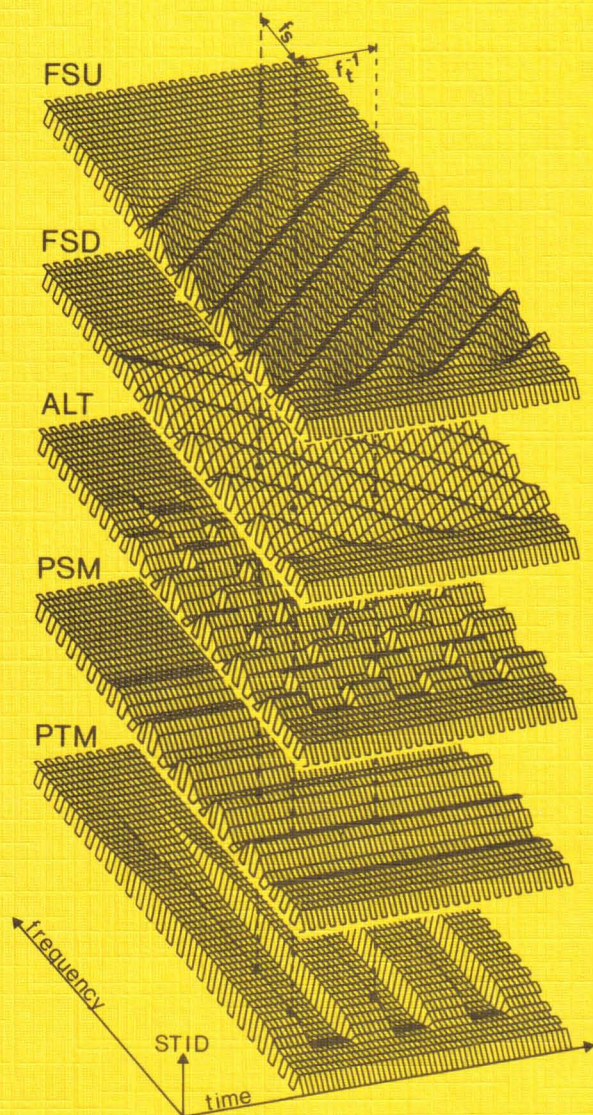


THE DETECTION OF TEMPORAL AND SPECTRAL MODULATIONS OF SOUND BY THE EAR



G.A. van Zanten

HOOFSTUK 7

INLEIDING, SAMENVATTING EN CONCLUSIES

INLEIDING

Horen is de omzetting van geluid, d.w.z. een trilling van de luchtdruk, in sensaties. De audiologie bestudeert dit omzettingsproces, zowel in normale toestand als bij stoornissen. Bij dit soort onderzoek wordt het geluid, de stimulus, beschreven in fysische termen en de informatie in termen van sensaties als luidheid, toonhoogte, klankkleur etc., of indirect in termen van gedragsveranderingEN teweeggebracht door het geluid. In dit proefschrift wordt een onderzoek beschreven, dat handelt over een detail van het horen, namelijk het waarnemen door norhaalhorende proefpersonen van kortdurende (éénmalige of herhaalde) veranderingen in een geluid. Bij de experimenten was de aard van de sensaties voor de proefpersonen onbelangrijk. Slechts van belang was of ze het verschil konden waarnemen (detecteren) tussen twee stimuli, de ene met en de andere zonder zo'n kortdurende verandering. De studie was gericht op (1)- de bepaling van het vermogen van het menselijk gehoor om kleine veranderingen in de sterkte (intensiteit) en de klankkleur (spectrum) van een ruis-signaal waar te nemen, en (2)- het vergelijken van dit vermogen met dat van een fysisch model dat optimaal functioneert in theoretische zin. Een indirect, maar zeker niet onbelangrijk, doel is het leveren van basiskennis voor het ontwikkelen van klinisch toepasbare testen van dit vermogen.

In de natuurkunde wordt een geluid in het algemeen beschreven met de geluidsdruk als functie van de tijd (de golfvorm) of met het spectrum van dat geluid. Het spectrum geeft weer welke trillingsfrequenties in het geluid voorkomen en in welke mate elke frequentie aan het geluid bijdraagt. Welke van deze twee beschrijvingswijzen in de natuurkunde wordt gekozen hangt af van welke de meest praktische is. Beide beschrijvingen zijn volledig en de één kan in de ander omgerekend worden met de Fourier-transformatie. Met betrekking tot het horen is noch de beschrijving op basis van alleen de golfvorm noch de beschrijving op basis van alleen het spectrum bevredigend. Het menselijk gehoor is in staat om gelijktijdig aspecten van zowel de golfvorm als van het spectrum waar te nemen. De eeuwenoude manier van

noteren van muziek op een notenbalk geeft aan dat dit gelijktijdig ervaren van de tijd- en de frequentie-dimensie al heel lang bekend is: de verticale positie van een noot op de notenbalk bepaalt de toonhoogte, welke sterk aan frequentie gerelateerd is, de horizontale positie bepaalt het moment van productie.

Om het verstaan van spraak te kunnen begrijpen bleek het nodig een woord niet als één geluid te zien, maar als een serie van sub-geluiden; een beschrijving met een serie van golfvormen of met een serie van spectra was nodig. De laatste beschrijving bleek het meeste inzicht te geven. De geluid-spectrograaf geeft geluid grafisch weer met een serie van spectra, waarbij, net als in het notenschrift, frequentie van beneden naar boven en tijd (=serienummer) van links naar rechts geschreven wordt (Koenig et al, 1946). De sterkte van iedere frequentiecomponent wordt met een grijstint weergegeven. Op deze manier wordt een gecombineerde "spectrale" en "temporele" beschrijving van het geluid gegeven, die beter aansluit bij wat we waarnemen dan de strict temporele en spectrale beschrijvingen. Gabor(1946) toonde aan dat (1)- de mate van detaillering (de resolutie) van het spectrum van zo'n stukje geluid steeds minder wordt, naarmate het stukje afneemt in tijdsduur, en (2)- de "spectro-temporele" beschrijvingswijze tot een fundamenteel begrip van de informatie-inhoud van geluid leidt. De hoeveelheid informatie is gequantiseerd en wordt beschreven met een geheel aantal logons; het logon is de fundamentele eenheid van geluidsinformatie. Het logon op zich wordt weergegeven met de relatie $B_s T_s = 1$, een relatie die verwant is aan de "onzekerheids-relatie" van Heisenberg in de quantumfysica. B_s geeft het (effectieve) frequentiebereik van een logon en T_s de (effectieve) tijdsduur ervan. Een logon kan dus in veel verschillende vormen worden aangeboden. Bijvoorbeeld, kortdurend, in de vorm van een tik, het frequentiebereik is dan groot, of, iets langer durend, in de vorm van een toonstootje, waarbij de bandbreedte kleiner is (want $T_s = 1/B_s$). De waarnemer (detector), welke de informatie uit het geluid moet halen, kan dat alleen optimaal als juist de frequenties die in het geluid voorkomen bemeten worden en ook juist gedurende de duur van het logon. Met andere woorden, de detector moet goed aangepast zijn aan het geluid. De theoretisch optimale detector wordt beschreven met de relatie $B_m T_m = 1$, waarin B_m en T_m nu respectievelijk de bandbreedte van het filter van de detector en de duur van de energie-meting

(integratie-tijd) door de detector zijn. Ook bij de ideale detector geldt dus: hoe kleiner de breedte van het spectrale integratiegebied (de bandbreedte), des te groter het temporele integratiegebied (snelle veranderingen in het geluid kunnen dan minder goed waargenomen worden) en andersom. Het produkt van bandbreedte en integratie-tijd blijft steeds gelijk aan 1.

Bij vele psycho-fysische luister-experimenten blijkt ons oor opgevat te kunnen worden als een verzameling van detectoren, die naast elkaar liggende spectrale integratiegebieden hebben ("kritische banden"). Deze detectoren, kunnen niet beter zijn dan de ideale detector en zijn waarschijnlijk minder perfect, hetgeen weergegeven kan worden met een product van bandbreedte en integratie-tijd groter dan 1. Uit experimenten, waarbij toonstootjes moesten worden waargenomen, concludeerde Gabor(1947) dat bij die detectietaak het product van filter-bandbreedte en integratietijd bij ons gehoor tussen de 1 en 2 ligt, dus bijna ideaal. Ook later is veel onderzoek gedaan aan het waarnemen van veranderingen in frequentie of intensiteit van zuivere tonen en ruis-signalen met een kleine bandbreedte. De gebruikte stimuli waren in een spectro-temporele representatie, zoals het spectrogram, eenvoudig en contrastrijk. Het spectrogram was bijna geheel blank met uitzondering van een horizontale donkere streep, welke de (bijna stabiele) zuivere toon weergeeft. De proefpersonen moesten veranderingen in de stimulus waarnemen, welke in het spectrogram overeenkomen met veranderingen van de positie, de omvang, de grijs tint of de richting van deze donkere streep in het frequentie-tijd vlak.

Ook in deze studie werd gezocht naar de grootte van de spectrale en temporele integratiegebieden welke een rol spelen bij de detectie van geluid door het gehoor. Daarbij werden (kunstmatige) ruis-signalen met een matige tot grote bandbreedte gebruikt als stimulus in psycho-fysische luister-experimenten. Deze signalen werden aangeboden aan proefpersonen, die getraind waren in het luisteren naar juist dit soort stimuli. Bij sommige stimuli bleek veel training nodig, bij andere niet. De detectie van veranderingen van de intensiteit, van veranderingen in het spectrum en van gelijktijdige veranderingen in beide aspecten zijn bestudeerd. In deze studie was het basisgeluid witte ruis. Het spectrogram ervan toont slechts één grijs tint over het gehele frequentie-tijd vlak. De proefpersonen moesten veranderingen in het geluid detecteren, welke in het spectrogram zichtbaar

zijn als afwijkingen van de standaard grijstint. De taak van de proefpersonen is op deze visuele manier weergeven, mede doordat de experimenten sterk geïnspireerd zijn door onderzoekingen naar het menselijk zien, bijvoorbeeld door Van Nes e.a. (1967) en door Koenderink en Van Doorn (1977) (waarbij proefpersonen de afwijkingen in de lichtsterkte van een grijs vlak moesten waarnemen).

Op de voorkant van dit proefschrift staan vijf van de taken van de proefpersonen gevisualiseerd op een quasi drie-dimensionale manier (frequentie naar achteren, tijd naar rechts en intensiteit naar boven). De vijf lagen in de figuur geven vijf verschillende stimuluscondities weer. In iedere conditie moest de proefpersoon het verschil tussen twee ruisgeluiden waarnemen, het ene was gemoduleerde ruis, gesymboliseerd door één van de vijf gestructureerde vlakken in de figuur, het andere witte ruis, gesymboliseerd door een plat vlak (niet in de figuur opgenomen). Bij de experimenten werd het minimale hoogteverschil (de drempel) in de structuur bepaald, welke nodig is voor detectie.

SAMENVATTING EN CONCLUSIES

In de hoofdstukken 2 en 3 wordt het juist-waarneembare verschil beschreven tussen "witte" ruis, dus zonder spectro-temporele structuur, en ruis met alléén een temporele structuur. In hoofdstuk 4 worden de resultaten verklaard met een theoretisch detector model.

In de experimenten beschreven in hoofdstuk 2 werd de temporele structuur verkregen door gedurende korte tijd, 0,1 tot 300 milli-seconde, de intensiteit van ruis te verhogen of verlagen. Als functie van de bandbreedte en centrale frequentie van de ruis en van de duur van de verandering werd de minimaal waarneembare verhoging en verlaging in intensiteit gemeten. Geconcludeerd kan worden dat het auditieve systeem de integratie-tijd kan aanpassen aan de tijdsduur van de stimulus zelf, maar niet tot minder dan 3 milli-seconde. Bij kortere veranderingen wordt verlaging van de intensiteit helemaal niet waargenomen en verhoging alleen als de energie-inhoud van de verhoging een bepaalde grenswaarde te boven gaat. In dit opzicht is het auditieve systeem lang niet optimaal.

In de experimenten beschreven in hoofdstuk 3 werd de temporele structuur verkregen door de intensiteit te moduleren, d.w.z. afwisselend van een

laag naar een hoog niveau te brengen (gesymboliseerd op de kft met de PTM-laag). Detectiedrempels voor de modulatiesterkte werden gemeten als functie van de bandbreedte en centrale frequentie van de ruis en van de modulatiefrequentie (het aantal hoog-laag perioden per seconde). Uit de resultaten kan men concluderen dat het gehoor het meest gevoelig is voor modulaties met een frequentie beneden 10 tot 20 Hz. Voor hogere modulatiefrequenties daalt de gevoeligheid langzaam naarmate de modulatiefrequentie stijgt. Verder blijkt dat de modulatiegevoeligheid van het gehoor niet zozeer van de bandbreedte van de ruis afhangt als wel van de centrale frequentie ervan. Bij de optimale detector is alleen de bandbreedte van belang en de centrale frequentie helemaal niet. Dit betekent dat het auditieve systeem ook in dit opzicht minder dan optimaal is.

In hoofdstuk 4 wordt een model gepresenteerd, waarmee bijna alle meetresultaten uit de hoofdstukken 2 en 3 kunnen worden verklaard. Dit model bemeet slechts een beperkt frequentiegebied van de stimulus. Het bandfilter dat het model gebruikt is gelijk aan dat van de hoogste "kritische band" die door het geluid gestimuleerd wordt. Volgens dit model gebruikt ons gehoor dus een beperkt frequentiegebied. Dit is op zich zelf een grote afwijking van de optimale detector, waarvan de bandbreedte zich aanpast aan het frequentiegebied van de stimulus. Maar, gegeven het beperkte frequentiegebied, werkt de auditieve detector volgens dit model bijna optimaal; het produkt van bandbreedte en integratie-tijd is 1 tot 2.

In de experimenten beschreven in de hoofdstukken 5 en 6 werd de waarneembaarheid bepaald van veranderingen in geluiden, welke grafisch zijn weergegeven in de bovenste drie lagen van de figuur op de kft. Het basisgeluid was nu spectraal-gemoduleerde ruis, d.w.z. ruis met cosinusvormige rimpels in het spectrum (de PSM-laag). Door de pieken in dit spectrum te laten bewegen ontstaat spectro-temporeel gemoduleerde ruis, zoals weergegeven in de bovenste drie lagen in de figuur op de kft. De pieken maakten een sprong in frequentie (de ALT-laag) of bewogen omlaag (de FSD-laag) of omhoog (de FSU-laag) langs de frequentie-as.

In hoofdstuk 5 wordt aangetoond dat het auditieve systeem even gevoelig is voor spectro-temporele modulaties van het ALT-, het FSD- en het FSU-type.

In hoofdstuk 6 wordt een experiment beschreven, waarin gestimuleerd

werd met ruisbanden, die een centrale frequentie van 0.5 kHz en van 2.0 kHz hadden. Bij verschillende bandbreedten van de ruis werd de gevoeligheid voor het ALT-type modulatie bepaald. De resultaten worden vergeleken met die van de optimale detector en ook met meetresultaten en modellen beschreven in de literatuur. De resultaten voor de ruisbanden met een centrale frequentie van 0.5 kHz kunnen kwalitatief verklaard worden met een model als beschreven in hoofdstuk 4. Maar bij de centrale frequentie van 2.0 kHz kan dat niet, daar lijkt een geheel ander mechanisme in de detectie (mede) werkzaam te zijn.