het begin van het eerste leerjaar een uitgebreid en intensief taalvaardigheidsprogramma doorwerken, en dat er in de loop van dat jaar blijvend aandacht en zorg aan die taalvaardigheid wordt besteed. Aangezien die vaardigheid noodzakelijk is voor alle leren, hoeft dat startprogramma volgens ons niet beperkt te worden tot de taallessen alleen: het rendement van het onderwijs in andere vakken zal voor deze groep toenemen als men in een eerste fase de tijd die normaal aan deze vakken wordt besteed, voor de opbouw van taalvaardigheid gebruikt.

Het feit dat in deze klassen vaak ook enkele leerlingen aanwezig zijn die op de TAL boven de 42-drempel scoren, hoeft geen belemmering te vormen voor het uitvoeren van een dergelijk programma. Voor leerlingen met een score tussen 42 en 45, is een dergelijk programma zelfs een goede zaak: we hebben immers vastgesteld dat deze leerlingen ook een verhoogd risico op tegenvallende leerprestaties lopen.

Met leerlingen in deze klassen die op de TAL 46 of meer scoren, kan men in principe twee richtingen uit. Men kan extra uren gebruiken om deze leerlingen al te laten starten met het gewone programma, of men kan ze bij het taalvaardigheidsprogramma betrekken. In het eerste geval bouwen deze leerlingen met betrekking tot met name aanvankelijk lezen, schrijven en rekenen een zekere

voorsprong uit. Die voorsprong kan men dan later weer laten aansluiten bij de differentiatiemogelijkheden die in de aanpak van aanvankelijk lezen, schrijven en rekenen zijn voorzien. In het tweede geval differentieert men in het taalvaardigheidsprogramma zelf, zodat deze leerlingen in dit programma geholpen worden om hun eigen vaardigheid verder uit te bouwen.

In elk geval is het zo dat ook deze taalvaardige leerlingen erbij gebaat zijn dat de klas als geheel taalvaardiger gemaakt wordt: als dit niet gebeurt, verzeilen zij immers in een klas met kinderen die de gestelde leertaken niet aankunnen. Het feit dat gedurende het jaar constant geprobeerd moet worden om die zwakkere kinderen mee te trekken, vermindert automatisch de efficiëntie van het onderwijs, ook voor de taalvaardige leerlingen.

Voor klassen met een gemiddelde taalvaardigheidsscore van 46 of meer, lijkt een taalvaardigheidsprogramma bij aanvang niet nodig. Hier lijkt het meer aangewezen om de leerlingen met een beperkte taalvaardigheid die in deze klassen zitten, via de remediëring een extra taalvaardigheidsimpuls te geven.

Voor klassen met een gemiddelde TAL-score tussen de twee drempels is het moeilijker te bepalen hoe het best gewerkt kan worden. Men kan besluiten dat een intensief taalvaardigheidsprogramma aangewezen is. Men kan ook beginnen met het standaardprogramma, en onderweg aandacht voor taalvaardigheid inbouwen. Veel zal hier afhangen van de concrete situatie op elke school. Een school die met een aanvankelijk leesprogramma werkt waar relatief veel aandacht aan taalvaardigheid gegeven wordt, kan er voor kiezen om met het gewone programma te beginnen. Een school daarentegen die een aanvankelijk leesprogramma gebruikt, waarin er niet aan taalvaardigheidsopbouw gewerkt wordt, doet er ons inziens best aan in elk geval een startprogramma taalvaardigheid te voorzien. Een school die over veel mogelijkheden tot differentiatie en remediëring beschikt, kan van die mogelijkheden gebruik maken om

de zwakkere leerlingen uit de groep een startprogramma voor taalvaardigheid aan te bieden, en terzelfdertijd met de klas als geheel met het gewone programma starten.

In elk geval is het zo dat men er in deze situatie op bedacht moet zijn dat er leerproblemen op kunnen treden als gevolg van een te beperkte taalvaardigheid. Dit betekent dat men de leerprestaties van de leerlingen in de gaten zal moeten houden, en als die onvoldoende blijken, meer aandacht aan die taalvaardigheid moet geven. Het gevaar dat hier bestaat, is dat men de taalvaardigheid van de leerlingen voldoende acht, en eventuele problemen later aan andere factoren gaat wijten. In dat geval komen we terecht in een situatie waarin leerlingen voor de verkeerde dingen geredigeerd worden.

Een laatste, belangrijke opmerking: de TAL mag zeker niet gebruikt worden voor het creëren van homogene groepen. In situaties waarin leerlingen aanzienlijk verschillen qua niveau van taalvaardigheid, wordt het werken met homogene groepen vaak als een optimale oplossing gezien. Onderzoek heeft echter uitgewezen dat het creëren van homogene groepen geen positieve effecten op het leerproces heeft (Slavin 1994: 138). Een uitzondering kan worden gemaakt voor lezen wanneer leerlingen volgens een bepaald plan (Joplin plan) over klassen heen gegroepeerd worden. Aangezien dit de enige uitzondering is, moet het werken met homogene groepen echter afgeraden worden. Bovendien gaan de taalvaardigheidsprogramma's die hier worden voorgesteld dan ook uit van heterogene groepen. Het lijkt dus niet zinvol om eerst homogene groepen te gaan maken op basis van TAL-scores om vervolgens te constateren dat dit niet aansluit bij de uitgangspunten van de taalvaardigheidsaanpak.

8. Conclusies

Wellicht is het de lezer opgevallen dat de TAL in een aantal opzichten een andere toets is dan wat men gewend is. Zo treft men in dit boek geen normen aan. Dit heeft te maken met het feit dat de TAL niet in de eerste plaats een toets is die aangeeft waar een leerling zich met betrekking tot taalvaardigheid bevindt ten aanzien van zijn leeftijdsgenoten. De TAL probeert in kaart te brengen of een leerling over voldoende taalvaardigheid beschikt om talig in het eerste leerjaar te kunnen functioneren.

Ook bevat de TAL geen diagnostische aanwijzingen in de zin dat wordt afgetast aan welk specifiek probleem een eventueel algemeen tekort aan taalvaardigheid toegeschreven moet worden. Taalvaardigheid wordt hier immers niet opgevat als een optelsom van deelvaardigheden en van elkaar af te grenzen vormen van kennis. De TAL geeft daarom ook geen indicaties over welke deelvaardigheid of welke vorm van kennis geredieerd zou moeten worden om leerlingen in staat te stellen hun achterstand weg te werken.

Het feit dat de TAL niet dit soort informatie verschaft, betekent natuurlijk niet dat hij een vrij nutteloos instrument zou zijn, of dat hij alleen gebruikt kan worden om het onvermijdelijke falen van leerlingen al vooraf te kunnen voorspellen. Wel integendeel, de TAL probeert de algemene taalvaardigheid van leerlingen in een klas in kaart te brengen, om leerkrachten in staat te stellen met die taalvaardigheid rekening te houden bij de planning van hun onderwijs.

Het is dus duidelijk niet de bedoeling om aan de hand van de TAL klassen te maken die homogeen zijn qua taalvaardigheid, of om kinderen door te verwijzen naar het Buitengewoon Onderwijs. Uit ervaring blijkt dat beide reacties geen oplossing bieden voor de taalvaardigheidsachterstand, maar veeleer een manier vormen om binnen het onderwijs met die achterstand te leren leven.

De TAL heeft als bedoeling de mogelijkheden te vergroten om in het lager onderwijs adequaat te werken met leerlingen van verschillende taalvaardigheidsniveaus. Hij geeft een dubbel signaal: de leerkracht moet rekening houden met de taalvaardigheid van zijn/haar leerlingen bij het bepalen van hoe zijn/haar onderwijs vorm wordt gegeven, en hij/zij zal in dat onderwijs maximaal mogelijkheden moeten opnemen voor de taalzwakkere leerlingen om hun achterstand in te lopen. Die mogelijkheden bestaan dan niet uit allerlei lesjes over allerlei aspecten van het taalsysteem, maar uit het aanbieden van situaties die stimuleren en motiveren om talig te functioneren, die rijk zijn aan relevant nieuw taalaanbod, en die mogelijkheden bieden om ondersteuning te krijgen bij het verwerken van dat aanbod. Het is onze vaste overtuiging dat een reactie op de TAL-resultaten die in deze lijn ligt, heel wat taalzwakke leerlingen in staat zal stellen om ondanks die initiële achterstand, uit te groeien tot leerlingen die kunnen profiteren van wat de lagere school hen aan leermogelijkheden biedt. En dat is precies de bedoeling.

Referenties

- Baker, D. (1989) *Language testing*. London: Edward Arnold.
- Carrol, B. (1982) *Testing communicative performance*. Oxford: Pergamon Press.
- Carter, R. & M. McCarthy (1989) *Vocabulary and language teaching*. London: Longman.
- Cito toetsen (1992) *Leerling volg systeem 3 & 4: Schaal Vorderingen in Spellingvaardigheid 1; Leerling volg systeem 3 & 4: Rekenen - Wiskunde 1; Leerling volg systeem 3, 4, 5, 6, 7 & 8: Drie-Minuten-Toets; Leerling volg systeem 3 & 4: Lezen met Begrip 1*. Arnhem: Cito.
- Cummins, J. (1979) Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49 (2), 221-255.

- Cummins, J. (1980) The cross-lingual dimension of language proficiency: implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14, 175-187.
- Cummins, J. (1984) Wanted: a theoretical framework for relating language proficiency to academic achievement among bilingual students. In: C. Rivera (ed.) *Language proficiency and academic achievement*. (pp. 2-19). Clevedon: Multilingual Matters.
- Cureton, E.E. (1951) Validity. In: E.F. Lindquist (ed.) *Educational measurement*. (pp. 621-694). Washington: American Council of Education.
- Davies, A. (1990) *Principles of language testing*. Oxford: Blackwell.
- De Haan, D. (1985) Cognitive academic skills: an attempt to operationalize the concept. In: G. Extra & T. Vallen (eds.) *Ethnic minorities and Dutch as a second language*. (pp. 147-165). Dordrecht: Foris Publications.
- De Haan, D. (1987) De verwerving van het Nederlands als schooltaal: kenmerken van taalgebruik van Surinaamse en Nederlandse kinderen. In: G. Extra, R. van Hout & T. Vallen (eds.) *Etnische minderheden. Taalverwerving, Taalonderwijs, Taalbeleid*. (pp. 65-83). Dordrecht: Foris Publications.
- Drenth, P.J.D. & K. Sijtsma (1990) *Testtheorie. Inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn Stafleu Van Loghum.
- Driessen, G. (1990) *De onderwijspositie van allochtone leerlingen*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Ellis, R. (1990) *Instructed second language acquisition*. Oxford: Blackwell.
- Extra, G. & L. Verhoeven (1985) Bias in intelligentie-onderzoek bij allochtone kinderen. *Pedagogische Studiën*, 62, 392-395.
- Henning, G. (1987) *A guide to language testing*. Cambridge (MA): Newbury House Publishers.
- Hudson, T. & B. Lynch (1984) A criterion-referenced measurement approach to ESL achievement testing. *Language Testing*, 1, 171-201.
- Hughes, A. (1989) *Testing for language teachers*. Cambridge: Cambridge University Press.
- Hulstijn, J. (1984) Vaardigheden in moedertaal en tweede taal. Cummins' theorie van wederzijdse afhankelijkheid. *Levende Talen*, 388, 18-23.

- Jaspaert, K. (1991) *Nederlands als tweede taal: een programmatekst*. Steunpunt NT2 K.U. Leuven.
- Jaspaert, K. & W. Schrooten (1995) *Taalcahiers deel 2*. Antwerpen: Plantijn.
- Klatter-Folmer, J. & T. Vallen (1995) *Een toets mag wel moeilijk zijn, maar niet partijdig: over talige en culturele bias voor allochtone leerlingen in Nederlandstalige toetsen*. (pp. 357-371). In: E. Huls & Klatter-Folmer (eds.) *Artikelen van de Tweede Sociolinguïstische Conferentie*, Delft: Eburon.
- Lado, R. (1961) *Language testing*. London: Longman.
- Litjens, P. (1989) Ontwikkeling van schoolse taalvaardigheden tijdens de zaakvakken in de bovenbouw van het basisonderwijs. In M. Kienstra (ed.) *Onderwijs aan anderstaligen*. (pp. 79-100). Dordrecht: Foris Publications.
- Litjens (1990) *Plaatsbepaling en karakterisering van het taalonderwijs Nederlands in meertalige onderwijssituatie*. Enschede: SLO
- McLaughlin, B. (1985) *Second-language acquisition in childhood: Volume 2. School-age children*. Hillsdale (NJ): Lawrence Erlbaum.
- Morrow, K. (1979) Communicative language testing: revolution or evolution?. In: C.J. Brumfit and K. Johnson (eds.) *The communicative approach to language teaching*. (pp. 143-166). Oxford: Oxford University Press.
- Oller (1979) *Language tests at school*. London: Longman.
- Oller (1983) 'g', what is it? In: A. Hughes en D. Porter (eds.) *Current developments in language testing*. (pp. 35-37). London: Academic Press.
- Popham, W.J. (1978) *Criterion-referenced measurement*. Englewood Cliffs (NJ): Prentice Hall.
- Slavin, R.E. (1994) School and classroom organization in beginning reading. In: R. E. Slavin, N.L. Karweit & B.A. Wasik (eds.) *Preventing early school failure*. (pp. 122-142). Boston: Allyn and Bacon.
- Uiterwijk, H. (1994) *De bruikbaarheid van de Eindtoets Basisonderwijs voor allochtone leerlingen*. Arnhem: Instituut voor Toetsontwikkeling.
- Van Berkel, S. (1995) *Stap over de drempel. Een onderzoek naar de validiteit van de Taalvaardigheidstoets Aanvang Lager onderwijs*. Doctoraalscriptie K.U. Brabant.

- Van Bon, W.H.J. (1982) *Taaltests voor Kinderen. Handleiding*. Lisse: Swets & Zeitlinger.
- Van Dongen, D. & H. Van Leent (1981) *Preventie van leesmoeilijkheden*. Nijmegen: Instituut voor Onderwijskunde.
- Verhoeven L., (1982) *Begrippentoets*. Arnhem: Instituut voor Toetsontwikkeling.
- Verhoeven L. & A. Vermeer (1989) *Diagnose van kindertaal*. Tilburg: Zwijzen.
- Vollmer, H.J. (1983) The structure of foreign language competence. In: A. Hughes & D. Porter (eds.) *Current developments in language testing*. (pp. 3-29). London: Academic Press.