

Bepaling van de woordidentificatiedrempel bij peuters

Th.A.M. Crul^{1, 2}, J.T.M. Braks³, A.F.M. Snik¹ en J.P.L. Brokx⁴

¹*Audiologisch Centrum, Academisch Ziekenhuis Nijmegen*

²*Medische Psychologie, Academisch Ziekenhuis Nijmegen*

³*Afdeling KNO, Academisch Ziekenhuis Nijmegen*

⁴*Audiologisch Centrum, Instituut voor Doven, St. Michielsgestel*

De PAS-test (Peuter Adaptieve Spraakdrempelbepaling) is een spraakaudiometrische test voor de bepaling van de woordidentificatiedrempel bij kinderen vanaf 2 jaar of bij ouderen met een overeenkomstige mentale leeftijd. Beschreven wordt de constructie, normering en validering van twee parallel-versies, A en B. De geluidssterkte waarbij 71% van de woorden goed wordt verstaan wordt met zorgvuldig geselecteerde en bewerkte woorden bepaald volgens een adaptieve procedure. Een respons wordt gegeven door een concreet object, behorend bij een woord, tussen andere aan te wijzen of aan te geven. Het blijkt dat de resultaten die verkregen zijn met de geautomatiseerde test een valide basis vormen voor het maken van een schatting van de gemiddelde toondrempel in het beste oor.

Inleiding

De gevolgen van slechthorendheid op de emotionele en educatieve ontwikkeling van het jonge kind zijn uitvoerig beschreven (zie o.a. Tucker & Nolan, 1984). De in de literatuur aangegeven frequentie van voorkomen van een gehoorverlies bij kinderen tijdens de belangrijke periode voor de spraak- en taalverwerving varieert sterk, afhankelijk van het gegeven of men zich beperkt tot het noemen van alleen de ernstige, meestal perceptieve gehoorverliezen (1 tot 2 op de 1000 op de leeftijd van 2 jaar) of dat ook vormen van tijdelijke lichte gehoorverliezen worden meegeteld (1 op de 25 twee-jarigen, Northern & Downs, 1991). Omdat ook vormen van een licht gehoorverlies, vooral als deze langer bestaat, consequenties kunnen hebben voor de ontwikkeling van een kind, wordt er algemeen veel

aandacht besteed aan de wens van een (vroeg)tijdige opsporing en interventie (van der Lem & Baart de la Faille, 1981; Crul, 1986). Buiten de audiologische screening- en diagnostische methoden van onderzoek naar de gehoordrempel bij een klein kind, bestaat sinds 1983 in Nederland ook een gestandaardiseerde spraaudiometrische test voor kinderen vanaf 3 jaar, de zogenaamde SAP-test (Sprak Audiometrie met Plaatjes, Crul 1983, 1984). Deze test is zowel geschikt voor de leeftijdsgroep van 3 tot 6-jarigen als voor ouderen die op een overeenkomstig mentaal niveau functioneren. De vraag kan worden gesteld of het ook mogelijk is om van kinderen die jonger zijn dan 3 jaar audiometrisch betrouwbare gegevens te verkrijgen over de spraakreceptie. Omdat 2-jarigen midden in het spraakverwervingsproces verkeren en zich door hun korte aandachtspan niet makkelijk lenen voor een relatief langdurig onderzoek dat veel concentratie vereist, moet een spraaudiometrische test voor deze groep aan andere eisen voldoen dan de gangbare tests voor oudere kinderen of volwassenen. In dit artikel wordt de constructie, normering en validering beschreven van een test ter bepaling van de woord-identificatiedrempel bij kinderen vanaf 2 jaar. Als voorbeeld gold hierbij de zogenaamde "McCormick Automated Toy Test" uit Engeland (McCormick, 1977, 1988; Ousey, Sheppard, Twomey & Palmer, 1989; Palmer, Sheppard & Marshall, 1991).

Het principe van de Automated Toy Test

De Automated Toy Test is afgeleid van een door Kendall (1953; 1954) ontworpen en later door McCormick (1977;1988) aangepaste screeningtest voor kinderen met een mentale leeftijd van 2 jaar en hoger. Bij de oorspronkelijke opzet van deze test moet het te onderzoeken kind alleen op het gehoor een concreet object tussen 15 andere uitkiezen wanneer het bijbehorende woord met normale stem in het vrije veld wordt uitgesproken op een geluidsterkte die de onderzoeker die tegenover het kind zit zelf moet variëren en controleren met behulp van een geluidsdrukmeter. De bedoeling is om een schatting te kunnen maken van de spraakreceptiedrempel van een kind. Als grens voor een normaal gehoor wordt hierbij traditioneel een kritische aanbiedingssterkte geaccepteerd van 40 dB (A). Het voordeel van de test is dat de woorden en voorwerpen geselecteerd zijn overeenkomstig het ontwikkelingsniveau van een 2-jarig kind. Een nadeel is onder andere dat de onderzoeker over een wel zeer goede stemcontrole moet beschikken bij het produceren van een betrouwbare geluidsterkte tijdens het uitspreken van een woord en dat de test niet echt is gestandaardiseerd en dus alleen als screeningtest kan functioneren. Deze nadelen bestaan niet meer bij de nieuwe versie van de Toy Test. De ontwerpers zorgden ervoor dat de afname geheel geautomatiseerd kan verlopen en kozen 14 meestal in één foneem van elkaar verschillende woordparen en bijbehorende concrete speelobjecten waarvan het te onderzoeken kind er telkens een moet aanwijzen wanneer het bijbehorende woord wordt gepresenteerd. Deze woorden zijn digitaal opgeslagen en

worden met controle van een computer na digitaal/analooq conversie op precies vastgestelde presentatie-niveaus via een luidspreker in het vrije veld aangeboden. De onderzoeker gebruikt hierbij zelf een klein aantal toetsen op een op afstand te bedienen paneel om de testprocedure te starten en het programma verder te laten verlopen. Hiermee wordt: 1) de set van woordparen gekozen overeenkomstig de passieve woordenschat van een kind; 2) aangegeven of een gegeven respons goed of fout is; 3) indien nodig, bijvoorbeeld in geval van een slechte concentratie, een herhaling van de laatste stimulus met een identieke of een grotere geluidssterkte gegeven. De woorden worden in een draagzin (bijvoorbeeld "Show me the man") gepresenteerd. Op een monitor wordt de geluidssterkte van ieder gepresenteerd woord aangegeven en ook telkens een voorlopig berekende drempelschatting op grond van de tot op dat moment aangeboden stimuli. Via eveneens in het geheugen opgeslagen ruis kan de test worden gecali-breed. Het eerste woord wordt op een comfortabele geluidssterkte van 72 dB (A), of groter als dat nodig is, aangeboden. In het begin van de test wordt na een correcte respons het niveau automatisch met 12 dB verzwakt, totdat een foute respons wordt gegeven. Na de eerste fout wordt het niveau weer met 12 dB verhoogd tot er een correcte respons volgt. Hierna wordt consequent een adaptieve two-down, one-up procedure gevolgd (Levitt, 1971): na twee correcte reacties wordt het geluidsniveau van de volgende stimulus automatisch met 6 dB verlaagd, na één incorrecte reactie wordt het niveau met 6 dB verhoogd. Een verandering van een hoger naar een lager niveau of vice versa heet een omslag. In de meest recente versie van de Automated Toy Test wordt de woord-identificatie-drempel bepaald door het gemiddelde te berekenen over de gemiddelde luidheidsniveaus van in totaal zes omslagen. Op deze wijze betreft de uiteindelijke schatting van de spraakdrempel het niveau waarop 71% van de woorden goed is verstaan.

De resultaten van deze spraakaudiometrische test voor kleine kinderen blijken betrouwbaar en valide, terwijl de test efficiënt binnen een korte tijd afgenomen kan worden. Bovendien kan een woord-identificatiedrempel beneden 40 dB worden gemeten en blijkt de gevonden drempel in de meeste gevallen de ligging van de toondrempel aan het beste oor goed te voorspellen (Ousey et al, 1989; Palmer et al, 1991).

I. Constructie van de Peuter Adaptieve Spraakdrempelbepaling test

Selectie van de woorden

Voor een Nederlandstalige versie van de Toy Test, verder hier genoemd de PAS-test (Peuter Adaptieve Spraakdrempelbepaling test), werden woorden geselecteerd die aan diverse eisen moesten voldoen (voor gedetailleerde gegevens: zie Braks, 1993; Alferink, 1993). Uit het CELEX-bestand van het Centrum voor Lexicale Informatie in Nijmegen werden alle eenlettergrepige zelfstandige naamwoorden die een concreet voorwerp representeren en een hoge frequentie van

voorkomen hebben geselecteerd. Hiervan werden vervolgens 54 woorden gekozen die objecten voorstellen, waarvan men mag aannemen dat zij behoren tot de onmiddellijke ervarings- en belevingswereld van een kind van ± 2 jaar. In een pilot-onderzoek (Alferink, 1993) werd bij 40 kinderen tussen 21 en 48 maanden oud nagegaan welke van deze woorden nagenoeg zeker deel uitmaken van hun passieve woordenschat. Van 36 woorden kon worden gezegd dat ze behoren tot het mentale lexicon van kinderen vanaf 24 maanden oud.

Het egaliseren van de drempelintensiteiten

In een luistertest waarbij woordmateriaal wordt gebruikt om nauwkeurige drempelmetingen te verrichten, wordt als uitgangspunt gesteld dat de items, wanneer ze op drempelniveau worden aangeboden, een onderling gelijke kans hebben om correct te worden herkend (Plomp en Mimpfen, 1979; Crul, 1983). Als drempelintensiteit geldt hier die geluidssterkte waarop een bepaald woord 50% kans heeft om goed te worden geïdentificeerd. Om dit bij de 36 geselecteerde woorden te bereiken is getracht de perceptuele drempelintensiteiten van de woorden onderling gelijk te maken, zodat aan dit criterium wordt voldaan.

Een spreekster zonder accent sprak de 36 woorden aan het einde van een draagzin in. De opname-sterkte en kwaliteit werden hierbij nauwkeurig gecontroleerd. Bovendien werden nog drie losse draagzinnen ingesproken die bestemd waren voor de uiteindelijke PAS-versie. De opnamen werden met een digitale recorder gemaakt en voor verdere bewerking op computerschijf overgezet met een sampling frequentie van 20 kHz. Via een segmentatieprocedure werden met behulp van het ILS-programma (ILS-PC, 1989) de 36 woorden van de draagzin geïsoleerd. Om de woordherkenningskans bij drempelmeting te maximaliseren werd de irrelevante akoestische energie in de laagste spraakfrequenties beneden 200 Hz weggefilterd (Wilson, Preece & Crowther, 1991). Vervolgens werden voor het uitvoeren van de drempellegalisering twee verschillende digitale luisterbanden vervaardigd waarop de 36 woorden in willekeurige volgorde waren gerangschikt, voorafgegaan door enkele oefenwoorden. Van ieder woord op de computerschijf werd met behulp van het ILS-programma de gemiddelde intensiteit bepaald door de Root Mean Square (RMS) te berekenen. De RMS wordt uitgedrukt in een dB waarde die staat voor de fysische gemiddelde intensiteit van een woord (Plomp & Mimpfen, 1979). Met behulp van deze 36 RMS-waarden is de bijstellingsprocedure van de respectievelijke perceptuele woord-identificatiedrempels gerealiseerd. Na ieder van drie opeenvolgende luisterexperimenten, waarbij de woorden in een open antwoordprocedure op drempelniveau per hoofdtelefoon aan drie verschillende groepen van audiometrisch normaal horende adolescenten (N=15 per groep) voor identificatie werden aangeboden, werd de gemiddelde intensiteit van een woord gehandhaafd, versterkt of verzwakt. Bij deze procedure werd voor een presentatie per hoofdtelefoon in plaats van in het vrije veld gekozen om de bepaling van de drempelintensiteit van de woorden zo nauwkeurig mogelijk te laten verlopen. De bijstelling van het gemiddelde intensiteitsniveau van een woord werd gerealiseerd door het digitale signaal met een factor

te vermenigvuldigen. Er werd zodoende voor ieder luisterexperiment een andere band vervaardigd. De gemiddelde intensiteit van een woord werd gehandhaafd als uit de resultaten van een van de luisterexperimenten op drempelniveau bleek dat het aantal correcte identificaties van dit woord bleef binnen het 95% betrouwbaarheidsinterval, gebaseerd op 50% correcte woordidentificaties en vastgesteld volgens binominaal criterium (Barton, 1975). Een woord werd versterkt of verzwakt alnaargelang er op grond van hetzelfde vastgestelde criterium te weinig of te veel correcte identificaties van bleken te zijn gegeven. Een positieve of negatieve bijstelling van de drempelintensiteit van een woord bedroeg na het eerste luisterexperiment 5 dB en in de twee volgende minder.

Definitieve samenstelling van de test

Na de stapsgewijze manipulatie van de drempelintensiteit van de woorden waarbij dit nodig bleek, bleven uiteindelijk 26 woorden over die volgens de experimentele resultaten een gelijke kans hebben om herkend te worden bij een luisterexperiment op drempelniveau. Tien woorden bleven ook na 3 bijstellingen van de gemiddelde intensiteit te gemakkelijk of te moeilijk herkenbaar als ze op drempelniveau werden aangeboden. Uit de 26 overgebleven woorden zijn 2 sets, A en B, samengesteld, ieder bestaande uit 4 verschillende woordparen. Elke set bestaat dus uit 8 verschillende woorden. De woordparen in iedere set zijn zodanig gekozen dat: 1) de beide woorden van een paar dezelfde klinker hebben, zodat het discrimineren rond de herkenningdrempel niet berust op het waarnemen van relatief duidelijke klinkerverschillen, maar vooral op medeklinkerherkenning; 2) de klinkers van de 4 woordparen zoveel mogelijk een extreme positie innemen op de contour van de klinker-formantendriehoek en ten opzichte van elkaar ver uiteen liggen. Hiermee wordt beoogd dat er ook een indirecte uitspraak kan worden gedaan over de klinkerwaarneming van de niet bij de test betrokken klinkers, omdat hierdoor het gehele perceptuele klinkerdomein bij de test is betrokken.

Na de definitieve keuze van de testwoorden werd de gemiddelde intensiteit van de drie opdrachtzinnen, op het einde waarvan deze woorden tijdens de adaptieve testprocedure zullen worden gepresenteerd, aangepast aan de gemiddelde sterkte van de 26 woorden na de laatste bijstelling. Hierdoor werd bereikt dat de gemiddelde geluidsterkte van de testwoorden overeenkomt met die van de opdrachtzinnen. Deze opdrachtzinnen zijn van zodanige syntactische opbouw dat ze willekeurig aan de 16 verschillende woorden kunnen worden gekoppeld. Tenslotte werden bij de 16 woorden de passende concrete objecten uitgezocht. Als criterium gold hierbij een duidelijke herkenbaarheid voor een klein kind en een praktische afmeting en manipuleerbaarheid voor de testafname.

In bijlage A worden de vier woordparen van de twee testversies en de drie opdrachtzinnen vermeld, naast een afbeelding van de objecten die bij de woorden van set A en set B horen.

II. Normering van de PAS-Test

Apparatuur

De elektronische apparatuur en de software voor het sturen van het adaptieve programma werden in samenwerking met het Instituut voor Doven in St. Michielsgestel ontwikkeld (zie voor uitvoerige beschrijving Braks, 1993). De 16 stimuluswoorden, de drie opdrachtzinnen en een smalbandruis rond 1000 Hz worden gegenereerd met behulp van digitaal opgeslagen informatie. Zij worden na digitaal-analoog conversie en via een digitaal gecontroleerde verzwakker (de Haan, 1988) door een geluidsbox en versterker in het vrije veld gepresenteerd. Het geheel wordt door een computer gestuurd. Software voor de controle van het verloop en de registratie van de adaptieve procedure werd ontworpen overeenkomstig de aanwijzingen van Ousey et al (1989). De testleider controleert het programma met behulp van een klein aantal cursortoetsen en een beeldscherm of met een muis waarmee: 1) de woordparen voor set A of set B of anders afzonderlijke paren in overeenstemming met de tevoren bepaalde woordenschat van een kind voor presentatie kunnen worden gekozen; 2) wordt aangegeven of een kind een correcte- of een foutieve respons heeft gegeven; 3) de laatst gepresenteerde stimulus in geval van onvoldoende concentratie op hetzelfde of een hoger luidheidsniveau kan worden herhaald, zonder dat dit invloed heeft op de drempelberekening (fig. 1). Verder wordt op het beeldscherm aangegeven welke woordparen voor aanbieding zijn geselecteerd, welk intensiteitsniveau de onderhavige trial heeft en na iedere successieve omslag een voorlopig berekende woordher-

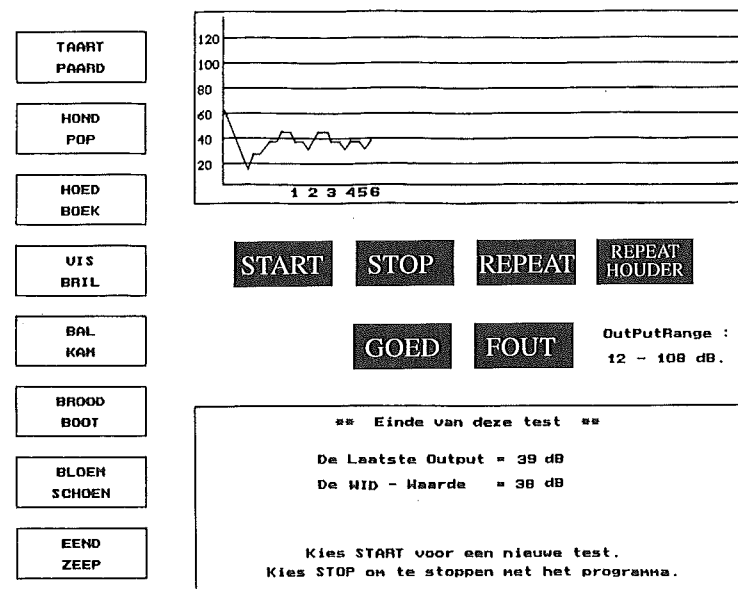


Fig. 1. Beeldscherm behorend bij de PAS-procedure.

kenningsdrempel. De resultaten met de adaptieve procedure worden bovendien ook grafisch voorgesteld. Afhankelijk van de ernst van de slechthorendheid kan uit twee output-ranges gekozen worden: een van 12 tot 108 dB (SPL) en een van 24 tot 120 dB (SPL). Ook deze keuze staat op het scherm. Indien een extreem groot geluidsniveau is gewenst bij de afname van de test, is de kwaliteit van de aangeboden woorden afhankelijk van de gebruikte versterker en luidspreker.

Proefpersonen

Aanvankelijk namen aan het normeringsonderzoek 71 kinderen tussen 1;8 en 7;2 jaar deel die afkomstig waren van diverse peuterspeelzalen, kleuterscholen of gezinnen van medewerkers. Het onderzoek werd op het Kinderaudiologisch Centrum verricht. Kinderen die jonger waren dan drie jaar kregen een audiometrisch onderzoek in het vrije veld met de hoofdwend-methode (VRA) en een warbletoon voor de voor spraak relevante octaaf frequenties tussen 500 en 4000 Hz. Van kinderen boven 3 jaar werd een toondrempelaudiogram per oor bepaald met een hoofdtelefoon en met behulp van spelaudiometrie (blokjestest, Tucker & Nolan, 1984). Overeenkomstig het klinisch-audiometrische criterium gold voor kinderen die het VRA onderzoek in het vrije veld ondergingen als uiterste toelatingscriterium een gemiddelde gehoordrempel van 20 dB (HL) en voor kinderen met spelaudiometrie een gemiddelde toondrempel van 15 dB (HL) voor het beste oor. Bij twijfel of constatering van een verminderd gehoor werd bovendien een impedantiemetrie-onderzoek verricht. Op grond van deze criteria werden 13 kinderen van het onderzoek uitgesloten. Bij het verdere onderzoek werden nog eens 6 kinderen uit de normeringsgroep verwijderd wegens onvoldoende medewerking. Uiteindelijk bleven 52 kinderen over die deel uitmaakten van de definitieve normgroep (Tabel 1).

Procedure

De akoestische output van de PAS-test werd aan het begin van elke onderzoeksdag in het vrije veld gec calibreerd met behulp van de digitaal opgeslagen ruis via de versterker en de luidspreker. Deze bevond zich hierbij ± 70 cm hoog en 1 m verwijderd van een Sound Level Meter (Brüel & Kjaer, 2203), waarvan de microfoon zich ter plaatse en op oorhoogte van de te onderzoeken kinderen bevond.

Tabel 1. Leeftijdverdeling van de kinderen met een normaal gehoor behorende tot de normgroep en van de kinderen met een gehoorverlies of die om andere redenen van het onderzoek zijn uitgesloten.

	1 - 2 jr	2 - 3 jr	3 - 4 jr	4 - 7 jr
normgroep	-	16	28	8
path. groep	1	4	6	2
uitgesloten groep	-	4	2	-

Na de calibratie kon het programma gestart en de outputrange geselecteerd worden. Tijdens de test zat het kind aan een laag tafeltje waarop de voorwerpen van een van de beide testversies uitgestald lagen. Aan de voorzijde van het tafeltje stond de luidspreker en zat ook de testleidster. Bij de kleinste kinderen stond een pop of een beer bovenop de luidspreker. Iemand anders bediende de programmacursors (zie Apparatuur), maar een ervaren akoepedist(e) kan de test gemakkelijk alleen afnemen. Na controle of alle voorwerpen overeenkomstig het bijbehorende woord herkend werden en een eenvoudige uitleg over de bedoeling van de test, werden eerst enkele testopdrachten met eigen stem gegeven, waarna duidelijk werd gemaakt dat de beer of luidspreker het verder zou overnemen en dat deze ook erg zacht kon gaan praten. Uiteindelijk werd het adaptieve programma gestart en werden de geselecteerde woorden, at random gekoppeld aan een van de drie opdrachtzinnen, gegenereerd totdat er zes omslagpunten rond de woordherkenningsdrempel waren verkregen. Na iedere keuze van een voorwerp werd dit weer tussen de andere op tafel teruggelegd. Een bepaald woord kwam nooit twee maal na elkaar voor, behalve als het nodig was dit te herhalen. Behalve de automatische registratie van de drempelwaarden werd de exacte duur van de testafname met de hand genoteerd. Van de normeringsgroep kregen 44 kinderen beide testversies A en B in gebalanceerde volgorde aangeboden. De jongste kinderen onder de 2¹/₂ jaar kregen maar één test aangeboden, waarbij de PAS-testafname voorafging aan de audiometrie.

Resultaten

Gemiddelde woordidentificatiedrempel

De gemiddelde woordidentificatiedrempels (WID) van de testversies A en B liggen zeer dicht bij elkaar en bedragen respectievelijk 28.5 dB SPL voor test A (SD = 4.6, range = 18-39 dB SPL) en 27.9 dB SPL voor test B (SD = 5.0, range = 16-40 dB SPL). Volgens een variantie-analyse is het verschil van 0.6 dB niet

Tabel 2. De gemiddelde woordidentificatiedrempel (WID) in dB SPL van beide testversies met bijbehorende standaarddeviaties (in dB) voor drie leeftijdsgroepen van kinderen en voor een groep volwassenen. Bij de kinderen is een significant afwijkende drempel aangegeven met een *.

leeftijd	test A			test B		
	gem. WID	SD	N	gem. WID	SD	N
2-3 jaar	30.0	5.5	14	31.4*	5.1	13
3-4 jaar	28.8	3.8	28	26.9	4.5	28
4-7 jaar	24.0*	2.5	7	24.5	3.0	6
totaal kinderen	28.5	4.6	49	27.9	5.0	47
volwassenen	19.0	1.4	5	19.4	3.4	5

significant [$F(1,41) = .70, p = .41$]. Hieruit kan worden geconcludeerd dat beide testversies een gelijke moeilijkheidsgraad hebben.

Uit tabel 2 blijkt dat, overeenkomstig de bevindingen met de Engelse Automated Toy Discrimination Test (Palmer et al, 1991), de identificatiedrempels hoger liggen al naar gelang de tests bij jongere kinderen worden afgenomen. Een one-way variantie analyse en een Tukey-Kramer post hoc vergelijking (Rietveld & van Hout, 1993) met betrekking tot de verschillen tussen de drie identificatiedrempels van de verschillende leeftijdsgroepen geven aan dat bij test A een significant verschil bestaat tussen de drempelscore van de twee jongste groepen tussen 2 en 4 jaar enerzijds en de oudste groep van 4-7 jaar anderzijds [$F(2,46) = 4.81, p < .05$]. Bij test B bestaat een significant verschil tussen de drempelscore van de jongste groep van 2 tot 3 jaar en die van de twee oudere groepen van 3 tot 7 jaar [$F(2,44) = 6.10, p < .05$].

Voor het vaststellen van "ideale" referentiegegevens zijn in tabel 2 ook de resultaten opgenomen van vijf volwassen proefpersonen. Zoals verwacht kan worden (Ousey et al 1989) liggen deze aanmerkelijk beneden de woordidentificatiedrempels van de kinderen.

Gemiddelde tijdsduur per test

Bij ieder kind werd vanaf het moment dat de eerste stimulus door de computer werd gegenereerd de tijdsduur van de testafname vastgelegd. Deze bedraagt gemiddeld voor test A 4 min 40 s (SD = 1 min 33 s; range 2 min- 9 min 31 s) en voor test B 5 min 2 s (SD = 1 min 54 s; range 2 min-11 min 54 s). Het verschil in gemiddelde afnameduur blijkt niet significant. Zoals het geval blijkt bij de verschillen tussen de gemiddelde drempelwaarde per leeftijdsgroep, is er ook een significant verschil in gemiddelde afnameduur van test A tussen respectievelijk de twee jongste groepen en de oudste groep [$F(2,45) = 6,74, p < .05$] en voor test B tussen respectievelijk de jongste groep en de twee oudste groepen [$F(2,44) = 8.07, p < .05$].

Oudere kinderen hebben minder tijd nodig voor de test dan jonge kinderen, zoals ook blijkt uit tabel 3. De resultaten van de afnameduur tonen aan dat de

Tabel 3. De gemiddelde testafnameduur (in seconden) van beide tests voor de drie leeftijdsgroepen met bijbehorende standaarddeviaties. Een * geeft aan dat de bijbehorende gemiddelde afnameduur significant afwijkt.

leeftijd	gem. tijd	test A		gem.tijd	test B	
		SD	N		SD	N
2-3 jaar	332.0	98.3	13	385.5*	143.7	13
3-4 jaar	279.6	83.1	28	286.2	77.1	28
4-7 jaar	188.0*	45.8	7	199.5	76.7	6
totaal	280.5	93.3	48	302.6	114.0	47

Tabel 4. Het gemiddeld aantal aangeboden stimuli per leeftijdsgroep en test met bijbehorende standaarddeviaties. De verschillen tussen de groepen zijn niet significant.

leeftijd	stimuli	test A		stimuli	test B	
		SD	N		SD	N
2-3 jaar	24.4	4.1	14	25.2	2.3	13
3-4 jaar	24.5	2.9	28	24.2	3.5	28
4-7 jaar	25.1	2.1	7	24.5	3.2	6
totaal	24.6	3.2	49	24.5	3.2	47

woordidentificatiedrempel met de adaptieve procedure echter ook relatief snel bij jonge kinderen kan worden bepaald.

Het aantal woordpresentaties per test

In de resultatenfiles van de kinderen werd automatisch vastgelegd hoeveel woorden er per test daadwerkelijk bijdroegen tot de drempelscore.

Uit tabel 4 blijkt dat het gemiddelde aantal gepresenteerde woorden voor versie A van 24.6 (SD = 3.2; range 19-33) nagenoeg gelijk is aan het gemiddelde aantal presentaties van 24.5 bij test B (SD = 3.2; range 17-31), terwijl ook de aantallen per leeftijdsgroep nauwelijks van elkaar verschillen. Er bestaat geen significant verschil tussen deze geobserveerde waarden. Uit deze resultaten kan worden geconstateerd dat de vastgestelde langere afnameduur bij jongere kinderen niet wordt veroorzaakt door een langduriger schommelen rond de woordidentificatiedrempel en het daarmee samenhangende aanbieden van meer woorden, maar eerder door een tragere reactie, een verminderde aandacht, motivatieverlies en de factor vermoeidheid. Uit de gegevens van de volwassenen, die niet in tabel 4 zijn opgenomen, blijkt het iets hoger liggende gemiddelde aantal presentaties bij de volwassenen voor versie A (M = 26.4, SD = 3.1) en voor versie B (M = 26.4, SD 4.3) evenmin significant van dat van de kinderen te verschillen. Het kleine verschil tussen het aantal presentaties van volwassenen en kinderen kan worden verklaard door het feit dat in de beginfase van beide tests aan de volwassenen meer stimuli moesten worden aangeboden omdat hun woordidentificatiedrempels lager waren dan van de kinderen.

III. Betrouwbaarheid en validiteit van de PAS-test

Proefpersonen

Voor de bepaling van de validiteit is met de PAS test een aantal slechthorende kinderen onderzocht. In dit deel van het onderzoek zijn ingesloten de resultaten van slechthorende kinderen die op beide subtests hebben gescoord. Twee kinderen met sterk discante gehoorverliezen zijn vervolgens uitgesloten.

De groep bestond uit 16 kinderen met een perceptief verlies en 14 kinderen met een geleidingsverlies. De kinderen met een perceptief verlies waren ouder dan de kinderen met een geleidingsverlies (5.8 ± 1.1 jaar vs. 3.8 ± 1.3 jaar) en hadden een ernstiger gehoorverlies (53 ± 11 dB HL vs. 34 ± 14 dB HL). In de eerste plaats werd onderzocht of de relatie tussen het gemiddelde gehoorverlies (van het beste oor), berekend bij 0.5, 1, 2 en 4 kHz, en de gemiddelde score van de beide PAS subtests bij kinderen met een perceptief verlies en een geleidingsverlies, een vergelijkbaar beeld vertoonde. Hierbij kan het verschil in leeftijd een ongewenst effect hebben, aangezien de leeftijd een significant effect heeft op de PAS drempel (zie o.a. Tabel 2). Vandaar dat een (eerste-orde) correctie is toegepast voor de leeftijd, gebaseerd op een lineaire regressie analyse van de PAS-drempel versus de leeftijd in de groep normaal horende kinderen. Hieruit bleek dat de PAS-drempel tussen 2 en 7 jaar gemiddeld afneemt met 1.92 dB (afgerond: 2 dB) per jaar. Met deze waarde zijn de PAS-drempels van de slechthorende kinderen gecorrigeerd teneinde het leeftijdeffect te minimaliseren.

Betrouwbaarheid van test A en B

Bij 71 van de normaalhorende en slechthorende kinderen werd zowel test A als test B afgenomen. De correlatie tussen de verkregen woord-identificatiedrempelwaarden van test A en B, met als hypothese dat twee parallelle testvormen met elkaar vergeleken worden (Drenth & Sijtsma, 1990), is hoog: $r = .98$. Hieruit blijkt dat naast de bevindingen uit het normeringsonderzoek beide testversies als equivalent kunnen worden beschouwd. Uitgaande van deze equivalentie is een test-hertest analyse uitgevoerd door het resultaat van de test B af te trekken van dat van test A. Een gemiddeld verschil (\pm de standaard deviatie) van 0.8 ± 4.0 dB en 0.3 ± 4.2 dB is gevonden respectievelijk voor de normaalhorende en slechthorende kinderen. Zoals verwacht wijken beide gemiddelden niet significant van 0 af. De (intra-individuele) standaard deviatie zegt wat over de onnauwkeurigheid van de testuitslag: met een 2 SD criterium (betrouwbaarheid van 95%) wordt gevonden dat de werkelijke drempel 8.4 dB of minder afwijkt van de gevonden testdrempel.

De betrouwbaarheid over N omslagpunten

De vraag kan worden gesteld hoe betrouwbaar het meetresultaat van de PAS-test is als deze voortijdig moet worden afgebroken en de schatting van de woordidentificatiedrempel moet worden gemaakt over minder dan zes omslagpunten. Hoeveel kan, uitgedrukt in dB, met enige zekerheid na een of meer omslagpunten de geobserveerde drempel afwijken van de uiteindelijke drempel bij zes omslagpunten? Voor de bepaling van het 95% betrouwbaarheidsinterval bij succesievelijk 1 t/m 6 omslagpunten is een berekening uitgevoerd volgens Plomp en Mimpen (1979) en Oursey (1989), gebaseerd op de geobserveerde verschillen tussen een voorlopige en een uiteindelijke identificatiedrempel bij de normaalhorende kinderen. Deze berekening levert het interval waarbinnen met 95% zekerheid de ware drempel geacht kan worden te vallen.

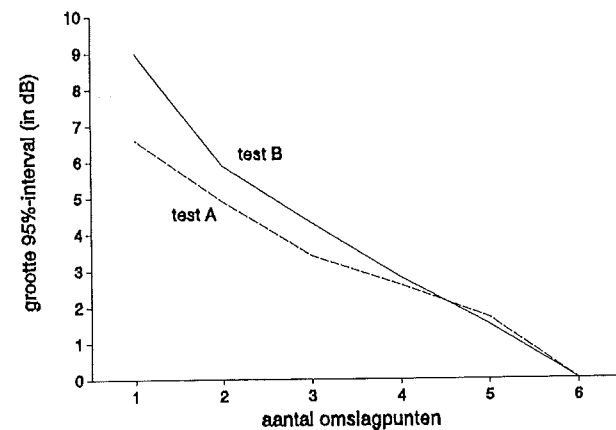


Fig. 2. Het 95% betrouwbaarheidsinterval van test A en B met betrekking tot de werkelijke identificatiedrempel in functie van het aantal gemeten omslagpunten.

Figuur 2 laat zien dat dit 95% betrouwbaarheidsinterval bij de versies A en B afhangt van het aantal omslagpunten dat bij de berekening van identificatiedrempels wordt betrokken. Overeenkomstig de verwachting blijkt het testresultaat betrouwbaarder naarmate meer omslagpunten bij de berekening van een drempel worden betrokken. Anderzijds blijkt de afwijking bij minder omslagpunten dan zes, ten opzichte van de drempelintensiteit bij afname van de gehele test, al na twee omslagen relatief klein en klinisch audiometrisch niet relevant te zijn, namelijk ± 2.5 dB bij test A en ± 3 dB bij test B. Gezien de jonge leeftijd van de kinderen steken deze bevindingen gunstig af tegen enerzijds het betrouwbaarheidsinterval van de spraakreceptiedrempel voor monosyllaben bij volwassenen volgens het DIN 45621 standaardisatie-onderzoek (Brinkmann, 1987), dat ± 3.5 dB bedraagt en anderzijds ten opzichte van het 99% betrouwbaarheidsinterval van de spraakreceptiedrempel van de SAP-test, dat bij 3 tot 4-jarigen ± 3.66 dB bedraagt (Crul, 1983).

Klinische validiteit

In Figuur 3 is de relatie weergegeven tussen de (leeftijdgecorrigeerde) PAS-drempel (gemiddelde score van de subtest A en B) en het gemiddelde gehoorverlies bij 0.5, 1, 2 en 4 kHz van het beste oor. Weergegeven zijn individuele data; de verschillende symbolen refereren naar de twee typen slechthorendheid. Aangezien op het oog de resultaten van de groep kinderen met een perceptief verlies en met een geleidingsverlies overlappen, zijn deze samengenomen in de navolgende analyse, tenzij anders vermeld.

De Pearson correlatie voor de data gepresenteerd in Figuur 3 bedraagt 0.91 waaruit geconcludeerd kan worden dat de variantie van de PAS-drempel ongeveer 83% gemeenschappelijk heeft met de variantie in de toondrempelwaarde. Wanneer de correlatie wordt berekend over de afzonderlijke groepen met een

perceptie- of geleidingsverlies, blijken de correlaties in dezelfde orde van grootte.

Op grond van de verwachting dat met het toenemen van het toondrempelverlies de PAS-drempel evenredig zal toenemen, is in de figuur de best passende lijn (bepaald met de kleinste kwadraten methode) met richtingscoëfficiënt van 1 getekend (de 45°-lijn). De individuele punten van beide subgroepen liggen gespreid rond deze lijn zodat geconcludeerd wordt dat inderdaad, zoals verwacht, de PAS-drempel evenredig toeneemt met het gehoorverlies. De figuur kan gebruikt worden om op grond van de PAS-drempel een voorspelling te doen over de gemiddelde toondrempel en wel volgens:

$$\text{toondrempel} = \text{PAS-drempel} - 16 - C$$

Hierin is C de leeftijdsafhankelijke correctie. Deze bedraagt voor tweejarigen 1 dB, voor driejarigen 3 dB, voor vierjarigen 5 dB etc. Op grond van de huidige gegevens kan gesteld worden dat dit verband geldt zowel voor kinderen met een perceptief verlies als met een geleidingsverlies, echter met een gehoorverlies van 65 dB HL of minder.

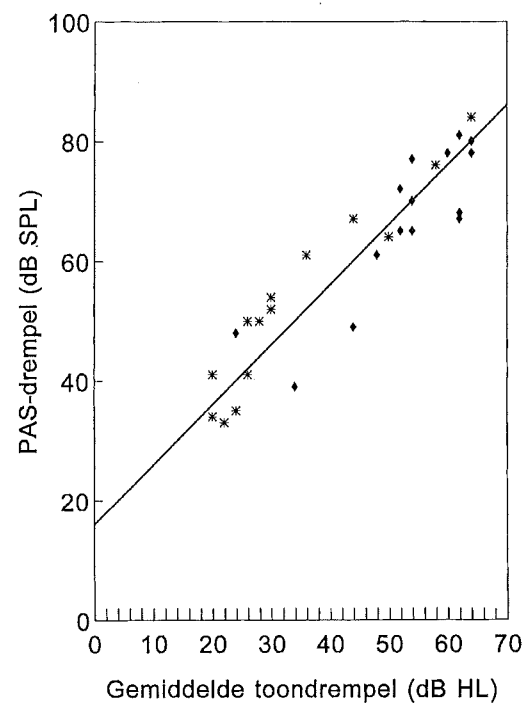


Fig. 3. De voor de leeftijd gecorrigeerde, over test A en B gemiddelde PAS-spraakdrempel als functie van de bijbehorende toondrempel van alle kinderen die beide tests aflegden en een geleidings- (*) of perceptief (◆) gehoorverlies hadden. De lijn is de regressielijn, berekend op grond van de data van kinderen met een gehoorverlies tussen 20 en 70 dB (HL).

Figuur 3 kan ook gebruikt worden om te schatten of een gevonden PAS-drempel als normaal of afwijkend moet worden gekwalificeerd. Als we aannemen dat een gemiddelde toondrempel van 20 dB nog net normaal is, dan vinden we dat de PAS-drempel kleiner of gelijk moet zijn aan 36 – C dB. Aangezien het 95% betrouwbaarheidsinterval van de PAS-drempelbepaling ruim 8 dB is (zie: Betrouwbaarheid van test A en B) komen we tot de slotsom dat met een waarschijnlijkheid van 95% een PAS-drempel van 44 – C dB pathologisch is.

Discussie

De PAS-test maakt gebruik van een procedure waarbij volgens een adaptieve methode en met eenvoudige eenlettergrepige woorden en bijbehorende objecten een betrouwbare schatting kan worden gemaakt van de spraakdrempel bij kinderen of ouderen met een mentale leeftijd van twee jaar of daarboven. Bovendien kan met behulp van de vastgestelde spraakdrempel een valide schatting worden gemaakt van de ligging van de toondrempel in het drempelaudiogram van het besthorende oor. Er zijn enkele belangrijke verschillen met de al langer bestaande SAP-test, die gebruik maakt van woorden en plaatjes in een meerkeuze-arrangement en die bedoeld is voor kinderen vanaf 3½ jaar (Crul, 1984). Met opzet is gekozen voor concrete objecten uit de leef- en belevingswereld van zeer jonge kinderen, omdat zij deze gemakkelijker herkennen dan meer abstracte afbeeldingen van objecten (Russel, 1993). Hierdoor kon de benedengrens van de leeftijd voor de doelgroep aanmerkelijk lager worden gesteld. Het jongste kind dat de test met goed resultaat heeft uitgevoerd was 21 maanden oud. Behalve de betere herkenbaarheid van de objecten oefenen deze ook een grotere aantrekkingskracht uit op jonge kinderen, waardoor de aandacht tijdens het verloop van de test gemakkelijker gevangen gehouden kan worden. In tegenstelling tot de SAP-test, die door middel van fonetisch gebalanceerde woordlijsten een volledige spraakverstaanbaarheidscurve in functie van de luidheid produceert, bepaalt de PAS-test alleen een woordidentificatiedrempel. Het voordeel bij kleine kinderen is dat een drempelbepaling volgens de gehanteerde adaptieve methode op speelse wijze in relatief korte tijd met beperkt materiaal kan worden verricht. De woordidentificatiedrempel geeft een inzicht hoe een kind auditief functioneert bij het luisteren naar spraak. Dit kan een belangrijk gegeven zijn bij het evalueren van de verschillen in pre- en postoperatief vastgestelde spraakdrempels, bijvoorbeeld de winst na het plaatsen van trommelvliesbuisjes of na een middenoor-operatie. Het kan verder een belangrijk gegeven zijn bij het instellen van de versterking van hoorapparatuur bij kleine kinderen en bij het vaststellen van het voordeel dat hoorapparatuur biedt bij het luisteren naar spraak.

De PAS-test meet geen discriminatieverlies bij het verstaan van spraak. Evenmin is het de bedoeling om met deze test een analyse te maken van wel of niet waargenomen spraakklanken of foneemkenmerken en onderlinge substituties. Voor dergelijke gegevens bestaan voor iets oudere kinderen andere middelen,

zoals de SAP-test en diverse Auditieve Discriminatie-tests (Crul & Peters, 1976; Crul 1984).

De PAS-test kan ook worden afgenomen bij geestelijk gehandicapte kinderen of ouderen, mits deze over een mentale ontwikkelingsleeftijd beschikken van tenminste 2 jaar. Aangezien met een beperkter aantal woorden en objecten dan in de test aanwezig gewerkt kan worden, kunnen de begrippen in de meeste gevallen van tevoren inge oefend worden. Diverse jonge kinderen met het syndroom van Down legden bij de validering de test met goed resultaat af. De test biedt bovendien het voordeel dat deze gemakkelijk door slechts één ervaren akoepe-dist of logopedist kan worden afgenomen, omdat alleen de goede- of foute responsen door indrukken van een toets in de onmiddellijke nabijheid van de testtafel geregistreerd worden. Omdat de testafname verder automatisch verloopt, kan meer aandacht aan het te onderzoeken kind worden geschonken. Door de objecten op tafel verder uit elkaar te plaatsen is het ook mogelijk om via observatie van oogbewegingen of hoofdwendingen de spraakdrempel van ernstig motorisch gehandicapte kinderen, die onvoldoende kunnen manipuleren, te bepalen.

Ofschoon met de PAS-test een belangrijk aspect van de spraakwaarneming als de spraakdrempel kan worden bepaald bij kleine kinderen die nog geen formele spraakaudiometrische test kunnen afleggen, geldt toch de aanbeveling om zo spoedig mogelijk te beschikken over meer volledige spraakaudiometrische gegevens. De test kan in twee gevallen een vertekend beeld geven van de werkelijkheid. Bij een discantverlies kan de gevonden spraakdrempel niet overeen komen met de gemiddelde toondrempel en levert dan meestal een gunstiger beeld. Omdat de test in het vrije veld wordt afgenomen, wordt een asymmetrisch gehoorverlies of eenorig horen niet opgespoord, omdat de drempel berust op de waarneming van het beste oor. Omdat eenorigheid als een wezenlijke handicap wordt beschouwd (Bess & Thorpe, 1988) dient hiermee bij alleen de afname van de PAS-test rekening te worden gehouden.

De Engelse groep van audiologen die de Automated Toy Test heeft ontwikkeld tracht hiervan gebruik te maken bij binaurale maskeerexperimenten (BMLD) bij kleine kinderen. Het doel is om op deze wijze te onderzoeken of de auditieve integratie van via beide oren binnengekomen informatie goed verloopt, bijvoorbeeld als er eerder sprake is geweest van auditieve deprivatie. In de nabije toekomst zal dergelijk onderzoek eveneens worden gedaan op de audiologische centra waar de auteurs werkzaam zijn.

Om verspreiding van de PAS-test mogelijk te maken, wordt momenteel moeite gedaan om de nodige software en hardware beschikbaar te kunnen stellen voor andere gebruikers. Over de test zal in de tweede helft van 1994 kunnen worden beschikt.

Summary

The PAS-test is an automated, sensitive and accurate procedure, constructed in order to

obtain word identification thresholds in quiet from children with mental ages of 2 years and above. The Dutch version is an elaborated construction of the McCormick Automated Toy Discrimination Test. In this paper the development and standardization of two parallel versions, A and B, are described. The test proves to be both reliable and valid, and provides a basis for estimating the mean elevation of pure tone-threshold of a child. The correlation between speech and pure-tone thresholds is high ($r = .91$). The average of the better ear pure-tone threshold at 0.5, 1, 2 and 4 kHz can be predicted from the word-discrimination threshold obtained with the PAS-test, with a 95% confidence interval of ± 8 dB.

Dankwoord

Wij danken Drs. M. Franken voor het inspreken van het woordmateriaal, Drs. P. Groenen, Dr. B. Maassen en Dr. A. Rietveld voor hun hulp bij het bewerken van de stimuli, en M. van der Aakster voor zijn bijdrage aan het computerprogramma. Bij het normeringsonderzoek werd audiometrische hulp verleend door M. Frösch en J. van der Stappen.

Literatuurlijst

- Alferink, M. (1993). *Kinder-Spraakaudiometrie bij kinderen van 2 tot 4 jaar: Eerste onderzoek naar woordmateriaal. Verdieping Spraakstoornissen; Vrije Studierichting Spraak & Taalpathologie.*
- Bess, F.H., & Thorpe, A.M. (1988). Performance and management of children with unilateral sensorineural hearing loss. *Scandinavian Audiology*, Suppl. 30, 75-79.
- Barton, D.P. (1975). Statistical significance in phonemic perception experiments. *Journal of Child Language*, 2, 297-298.
- Braks, J.T.M. (1993). *PAS: Peuter Adaptieve Spraakdrempelbepaling.* Afstudeerscriptie Vrije Studierichting Spraak & Taalpathologie, KUN.
- Brinkman, K. (1987). The German Path to Standardization in Speech Audiometry. In M. Martin (Ed.), *Speech Audiometry*, pp. 89-107. London: Whurr Publishers.
- Crul, Th.A.M. (1983). *SAP, een gestandaardiseerde spraakaudiometrie test met plaatjes voor peuters, kleuters en geretardeerde kinderen.* Intern rapport Katholieke Universiteit, Nijmegen no. 83 WTS 02.
- Crul, Th.A.M. (1984). Spraakaudiometrie met plaatjes. *Logopedie en Foniatrie*, 56, 2-6.
- Crul, Th.A.M. (1986). Het belang van vroegtijdige onderkenning en interventie voor het slechthorende kind. *Medisch Contact*, 42, 1359-1360.
- Crul, Th.A.M., & Peters, H.F.M. (1976). *Auditieve Discriminatie Test: ADIT.* Amsterdam: Swets en Zeitlinger.
- Drenth, P.J.D., & Sijtsma, K. (1990). *Testtheorie: Inleiding in de theorie van de psychologische test en zijn toepassingen.* Houten: Bohn Stafleu Van Loghum.
- Haan, de R. (1988). *S.O.B.: Sound Output Box.* Rapport Instituut voor Doven, St. Michielsgestel, afdeling Research & Development.
- ILS-PC, V6.1. (1989). *Interactive Laboratory System.* Goleta: Signal Technology.
- Kendall, D.C. (1953). Audiometry for young children. *Teacher of the Deaf*, 51, 171-177.
- Kendall, D.C. (1954). Audiometry for young children. *Teacher of the Deaf*, 52, 18-23.
- Lem, G.J. van der, & Baart de la Faille. (1981). Vroegtijdige opsporing van gehoorstoornissen. *Nederlands Tijdschrift voor Geneeskunde*, 51, (125) 2104-2109.
- Levitt, H. (1971). Transformed up-down methods in psycho-acoustics. *The Journal of the Acoustical Society of America*, 49, 467-477.

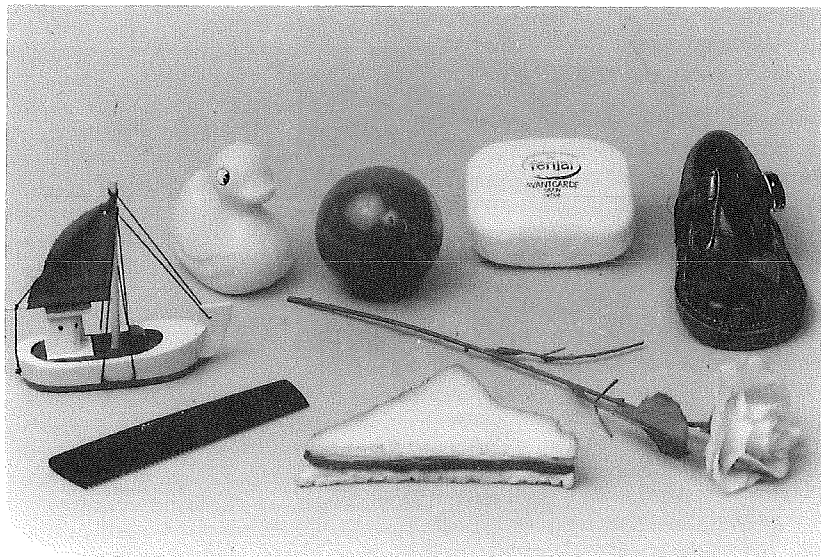
- McCormick, B. (1977). The Toy Discrimination Test: an aid for screening the hearing of children above a mental age of two years. *Public Health*, 91, 67-69.
- McCormick, B. (1988). *Screening for hearing impairment in young children*. Beckenham: Croom Helm.
- Northern, J.L., & Downs, M. (1991). *Hearing in Children*. Baltimore, MA: The Williams & Wilkins Company.
- Ousey, J., Sheppard, S., Twomey, T., & Palmer, A.R. (1989). The IHR-McCormick Automated Toy Discrimination Test – description and initial evaluation. *British Journal of Audiology*, 23, 245-249.
- Palmer, A.R., Sheppard, S., & Marshall, D.H. (1991). Prediction of hearing thresholds in children using an automated toy discrimination test. *British Journal of Audiology*, 25, 351-356.
- Plomp, R., & Mimpen, A.M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, 43-52.
- Rietveld, A.C.M., & Hout van, R. (1993). *Statistical techniques in the study of language and language behaviour*. Berlin: Mouton/De Gruyter.
- Russel, R.C. (1993). Influence of illustration type on picture naming. *Annual Convention of the American Speech-Language-Hearing Association*, p. 171.
- Tucker, J., & Nolan, M. (1984). *Educational Audiology*. London: Croom Helm.
- Wilson, R.H., Preece, J.P., & Crowther, C.S. (1991). Enhancement of word-recognition performance with a filtering technique. *Journal of Speech and Hearing Research*, 34, 1436-1438.

Bijlage A

Objecten behorende bij test A



Objecten behorende bij test B



Woorden bij test A: taart - paard; hond - pop; hoed - boek; vis - bril.

Woorden bij test B: bal - kam; brood - boot; bloem - schoen; eend - zeep.

Draagzinnen: - geef mij maar een; - pak maar een; - zoek maar een