

## SAMENVATTING

De auditieve verwerking van spraak is een complex proces waarin onder meer het frequentie-oplossend vermogen van het oor een rol speelt. Dit suggereert een relatie tussen de auditieve frequentieresolutie en spraakverstaan. Omdat spraak in stilte zeer redundant is, wordt het grote nut van de auditieve frequentieresolutie gezocht in ons vermogen onder moeilijke luisteromstandigheden nog spraak te verstaan. Hoewel in de afgelopen twintig jaar het frequentie-oplossend vermogen van het auditieve systeem uitvoerig is onderzocht, en er vaak gewezen is op de relatie met het verstaan van spraak, is de vraag in hoeverre de verstaanbaarheid van spraak ook daadwerkelijk afhangt van het frequentie-oplossend vermogen nog vrijwel niet onderzocht. De moeilijkheid is dat de frequentieselectiviteit van het oor niet rechtstreeks kan worden gevarieerd.

Het frequentie-oplossend vermogen van het perifere auditieve systeem wordt meestal voorgesteld als een serie overlappende filters. Het signaal aan de uitgang van zo'n auditief filter met een vaste centrale frequentie wordt bepaald door de vorm van het filter toegepast op het energiespectrum van het binnenkomende signaal. Verbreding van de auditieve filters leidt dus tot een verminderd frequentie-oplossend vermogen. Dit zou op grofweg twee manieren het verstaan van spraak kunnen beïnvloeden. Ten eerste geven bredere auditieve filters de spectrale eigenschappen van spraak slechter door, waardoor het moeilijker wordt spraakklanken op basis van spectrale eigenschappen te onderscheiden. Wat betreft deze spectrale eigenschappen is het vervagen van de spectrale omhullende waarschijnlijk de belangrijkste factor. Ten tweede zullen bredere auditieve filters in een situatie waarbij spraak moet worden verstaan tegen een achtergrond van stoorlawaai meer lawaai doorlaten, waardoor het verstaan moeilijker wordt. Een verminderd frequentie-oplossend vermogen is dan ook regelmatig in verband gebracht met gehoorverlies voor spraakverstaan in ruis, zoals kan optreden bij slechthorendheid.

Een methode om zonder de frequentieresolutie van het oor te veranderen toch iets te weten te komen over het belang van deze eigenschap voor het verstaan van spraak, is de spectrale informatie in het *spraaksignaal* dat aan het oor wordt aangeboden te degraderen en het effect van deze degradatie op de verstaanbaarheid te meten. In deze aanpak worden dus uit het oogpunt van auditieve frequentieresolutie die aspecten van het spraaksignaal gevarieerd die belangrijk geacht worden voor het verstaan van spraak.

Het doel van dit onderzoek is het belang van spectrale contrasten voor de verstaanbaarheid van spraak te bepalen. Afgezien van het wetenschappelijke belang een beter inzicht te krijgen in de spectrale eigenschappen van spraak en de rol van auditieve frequentieresolutie in spraakverstaan, kunnen de gegevens ook gebruikt worden voor meer praktische toepassingen, zoals de codering en transmissie van spraaksignalen, en de signaalbewerking ter verbetering van de verstaanbaarheid van spraak voor slechthorenden.

De in deze studie toegepaste signaalbewerking om spectrale contrasten te verminderen is gebaseerd op een analyse- en resynthesesysteem dat bestaat uit Fourier transformaties op korte segmenten van het spraaksignaal en na bewerking een overlappende optelling van die segmenten om een continu signaal te reconstrueren. De spectrale contrasten worden verminderd door een convolutie van de omhullende van het energiespectrum met een Gaussisch filter op een logaritmische frequentie-as. Op deze manier kunnen de spectrale contrasten geleidelijk worden verminderd, terwijl de harmonische structuur en de daarbij behorende faserelaties in het spraaksignaal bewaard blijven.

Als maat voor de verstaanbaarheid van spraak is de drempel voor het verstaan van eenvoudige zinnen tegen een achtergrond van stoorspraak, de "speech-reception threshold" of SRT, genomen. De SRT is gedefinieerd als die signaal-ruisverhouding in decibels waarbij een luisteraar 50% van de zinnen zonder fouten kan reproduceren. Spectrale contrasten werden zowel in de zinnen als in het stoorsignaal verminderd.

Na een algemene inleiding in hoofdstuk 1, wordt in de hoofdstukken 2 en 3 een serie luisterproeven met in totaal veertig normaalhorende luisteraars beschreven. In deze experimenten wordt de mate van spectrale versmering systematisch gevarieerd. Het primaire doel van deze serie experimenten is het effect van verminderde spectrale contrasten op de verstaanbaarheid van zinnen onder moeilijke luisteromstandigheden te bepalen. Hiertoe wordt de SRT in ruis gemeten als functie van de mate van spectrale versmering, zowel voor zinnen uitgesproken door een vrouwenstem als voor zinnen uitgesproken door een mannenstem. De bandbreedte van de versmering varieert van 1/8 tot 4 octaven. De resultaten laten zien dat versmering van de spectrale omhullende over een bandbreedte tot 1/3 octaaf geen invloed heeft op het spraakverstaan. Echter, voor versmering over meer dan 1/3 octaaf neemt de zinverstaanbaarheid progressief af, dat wil zeggen, de SRT in ruis neemt toe ten opzichte van de SRT voor onversmeerde spraak. Verder laten de resultaten zien dat

de verstaanbaarheid van mannen- en vrouwenspraak niet verschillend te lijden heeft onder versmering van de spectrale omhullende, ondanks de verschillen tussen mannen- en vrouwenspraak in spectrale contrasten in het spraaksignaal, inclusief de positie, amplitude, en bandbreedte van de formantpieken en de diepte van de dalen daartussen. Voor beide stemmen neemt de SRT in ruis op dezelfde wijze toe voor versmering over meer dan 1/3 octaaf.

Een tweede doel van deze experimenten is te bepalen hoe goed individuele klinkers en medeklinkers bestand zijn tegen een vermindering van spectraal contrast. Zoals valt te verwachten, blijkt de identificatie van losse klinkers veel gevoeliger te zijn voor ernstige versmering van de spectrale omhullende dan de identificatie van begin- en eindmedeklinkers in betekenisloze woordjes van het type medeklinker-klinker-medeklinker. Uit de verwarringsmatrix verkregen bij sterke versmering van de spectrale omhullende (versmering over 2 octaven) blijkt dat een groot aantal klinkers verward wordt met de klinkers /ɔ/ en /u/ (als in de woorden "bot" en "voet"). Over het algemeen worden de klinkers verward met een achterklinker waarvan de eerste formant vergelijkbaar is met die van de te identificeren klinker. De resultaten suggeren dat een correcte identificatie van klinkers gewaarborgd is zolang de modulaties die de eerste en tweede formant aangeven in de spectrale omhullende behouden blijven. Medeklinkers worden vooral verward wat betreft de plaats van articulatie. Dus, gesteld in productietermen, plofklanken, wrijfklanken, nasalen, en halfklinkers worden voornamelijk verward met medeklinkers binnen dezelfde klasse. Verder hangt de mate waarin een medeklinker gevoelig is voor spectrale versmering zowel af van de mate waarin die medeklinker spectraal bepaald is als van de manier waarop de energie van die medeklinker verspreid is over het spectrum.

Tenslotte wordt de bijdrage van spectrale contrasten aan het nogal grote verschil in maskering tussen spraak en ruis als stoorder onderzocht (spraak wordt makkelijker verstaan tegen een achtergrond van stoorspraak dan stoorspraak). De resultaten laten zien dat dit verschil afneemt als functie van de mate van versmering van de spectrale omhullende. Dit wijst er op dat spectrale contrasten bijdragen aan het verschil in maskering voornamelijk omdat de spectrale dalen in de stoorspraak de te volgen spraak helpen detecteren door hun gunstige signaal-ruisverhouding.

Een alternatieve manier om het effect van versmering van de spectrale omhullende uit te drukken is in termen van de dichtheid van de modulaties (in perioden per octaaf) in de spectrale omhullende van het signaal voor en na bewer-

king. Hierin uitgedrukt, geven de resultaten uit de hoofdstukken 2 en 3 aan dat spectrale modulaties tot ongeveer 1,5 periode/octaaf al voldoende zijn voor de verstaanbaarheid van spraak onder moeilijke luisteromstandigheden (inclusief identificatie van klinkers en medeklinkers). Een voorbeeld kan het belang van een hiermee corresponderende auditieve frequentieresolutie duidelijk maken. Het verstaan van spraak in aanwezigheid van een storende spreker, met een signaal-stoorverhouding van 0 dB, is een alledaagse situatie, waarmee een goedhorende luisteraar geen enkel probleem heeft. Echter, als de mate van auditieve frequentieresolutie van deze luisteraar zou overeenkomen met een filter van 2 octaven, dan zou deze luisteraar in de gegeven situatie onmogelijk de bedoelde spraak kunnen verstaan (vgl. hoofdstuk 3, figuren 3 en 4).

Een interessante vraag is of een verminderd vermogen om spectrale contrasten op te lossen het gehoorverlies voor spraakverstaan in ruis zoals dat voorkomt bij slechthorendheid zou kunnen verklaren. Hoofdstuk 4 onderzoekt deze vraag bij vijftien luisteraars met een beperkt, vlak perceptief gehoorverlies en acht normaalhorende luisteraars. Als maat voor het auditief frequentie-oplossend vermogen, wordt de vorm en de breedte van het auditieve filter gemeten met behulp van een "notched-noise" maskeerparadigma bij de middenfrequenties 0,8, 1,6, and 3,2 kHz. De spectrale contrasten die nodig zijn voor het spraakverstaan worden bepaald door de spectrale contrasten in spraak te verminderen en die versmering te bepalen waarbij de verstaanbaarheid begint af te nemen. Hoewel de gemiddelde bandbreedte van het auditieve filter bij de slechthorende luisteraars varieert van normaal tot 2,5 keer breder dan normaal, is de relatie tussen de resolutie van spectrale contrasten nodig voor spraakverstaan en het gehoorverlies voor spraakverstaan in ruis zwak.

Samenvattend, tonen de resultaten van dit proefschrift aan dat er een duidelijke relatie is tussen spectrale contrasten en de verstaanbaarheid van spraak onder moeilijke luisteromstandigheden. Wat betreft gehoorverlies voor spraakverstaan in ruis lijkt een verminderd vermogen om spectrale contrasten op te lossen een factor van ondergeschikte betekenis te zijn.

## REFERENCES

- Allen, J.B., and Rabiner, L.R. (1977). "A unified approach to short-time Fourier analysis and synthesis," *IEEE* **65** (11), 1558-1564.
- Beddor, P.S., and Hawkins, S. (1990). "The influence of spectral prominence on perceived vowel quality," *J. Acoust. Soc. Am.* **87**, 2684-2704.
- Bonding, P. (1979). "Frequency selectivity and speech discrimination in sensorineural hearing loss," *Scand. Audiol.* **8**, 205-215.
- Bosman, A.J. (1989). "Speech perception by the hearing impaired," Ph.D. thesis, Utrecht, The Netherlands.
- Carhart R., and Tillman T.W. (1970). "Interaction of competing speech signals with hearing losses," *Arch. Otolaryng.* **91**, 273-279.
- Celmer, R.D., and Bienvenue, G.R. (1987). "Critical bands in the perception of speech signals by normal and sensorineural hearing loss listeners," in *The psychophysics of speech perception* edited by M.E.H. Schouten (NATO ASI Series, 1987), Series D, Behavioural and Social Sciences 39, pp. 473-480.
- Childers, D.G., and Wu, K. (1991). "Gender recognition from speech. Part II: Fine analysis," *J. Acoust. Soc. Am.* **90**, 1841-1856.
- Delattre, P., Liberman, A.M., Cooper, F.S., and Gerstman, L.J. (1952). "An experimental study on the acoustic determinants of vowel color; Observations on one- and two-formant vowels synthesized from spectrographic patterns," *WORD* **8** (3), 195-210.
- Dixon, W.J. (1985). *BMDP Statistical software* (University of California Press, Berkeley).
- Dorman, M.F., Marton, K., and Hannley, M.T. (1985). "Phonetic identification by elderly normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **77**, 664-670.
- Dreschler, W.A., and Plomp, R. (1985). "Relations between psychophysical data and speech perception for hearing-impaired subjects. II," *J. Acoust. Soc. Am.* **78**, 1261-1270.
- Dubno, J.R., and Dorman, M.F. (1987). "Effects of spectral flattening on vowel identification," *J. Acoust. Soc. Am.* **82**, 1503-1511.
- Dubno, J.R., and Schaefer, A.B. (1989). "Frequency resolution for broadband-noise masked