

Evaluatie Kinderhoortest

Augustus 2014

Uitgevoerd in opdracht van:



Datum	Augustus 2014
Versie	2.0
Status	Eindrapportage
Auteurs	Marya Adluni-Sheikh Rashid, MSc Prof.dr.ir. Wouter A. Dreschler, AMC
Trefwoorden	Kinderhoortest, kinderen, spraak-in-ruis, screening, inventarisatie
Opmerkingen	Een evaluatie van de Kinderhoortest.

Inhoudsopgave

Inhoudsopgave	1
Beschrijving van de originele Kinderhoortest	2
Doelstelling van de studie	3
Populatiekenmerken en testuitslagen	4
De aanloop naar de meting en de lengte van de test	10
Woordverwisselingen.....	13
Woord-specifieke discriminatiefuncties.....	14
Aanbevolen aanpassingen van de Kinderhoortest.....	17
Bijlage 1. Discriminatiefuncties van de woorden	19

Beschrijving van de originele Kinderhoortest

De Nationale Hoorstichting heeft verschillende hoortesten gelanceerd om mensen de gelegenheid te geven op een laagdrempelige manier hun gehoor te testen. Dit onderzoek richt zich op de evaluatie van de Kinderhoortest (KHT).

De Kinderhoortest is een spraak-in-ruis screeningstest. De test is alleen gevoelig voor perceptieve gehoorverliezen, waarbij de selectiviteit van het gehoor wordt aangetast en niet voor geleidingsverliezen, zoals die vaak voorkomen bij jonge kinderen. Bij oudere kinderen in de leeftijd van 6 tot 12 jaar (de doelgroep van de Kinderhoortest) komt geleidingslechl-horendheid minder frequent voor dan op jongere leeftijd, maar het blijft één van de mogelijke oorzaken. Het is dan ook van belang in de communicatie rondom de test te benadrukken dat een spraak-in-ruis test niet gevoelig is voor geleidingsverliezen.

Voor de Kinderhoortest worden negen verschillende woorden gebruikt, die steeds te zien zijn op het scherm in een vierkant schema met de plaatjes van de negen woorden (figuur 1). Bij het doen van de Kinderhoortest krijgen de respondenten achtereenvolgens 27 keer gerandomiseerd één van deze negen woorden te horen, waarbij de intensiteitsverhouding tussen spraak en ruis (de signaal-ruis verhouding in dB) varieert. Per keer moet de respondent het plaatje van het gehoorde woord aanklikken. Als de respondent het woord niet verstaan heeft, kan hij/zij een icoon met een vraagteken (niet verstaan) aanklikken. De eerste zeven aanbiedingen worden als aanloop van de meting beschouwd en de laatste 20 als de eigenlijke meting.

Sinds 15 januari 2010 wordt er gebruik gemaakt van een nieuwe maskeerruis met gewijzigde afkapwaarden, gelijk aan die van de Oorcheck. Daarvoor werd er gebruik gemaakt van een breedbandige ruis. Het is niet helemaal duidelijk met welk doel dit is gebeurd. Ook is niet onderzocht of de afkappunten van de Oorcheck zondermeer konden worden overgenomen voor de Kinderhoortest. Het is niet waarschijnlijk dat deze afkappunten exact hetzelfde zijn, mede omdat er bij de Oorcheck kleine niveau-aanpassingen hebben plaats gevonden (het zogenaamde homogeniseren) en dit is bij de Kinderhoortest (nog) niet gebeurd.

Overwogen moet worden of de ruis niet beter weer kan worden veranderd naar de breedbandige versie (**aandachtspunt 1**).



Figuur 1. Scherm Kinderhoortest

Doelstelling van de studie

Ten behoeve van de optimalisatie van de Kinderhoortest is een analyse verricht op testuitslagen uit het verleden. Het doel was om inzicht te verkrijgen in diverse aspecten van de test om zodoende aangrijpingspunten te vinden ter verbetering van de Kinderhoortest. In de analyse stonden de volgende aspecten centraal:

- Karakteristieken en testuitslagen van de deelnemers.
- De invloed van de aanloop naar de daadwerkelijke meting en de lengte van de test op de verdeling van de testuitslagen.
- Het patroon van woordverwisselingen.
- De ligging en de steilheid van de woord-specifieke discriminatiecurves.

In de volgende paragrafen zullen de verschillende aspecten behandeld worden. In de laatste paragraaf van dit hoofdstuk zal een overzicht gegeven worden van de aanbevolen aanpassingen naar aanleiding van de bevindingen van het inventariserend onderzoek.

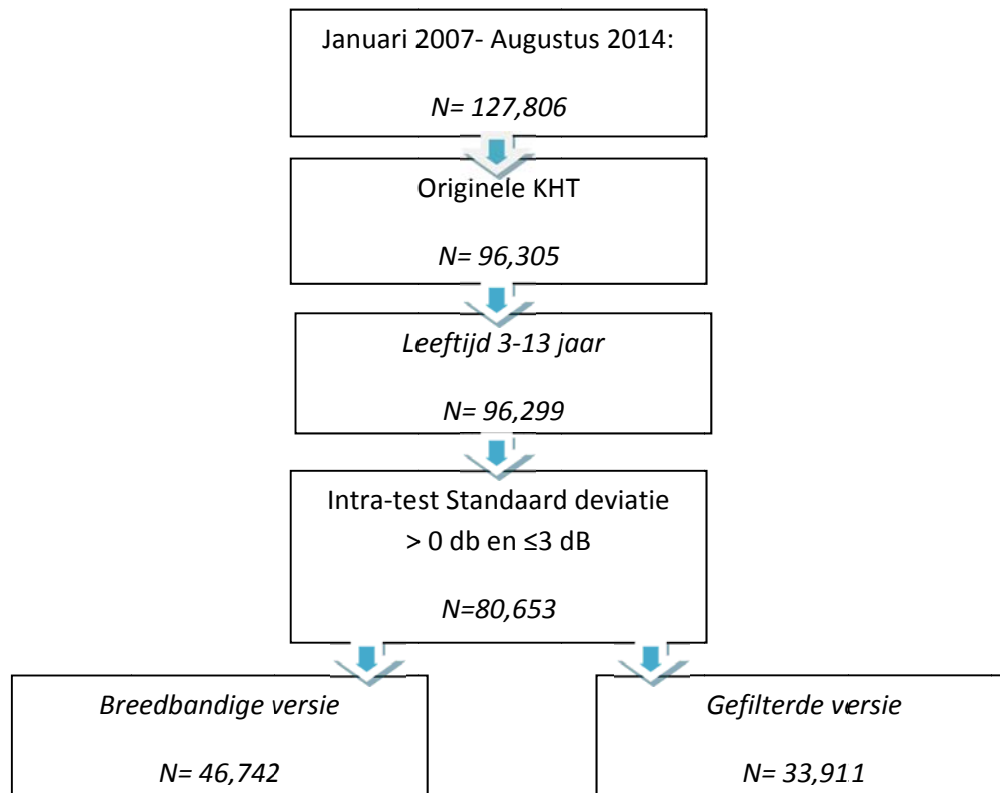
Populatiekenmerken en testuitslagen

Gegevens van Kinderhoortest-uitslagen, verzameld in de periode van 5 januari 2007 tot en met het heden (5 augustus 2014), werden geanalyseerd (tabel 1). Niet alle resultaten bleken geschikt te zijn voor analyse. Er bestaan verschillende versies van de Kinderhoortest met diverse populaire sprekers (Bibi, Edward, Sita). Om de nauwkeurigheid van de test te waarborgen, is alleen de Kinderhoortest met de originele spreker geanalyseerd. Verder werden alleen de uitslagen geselecteerd van respondenten tussen de 3 en 13 jaar. Tenslotte werden onbetrouwbare uitslagen uitgesloten. De betrouwbaarheid van de testresultaten werd bepaald aan de hand van de standaarddeviatie van de adaptieve procedure. Een standaarddeviatie van 0 dB of groter dan 3 dB werd als onbetrouwbaar beschouwd. Een standaarddeviatie groter dan 3 dB kan het gevolg zijn van gebrek aan aandacht.

Tabel 1. Aantal respondenten per jaar

Jaar	Aantal
2007	34,114
2008	24,235
2009	17,342
2010	14,357
2011	14,416
2012	10,671
2013	7,732
2014	4,939
Totaal	127,806

Na het uitsluiten van niet-originele versies van de test, respondenten jonger dan 3 en ouder dan 13 jaar en onbetrouwbare resultaten bleven er ruim 80, 000 testuitslagen over voor analyse (zie figuur 2).



Figuur 2. Flowchart van aantal resultaten van de KHT bruikbaar voor analyse

Om de resultaten correct te kunnen analyseren, zijn de uitslagen gesplitst per ruis type. Zie tabel 2 voor de uitslagcategorïeën en bijbehorende afkapwaarden voor beide versies.

Tabel 2. Uitslagcategorïeën en afkapwaarden voor de breedbandige en gefilterde KHT

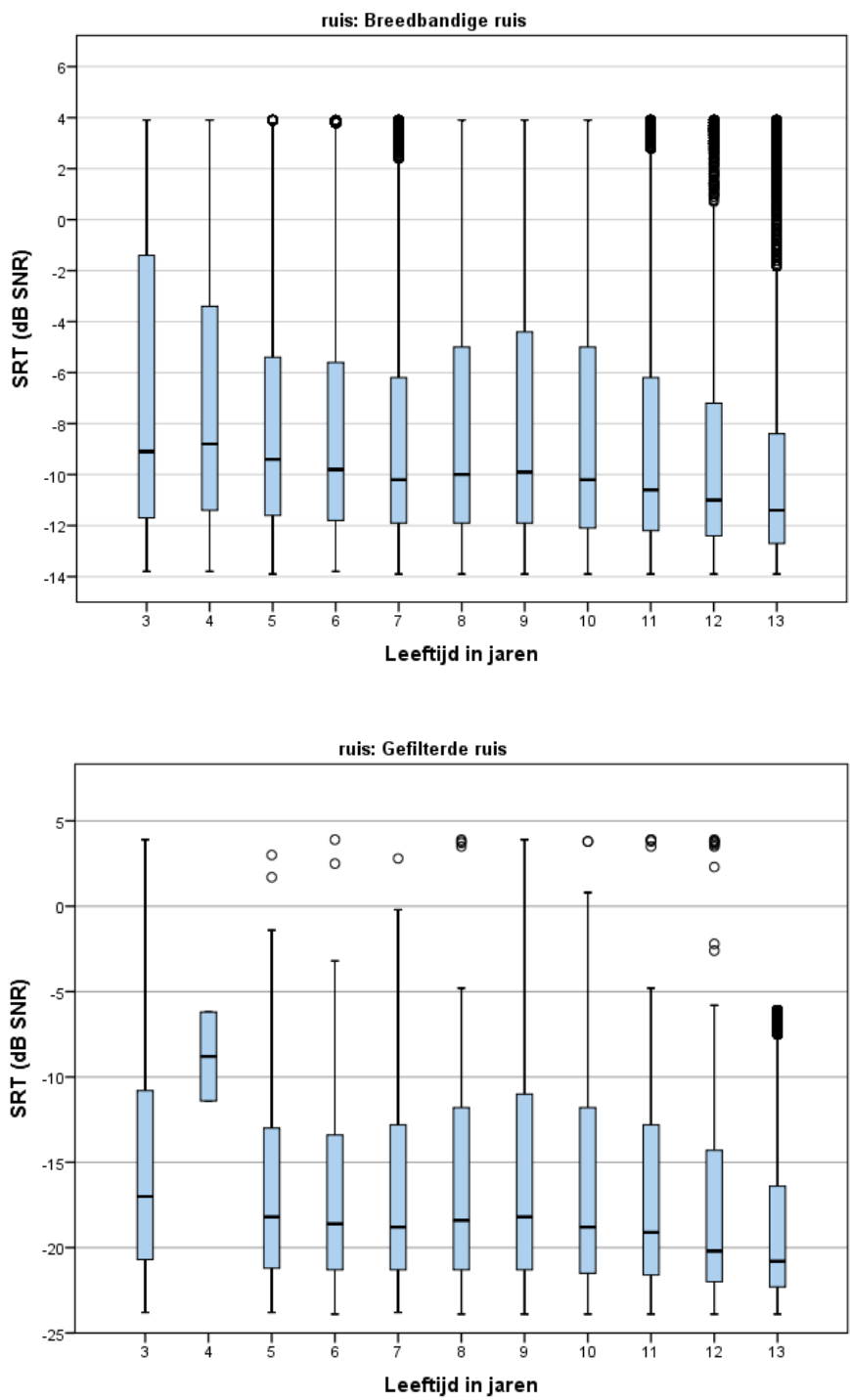
Uitslag categorie	Breedbandige versie	Gefilterde versie
Goed	≤ -8 dB	$\leq -18,4$ dB
Minder/onvoldoende	> -8 dB en < -4 dB	$> -18,4$ dB en $< -12,7$ dB
Slecht	≥ -4 dB	$\geq -12,7$ dB

Door middel van beschrijvende statistiek is er inzicht verkregen in de karakteristieken en testuitslagen van deelnemers aan de Kinderhoortest. De meeste variabelen zijn in percentages en/of gemiddelden gerapporteerd. De gemiddelde leeftijd bedroeg 9,23 jaar (standaard deviatie 3,01 jaar). Tabel 3 geeft de overige karakteristieken van deelnemers aan de KHT per ruisstype weer.

Tabel 3. Karakteristieken van deelnemers aan de KHT, gesplitst per ruis

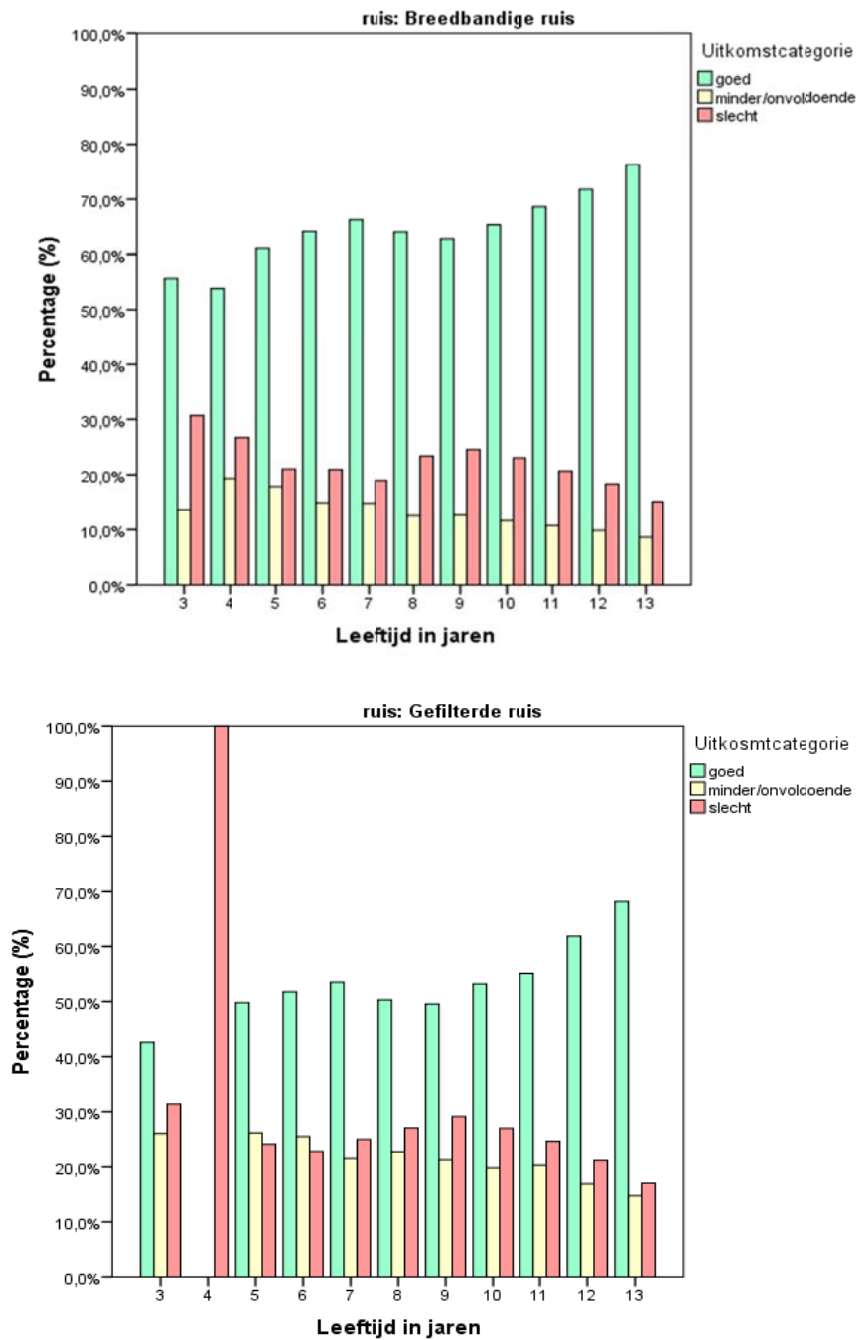
Karakteristieken	Breedbandige versie	Gefilterde versie
Gemiddelde leeftijd (sd)	9,08 (2,96)	9,43 (3,07)
Geslacht		
Jongen	7,548 (16,1%)	12,840 (37,9%)
Meisje	12,505 (26,8%)	21,062 (62,1%)
Niet ingevuld	26,689 (57,1%)	8 (0,0%)
Eigen inschatting van het gehoor		
Goed	33,368 (71,4%)	24,134 (71,2%)
Minder	10,355 (22,2%)	7,660 (22,6%)
Slecht	3,019 (6,5%)	2117 (6,2%)
Gemiddelde testuitslag (SNR in dB)	-8,04 (5,49)	-17,08 (5,70)

Het waren vooral meisjes die de Kinderhoortest deden (voor de breedbandige versie zijn er veel missende gegevens wat betreft geslacht. Waarschijnlijk was het de eerste jaren nog niet verplicht om je geslacht in te vullen). De meeste respondenten schatten hun eigen gehoor goed in (71%). De gemiddelde testuitslag over alle deelnemers is -8,04 dB (sd= 5,49) voor de breedbandige testversie en -17,08 dB (sd= 5,70) voor de gefilterde testversie. Voor de breedbandige versie komt dit overeen met een goede uitslag en voor de gefilterde versie met een onvoldoende uitslag. In figuur 3 worden boxplots weergegeven van de resultaten van de twee ruisversies van de KHT gesplitst voor leeftijdsjaren. De score lijkt iets te verbeteren naarmate de leeftijd toeneemt. Dit beeld is in overeenstemming met het onderzoek van de pDIN, dat op het VUmc is uitgevoerd (Koopmans, 2014).



*Figuur 3. Boxplots van SRT's per leeftijdsjaar, per ruistype
 (*Gefilterde ruis: het aantal respondenten met een leeftijd van 4 jaar is 2)*

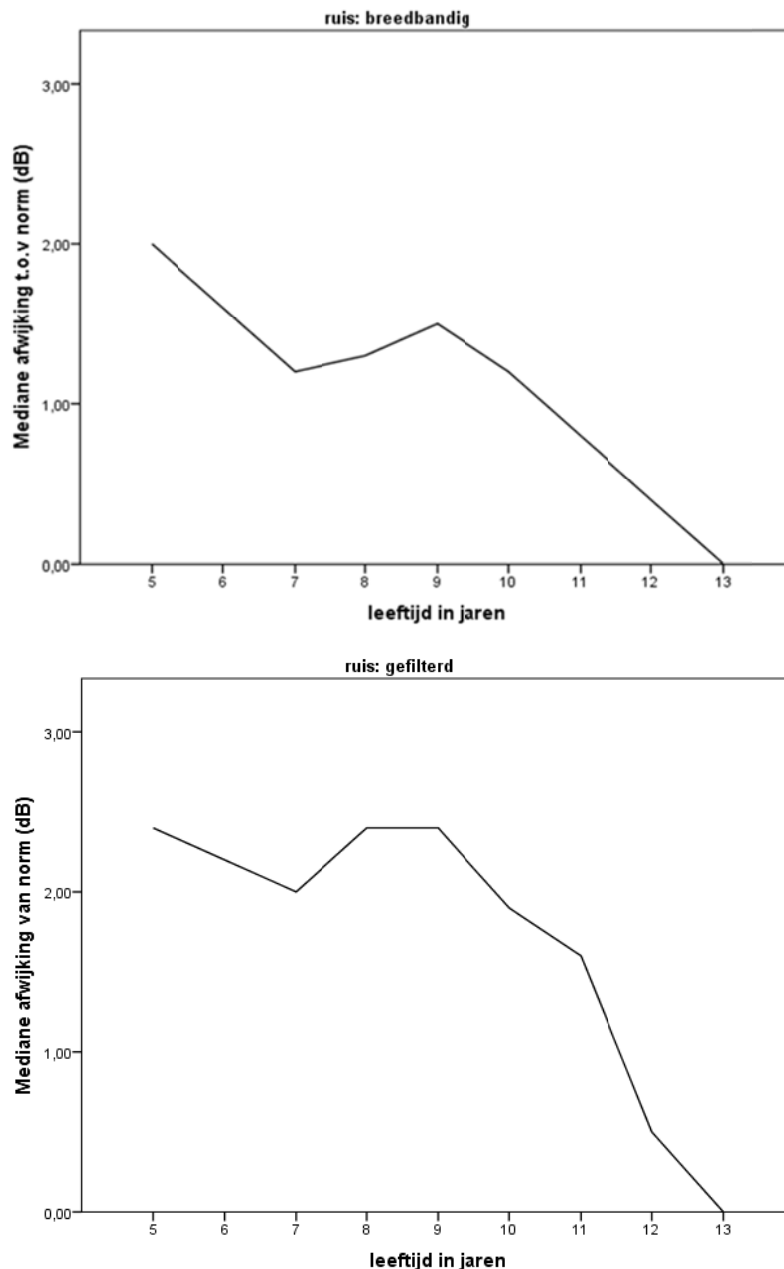
Figuur 4 geeft de uitslag van de gehoortest in categorieën weer, per leeftijdsjaar (Er waren maar twee 4-jarigen die de gefilterde versie van de test hebben uitgevoerd).



Figuur 4. Testuitslag (in categorieën) per leeftijdsjaar

Een gevolg van de leeftijdsafhankelijkheid van de uitslag is dat er – bij gebruik van een vast criterium - vaker goed wordt gescoord naarmate de respondenten ouder worden. Om een foutieve interpretatie van deze uitkomsten te voorkomen zou een leeftijdsafhankelijk criterium overwogen kunnen worden (**aandachtspunt 2**). Wat voor een 5-jarige een normale uitslag is, zou namelijk voor een 12-jarige een afwijkende uitslag kunnen betekenen.

Figuur 5 geeft per leeftijd de afwijking van de mediaan voor 5 tot 12 jarigen t.o.v. de mediaan van 13-jarigen (de 13-jarigen worden als oudste leeftijdsgroep als norm beschouwd). Zo kan de leeftijdsafhankelijkheid van de test vergeleken worden met de leeftijdsafhankelijkheid die bij de pDIN is gevonden (Koopmans, 2014). Er is bewust gekozen om de mediaan te nemen voor het leeftijdseffect. In dat geval heeft het kleine percentage slechthorenden in iedere leeftijdsgroep geen effect op de uitkomst. Uit de figuur blijkt dat er een leeftijdseffect is van ongeveer 2 dB voor de breedbandige ruis en een effect van bijna 2,5 dB voor de gefilterde ruis.



Figuur 5. Leeftijdsafhankelijke afwijking t.o.v. de norm (13 jaar)

De aanloop naar de meting en de lengte van de test

Zoals eerder beschreven worden bij de originele KHT 27 woorden aangeboden. Vervolgens worden de eerste 7 stimuli als aanloop naar de meting beschouwd. Van de laatste 20 presentatieniveaus die voor de meting gebruikt worden, wordt de SRT bepaald. Leereffecten en concentratie kunnen een rol spelen bij de bepaling van een testuitslag. Om deze reden is bestudeerd wat de invloed van de aanloop naar de daadwerkelijke meting is op de resultaten. Tevens is gekeken naar de invloed van de lengte van de test. Gemiddelde SRT's en intra-test standaarddeviaties voor verschillende combinaties van aanloop en lengte zijn weergegeven in tabel 4. Voor deze analyse werden ook de testuitslagen met een intra-test standaarddeviatie groter dan 3 dB bekeken (N=94,252).

Tabel 4. Invloed van aanloop/lengte meting op testuitslag

Omschrijving aanloop/lengte	Te middelen stimuli	SRT(dB)	Intra-test sd (dB)
Breedbandig			
<i>Huidige meting</i>	8 t/m 27	-7,79	2,06
<i>Langere aanloop</i>	11 t/m 27	-7,96	1,94
<i>Zeer lange aanloop</i>	16 t/m 27	-8,11	1,77
<i>Kortere lengte 1</i>	8 t/m 20	-7,57	1,94
<i>Kortere lengte 2</i>	8 t/m 23	-7,67	2,00
<i>Kortere lengte 3</i>	8 t/m 25	-7,74	2,03
<i>Langere aanloop + korte meting</i>	11 t/m 20	-7,78	1,79
Gefilterd			
<i>Huidige meting</i>	8 t/m 27	-16,98	2,24
<i>Langere aanloop</i>	11 t/m 27	-17,23	2,08
<i>Zeer lange aanloop</i>	16 t/m 27	-17,48	1,83
<i>Kortere lengte 1</i>	8 t/m 20	-16,63	2,09
<i>Kortere lengte 2</i>	8 t/m 23	-16,79	2,16
<i>Kortere lengte 3</i>	8 t/m 25	-16,89	2,20
<i>Langere aanloop + korte meting</i>	11 t/m 20	-16,95	1,91

In Tabel 5 zijn de resultaten uit tabel 4 anders weergegeven om de effecten eenvoudiger te kunnen interpreteren. De resultaten van de bestaande testen zijn grijs gearceerd en gelden als uitgangspunt.

Tabel 5. Analyse van de invloed van aanloop/lengte meting op testuitslag

Breedbandige ruis

Lengte Aanloop	Absolute testlengte >>			
	27	25	23	20
7	-7,79 (sd 2,06)	-7,74 (sd 2,03)	-7,67 (sd 2,00)	-7,57 (sd 1,94)
10	-7,96 (sd 1,94)			-7,78 (sd 1,79)
15	-8,11 (sd 1,77)			

Gefilterde ruis

Lengte Aanloop	Absolute testlengte >>			
	27	25	23	20
7	-16,98 (sd 2,24)	-16,89 (sd 2,20)	-16,79 (sd 2,16)	-16,63 (sd 2,09)
10	-17,23 (sd 2,08)			-16,95 (sd 1,91)
15	-17,48 (sd 1,83)			

Bij de veranderingen van de aanloop of van de absolute lengte (of een combinatie daarvan) zien wij consistente effecten, die optreden voor de breedbandige ruis zowel als voor de gefilterd ruis. Het is belangrijk te benadrukken dat dit onafhankelijke metingen zijn in twee verschillende groepen van grote omvang.

In de verticale richting bewegend (naar beneden) wordt de aanloop langer (bij een gelijke absolute testlengte) en vindt de middeling dus plaats over minder meetpunten. Het effect is dat de drempel beter wordt en de standaard deviatie afneemt. Dit wijst er op dat de aanloop niet te kort genomen mag worden om te voorkomen dat de onzekerheden van de aanloop doorklinken in de resultaten.

In de horizontale richting bewegend (naar rechts) wordt de absolute testlengte korter (bij een gelijkblijvende aanloop) en vindt de middeling dus ook over minder meetpunten plaats. Het effect is voor de standaard deviatie vergelijkbaar: de standaard deviatie neemt af. Dit wijst er op dat de meting in absolute zin niet te lang mag zijn om de voorkomen dat concentratie, vermoeidheid of verveling gaan doorklinken in de resultaten. Maar de drempel neemt toe, mogelijk doordat er bij een korte meting de drempel sterker afhankelijk is van een leereffect.

Er lijken dus tegengestelde effecten van invloed op het eindresultaat. Een lage drempel wordt bereikt door een lange aanloop en een lange absolute testduur, maar voor de nauwkeurigheid (afgeleid van de standaard deviatie binnen de meting) mag de totale lengte niet te groot worden. Een aantrekkelijk compromis tussen deze tegengestelde effecten lijkt een aanloop van 11 en een totale absolute lengte van 20 aanbiedingen. De resultaten zijn geel gearceerd en tonen dat gemiddeld over grote groepen een vergelijkbaar resultaat wordt behaald als in de oorspronkelijke test met een duidelijk gereduceerde standaard deviatie. Dit geldt zowel voor de breedbandige ruis als voor de gefilterde ruis.

De verschillen zijn significant volgens GLM repeated measures, pairwise comparisons analyse. Bovenstaande analyses wijzen er op dat de keuzen voor aanloop en testlengte in de oorspronkelijke testen niet optimaal zijn geweest. Met een kortere test worden zelfde resultaten behaald, maar met een kleinere standaarddeviatie. Overwogen moet worden om dit aan te passen.

Woordverwisselingen

Als bepaalde stimulus-woorden vaak verkeerd worden verstaan of regelmatig met een specifieke verkeerde respons worden aangeduid, kan dit wellicht duiden op een testprobleem in plaats van een gehoorprobleem. Om hier zicht op te krijgen is een woordverwisselingsmatrix opgesteld (voor de breedbandige versie). Voor deze woordverwisselingsmatrix zijn voor de laatste 20 aangeboden woorden in iedere test de stimulus en de respons tegen elkaar uitgezet (tabel 6). Uit deze matrix is af te lezen welke stimulus woorden vaak goed en welke vaak verkeerd worden gehoord.

Er werden opvallend veel fouten gemaakt in de woorden 'tong', 'boek' en 'leeuw'. Deze woorden werden in meer dan de helft van de gevallen niet verstaan of verward met een ander woord. 'Boek' werd opvallend veel verward met 'kip' (34,5%), maar 'kip' andersom niet met 'boek' (4,4%). Er was geen sprake van hoge woordverwisselingen tussen naastliggende respons plaatjes.

Tabel 6. Woordverwisselingen in percentages (voor de breedbandige ruis versie)

Stimulus \ Respons	Duim	Vuur	Maan	Boek	Roos	Kip	Leeuw	Geit	Tong	Niet verstaan	Totaal
Duim	65,2	2,4	6,2	2,2	2,6	2,3	1,8	4,5	2,2	10,7	100
Vuur	2,4	54,3	2,3	4,9	3,1	6,0	3,6	2,4	2,7	18,4	100
Maan	2,4	2,3	72,6	1,9	2,3	3,4	1,5	2,0	1,7	9,9	100
Boek	2,0	2,8	1,8	33,3	3,0	34,5	2,5	2,2	2,9	14,9	100
Roos	3,1	2,4	2,1	3,4	57,4	2,7	2,9	3,7	8,1	14,3	100
Kip	1,9	3,5	1,9	4,4	2,6	65,3	2,1	2,1	2,2	13,9	100
Leeuw	2,9	10,6	2,8	4,9	4,9	7,2	35,4	3,0	4,2	24,3	100
Geit	4,1	1,9	2,0	2,0	2,5	2,9	2,4	68,5	2,3	11,4	100
Tong	4,1	6,4	3,6	6,7	5,3	12,1	4,2	4,0	18,5	35,2	100

Woord-specifieke discriminatiefuncties

Woord-specifieke discriminatiefuncties ontstaan als per stimuluswoord het percentage correcte scores wordt bepaald als functie van het aanbiedingsniveau. Om de aanbiedingsniveaus over proefpersonen te kunnen combineren wordt gewerkt met het relatieve aanbiedingsniveau ten opzichte van de individuele gehoordrempel.

Voor ieder woord is er een apart discriminatiecurve bepaald, voor de breedbandige testversie (bijlage 1). In de figuren staat op de x-as het presentatieniveau t.o.v. de individuele drempel en op de y-as het percentage correct beantwoord. Als het presentatieniveau en de individuele drempel gelijk zijn, wordt er verwacht dat in 50% van de gevallen er een juiste respons is. Als dit percentage hoger dan de individuele drempel ligt dan is het woord perceptief relatief makkelijker en als het percentage lager ligt is het woord relatief moeilijker.

Op basis van de verschillende figuren kan geconstateerd worden dat de woorden 'tong', 'boek' en 'leeuw' lastige woorden zijn (dit blijkt ook uit de matrix), terwijl 'maan' en 'geit' de makkelijkste stimuli blijken te zijn. De woord-specifieke discriminatiefuncties zijn tevens geschat met behulp van logistische regressie analyse. Dit model vat de discriminatiefunctie samen met twee getallen:

1. De verschuiving van het 50%-punt in dB (ten opzichte van 0 dB), die samenhangt met de moeilijkheidsgraad van het woord (shift). Een hogere positieve waarde komt overeen met een moeilijker woord.
2. De helling in %/dB, die samenhangt met het discriminerend vermogen van het betreffende woord (slope). Hoe hoger de waarde, des te steiler de curve en des te groter het discriminerend vermogen.

Deze parameters van de woord-specifieke discriminatiefuncties worden in tabel 7 samengevat. De woorden 'duim', 'roos' en 'tong' hebben discriminatiecurven met de steilste helling. De curve van het woord 'tong' heeft zelfs een zeer steile helling t.o.v. de andere woorden, namelijk 15,5%/dB. De woorden 'vuur' en 'geit' zijn de zwakste woorden met een helling kleiner dan 10%/dB. Verder bestaan er grote verschillen in moeilijkheidsgraad tussen de woorden. De verschuiving van het 50%-punt in dB ten opzichte van 0 dB varieert van -2,55 dB voor het eenvoudigste woord ('maan') tot +2,71 dB voor het moeilijkste woord ('tong').

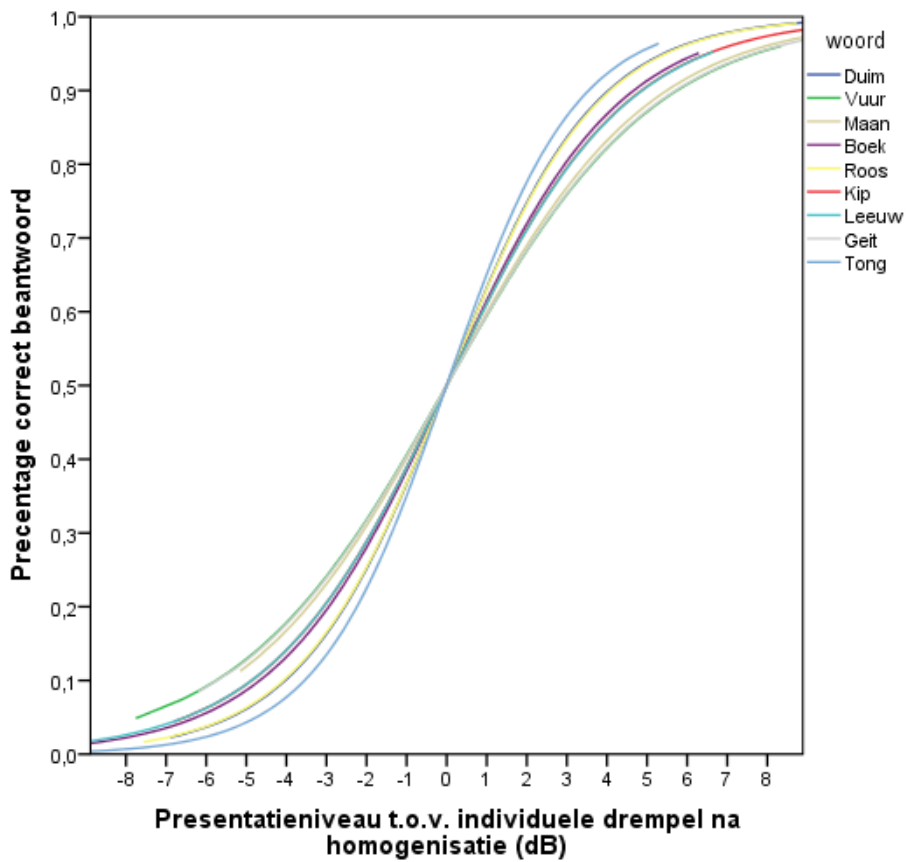
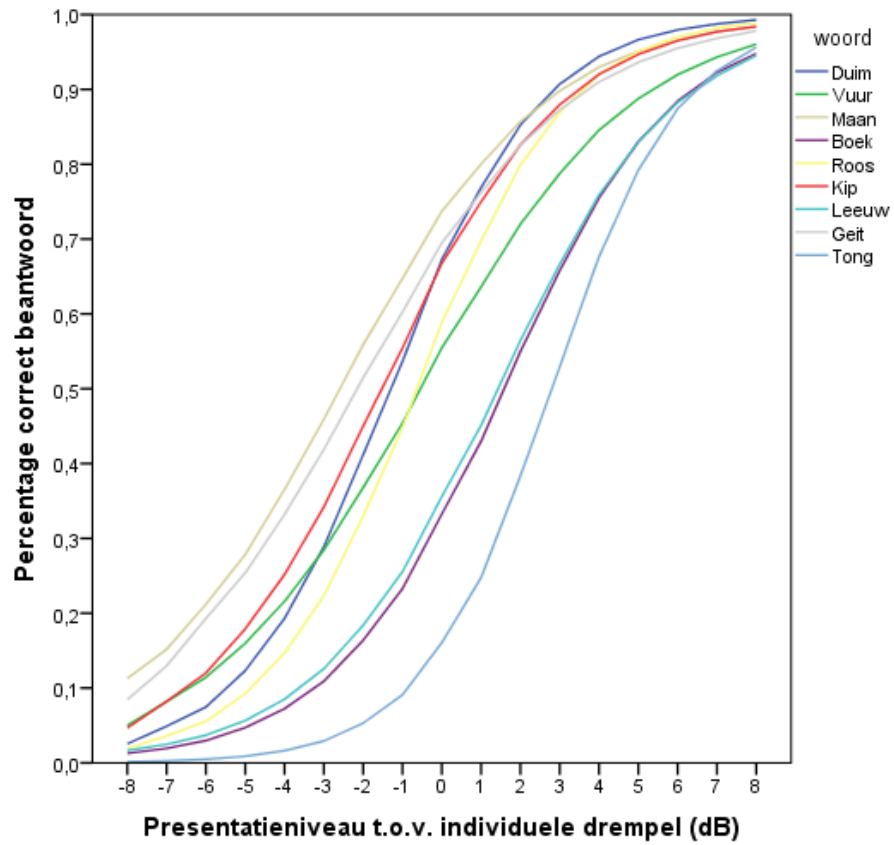
Tabel 7. Ligging en steilheid van de woord-specifieke discriminatiefuncties voor de breedbandige KHT

Woord	Shift (dB)	Slope (%/dB)	% correct bij 0 dB*
Duim	-1,30	13,6	67
Vuur	-0,54	9,6	55
Maan	-2,55	10,0	73
Boek	+1,51	11,8	33
Roos	-0,64	13,4	42
Kip	-1,51	11,3	67
Leeuw	+1,35	11,2	35
Geit	-2,11	9,6	69
Tong	+2,71	15,5	16

*presentatieniveau t.o.v. individuele drempel

Voor een nauwkeurige test is het van belang dat het spraakmateriaal zo homogeen mogelijk is. Dit betekent twee zaken:

- a) De shift van de verschillende woorden dient zo mogelijk te worden gecorrigeerd (**aandachtspunt 4**). Zie figuur 6 met alle discriminatiecurven door elkaar, voor en na correctie voor de shifts.
- b) De geselecteerde woorden dienen een zo hoog mogelijke steilheid te hebben, en bij voorkeur niet te veel verschil tussen de verschillende woorden (**aandachtspunt 5**).



Figuur 6. 'Woorden-waaier' met alle discriminatiecurven voor (boven) en na (onder) correctie voor de verschuivingen

Aanbevolen aanpassingen van de Kinderhoortest

Dit inventariserend onderzoek naar de originele Kinderhoortest heeft het doel om testuitslagen uit het verleden te evalueren en punten van verbetering op te sporen, om zodoende de kwaliteit van de test te verhogen.

In het voorgaande zijn vijf aandachtspunten gesignaleerd en gemarkeerd, die wij – gegeven de analyses van eerder uitgevoerde testen - van nadere overwegingen en aanbevelingen willen voorzien.

Aandachtspunt 1: de ruis

Overwegingen

De gefilterde ruis is bij de Oorcheck en de Bedrijfsoorcheck geïntroduceerd om de test gevoeliger te maken voor hoge-tonen verliezen ten gevolge van lawaai. Dit is niet de belangrijkste doelstelling in de groep van kinderen en daarom dient er een nieuwe afweging te worden gemaakt tussen de voordelen van de gefilterde ruis (nl. een verbeterde gevoeligheid voor schade door lawaai) en de voordelen van een breedbandige ruis (nl. een hogere meetnauwkeurigheid).

Aanbeveling

Er wordt aanbevolen om de test aan te bieden met een breedbandige ruis i.p.v. een gefilterde ruis.

Aandachtspunt 2: een leeftijdsafhankelijk criterium

Overwegingen

Uit bovenstaande analyses blijkt – evenals uit onderzoek op het VUmc – dat er een duidelijke leeftijdsafhankelijkheid is voor de drempels van spraak-in-ruis testen bij normaalhorende kinderen. Het effect is 2 – 3 dB en dat is dermate groot dat de interpretatie van de uitkomsten op groepsniveau wordt gehinderd als er geen rekening wordt gehouden met deze leeftijdsafhankelijkheid.

Ook is het onduidelijk in hoeverre de testen betrouwbaar kunnen worden uitgevoerd door kinderen van 3, 4, of 5 jaar oud.

Aanbeveling

De toepasbaarheid van de test bij jonge kinderen en de leeftijdsafhankelijkheid van de “normale” drempels voor spraakverstaan in ruis dienen in een populatie van normaalhorende kinderen van verschillende leeftijden te worden vastgesteld. Dit validatie onderzoek kan parallel worden uitgevoerd met de validatie van een Kinderoorcheck v2, die op basis van de aanbevelingen in dit rapport is bijgesteld.

Aandachtspunt 3: de aanloop en testduur

Overwegingen

In het onderzoek is aangetoond dat de aanloop niet te kort mag zijn en de absolute testduur niet te lang. Het blijkt mogelijk om de test als geheel in te korten, maar binnen de test dan

toch nog een langere aanloop (en dus een kortere meting) te kiezen. Het inkorten van de test als geheel is natuurlijk aantrekkelijk.

Het werkt prettiger en de test duurt niet te lang, zodat kinderen goed hun aandacht erbij kunnen houden. Tevens krijgen de kinderen zo langer de gelegenheid om de test door te krijgen voordat de daadwerkelijke meting start.

Aanbeveling

Op basis van onze bevindingen wordt voorgesteld de Kinderhoortest in te korten naar 20 woorden, met een aanloop van 10 stimuli.

Aandachtspunt 4: het homogeniseren van de aanbiedingsniveaus

Overwegingen

Er blijken aanzienlijke verschillen in moeilijkheidsgraad tussen de woorden te bestaan. De verschuiving van het 50%-punt in dB ten opzichte van 0 dB varieert van -2,55 dB ('maan') tot +2,71 dB ('tong'). De homogeniteit van het testmateriaal kan worden verbeterd door de onderlinge relaties in aanbiedingsniveaus van de verschillende woorden zo te wijzigen dat de woordspecifieke discriminatiecurven optimaal over elkaar liggen (het compenseren van de geconstateerde verschuivingen door versterkingen en verzwakkingen). Een moeilijker woord, zoals 'boek', zou met iets meer volume aangeboden moeten worden.

Als de testwoorden perceptief meer gelijk zijn, zal dit leiden tot een hogere steilheid van de discriminatiecurve op lijstniveau en tot een betere test-retest betrouwbaarheid van de test.

Aanbeveling

De onderlinge perceptieve verschillen in moeilijkheidsgraad dienen te worden geminimaliseerd door niveau-correcties toe te passen conform de "shift"-waarden, zoals beschreven in tabel 6. Overigens dient dan ook de toe te passen ruis opnieuw te worden samengesteld, zodat het spectrum overeen komt met het gemiddelde spectrum van de testwoorden (in hun gewijzigde samenstelling).

Aandachtspunt 5: het spraakmateriaal

Overwegingen

De verschillende versies van Kinderhoortest met diverse populaire sprekers (Bibi, Edward, Sita) dragen naar verwachting niet bij aan de betrouwbaarheid van het testresultaat. De "alternatieve" sprekers zijn geen geoefende sprekers en de testwoorden zijn ook niet goed gevalideerd en/of gehomogeniseerd.

Er bestaan forse verschillen in steilheid van de woordspecifieke PI-functies. Het woord 'tong' blijkt een buitenbeen te zijn met een helling van 15,5%/dB.

Aanbeveling

Er wordt aanbevolen om alleen de versie aan te bieden met de originele, geoefende spreker. Verder adviseren wij het woord 'tong' te verwijderen uit de testset. Het vraagteken zou dan in het midden kunnen worden gepositioneerd en het woord 'roos' rechtsonder.

Wij adviseren tenslotte te zoeken naar een alternatief symbool voor het vraagteken.

Bijlage 1. Discriminatiefuncties van de individuele woorden

