

NEDERLANDS AKOESTISCH GENOOTSCHAP

DE MODULATIE-OVERDRACHTSKARAKTERISTIEK ALS VOORSPELLER VAN
SPRAAKVERSTAANBAARHEID IN RUIMTEN

Ir. T. Houtgast en H.J.M Steeneken, ing.
Instituut voor Zintuigfysiologie TNO, Soesterberg

THE MODULATION TRANSFER FUNCTION IN ROOM ACOUSTICS AS A PREDICTOR OF
SPEECH INTELLIGIBILITY

Summary

Generally, when a signal is passed through an enclosure, the envelope of the received signal is a smoothed version of the original envelope. Traditionally, this effect of the enclosure is quantified by its reverberation curve. We suggest that the relation between the input and output envelopes can also be described by the action of a filter characteristic, effective upon the input envelope. In the case of smoothing, the filter is essentially a low-pass filter. This filter characteristic (Modulation Transfer Function, MTF) is a specific feature of a given enclosure, applicable to the envelope of any input signal. We investigated the relevance of this MTF for quantifying speech intelligibility for a great number of transmission paths. It appeared that, for each condition, a single value could be deduced from the MTF which correlates highly with the PB-word score.

Inleiding

Bij het beschouwen van stationaire klanken in besloten ruimten, zoals in het voorgaand verhaal, is het zinvol de transmissieweg tussen bron en waarnemer te beschouwen als een frekwentiekarakteristiek met een zeer onregelmatig amplitude- en faseverloop. Voor de transmissie van dynamische geluiden (spraak, muziek) impliceert een dergelijke scherp-gepiekte frekwentiekarakteristiek een zekere mate van traagheid en versmering in de responsie. Voor dergelijke geluiden is het dan ook zinvol de transmissieweg te beschouwen als een "omhullende-versmeerder", zoals b.v. wordt weergegeven door de nagalmkurve. Nagalm introduceert een zekere mate van "onscherpte in de tijd", waardoor de omhullende (intensiteit als functie van de tijd) van het signaal bij de waarnemer een vervaagde en vervlakte versie is van de omhullende van het oorspronkelijke signaal.

In dit verhaal wordt de "onscherpte" bij het reproduceren van de omhullende van het oorspronkelijke signaal, zoals veroorzaakt door

nagalm, enkele echo's en ook door stoorlawaai, gemeten op een wijze die grote overeenkomst vertoont met het bepalen van "onscherpte" bij de afbeelding door optische systemen. Nagegaan is in hoeverre, voor een 60-tal laboratoriumcondities, deze fysich gemeten "onscherpte" gecorreleerd is met subjectief bepaalde verstaanbaarheid.

Hier wordt volstaan met een beknopte uiteenzetting van deze aanpak. Voor meer uitvoerige informatie wordt verwezen naar (1) en (2).

Modulatie Transmissie Functie

De welbekende nagalmkurve van een transmissieweg tussen bron en waarnemer geeft direkt aan hoe een korte impuls van geluidenergie door de transmissieweg wordt vervaagd. Dit beschrijft volledig de "onscherpte" die door de transmissieweg wordt geïntroduceerd. [Mathematisch geldt dat, voor elk willekeurig signaal, de omhullende (intensiteit als functie van de tijd) bij de waarnemer gelijk is aan de convolutie van de oorspronkelijke omhullende bij de bron en de knalresponsie van de transmissieweg.] Het blijkt echter in vele gevallen moeilijk om uit de nagalmkurve van een transmissieweg één getal af te leiden met een zekere voorspelkracht t.a.v. de spraakverstaanbaarheid, vooral wanneer de nagalmkurve geen ideaal, exponentieel, verloop vertoont. Daarom is gezocht naar een andere wijze voor het kwantificeren van de mate van "omhullende vervlaking" bij transmissie van dynamisch signalen in een ruimte, die misschien beter aansluit bij de invloed van de ruimte op de spraakverstaanbaarheid.

Een alternatieve wijze voor het beschrijven van het verband tussen het intensiteitverloop bij de bron en dat bij de waarnemer is door middel van een filterkarakteristiek: de omhullende bij de waarnemer is een gefilterde versie van de omhullende bij de bron. Daar juist de snelle intensiteitfluctuaties bij transmissie van bron naar waarnemer worden vervlakt, heeft het "omhullendefilter" in het algemeen een laag-doorlaat karakter. Het is gemakkelijk in te zien dat het "omhullende-filter", evenals de nagalmkurve, een vast gegeven is voor een bepaalde transmissieweg in een ruimte. Voor elk willekeurig signaal is de omhullende (intensiteit als functie van de tijd) bij de waarnemer bepaald door op de oorspronkelijke omhullende het omhullende-filter toe te passen.

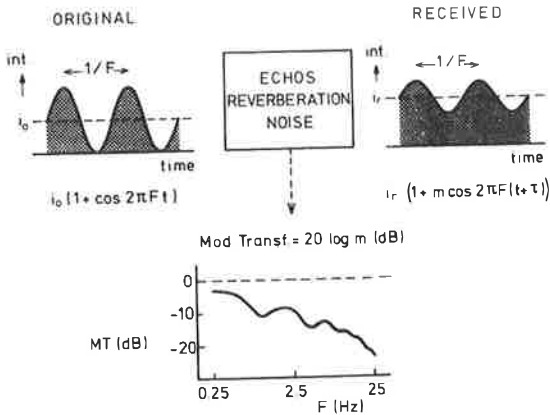


Fig. 1. Het effect van een transmissieweg (echo's, nagalm, stoorlawaai) op de modulatie diepte wordt weergegeven als een verzwakking in dB. Deze verzwakking als functie van de modulatie frequentie is de Modulatie Transmissie Functie MTF.

De karakteristiek van het omhullende-filter voor een bepaalde conditie kan op verschillende wijzen worden bepaald. Het kan, mathematisch, worden bepaald uit de nagalmcurve, mits deze voldoende nauwkeurig bekend is. Het kan ook worden bepaald door bij de bron een onregelmatig fluktuuerend signaal te introduceren (b.v. lopende spraak) en het spectrum van de omhullende bij de bron te vergelijken met het spectrum van de omhullende bij de waarnemer (zie hiervoor (2)). Het filter kan echter ook rechtstreeks bepaald worden door toepassing van een testsignaal bestaande uit ruis met een sinus-vormig gemoduleerde intensiteit, zoals weergegeven in Fig. 1. Het is gemakkelijk in te zien dat onder invloed van nagalm, enkele echo's en ook van stoorlawaai het sinus-vormig karakter van de omhullende behouden blijft; alleen de fase en amplitude van de modulerende sinus bij de waarnemer kunnen afwijken van die bij de bron. We beschouwen alleen het effect van de transmissieweg op de amplitude van de modulerende sinus. De verzwakking van de oorspronkelijke modulatie diepte één tot de resulterende waarde m , in termen van $i_{gem} (1 + m \cos 2\pi Ft)$, wordt uitgedrukt in de modulatie transmissie MT volgens $MT(\text{dB}) = 20 \log m$. Deze verzwakking als functie van de frequentie F van de modulerende sinus geeft de Modulatie Transmissie Functie, MTF. De MTF is dus de amplitudekarakteristiek van het bovengenoemde omhullende-filter.

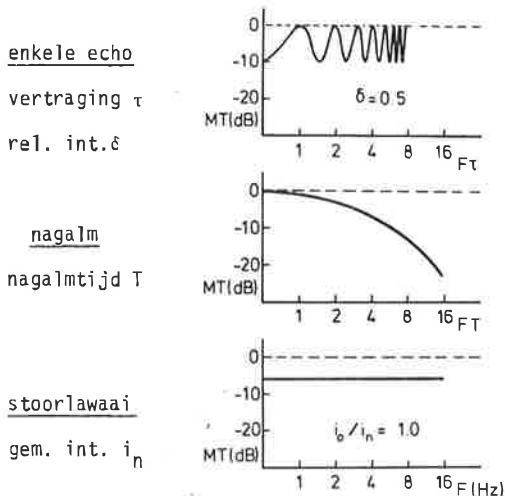


Fig. 2. Enkele typen storingen met bijbehorende Modulatie Transmissie Functies. (F is de modulatie-frekwentie en i_0 de gemiddelde intensiteit van het testsignaal, zie Fig. 1.)

Fig. 2 is voor enkele typen storingen de bijbehorende MTF gegeven. Voor een enkele echo, met vertraging τ en relatieve intensiteit δ , is de resulterende modulatie diepte m , en daarmee MT, bepaald door δ en het produkt $F\tau$. Voor ideale, exponentiële, nagalm met nagalmartijd T is het produkt FT bepalend. Voor stoorlawaai is bepalend de verhouding van de gemiddelde intensiteit van het testsignaal en de gemiddelde intensiteit van het stoorlawaai. Deze verschillende typen storingen laten zich dus alle vangen binnen het kader van de MTF: een filterkarakteristiek die toegepast kan worden op de omhullende van elk oorspronkelijk signaal teneinde de omhullende bij de waarnemer te verkrijgen. Alleen voor het geval dat stoorlawaai mede bepalende is geldt de beperking dat de MTF slechts relevant is voor die signalen waarvan de gemiddelde intensiteit gelijk is aan de gemiddelde intensiteit van het testsignaal waarmee de MTF bepaald is.

MTF en spraakverstaanbaarheid

Nagegaan is in hoeverre het kwantificeren door de MTF van de "versmerende" invloed van een transmissieweg aansluit bij het effect van de transmissieweg op de spraakverstaanbaarheid. Daartoe werd van een zestigtal transmissiewegen, onderworpen aan nagalm, enkele echo's, stoor-

lawaai of combinaties daarvan, zowel de MTF als de spraakverstaanbaarheid bepaald. Slechts enkele aspecten van de metingen worden hier onder de aandacht gebracht.

Nagalm werd geïntroduceerd door het signaal (spraak of testsignaal) door een nagalmkamer te voeren d.m.v. luidspreker en mikrofoon. De nagalmtijd T varieerde van 0,5 tot 2,4 sec. De enkele echo had steeds een relatieve intensiteit van 0,5, met een vertraging variërend van 50 tot 200 msec. Als stoornis werd thermische ruis gebruikt. In elke conditie ondergingen het testsignaal en het spraaksignaal precies dezelfde storingen.

In het algemeen zullen storingen van het hier beschouwende type min of meer frekwentieafhankelijk zijn. Dit zou betekenen dat er niet sprake is van één MTF, geldende voor het gehele frekwentiegebied, doch dat een aantal MTF's bepaald zouden moeten worden, b.v. voor een aantal opéenvolgende oktaafbanden. Voor dit oriëntierende onderzoek is deze complicerende faktor buiten beschouwing gelaten: de nagalm, de sterkte van de enkele echo's en de signaal/ruis verhoudingen zijn zodanig gekozen dat deze frekwentieonafhankelijk waren (tussen ca. 200 en 6000 Hz). Voor dat frekwentiegebied geldt dan dus één MTF.

Opdat de gemeten MTF relevant is voor de gebruikte spraaksignalen werd de gemiddelde intensiteit van het testsignaal (ruis met een sinus-vormig gemoduleerde intensiteit) aangepast aan de gemiddelde intensiteit van de spraaksignalen gebruikt bij de verstaanbaarheidsmetingen. Zoals vermeld is dat alleen van belang voor die gevallen waarin stoornis mede bepalend is voor de kwaliteit van de transmissieweg.

De MTF werd bepaald voor modulatiefrekventies van het testsignaal van 0,25 tot 25 Hz, in stappen van 1/3-oktaaf.

Bij de verstaanbaarheidsmetingen werd gebruik gemaakt van logatomen (éénlettergreppige woorden), steeds voorafgegaan en gevolgd door een neutraal woord. Bij de metingen waren vier sprekers en vier luisteraars betrokken.

Het (gestoorde) spraaksignaal werd de luisteraar aangeboden via beide elementen van een hoofdtelefoon. Bij dit onderzoek zijn mogelijke binaurale aspecten, die van belang kunnen zijn bij een luisteraar die zelf in een ruimte geplaatst is, geheel buiten beschouwing gelaten.

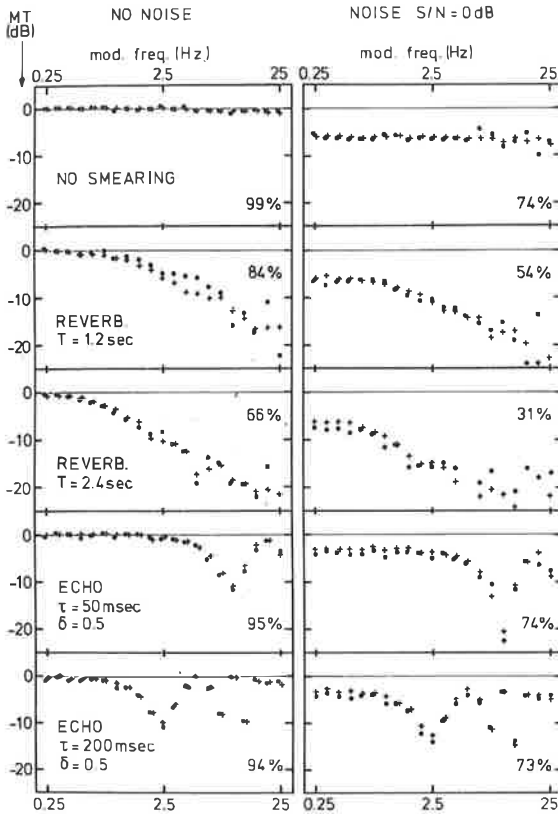


Fig. 3. Voor tien condities zijn de resultaten van modulatie transmissie metingen en de percentages logatoomverstaanbaarheid gegeven. (De MTF werd bepaald voor testsignalen binnen de oktaafbanden rond 500 Hz en rond 2000 Hz). In totaal werden 68 condities onderzocht.

In Fig. 3 zijn, voor tien condities, de logatoomverstaanbaarheid en de bijbehorende MTF gegeven. Het totaal aantal condities was 68. Nagegaan is of uit de MTF voor elke conditie één getal kon worden afgeleid dat goed aansluit bij de verstaanbaarheid. Dat getal is genoemd de Gewogen Modulatie Transmissie Functie (WMTF). Op de problematiek van het optimaliseren bij de herleiding van 68 MTF's tot 68 getallen (WMTF's) die zo goed mogelijk correleren met de 68 verstaanbaarheidspercentages wordt hier niet ingegaan.

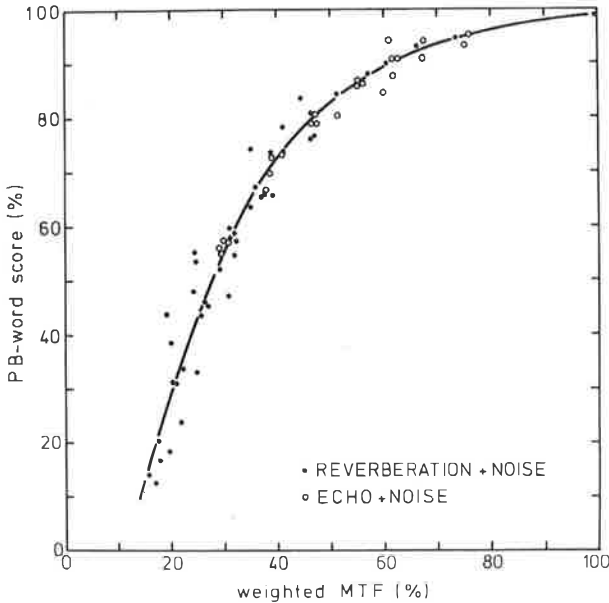


Fig. 4. Voor elk van de 68 onderzochte condities werd uit de MTF één getal afgeleid, de gewogen MTF. Deze figuur geeft het verband tussen de gewogen MTF en de logatoomverstaanbaarheid.

Het resultaat is weergegeven in Fig. 4, gebaseerd op de volgende herleiding:

$$WMTF = \frac{1}{21} \sum_{n=1}^{21} (1 - 0,3MT(n))^{-1} \times 100\%$$

waarin de index n refereert naar de 21 discrete waarden voor de modulatiefrequentie van het testsignaal (van 0,25 tot 25 Hz).

Uit Fig. 4 blijkt dat de WMTF een voorspelling van de logatoomverstaanbaarheid toelaat met een nauwkeurigheid van ca. 4%. Daarmee is de relevantie van de MTF voor spraakverstaanbaarheid in ruimten, althans binnen de beperkingen van dit oriënterend onderzoek, aangetoond.

Literatuur

- (1) Houtgast, T. and Steeneken, H.J.M., The modulation transfer function in room acoustics as a predictor of speech intelligibility, *Acustica* 28, (1973), 66-73.
- (2) Houtgast, T. and Steeneken, H.J.M., Envelope spectrum and intelligibility of speech in enclosures, *IEEE-AFCRL 1972 Speech Conference*, paper K2.