

Ing. H.J.M. Steeneken en Dr. -Ir. T. Houtgast

Instituut voor Zintuigfysiologie TNO, Soesterberg

A physical method for measuring the quality of speech transmission channels is presented. The method is based on the intensity fluctuations of running speech. The decrease of these fluctuations due to disturbances can be expressed by a single value which appears to be a good predictor for the intelligibility that can be obtained by using the system.

INLEIDING

De meest voor de hand liggende meetmethode voor het bepalen van de kwaliteit van spraaktransmissiekanalen is gebruik te maken van die signalen waarvoor de te testen verbinding is bedoeld, dus spraak of spraakklanken met een spreker aan de zenzijde en een luisteraar aan de ontvangzijde. Als spraakmateriaal zou men gebruik kunnen maken van berichten of zinnen. We meten dan de zinsverstaanbaarheid. Normale berichten of zinnen bezitten echter een slecht gedefinieerde redundantie, d.w.z. niet alle woorden van het bericht zijn nodig om het bericht te verstaan. Voor reproduceerbare, goed gedefinieerde meetresultaten maakt men bij voorkeur gebruik van logatomen. Dit zijn merendeels betekenisloze woorden bestaande uit een medeklinker-klinker-medeklinker combinatie. Men gebruikt lijsten van 50 verschillende woorden. Per lijst is er voor gezorgd dat de frequentie van voorkomen van de verschillende fonemen representatief is voor de taal. Deze methode is echter tijdrovend, te meer daar voor een betrouwbare meting verschillende sprekers en luisteraars nodig zijn. Voor één meetpunt moet men dan ook op ca. 2 uur meettijd rekenen.

Het ligt voor de hand dat er gezocht is naar methoden waaraan de bezwaren van een lange meettijd en veel personeel niet kleven, en dat liefst met behoud van nauwkeurigheid.

Om hieraan te kunnen voldoen zullen we in moeten gaan op de fysica van de spraak.

HET OMHULLENDE-SPEKTRUM

Lopende spraak is in feite een opeenvolging van signalen met een verschillend frekwentiespektrum. Voor een goede verstaanbaarheid via een transmissiekanal moet dit kanaal dus in staat zijn deze spektrale verschillen over te brengen. Hiertoe beschouwen we de omhullende-functie van lopende spraak (Houtgast en Steeneken, 1972). Een voorbeeld van een dergelijke functie is in de bovenste helft van Fig. 1 gegeven. De mate waarin de door deze functie gegeven fluktuaties worden behouden kan als essentieel voor de verstaanbaarheid worden beschouwd, daar deze fluktuaties direkt zijn gekoppeld aan de verschillen in spektrum tussen de opeenvolgende spraakklanken. Als referentie voor het bepalen van de relatieve sterkte van de fluktuaties wordt de gemiddelde intensiteit gebruikt. Deze gemiddelde intensiteit is in Fig. 1 weer gegeven door de gestippelde lijn.

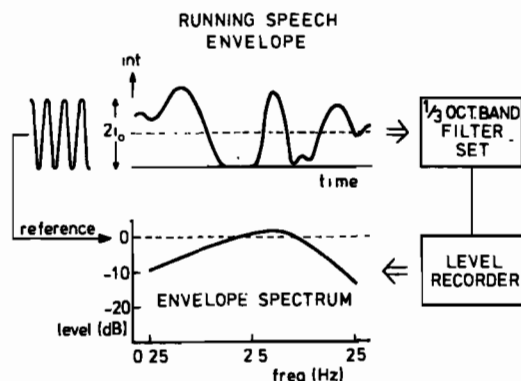


Fig. 1. Voorbeeld van een omhullende-functie en omhullende-spektrum voor lopende spraak.

De omhullende-functie wijzigt als de spraak verstoord wordt. Zo zal een achtergrondgeruis de gemiddelde intensiteit doen stijgen, begrenzing (peak-clipping) zal de pieken van de omhullende afkappen, nagalm zal een versmeringseffect hebben (met name voor snelle fluktuaties), terwijl een automatische versterkingsregeling de langzame fluktuaties nivelleert.

Als we de fluktuaties in omhullende-functie willen gebruiken voor meetdoeleinden zullen we er een eenduidige beschrijving van moeten hebben. De onderste helft van Fig. 1 geeft het spectrum van de omhullende-functie: het omhullende-spectrum (envelope spectrum). De gemiddelde intensiteit is weer als referentie gebruikt; een spectraal niveau van 0 dB betekent een modulatie diepte van de omhullende van 100%.

Daar de invloed van de meeste storingen frekwentieafhankelijk is dient dit omhullende-spectrum voor verschillende frekwentiebanden van het spraakspectrum te worden bepaald. Het blijkt dat een analyse van de oktaafbanden met middenfrekwenties van 250 Hz t/m 4000 Hz voldoende nauwkeurig is.

De volgende punten zijn van belang indien we dit omhullende-spectrum als basis voor het bepalen van de spraakverstaanbaarheid willen gebruiken:

- Hoe meten we het omhullende-spectrum?
- Is het omhullende-spectrum reproduceerbaar?
- Hoe wijzigt het omhullende-spectrum t.g.v. verschillende storingen op een transmissiekanaal?
- Kunnen we aan het omhullende-spectrum één eenduidige maat ontleen die goed gerelateerd is aan de subjectieve verstaanbaarheid?

Het eerste punt - hoe meten we het omhullende-spectrum - kan aan de hand van Fig. 2 worden verklaard. Lopende spraak, gefilterd via een oktaafilter, wordt gekwadeerd door een analoge vermenigvuldiger (om de intensiteit te verkrijgen) en daarna via een laagdoorlatend filter geleid. Het verkregen signaal wordt bemonsterd met een analoog-digitaalomzetter en opgeslagen in het geheugen van een rekenmachine. Na bemonstering van ca. 40 sec lopende spraak wordt het opgeslagen signaal via een digitaal-analoogomzetter geregenereerd,

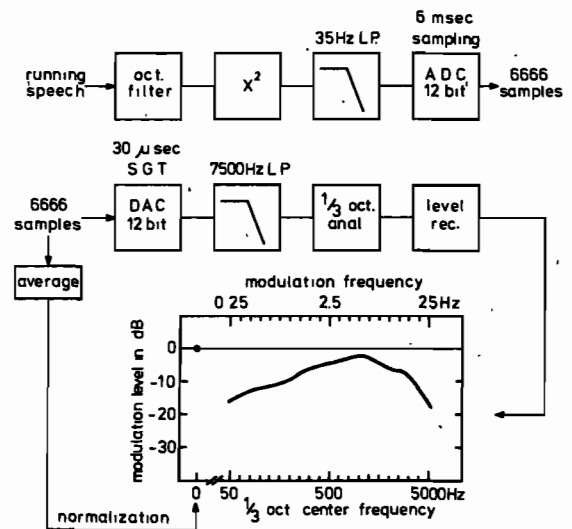


Fig. 2. Overzicht van de meetopstelling voor het meten van een omhullende-spectrum.

echter met een tweehonderdmaal hogere snelheid dan waarmede werd bemonsterd. Het omhullende-spectrum, met frekwenties van ca. 0,25 tot 25 Hz, omvat nu het gebied 50 tot 5000 Hz. Dit kan met standaard audioapparatuur worden geanalyseerd. De rekenmachine berekent uit de verkregen informatie eveneens de gemiddelde intensiteit.

Het tweede punt, de reproduceerbaarheid, kan worden onderzocht door een aantal omhullende-spektra te bepalen voor verschillende kondities zoals:

- verschillende stukjes spraak van eenzelfde spreker,
- verschillende sprekers,
- verschillende spreekwijzen (normaal, luid, fluisteren),
- verschillende oktaafbanden.

In Fig. 3 wordt ter illustratie de reproduceerbaarheid getoond van de omhullende-spektra van vijf verschillende stukjes spraak van 40 sec in de oktaafband 1000 Hz afkomstig van dezelfde spreker. De getrokken lijn geeft het referentieniveau. Ter illustratie van de sterkte van de (statistische) amplitude-fluktuaties van een stoorruis is met de stippellijn het omhullende-spectrum van een stationaire stoorruis gegeven, gemeten op geheel analoge wijze als bij de lopende spraak. In feite is het gebied tussen beide omhullende-spektra een maat voor het overbrengen van signalen met variërende spektra.

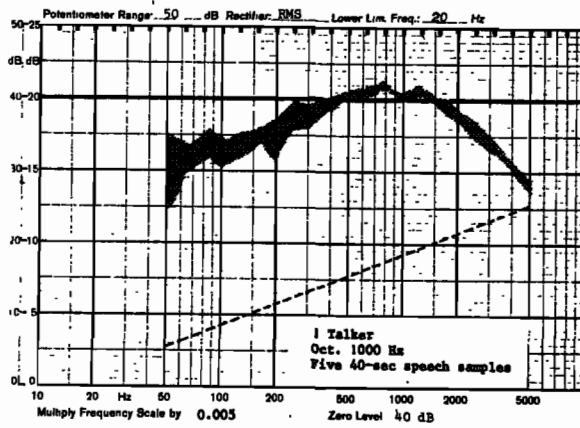


Fig. 3. Omhullende-spectrum voor vijf verschillende monsters van 40 sec lopende spraak voor één spreker in de oktaafband 1000 Hz.

Fig. 4 illustreert in hoeverre het toevoegen van ruis het omhullende-spectrum beïnvloedt. De linker kolom geeft het omhullende-spectrum voor drie verschillende signaal-ruisverhoudingen. De dunne lijn geeft het omhullende-spectrum voor de ongestoorde verbinding, de dik getrokken lijn voor de oktaafband 500 Hz, en de gestippelde lijn voor de oktaafband 2000 Hz. Omdat voor de stoornis witte ruis werd gebruikt, is de afname van

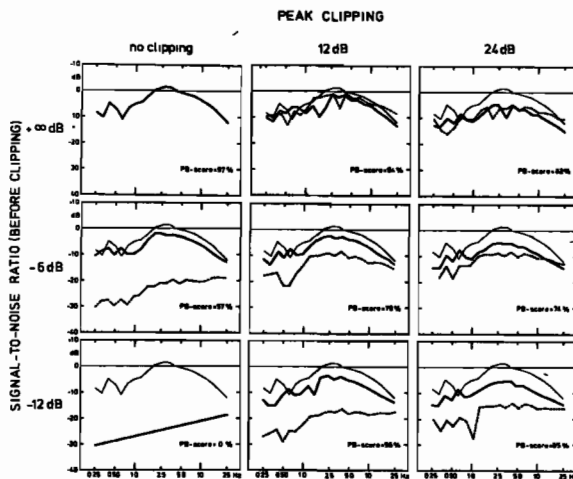


Fig. 4. Omhullende-spectrum voor drie verschillende kondities van amplitudebegrenzing in combinatie met drie signaalruisverhoudingen. De dunne lijn geeft het omhullende-spectrum voor ongestoorde spraak, terwijl de dikke lijn het bij de kondities behorende omhullende-spectrum geeft voor de oktaafband 500 Hz en de gestippelde lijn voor de oktaafband 2000 Hz.

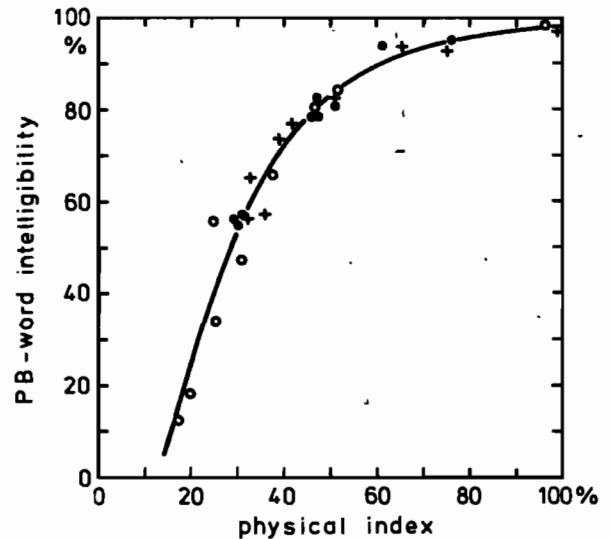


Fig. 5. Relatie tussen de fysische index en de logatoomverstaanbaarheid voor 27 verschillende verbindingen.

het omhullende-spectrum voor de oktaafband 2000 Hz groter dan voor 500 Hz. Bij de middelste en rechter kolom is eenzelfde signaal-ruisverhouding gekombineerd met amplitudebegrenzing van het spraaksignaal.

Teneinde deze benadering te evalueren als een betrouwbare voorspeller van de kwaliteit van een transmissiekanaal werd een groot aantal verschillende verbindingen onderzocht. Van deze verbindingen werd naast het omhullende-spectrum voor de vijf oktaafbanden ook de logatoomverstaanbaarheid bepaald.

De totale afname van de vijf omhullende-spectra per verbinding werd uitgedrukt in één getal. De manier van berekening van dit getal werd op zodanige wijze geoptimaliseerd dat de korrelatie met de subjectief bepaalde verstaanbaarheid maximaal was. Voor een aantal verbindingen is de relatie tussen de fysische en de perceptieve maat gegeven in Fig. 5. De nauwkeurigheid van deze fysische index wordt bepaald door de verticale spreiding van de punten. De standaarddeviatie voor de hier genoemde 27 meetpunten is 4,8% logatoomverstaanbaarheid.

De afname van het omhullende-spectrum van spraak t.g.v. storing op een transmissiekanaal kan beschreven worden in termen van de zgn. modulatie-overdracht karakteristiek. Een andere en meer directe methode voor het

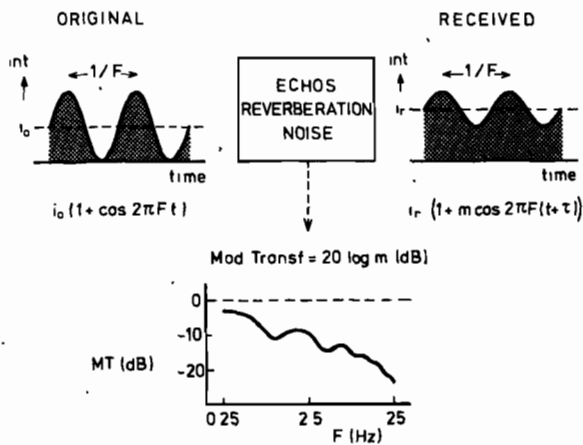


Fig. 6. Basisprincipe voor het bepalen van de modulatie-overdracht karakteristiek.

bepalen van de modulatie-overdracht karakteristiek is gegeven in Fig. 6 (Houtgast en Steeneken, 1973). Het test-signaal heeft een (variabele) modulatiefrequentie (100% sinusvormige intensiteitsmodulatie). Via het transmissiekanaal kan t.g.v. storing de modulatie diepte afnemen, hetgeen overeenkomt met de afname die het omhullende-spectrum van lopende spraak bij die frequentie vertoont. De praktisch realisatie van de benodigde meetopstelling waarmee de modulatie-overdracht karakteristiek kan worden bepaald is gegeven in Fig. 7. Aan de zenzijde wordt

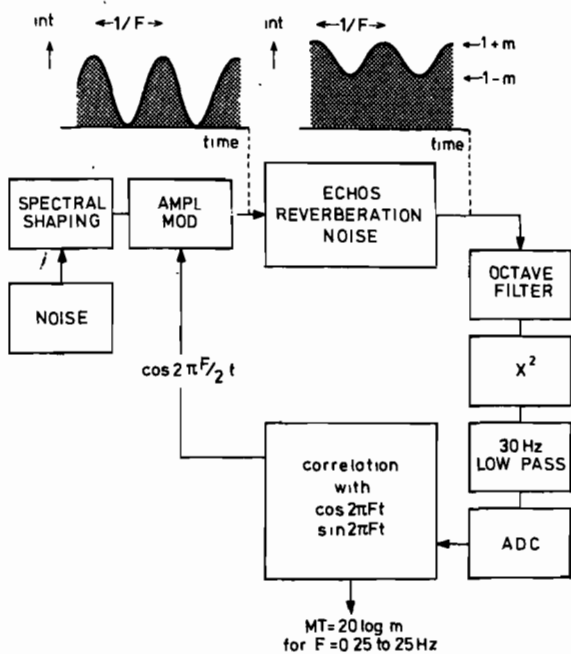


Fig. 7. Overzicht van de meetopstelling voor het meten van de modulatie-overdracht karakteristiek.

d.m.v. een amplitudemodulator het testsignaal gegenereerd, waarbij het modulatiesignaal wordt geleverd via een digitaal-analoogomzetter door een rekenmachine. Na transmissie, filtering en omhullende-detectie wordt het signaal via een analoog-digitaalomzetter naar dezelfde rekenmachine teruggevoerd. Door correlatie met de sinus- en cosinusfunctie van het modulatiesignaal en bepaling van de gemiddelde intensiteit is de modulatie diepte van het ontvangen signaal te bepalen en de modulatieoverdracht bekend. De meting kan worden uitgevoerd voor modulatiefrequenties van 0.25 tot 25 Hz.

In Fig. 8 is een schematisch overzicht gegeven van de wijze waarop thans in de praktijk een transmissiekanaal wordt bemeaten; Het bovenste gedeelte geeft de niveaus van het testsignaal voor de vijf oktaafbanden. Een dezer banden (in dit geval 2000 Hz) wordt qua intensiteit sinusvormig gemoduleerd en tegelijkertijd ge-

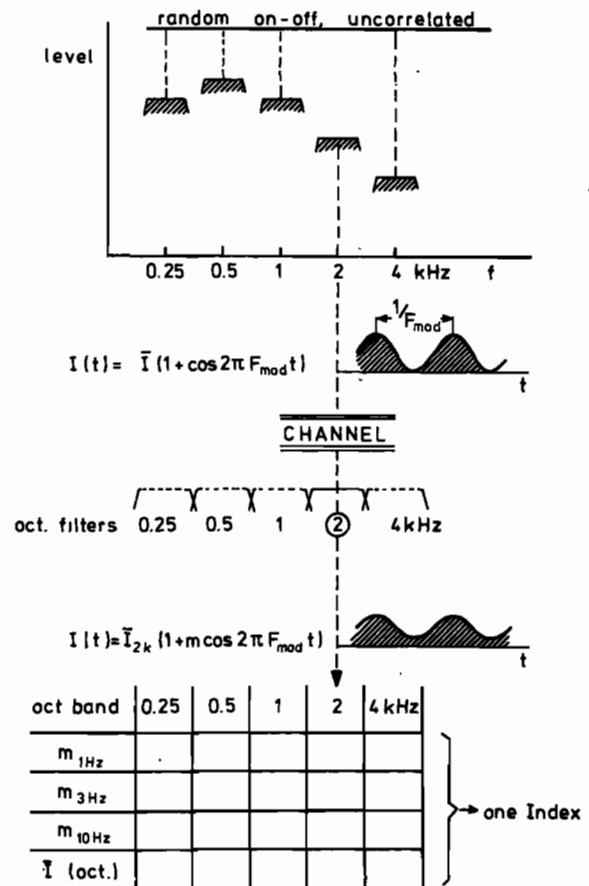


Fig. 8. Schematisch overzicht van de meetprocedure voor het bepalen van de modulatie-overdracht vereist voor het beoordelen van een transmissiekanaal.

analyseerd. Over het algemeen wordt de modulatieoverdracht slechts voor de modulatiefrequenties 1 Hz, 3 Hz en 10 Hz bepaald. Tevens wordt de gemiddelde intensiteit bepaald. Op het moment dat een der oktaven wordt bemeaten, worden in de overige oktaafbanden signalen gegeneerd met een willekeurige omhullende waarvan het gemiddelde omhullende-spektrum overeenkomt met dat van lopende spraak. Bij het onderzoek van lineaire transmissiekkanalen is de aanwezigheid van deze signalen in de niet-bemeaten oktaafbanden niet nodig. Bij kanalen waarbij amplitudebegrenzing of andere niet-lineaire bewerkingen worden toegepast die vervormingsprodukten in de te onderzoeken oktaafband veroorzaken, is echter de aanwezigheid van deze "spraakachtige" simulaties noodzakelijk. Het afwerken van deze meetprocedure voor een meetpunt duurt ca. 70 sec. In vergelijking met een verstaanbaarheidsmeting met sprekers en luisteraars (ca. 2 uur) is dit een aanzienlijke verbetering.

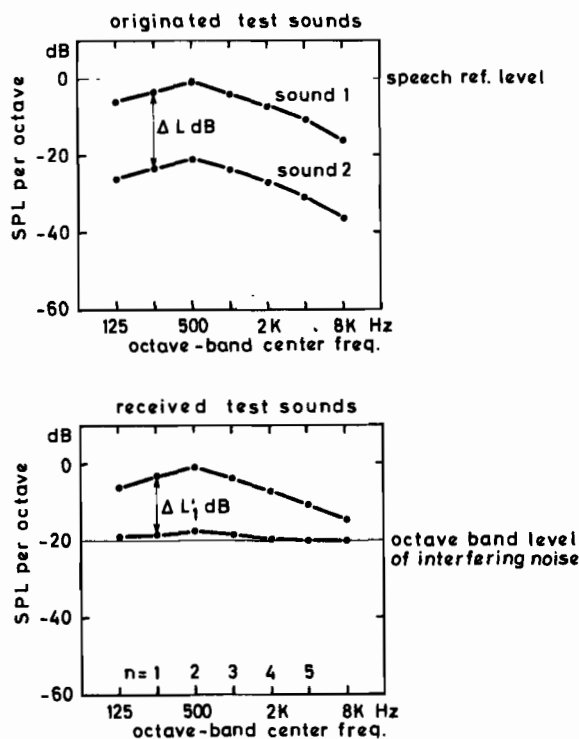


Fig. 9. De bovenste grafiek geeft een vereenvoudigd testsignaal voor het meten van de modulatieoverdracht bij één modulatie frequentie: "Sound 1" en "Sound 2" worden afgewisseld met een ritme van 3 Hz. Bij de onderste grafiek is op het testsignaal een stoorsignaal met een niveau van -20 dB werkzaam.

#### SAMENHANG MET EEN EERDER GEBRUIKTE METHODE

Reeds eerder is voor dit Genootschap een voordracht gehouden waarbij het meten van verstaanbaarheid langs fysische weg het onderwerp was (Houtgast en Steeneken, 1970). Het ging toen om het meten van radioverbindingen waarbij vooral ruis en bandbreedtebeperking als storingen optraden.

Het in dat geval gebruikte testsignaal is in Fig. 9 gegeven. Het is een afwisseling tussen signaal 1 en 2 met een ritme van 3 Hz. In feite werd de modulatieoverdracht alleen voor een modulatiefrequentie van 3 Hz gemeten. De vijf oktaafbanden werden tegelijkijd (in fase) gemoduleerd zodat geen synchronisatieproblemen kunnen ontstaan tussen zend- en ontvangzijde. In de on-

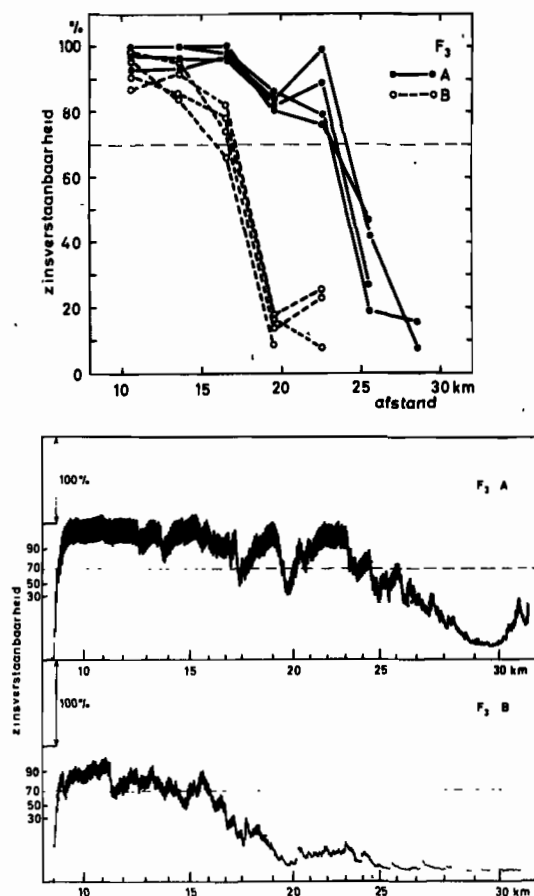


Fig. 10. Vergelijking tussen afstandbereikmetingen van een zendontvanginstallatie voor twee typen antennes. Het bovenste deel heeft betrekking op metingen met sprekers en luisteraars, terwijl het onderste deel de registraties van de fysische index weergeeft.

derste helft van deze figuur is geïllustreerd hoe een stoorruis het testsignaal beïnvloedt. Signaal 2 wordt gedeeltelijk gemaskeerd en de modulatiediepte neemt af.

Met dit systeem zijn veel onderzoeken (vooral t.b.v. de KL) uitgevoerd waarbij o.a. het afstandbereik werd gemeten van de huidige VHF/FM radioapparatuur en de EZB apparatuur. Als parameters werden o.a. gebruikt: antennetype, frekwentie, dag/nacht en terreinkondities. Fig. 10 is een illustratie van een afstandbereikmeting voor een zend-ontvanginstallatie met twee antenntypen (A en B), uitgevoerd met sprekers en luisteraars op vaste punten en met het fysisch meetsysteem. In de bovenste helft van de figuur zijn de resultaten per spreker gegeven. De onderste helft geeft de momentane registratie van het fysische meetsysteem.

#### SLOTOPMERKINGEN

Tenslotte nog enige opmerkingen over de huidige stand van zake van het eerder besproken meetsysteem met de modulatie-overdrachtfunctie. Kalibraties en metingen zijn uitgevoerd voor de volgende typen transmissiekanaalen.

<u>Type storingen en aantal</u>	<u>aantal combinaties</u>
Ruis (4) + Bandbreedte beperking (BP)(3)	49
Peakclipping (3), + BP(2) + Ruis (2)	54
AVC (4) + BP (2) + Ruis (1)	36
Delta Mod. (3) + Dig. Ruis	23
PCM (2x2) + Dig. Ruis	16
Telefoons (3) + Ruis	21
Mikrofoons (4) + Ruis	80

Uit bovenstaande kon worden gekonkludeerd dat dit meetsysteem

- snel is t.o.v. spreker/luisteraars
- reproduceerbaar is
- geoptimaliseerd is voor de meeste storingen
- weinig personeel behoeft.

Hoewel de relevantie van deze benadering voor een grote verscheidenheid van storingen is afgestemd, is het niet zo dat deze methode "blindelings" kan worden toegepast. Steeds zal men een open oog moeten hebben voor de mogelijkheid dat specifieke typen storingen niet op de juiste wijze door de huidige methode worden geëvalueerd.

#### REFERENTIES

- Houtgast, T. en Steeneken, H.J.M.: Beoordelen van spraakcommunicatiekanalen langs fysische weg. De Ingenieur jr. 82, (1970), pag. ET154-157.
- Houtgast, T. and Steeneken, H.J.M.: Evaluation of Speech Transmission channels by Using Artificial Signals, Acustica 25 (1971), 355.
- Houtgast, T. and Steeneken, H.J.M.: Envelope spectrum and Intelligibility of Speech in Enclosures. IEEE-AFCRL (1972), Speech Conference.
- Houtgast, T. and Steeneken, H.J.M.: The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility, Acustica, 28, (1973), 66.
- Steeneken, H.J.M. and Houtgast, T.: Intelligibility in Telecommunication derived from Physical Measurements, in Proc. F.A.S.E. (1973), Liège, 316.

Voordracht gehouden op 10 april 1975 in de KMA te Breda, op de 245ste werkvergadering van het NERG.