BINAURAL ASPECTS OF SPEECH PERCEPTION IN NOISE

STELLINGEN

1. De verbetering van spraakverstaan in rumoer, veroorzaakt door de bij een ruimtelijke scheiding van spraakbron en stoorbron optredende interaurale verschillen in aankomsttijd en geluiddrukniveau, is minder dan de som van de verbeteringen die door beide interaurale verschillen afzonderlijk worden veroorzaakt.

   (Hoofdstuk 2 van dit proefschrift)

2. Het begrip "cocktail party effect", dat gebruikt wordt als aanduiding voor het voordeel van binauraal, vergeleken met monauraal, spraakverstaan in een omgeving met storende sprekers, is ongelukkig gekozen. Het binaurale voordeel neemt namelijk af bij toenemend aantal storende sprekers en is dus juist in het geroezemoes van een "cocktail party" relatief klein.

   (Hoofdstuk 4 van dit proefschrift)

3. Het verdient de voorkeur bij een klinische test van spraakverstaan in lawaai niet alleen stationair maar ook fluctuerend stoorgeluid te gebruiken, dat bovendien niet alleen vanuit dezelfde richting als de spraak wordt aangeboden, maar ook vanuit andere richtingen.

   (Hoofdstuk 5 van dit proefschrift)

4. De bewering van Carhart et al. (1968), dat er "perceptual masking" optreedt, namelijk dat spraak meer wordt gemaskeerd door stoorsignalen die spraak bevatten dan door andere gemoduleerde stoorsignalen, wordt onvoldoende door hun meetgegevens ondersteund.

   (R. Carhart et al. "Perceptual masking in multiple sound background," J. Acoust. Soc. Am. 45, 694-703, 1968).

5. De bij de berekeningen van de "Articulation Index" en de "Speech Transmission Index" gemaakte aanname, dat spraakverstaanbaarheid optimaal beschreven wordt door een gewogen som over frequentiebanden van in decibels uitgedrukte signaal/ruis verhoudingen, verdient nader onderzoek.

6. De in de literatuur beschreven methoden om "frequentie-specifieke" AEPs (auditory evoked potentials) te verkrijgen, gebruik makend van stimuli bestaande uit korte toonstootjes, hebben weinig waarde als niet tevens wordt onderzocht, welke gedeelten van het basilair membraan bijdragen tot het ontstaan van de gemeten potentialen.

   (bv. H. Davis et al. "Threshold sensitivity and frequency specificity in auditory brainstem response audiometry," Audiology 24, 54-70, 1985).

7. Het verdient aanbeveling de in dit proefschrift gebruikte methode - het aanbieden van binaurale stimuli waarin een scheiding van interaurale verschillen in aankomsttijd en geluiddrukniveau is aangebracht - tevens bij het onderzoek van het richtinghoren toe te passen. Hierdoor kan meer inzicht worden verkregen in het richtinghoren bij slechthorendheid en bij het dragen van gehoorbeschermers.

8. Het toevoegen van achtergrondmuziek aan spraak, wat bij radio- en televisieuitzendingen een gewoonte dreigt te worden, heeft tot gevolg dat grote groepen ouderen en slechthorenden niet of nauwelijks in staat zijn de gesproken tekst te verstaan.

9. Het is onbegrijpelijk dat de Nederlandse overheid enerzijds streeft naar beëindiging van de verkoop van produkten die CFK's (chloorfluorkoolwaterstoffen) bevatten, terwijl zij anderzijds toestaat dat circa 10% van de wereldproductie van CFK's op Nederlands grondgebied plaatsvindt.

10. Gebruikers van tekstverwerkers letten bij het afdrukken van teksten vaak te veel op het uiterlijk en te weinig op de leesbaarheid.

VRIJE UNIVERSITEIT TE AMSTERDAM

BINAURAL ASPECTS OF SPEECH PERCEPTION IN NOISE

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Vrije Universiteit te Amsterdam,
op gezag van de rector magnificus
dr. C. Datema,
hoogleraar aan de faculteit der letteren,
in het openbaar te verdedigen
ter overstaan van de promotiecommissie
van de faculteit der geneeskunde
op vrijdag 18 mei 1990 te 15.30 uur
in het hoofdgebouw van de universiteit,
de Boelelaan 1105

door

ADELBERT WILLEM BRONKHORST

geboren te Horn

Bariet, Ruinen

Promotor : prof.dr.ir. R. Plomp
Referent : dr.ir. T. Houtgast

VOORWOORD

Het onderzoek dat in dit proefschrift is beschreven heeft grotendeels plaatsgevonden gedurende de tijd dat ik als klinisch audioloog werkzaam was bij de vakgroep Keel- Neus- en Oorheelkunde van de Vrije Universiteit te Amsterdam. Een klein gedeelte van het onderzoek, en een groter gedeelte van het schrijfwerk, is verricht bij de afdeling Audiologie van het Instituut voor Zintuigfysiologie te Soesterberg, mijn huidige werkkring. Voor de ruimte en ondersteuning die mij op beide plaatsen geboden is ben ik bijzonder erkentelijk.

Een aantal personen wil ik met name bedanken voor hun bijdragen aan de totstandkoming van dit proefschrift. Mijn dank gaat in de eerste plaats uit naar mijn promotor Reinier Plomp. Niet alleen wegens het feit dat hij het promotieonderzoek heeft geïnitieerd en zorgvuldig heeft begeleid, maar ook omdat hij mij immer als stimulerende en inspirerende kracht terzijde heeft gestaan. Verder zou ik Joost Festen willen danken voor de vele nuttige discussies en voor zijn hulp bij het aanmaken van het testmateriaal voor mijn laatste experiment. Theo Kapteyn en Tammo Houtgast dank ik voor de vele tijd die zij mij ruimhartig voor mijn onderzoek gegund hebben en voor de aandacht die zij als commissieleden aan het proefschrift hebben besteed. Ook wil ik Prof. G.B. Snow danken voor de door hem geboden mogelijkheden en ondersteuning. Verder ben ik Guido Smoorenburg erkentelijk voor zijn commentaar op een eerdere versie van het manuscript. Graag wil ik ook de overige commisieleden, Prof. E. de Boer, Prof. H. Schneider en Prof. N. Westerhof, bedanken. Fred Snel en Ton Houffelaar wil ik danken voor hun technische ondersteuning. Tanja Manvis, Haggis Matray en de mensen van het audiologisch centrum dank ik voor hun hulp bij het selecteren en uitnodigen van proefpersonen. Verder ben ik natuurlijk de proefpersonen zelf dank verschuldigd - zonder hen was dit werk niet mogelijk geweest. Tenslotte wil ik Titia bedanken voor haar steun, die voor mij onontbeerlijk is geweest.

De uitgave van dit proefschrift is financieel ondersteund door
Veenhuis Medical Audio B.V.

Omslagontwerp: B. Koning, Amsterdam

Aan Titia,
aan mijn ouders

CONTENTS

Chapter	1	GENERAL INTRODUCTION	
	1.1	Objective of this study	11
	1.2	General aspects of the experimental design	12
	1.3	Summary of results	13
	1.4	Towards a model for binaural speech perception in noise	15
Chapter	2	THE EFFECT OF HEAD-INDUCED INTERAURAL TIME AND LEVEL DIFFERENCES ON SPEECH INTELLIGIBILITY IN NOISE	19
	2.1	Introduction	20
	2.2	Material	21
	2.3	Experiment I	25
		A. Method	25
		B. Results	27
	2.4	Experiment II	28
		A. Material and Method	29
		B. Results	30
	2.5	Discussion	33
Chapter	3	BINAURAL SPEECH INTELLIGIBILITY IN NOISE FOR HEARING-IMPAIRED LISTENERS	37
	3.1	Introduction	38
	3.2	Material and Method	40
	3.3	Experiment I: Symmetrical impairment	
		A. Conditions	41
		B. Subjects	42
		C. Results	43
	3.4	Experiment II: Asymmetrical impairment	
		A. Conditions	47
		B. Subjects	47
		C. Results	48
	3.5	Discussion	51
		A. SRT for monotic and diotic stimulation	52
		B. Effect of spatial separation of speech and noise sources	52
		C. Effect of interaural differences in overall presentation level	55
	3.6	Conclusion	56
Chapter	4	BINAURAL SPEECH INTELLIGIBILITY FOR SIMULATED COCKTAIL-PARTY CONDITIONS	57
	4.1	Introduction	58
	4.2	Method	
		A. Material	59

B. Procedure 61
C. Subjects 63
4.3 Results 63
4.4 Discussion 68
A. Effects of envelope fluctuations on masking efficiency 69
B. Effect of the binaural cues ITD and ILD 69
C. Performance of the hearing-impaired 71
D. Implications for practical situations 72
4.5 Conclusion 72

Chapter 5 A CLINICAL TEST FOR THE ASSESSMENT OF BINAURAL SPEECH PERCEPTION IN NOISE 75
5.1 Introduction 76
5.2 Method
A. Material 77
B. Procedure 77
C. Subjects 79
5.3 Results and Discussion
A. Normal-hearing listeners 80
B. Hearing-impaired listeners 82
C. Optimization of the test 85
5.4 Conclusion 86

GENERAL CONCLUSION 87

SAMENVATTING 89

REFERENCES 93

1. GENERAL INTRODUCTION

1.1 Objective of this study

Auditory research has firmly established that there are several important advantages of binaural, as compared to monaural, listening. One advantage is the ability to localize sound sources, which is based on complex subconscious processing of spectral cues and interaural signal differences in terms of level and arrival time, in combination with position information of the body and head. This remarkable ability has been an object of study for at least two centuries, and it still remains of great interest to engineers and researchers, for instance in the field of stereophonic sound reproduction.

A less well-known advantage of binaural listening is that it may considerably improve detection of signals and perception of speech in noisy environments. In spite of the fact that this effect can be easily verified, for instance by plugging one ear when listening to a conversation in a noisy cafeteria or cocktail party, it was not until relatively recently that it became object of scientific study (Licklider, 1948; Hirsh, 1948, 1950; Koenig, 1950; Kock, 1950). The results of these and other studies show that this second advantage consists, in effect, of two separate contributions, one caused by interaural differences in signal-to-noise (S/N) ratio, occurring as a result of headshadow, the other caused either by interaural differences in arrival time, occurring when a sound source is closer to one ear than to the other, or by decorrelation of the signals at the two ears, taking place when sounds are reproduced in a reverberant environment. The second contribution is commonly referred to as "unmasking", "release from masking", or "masking-level difference" (MLD); it results from some kind of internal noise suppression, taking place at an early stage in the auditory pathways.

Since its discovery, a considerable amount of research has been devoted to this noise-suppression ability of the binaural system (extensive reviews for normal and impaired hearing are given in Durlach and Colburn, 1978, and Durlach *et al.*, 1981, respectively). In most experiments, earphone presentation of the sound stimuli was used, so as to enable easy manipulation of interaural signal differences. This had the additional advantage, from the point of view of the investigators, that there was no confounding effect of headshadow. However, when one wishes to gain insight into the advantages of binaural listening in everyday situations, the contribution of headshadow has to be taken explicitly into account. Though there are several studies that have employed actual sound field stimuli (e.g., Hirsh, 1950; Carhart, 1965; Dirks and Wilson, 1969a; Plomp and Mimpen, 1981), interpretation of the obtained results is complicated by the diversity of methods, stimuli, and test environments used. A further confounding factor is the concurrence of headshadow and interaural time delay during normal sound field listening, which precludes strict separation of their contributions. It is furthermore difficult to apply the results to everyday listening conditions, as most studies only considered an anechoic environment with a single interfering sound source. Finally, it appears that relatively little is known about binaural speech perception in noise of unaided and aided hearing-impaired listeners. It was, therefore, felt that there was need for a systematic investigation into the relationship between interaural signal differences and the advantage of binaural listening experienced by normal-hearing and hearing-impaired listeners in noisy situations. The research reported in this thesis represents an attempt to satisfy this need.

1.2 General aspects of the experimental design

An important consideration in the design of the experiments described in the following chapters was the abovementioned concern with binaural performance in everyday circumstances. The approach used was to model, as far as possible, the stimuli, conditions and procedures after daily-life situations. This resulted in the following design characteristics, employed throughout this study:

(1) Speech-reception thresholds (SRT) are determined using lists of short sentences, presented according to a simple up-down procedure that estimates the 50% intelligibility level (Plomp and Mimpen, 1979). The sentences (130 in total) are representative of everyday conversation. In the adaptive procedure, a faultless replication of the entire sentence is required for a correct response. The advantage of this method is that it allows fast and accurate measurement of the SRT. Due to the strictness of the criterion, however, it is not suitable for use with severely hearing-impaired subjects.

(2) A noise masker is used with the same long-term average spectrum as the sentences. The noise is either steady-state or modulated according to the envelope fluctuations of connected discourse (Festen, 1990). The use of a speech-like masker is motivated by the relative importance, in everyday life, of social activities, where speech is the major interfering sound. Matching the long-term average spectra of noise and speech has the additional advantage of reducing the speaker dependence of the SRTs in noise.

(3) The speech is presented always from the front, this being the most natural situation, especially for hearing-impaired listeners who normally rely also on visual cues. The noise is presented either from the front, or at a certain azimuth.

Another main consideration was the aim to evaluate separately the contributions to the binaural advantage of interaural time delay (ITD) and interaural level differences (ILD) due to headshadow. This is of interest because these contributions rely on completely different processes, and thus will differ in their dependence on, e.g., hearing loss or presentation level. Separation of both cues can only be achieved when the stimuli are presented through headphones. At the same time, however, actual sound field listening should be simulated as closely as possible, to comply with the former consideration.

The solution chosen was to make use of artificial-head recordings, which were computer-processed to realize separation of ITD and ILD. This method has several additional advantages: (1) it greatly facilitates testing, as no regular access to a suitable listening environment with proper reproduction equipment is required, (2) it allows independent setting of the presentation levels at both ears, to simulate or compensate for hearing loss, and (3) it prevents confounding effects of head movements. The main disadvantage is that it relies completely on the quality of the artificial head used. The stimuli used presently were recorded using a KEMAR manikin (Burkhard and Sachs, 1975), consisting of a head and torso, and equipped with two Brüel & Kjær (B&K) 4157 ear simulators. This is a standardized manikin with carefully chosen dimensions, taken in between mean values for males and females. The recordings were made in an anechoic room. The resulting quasi free-field stimuli and the derived signals were used in all experiments except the last one (described in chapter 5). This experiment employed true sound field stimuli to enable evaluation of the effects of reverberation and to allow for the use of hearing aids.

1.3 Summary of results

The *first study*, described in Chapter 2, was aimed at determining the relative contributions of headshadow and ITD to the advantage of binaural listening, as experienced by normal-hearing listeners in a free field. The experimental configuration studied included a speech source, located in front (0°), and one noise source, located at an azimuth Φ, ranging from 0° to 180° in steps of 30°. SRTs were determined for a group of 17 listeners using both the unprocessed KEMAR free-field recordings, and the derived signals containing either only ILD or only ITD. These signals were only generated for azimuths 30° through 150°, as no interaural differences are present when the sound source is located at 0° or 180°. The experiment yielded a mean SRT of -6.4 dB for the 0° noise azimuth, and, for the unprocessed stimuli, an SRT advantage that increased by up to 10 dB (for Φ=90°), and then decreased to less than 1 dB (for Φ=180°) as a function of noise azimuth. An important result was that the effects of ILD and ITD were not additive. For the 90° azimuth, for instance, the gain caused by ILD was 7.8 dB, and the unmasking due to ITD was 5 dB, so that almost 3 dB got lost during simultaneous presentation, as a result of some kind of interaction between both cues.

In a second experiment, conducted with 17 other normal-hearing subjects, several aspects were investigated further. A comparison was made between the effect of the head-induced (frequency-dependent) ITD, and that of a fixed ITD. The monaural component of the headshadow advantage was determined by presenting stimuli monaurally. Finally, a one-sided (conductive) hearing loss was simulated by introducing a 20-dB attenuation in one channel. The following results were found. First, it appeared that the unmasking caused by the head-induced ITD depends mainly on the delay for the low frequencies (250-500 Hz). Second, the monaural results showed that the headshadow advantage relies entirely on the ear presented with the most favorable S/N ratio, which means that it is, in principle, a monaural effect. Finally, it was found that the one-sided attenuation hardly affected unmasking due to ITD, but considerably reduced the headshadow advantage when it was introduced in the channel with the most favorable S/N ratio.

In the *second study*, presented in Chapter 3, the same stimuli were used as in the first study, but subjects were now two groups of 17 hearing-impaired listeners with either symmetrical or asymmetrical sensorineural hearing losses. Most results were obtained for just two noise azimuths: 0° and 90°. Again, conditions were included with an interaural difference in overall presentation level, now in order to simulate the effect of a monaural hearing aid. In all conditions tested, average SRTs appeared to be significantly higher than corresponding results for normal-hearing listeners. However, the relative advantage due to ITD, both with and without simultaneous presence of ILD, was found to be nearly normal. The relative headshadow advantage appeared to vary considerably among subjects: it ranged from 0 dB to more than 8 dB. Further analysis of the data showed that this advantage decreases as a function of the high-frequency hearing loss at the ear presented with the most favorable S/N ratio. The explanation of this dependence is that, because headshadow increases with frequency (see Chapter 2, Fig. 2), the headshadow advantage depends mainly on perception of the high-frequency speech components. The effect of the interaural difference in overall presentation level was similar as for the normal-hearing listeners: unmasking due to ITD was hardly affected, and the headshadow advantage decreased when the channel having the more favorable S/N ratio was attenuated. The latter finding has the interesting implication that a hearing aid may in certain situations increase the headshadow advantage by raising the signal level at the favorably located ear.

The *third study*, presented in Chapter 4, again employed the stimuli recorded with the KEMAR manikin, but the noise signals were now modulated like speech and added in various combinations in order to simulate a multiple-talker environment. The configuration studied comprised a speech source in front, and up to six independently fluctuating maskers either also in front, or positioned in a symmetrical or asymmetrical arrangement around the listener. Subjects were 17 normal-hearing listeners and 17 listeners with symmetrical sensorineural hearing losses. Results showed that the normal-hearing listeners can gain considerably from both headshadow and masker fluctuations: with a single masker, they had an headshadow advantage of 8 dB, and an advantage due to masker fluctuations of about 5 dB. Both gains, however, decreased rapidly as a function of the number of maskers. In contrast, ITD was found to yield a constant advantage of 2-3 dB, independent of the number of maskers. The hearing-impaired listeners had mean SRTs that were 4 to 10 dB poorer than corresponding normal results. A remarkable finding is that they experienced hardly any advantage due to masker fluctuations. They showed furthermore a reduced headshadow advantage compared with the normal hearing, but about the same gain due to ITD. An implication of the results for actual multiple-talker situations is that, when all voices are equally loud, speech remains intelligible for normal-hearing listeners even when there are as many as six competing talkers, while the hearing impaired can not tolerate more than three competing talkers.

The *fourth, final, study* is described in Chapter 5. Its aim was the development of a clinically feasible test for binaural speech perception in noise. The test configuration chosen comprises a speech source and a single noise source, located in a reverberant room, with listener either in the direct or the indirect field of the speech and/or noise source. This enables evaluation of the effects of reverberation (i.c. speech distortion and interaural signal decorrelation) on the SRT. Directional effects are assessed by presenting the noise from the left or right side in the conditions where both sound sources are close to the listener. The influence of masker fluctuations is determined by using either steady-state or speech-like modulated noise. The test was presented to a group of normal-hearing subjects, so as to obtain reference data, and to 14 unaided and 4 aided hearing-impaired subjects with sensorineural hearing losses. Results showed that the performance of individual hearing-impaired listeners deviated significantly from normal for at least two of the following aspects: (1) perception of speech in steady-state noise, (2) relative binaural advantage due to directional cues, (3) relative advantage due to masker fluctuations. In contrast, it appeared that both the hearing loss for reverberated speech and the relative binaural advantage due to interaural signal decorrelation were essentially normal for almost all hearing impaired tested. The results lead to the conclusion that, in order to prevent underestimation of the hearing handicap, a clinical test of speech perception in noise should employ both fluctuating and steady-state interfering sounds, and should use presentation of speech and interfering sound from both the same and from different directions.

1.4 Towards a model for binaural speech perception in noise

A possible extension of the experimental work presented in this thesis is the development of a model for binaural speech perception in noise. Such a model would be of interest because performance can then be predicted for many conditions, whereas the experimental results only apply to a restricted set. Recently, a model was proposed by Zurek (1987) in an unpublished paper. The model only applies to free-field conditions with a single steady-state noise source. Parameters are the azimuths of the speech and the noise sources, and the speech and noise spectra. The model takes the two following factors into account:

(1) Effect on speech intelligibility of spectral differences between speech and interfering sound, occurring at the ears. These differences can already be present in the speech and interfering sounds themselves. They can also be caused by headshadow, in situations where the sound sources are located at different azimuths. Prediction is based on Kryter's (1962) modification of the articulation theory of French and Steinberg (1947), and on Shaw's (1974) data on the azimuthal dependence of the sound-pressure transformation from free field to ear drum.

(2) Unmasking of speech due to interaural time delays present in the speech and/or the interfering sound. A formula based on Colburn's (1977) theory of

binaural interaction is used here. This formula also accounts for the effect of interaural differences in S/N ratio.

The model appears to be reasonably accurate in predicting a variety of experimental data, obtained by, e.g., Dirks and Wilson (1969a), Tonning (1971) and Plomp and Mimpen (1981). An example of an interesting application of the model is the calculation of the optimal head orientation for various locations of the speech and noise sources. A limitation of the model is that it can only be applied to conditions with a single steady-state noise source in a free field, and not to the more realistic conditions investigated in Chapters 4 and 5 of this thesis. A more general model should take the following factors into account as well:

(3) Effects of reverberation. First, the reduction in intelligibility of reverberated speech. Second, the unmasking of speech due to interaural decorrelation of speech and/or interfering sound. Third, the reduction of the headshadow advantage caused by fact that the indirect sound adds to the interfering sound.

(4) Effect of increasing the number of maskers. The S/N ratios present at both ears will depend on the number of maskers, on their spectral characteristics, and on their locations. In addition, the unmasking due to ITD changes in multiple-masker conditions (Carhart, 1969a).

(5) Gain in intelligibility caused by (speech-like) masker fluctuations. This is predominantly a monaural effect, but, in situations with multiple fluctuating maskers, there will also be a binaural contribution due to independently varying S/N ratios at both ears (see Section 4.4 B).

As to the effects of reverberation, the influence of the time-domain distortions on speech intelligibility can be adequately predicted from the Modulation Transfer Function (MTF) and the derived Speech-Transmission Index (STI) (Houtgast and Steeneken, 1973, 1985). Prediction of the unmasking due to interaural decorrelation can be based on the abovementioned theory of Colburn (1977). The latter effect of reverberation, i.e. the change in S/N ratio due to the contribution of the indirect sound, can be incorporated by modifying the speech and noise spectra that are used as input for the Articulation Index or STI calculation.

Extension of the model to multiple-masker conditions may be difficult. Not because of the changes in S/N ratio occurring at both ears - these can be easily calculated by adding the contributions of all individual maskers. The problem lies in the prediction of unmasking due to multiple maskers with different ITDs and ILDs. Such conditions have as yet hardly been studied, let alone accounted for by any model. It is possible that the unmasking for multiple maskers can be estimated with sufficient accuracy by taking the value for a single masker. This seems to be supported by the data presented in Chapter 4, which show approximately equal unmasking due to ITD for conditions with different numbers of maskers.

The prediction of the masking efficiency of fluctuating sounds presents the greatest difficulty. A possible approach would be to apply the modified STI calculation scheme proposed recently by Ludvigsen *et al.* (1990), which is based on linear regression analysis of the envelopes of speech and speech plus interfering sound. Festen (1990) has, however, shown that this method does not yield a satisfactory prediction of the masking efficiencies of speech or speech-like modulated maskers. Another approach, currently pursued by the present author, is to integrate the S/N ratio (or another measure of intelligibility, e.g., the modulation index *m*, from which the STI is derived) over a certain time interval. The S/N ratio is calculated using the intensity envelope function, which is first lowpass filtered to account for the limited auditory temporal resolution. It appears that this method is quite promising, but underestimates the masking efficiency of relatively slowly fluctuating signals (with modulation frequencies below 4 Hz). As such modulation frequencies are also present in speech, adequate prediction of its masking efficiency requires that further modifications to the calculation scheme have to be made.

The above considerations indicate that virtually all elements required for a general model of binaural speech perception in noise are available. The development of a firmly based model will, however, still involve a considerable amount of work.

2. THE EFFECT OF HEAD-INDUCED INTERAURAL TIME AND LEVEL DIFFERENCES ON SPEECH INTELLIGIBILITY IN NOISE

A.W. Bronkhorst and R. Plomp
Journal of the Acoustical Society of America 83, 1508-1516 (1988).

ABSTRACT

A study was made of the effect of interaural time delay (ITD) and acoustic headshadow on binaural speech intelligibility in noise. A free-field condition was simulated by presenting recordings, made with a KEMAR manikin in an anechoic room, through earphones. Recordings were made of speech, reproduced in front of the manikin, and of noise, emanating from seven angles in the azimuthal plane, ranging from 0° (frontal) to 180° in steps of 30°. From this noise, two signals were derived, one containing only ITD,the other containing only interaural level differences (ILD) due to headshadow. Using this material, speech reception thresholds (SRT) for sentences in noise were determined for a group of normal-hearing subjects. Results show that (1) for noise azimuths between 30° and 150°, the gain due to ITD lies between 3.9 and 5.1 dB, while the gain due to ILD ranges from 3.5 to 7.8 dB, and (2) ILD decreases the effectiveness of binaural unmasking due to ITD (on the average, the threshold shift drops from 4.6 to 2.6 dB). In a second experiment, also conducted with normal-hearing subjects, similar stimuli were used, but now presented monaurally or with an overall 20-dB attenuation in one channel, in order to simulate hearing loss. In addition, SRTs were determined for noise with fixed ITDs, for comparison with the results obtained with head-induced (frequency dependent) ITDs. Results show that (1) for noise with only ILD, the gain relies on the ear presented with the most favorable signal-to-noise ratio, but decreases when the noise presented to the other ear becomes relatively loud, (2) the effect of ITD is fairly insensitive to a 20-dB attenuation of either channel, also when ITD is introduced in noise with ILD, and (3) the unmasking caused by a head-induced ITD can be predicted from the delay for low frequencies (250-500 Hz).

2.1 INTRODUCTION

It is well established that differences in arrival time at the two ears play a role not only in localization, but also in the discrimination of speech in noise. In a number of studies, binaural speech intelligibility in noise was measured as a function of interaural time delay (ITD) in the noise masker or the speech signal (Kock, 1950; Schubert, 1956; Carhart *et al.*, 1967; Levitt and Rabiner, 1967a). It appeared that the introduction of ITD caused the intelligibility function to shift towards lower signal-to-noise (S/N) ratios. A measure of this effect is the gain in intelligibility at the 50% level, the binaural intelligibility level difference (BILD). Results reported in the abovementioned studies indicated that interaural delays of the same magnitude as those occurring in normal binaural listening, yield maximum BILDs of 3-9 dB, depending on the listening task involved. In comparing these results with free-field measurements of speech intelligibility in noise (Platte and vom Hövel, 1980; Plomp and Mimpen, 1981) a difficulty arises because the effect of ITD is then confounded with that of interaural level differences (ILD) due to headshadow. Moreover, it is unknown to what amount the BILD is affected by a dependence of ITD on frequency, such as caused by the human head (Kuhn, 1977). As a consequence, in the interpretation of the free-field results, a separate evaluation of the contributions of ILD, ITD and their interaction is not possible. Such a separation would be of interest since we expect the gain due to ILD to depend mainly on monaural ("best ear") performance, whereas the unmasking due to ITD represents purely binaural interaction.

In the first of the two experiments described in this study, the aim was to determine separately the effects of ITD and ILD (present in a free field) on the BILD for normal-hearing subjects. A spatial configuration was chosen with speech presented frontally, and the noise from seven directions in the horizontal plane. The second experiment, which was also conducted with normal-hearing subjects, contained conditions where hearing loss was simulated by presenting the signals monaurally or with an overall 20-dB attenuation in one channel. In addition, conditions were included that probed the effect of a fixed ITD, in contrast with the frequency- dependent ITDs in the first experiment.

In both experiments, the test material was presented through headphones. In this way, the ITDs and S/N ratios on either ear could be kept completely under control. However, when a comparison is to be made between results obtained with headphones and measurements in a real free-field environment, the typical delays and diffractions as caused by the human head also have to be present in the signals presented through headphones. We, therefore, decided to make use of an acoustical manikin. This technique of using artificial head recordings to measure binaural intelligibility in noise has been used before (Hirsh, 1950; Nordlund and Fritzell, 1963; Dirks and Wilson, 1969a). At present, the choice of a suitable artificial head is facilitated by the availability of the KEMAR manikin, which closely simulates the average human head and torso (Burkhard and Sachs, 1975). When it is required that dummy-head recordings, reproduced over headphones, produce an extracranial sound image with the correct localisation cues, proper equalization is important and care has to be taken that no phase shifts are introduced. In our case, this was not critical because neither ITD is affected, as long as both channels are identical, nor S/N ratio, since both speech and noise would be recorded with the manikin. Thus we assumed that by making use of a KEMAR manikin with a simple correction filter, the simulation of the free-field condition would be sufficiently accurate for our purposes. Starting with the recordings made with the manikin, the separation of ITD and ILD could be realized by digitizing the signals and applying FFT techniques. Both experiments described below made use of both the original recordings and the artificial signals generated in this way.

2.2 MATERIAL

The speech material used consisted of 130 short Dutch meaningful sentences of eight or nine syllables, typical of everyday conversation (Plomp and Mimpen, 1979). The sentences had all been read by the same female speaker and adjusted in level for equal intelligibility. The noise masker used had a spectrum equal to the long-term average spectrum of the sentences.

Recordings of both speech and noise were made in an anechoic room with a KEMAR acoustical manikin, consisting of a head and torso, placed at a distance of 1.5 m from a Quad electrostatic loudspeaker. The amplitude-frequency response of the loudspeaker and power amplifier (Quad 33) was verified using a Brüel & Kjær (B&K) 2118 audio test station. The response curve, measured with the manikin removed and the test microphone placed at the position of the center of the manikins head, appeared to be flat within ± 4 dB in the range from 125-8000 Hz.

The manikin, fitted with the standard size ear replica, included two B&K 4157 ear simulators with B&K 4134 microphones. After passing through a B&K 2610 amplifier both microphone signals were fed into a Tascam 44 tape recorder. The speech material was recorded only for a frontal position of the loudspeaker relative to the manikin. Recordings of the noise were made for seven azimuthal angles ranging from $\Phi=0°$ (frontal) to $\Phi=180°$ in steps of 30°. Both head and torso of the manikin were rotated while the position of the loudspeaker remained fixed.

When these recordings are played back over headphones, the ear-canal resonance of the listener adds to a similar resonance introduced by the ear simulators in the manikin. Thus a response peak of approximately 15 dB at 2700 Hz is introduced which causes a manifest deterioration of the sound quality. This problem was overcome by using an equalization filter as described by Killion (1979). The phase shift caused by this filter had no consequence in our case, since both channels were affected equally.

The recordings of the 130 sentences were split up in 17 lists of 8 sentences each. Thus six sentences (always having the second position in a list) were used twice. Of the recordings of the noise, 5-s intervals were sampled at 20 kHz with

12-bits resolution. For this purpose we used a PDP 11/73 computer equipped with a GTSC high-speed multichannel AD converter. Before separating ITD and ILD, a numerical analysis was applied to the digitized signals in order to quantify both effects.

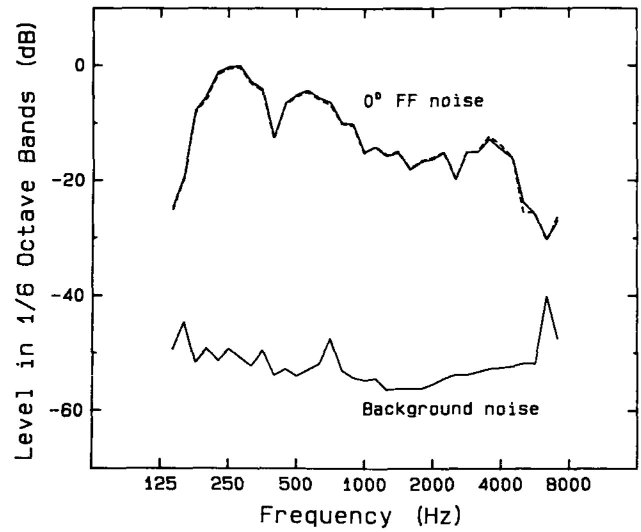

FIG. 1. Power spectra of the noise, recorded through both channels of KEMAR at an azimuth of 0°. The spectrum of a silent passage in the sentence list (also recorded at $\Phi=0°$) is plotted as well. It represents the sum of background and tape noise.

In order to calculate the ILDs, both the right and the left channels of the sampled noise intervals were broken up into 25 segments, each consisting of 4096 samples. Using an FFT algorithm, a separate power spectrum was calculated for each segment. (Here, as well as in the procedure to generate the artificial stimuli, the FFT was calculated without prior application of a window function. The spectra shown in Figs. 1 and 2 are hardly affected by the use of a window function. With a $\cos^2$ window, for instance, the differences are within ± 1 dB.) By averaging the power contained in 1/6 octave bands over all segments, a mean 1/6 octave power spectrum was obtained. In Fig. 1, both 0° power spectra are shown, as well as the spectrum of a silent passage in the recorded sentence list. The latter represents the sum of background and tape noise, present in the recordings. In Fig. 2, normalized response curves are plotted for the nonzero noise azimuths. These were obtained by subtracting the 0° power spectrum from the other power spectra. For each azimuth, there are two curves: one representing the ipsilateral ear with respect to the loudspeaker, the other the contralateral ear. Comparing these data with results reported by Blauert (1974, p. 74), a close correspondence can be observed except for the responses of the contralateral ear at backward angels. In those cases, the KEMAR manikin yields 5-15 dB less attenuation in the frequency region above 2 kHz.

A further analysis was performed on the digitized noise signals in order to obtain the ITDs introduced by the manikin. The dependence on frequency was determined by carrying out separate calculations for four octave bands spanning the range from 250-4000 Hz. These bands were extracted from the original noise signals using a 128- (250-500 Hz) or a 64-point (other bands) digital finite impulse response (FIR) filter with a slope of 36 dB/oct. For each filtered noise sample, three seperate calculations of the interchannel cross-correlation function were made by means of convolutions with different 128-point intervals. As estimate of the ITD, we took the average of the locations of the maxima of these functions. In each case, the maxima differed by no more then a single sample interval (50 μsec). The results, plotted in Fig. 3, show a gradual decline of delay as a function of frequency. They are in good agreement with data published by Abbagnaro *et al.* (1975, live subjects and manikin) and Kuhn (1977, KEMAR manikin) The results for $\Phi=0°$ and $\Phi=180°$ are also included in Fig. 3, these served to insure that no delay artifacts were present.

With the aid of the difference spectra shown in Fig. 2, two classes of artificial signals were generated from the original free-field noise samples (hence called FF noise). Just as in the calculation of the spectra, the noise samples were

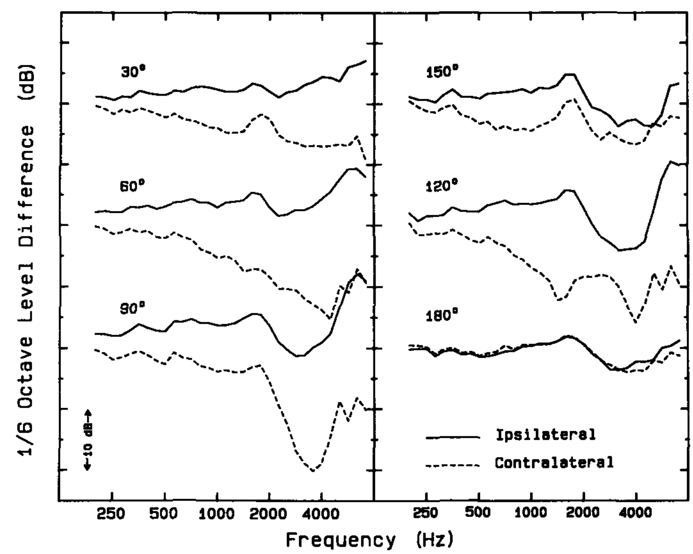

FIG. 2. Normalized response curves of KEMAR, obtained by subtracting the power spectrum of the noise recorded at $\Phi=0°$ from the spectra of the noise recorded at other angles.

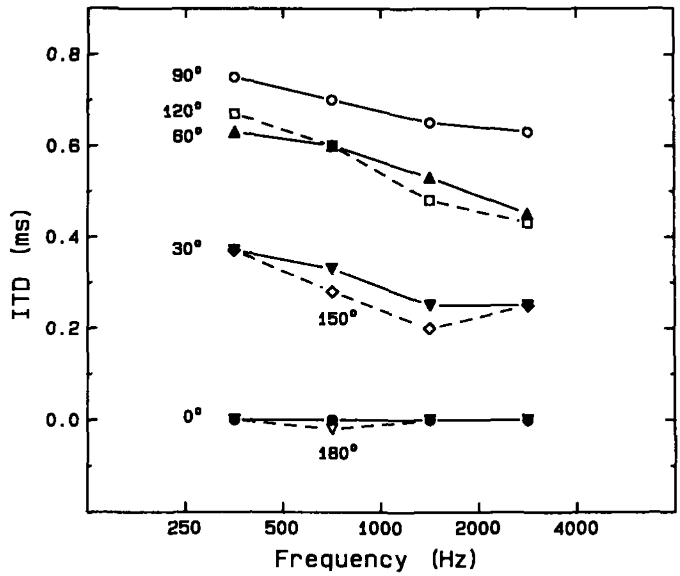

FIG. 3. Interaural time delays introduced by KEMAR for different angles of incidence in the azimuthal plane. The points represent the delay within octave bands.

divided into 25 segments of 4096 points, and the FFT was determined for all segments. Noise with only ITD, pertaining to a certain azimuth, and no ILD (dT noise), was obtained from the FF noise, recorded at that azimuth, by shaping the Fourier transforms according to the 0° spectrum and subsequent inverse transformation. Noise with only ILD and no ITD (dL noise) was generated from the 0° noise by modifying it according to the spectral characteristics of the lateral angles. In these operations, only the length of the spectral vectors was changed to mimimize distortions of the phase. To illustrate the effect of the spectral shaping, small portions of the 0° and 90° FF noise signals are displayed in Fig. 4 together with the derived 90° dL and dT noises. It turned out that the processing caused small but audible offset differences between the subsequent segments (due to the fact that no window function was applied). We, therefore, included a 128-point overlap in the segments and combined them, after processing, by multiplication with a trapezoid window and subsequent addition.

This calculation resulted in a dL noise and a dT noise for each of the five lateral angels and thus, together with the seven original (FF) noise samples we had a total of 17 different noise types. In order to check the properties of the dL and dT noises, average power spectra were determined and the same ITD calculation was applied, as described above for the case of the FF noise. The results showed that the differences between the dL and FF spectra, as well as those between the dT spectra and the 0° FF spectrum, were within ± 1 dB in the frequency range from 150-6000 Hz. The ITDs for dT noise were within 50 μsec of the corresponding values for FF noise; all ITDs for dL noise were less than 50 μsec.

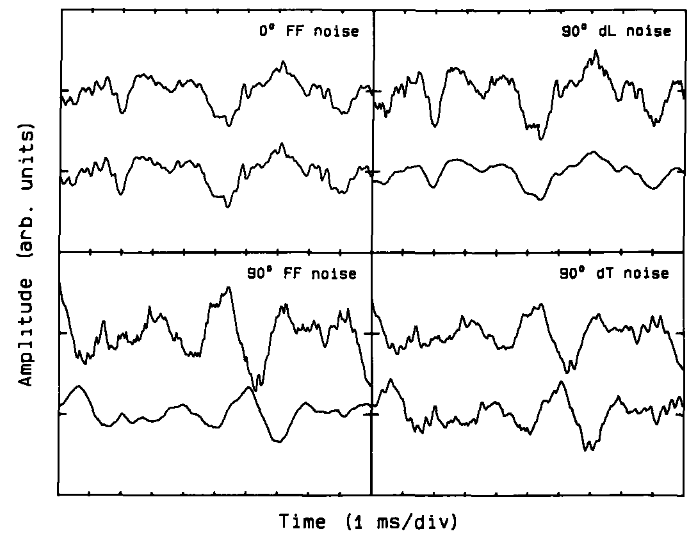

FIG. 4. Small parts of the recorded noise signals (left panels) and derived signals (right panels) illustrating the effect of the spectral shaping. The upper traces correspond to the KEMAR ear ipsilateral to the loudspeaker, the lower traces to the contralateral ear.

2.3 EXPERIMENT I

In this experiment, binaural speech reception thresholds were determined using the speech material and the 17 different noise types prepared according to the above description.

2.3 A. Method

Speech-reception thresholds were determined by varying the presentation level of the sentences using a simple up-down procedure with as criterion the correct replication of the entire sentence (Plomp and Mimpen, 1979). The first sentenceof each list was presented initially below threshold, and then raised in level in steps of 4 dB until it was reproduced correctly. The following seven sentences were presented only once, at a level that was lowered by 2 dB after a correct response, and raised by 2 dB after an incorrect response. The SRT was taken as the average presentation level after the second sentence. Plomp and Mimpen presented a graph of the measurement error as a function of the number of sentences used in the calculation of the SRT. According to these data, the measurement error will be about 1.3 dB in our case.

Seventeen copies of the list of 136 sentences (each consisting of 17 sublists of 8 sentences) were made on tape using the 4-track Tascam taperecorder. Through a 2-channel DA converter, the 17 noise signals were then added on separate tracks according to a 17 X 17 Latin square. Since no uneven digram-balanced Latin squares exist, an "optimally balanced" Latin square was calculated with a computer. During reproduction, the output signals of the recorder were mixed and amplified in two Madsen OB 822 clinical audiometers and presented over a set of Beyer DT-48 stereophonic headphones.

The experiment was performed in a double-walled sound-proof booth. Seventeen subjects, seven male and ten female, aged 20-29 years, participated in the experiment. They were paid for their services. For each subject, a pure-tone air-conduction audiogram was determined for both ears at octave frequencies

Table I. Results from experiment I. For each condition, the mean and standard deviation of the SRT is given. The BILDs in the last column were obtained by subtracting the mean SRTs from the mean result for 0° FF noise.

Noise Type	Noise Azimuth (°)	SRT (dB) mean	SRT (dB) s.d.	BILD (dB) *re:* 0° azimuth
FF	0	-6.4	1.5	
FF	30	-12.7	1.5	6.3
FF	60	-15.9	1.6	9.4
FF	90	-16.6	2.0	10.1
FF	120	-15.5	1.9	9.0
FF	150	-12.2	1.4	5.8
FF	180	-7.1	1.3	0.7
dL	30	-9.9	1.2	3.5
dL	60	-12.5	1.6	6.0
dL	90	-14.3	2.0	7.8
dL	120	-12.9	0.9	6.4
dL	150	-10.1	0.8	3.7
dT	30	-10.3	1.7	3.9
dT	60	-11.6	1.6	5.1
dT	90	-11.4	1.6	5.0
dT	120	-11.4	1.3	4.9
dT	150	-10.4	1.5	4.0

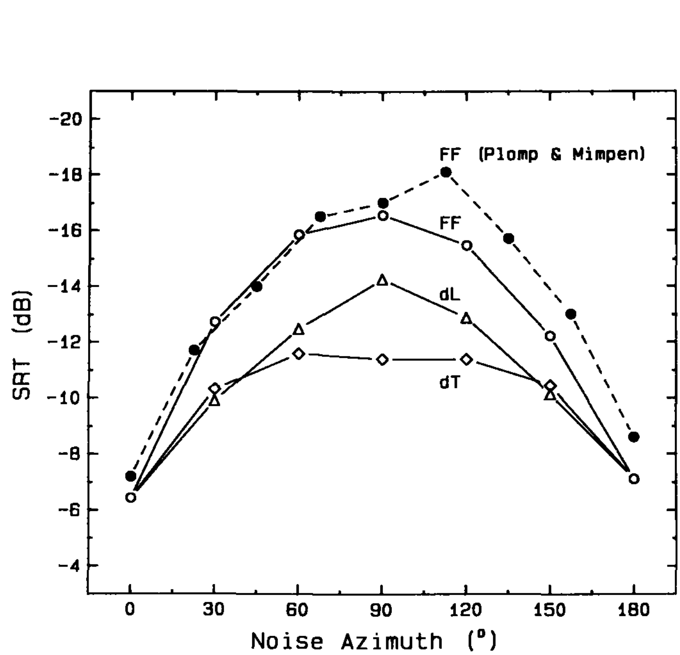

FIG. 5. Mean speech-reception thresholds obtained in experiment I for three different noise types: FF (free field), dL (headshadow only), and dT (ITD only). The closed data points represent results of Plomp and Mimpen (1981), obtained in a free field.

from 250-8000 Hz. All hearing levels were 20 dB or better except for two subjects, who had hearing levels of respectively 25 dB and 30 dB at 8000 Hz for one ear. Next, the sentences and noise maskers were presented as recorded on one of the 17 tapes. The measurement of the audiogram and the presentation of the sentence lists took about 45 min altogether. The reference presentation level of the noise was 60 dB(A) for the 0° noise, as measured with a B&K 4152 artificial ear. Normal and reverse headphone positionings were alternated between subjects.

2.3 B. Results

Means and standard deviations of the speech reception thresholds for the 17 conditions are presented in Table I. The last column shows the binaural intelligibility level difference relative to the SRT for $\Phi=0°$. According to an analysis of variance, applied to all conditions, 76.7% of the variance is contained in the conditions, and only 0.6% in the subjects. The error estimate is 1.5 dB. Further statistical analysis indicated that all BILDs were significant (*t* test on paired observations, $p<0.0001$) with the exception of the BILD of 0.7 dB obtained for $\Phi=180°$.

The mean SRT values are plotted in Fig. 5. The dashed line represents data of Plomp and Mimpen (1981) obtained with ten listeners in an anechoic room, using the same speech material. They agree well with our results for FF noise, especially at forward angles. Comparing the results for dL noise and dT noise, we see that ILD yields on the average a greater gain than ITD. Surprisingly, an almost constant BILD is obtained for all conditions with dT noise. Analysis of variance, applied to these conditions, indicates that the differences have only limited significance [$F(4,64)=3.01$, $p=0.024$], whereas the variances contained in the results for FF noise [$F(4,64)=23.4$, $p<0.001$] and dL noise [$F(4,64)=32.7$, $p<0.001$] are strongly significant. Moreover, it can be seen in Fig. 5 that the dependence of SRT on noise azimuth for FF noise and dL noise is very similar. In other words, the unmasking caused by introducing ITD seems independent of azimuth (within the range from 30°-150°), whether or not there is an ILD in the noise masker. This is tested statistically by evaluating the interaction of noise type versus noise azimuth in an analysis of variance, applied to the results for FF and dL noise. The analysis shows that there is no significant interaction [$F(4,64)=0.86$, $p=0.50$].

Averaging the BILDs over all five lateral angles yields mean intelligibility level differences for the conditions with FF, dL and dT noise of 8.1 dB, 5.5 dB and 4.6 dB, respectively. Comparison of these figures shows that the gains due to ILD and ITD do not add up.

2.4 EXPERIMENT II

The results of experiment I raised several questions which could not be answered satisfactorily without further data. We, therefore, decided to conduct a second experiment, using similar material and methods as in experiment I. The object of experiment II was threefold.

(1) We wanted to investigate whether the fact that the BILD due to ITD hardly depended on noise azimuth had been caused by the dependence of the head-induced ITDs on frequency. That such a dependence can have an effect on the BILD is demonstrated in the case of phase-inversion (corresponding to an $1/f$ dependence on frequency), which is known to yield greater unmasking than any fixed ITD (Carhart *et al.*, 1967). We, therefore, planned three conditions with a fixed ITD of 0.25, 0.50 and 0.75 ms, respectively.

(2) Concerning the results for dL noise, the question was whether the BILDs were entirely due to the ear with the most favorable S/N ratio, or whether there was also a contribution of the other ear. An analog in the case of unmasking of pure tones is the considerable masking level difference (MLD) caused by adding contralateral homophasic noise to monaural signal plus noise (Egan, 1965). To solve this point, we included two monaural conditions with 90° FF noise.

(3) To provide a link with future experiments with hearing-impaired subjects, we decided to include conditions in which the entire stimulus in one ear was

attenuated. These conditions would additionally serve as an interpolation between the normal binaural conditions and the monaural conditions. To limit the number of conditions, we restricted ourselves to one noise azimuth (90°) and one attenuation (20 dB). According to the literature on the effect of interaural differences in overall presentation level on the BILD (Wilson *et al.*, 1985, antiphasic speech in homophasic noise) and on the MLD for antiphasic pure tones (Egan, 1965; McFadden, 1968) a 20-dB attenuation already yields a significant decrease of unmasking. We did not choose a larger attenuation in order to prevent contamination of the results by threshold effects and cross-hearing (the FF and dL noise already having a maximum ILD of about 20 dB, cf. Fig. 2). Since the 90° FF, dL and dT noise signals are all dichotic, attenuating either the left or the right channel results in six different conditions. We decided to include the original 90° conditions as well. In this way, we would be able to analyse the experimental results more accurately (in an analysis of variance, the inter-subject variance can be subtracted).

Summing up all conditions, three with a fixed ITD, two monaural, six with and three without a 20-dB attenuation of one channel, and adding both the monaural and the binaural condition for 0° FF noise as reference and the 30° dT noise condition (for comparison with the fixed-ITD conditions) results in a total of 17 conditions, 5 of which are repetitions from experiment I.

2.4 A. Material and Method

In experiment II, we made use of the same speech material and part of the noise material as prepared for experiment I. In order to generate the noise with fixed ITD (hence called "constant time", CT noise), one channel of the (digitized) 0° FF noise was shifted in time with respect to the other. Since we had used a 20 kHz sample rate, the interchannel delay could be set in steps of 0.05 ms. Choosing delays of 0.25, 0.50 and 0.75 ms, we covered the range of ITDs present in the FF noise recordings (cf. Fig. 3). For the monaural conditions and the conditions with an attenuation of one channel, there was no need to modify the noise material. These conditions could be realized by switching off or attenuating one channel during reproduction.

Having the same number of conditions as in experiment I, the same procedure as described above could be followed for experiment II. We again used the list of 136 sentences with 8 sentences per condition and added the noise according to the same Latin square design. The experimental setup and calibration were identical as well.

A new group of 17 normal-hearing subjects was employed for experiment II. The subjects, ten male and seven female, aged 19-33 years, were paid for their participation. All except one had pure-tone hearing levels less than 20 dB for both ears at octave frequencies from 250-8000 Hz. One subject had a hearing level of 35 dB at 8000 Hz for one ear. The reference presentation level of the noise, in experiment I 60 dB(A) for the 0° FF noise as measured on an artificial

Table II. Results from experiment II. In the description of the noise type, "ipsi" and "contra" refer to the channel with the noise recorded on the ipsilateral and contralateral side of KEMAR with respect to the loudspeaker, respectively. For each condition, the mean and standard deviation of the SRT are given. The BILDs in the last column were obtained by subtracting the mean SRTs from the mean result for 0° FF noise.

Noise Type	Noise Azimuth (°)	SRT (dB) mean	SRT (dB) s.d.	BILD (dB) *re:* 0° azimuth
FF	0	-6.9	1.2	
FF monaural	0	-5.6	1.2	-1.3
CT 0.25 ms		-9.9	1.5	3.0
CT 0.50 ms		-11.0	2.0	4.1
CT 0.75 ms		-11.5	1.5	4.6
FF	90	-16.3	1.7	9.4
FF 20 dB att. ipsi	90	-16.1	1.3	9.2
FF 20 dB att. contra	90	-12.1	1.9	5.2
FF monaural contra	90	-13.8	1.3	6.9
FF monaural ipsi	90	-3.1	2.0	-3.8
dL	90	-13.5	2.3	6.5
dL 20 dB att. ipsi	90	-13.3	1.5	6.4
dL 20 dB att. contra	90	-10.0	1.9	3.0
dT	30	-10.6	1.1	3.7
dT	90	-11.6	1.4	4.7
dT 20 dB att. ipsi	90	-9.9	1.4	3.0
dT 20 dB att. contra	90	-10.9	1.2	4.0

ear, was now set at 65 dB(A), to prevent threshold effects in the conditions where one channel was attenuated by 20 dB. As in experiment I, normal and reverse headphone positionings were alternated between subjects.

2.4 B. Results

In Table II, means and standard deviations of all measured speech reception thresholds are presented, as well as the BILD with respect to the result for $\Phi=0°$. An overall analysis of variance showed that 82.3% of the variance is contained in the conditions, and 2.2% in the subjects. The analysis yielded an error estimate of 1.5 dB. The t tests on paired observations indicate that all BILDs are significant ($p=0.007$ for the 0° monaural condition, $p<0.0005$ for all others). Pooling the subjects, the results for the five duplicated conditions were compared with those obtained in experiment I. An analysis of variance showed that the results from both experiments are highly similar ($p=0.996$).

Mean SRTs for CT and dT noise are displayed in Fig. 6. The BILD for CT noise is seen to increase as a function of ITD, but not proportionally. The results for dT noise, plotted as horizontal lines that indicate the ITD range over frequency (cf. Fig. 3), correspond well with those for CT noise on the high-ITD (low-frequency) side. Thus the results indicate that there is no significant extra unmasking due to the dependence of ITD on frequency.

Mean SRTs obtained for the remaining conditions are shown in Fig. 7. The points with an abscissa value of 0 dB represent the conditions, which were also part of experiment I. Plotted to the right and left of these points are results for conditions where the channel with the ipsilateral and contralateral noise, respectively, was attenuated. In the case of noise with ITD, this corresponds to the channels with the leading and lagging noise, respectively. The monaural conditions are at the extremes. The noise azimuth is 90° in all cases except for the binaural and monaural conditions with 0° FF noise (indicated by closed circles).

Focusing first on the results for dT noise, we see that the unmasking is not strongly affected by the 20-dB attenuation. It drops by only 1.7 dB and 0.7 dB for attenuation at the side with the leading and lagging noise, respectively. The difference of 1 dB between both results is statistically significant [$F(4,64)=9.27$, $p<0.008$]. Since both channels of the dT noise have the same power spectrum as the 0° FF noise, the SRT converges for increasing attenuation towards the monaural result for 0° FF noise.

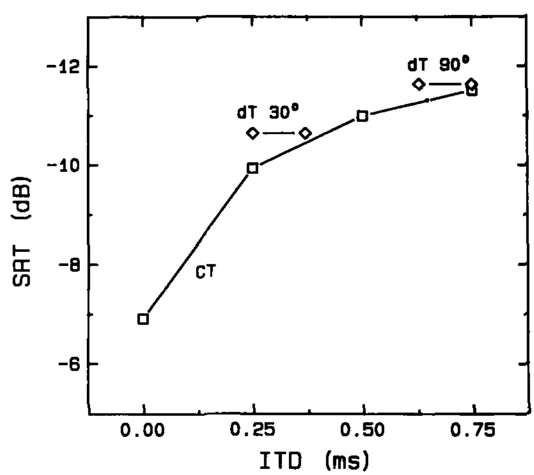

FIG. 6. Mean speech-reception thresholds obtained in experiment II for CT noise (fixed ITD) and dT noise (frequency-dependent ITD introduced by KEMAR). The latter results are plotted as lines indicating the ITD range.

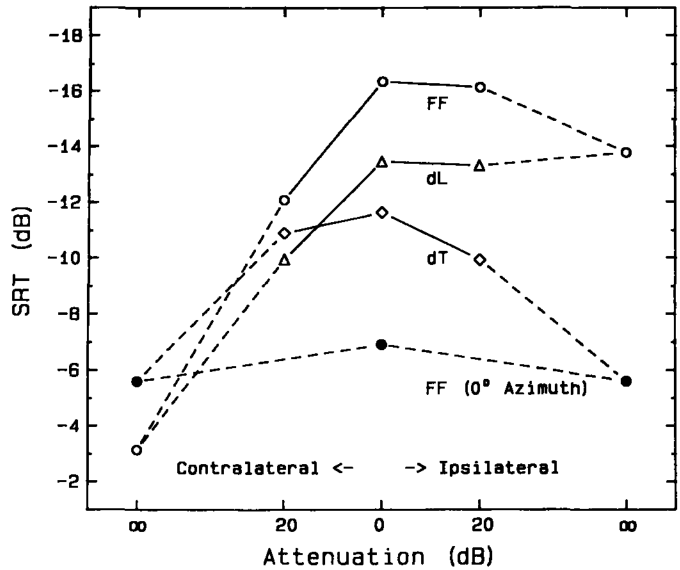

FIG. 7. Mean speech-reception thresholds obtained in experiment II for the monaural conditions, the conditions with an overall 20-dB attenuation in one channel, and the 0° and 90° conditions, which were also part of experiment I. The data points in the right and left halves of the figure were obtained for attenuation of the channel with the ipsilateral and contralateral noise, respectively. The open symbols corespond to the 90° noise azimuth, the closed symbols to the 0° azimuth.

The SRTs obtained for dL noise indicate that attenuation of the channel with the ipsilateral noise (the channel with the worst S/N ratio) does not result in a significant threshold shift. This is supported by an analysis of variance, applied to the monaural condition and the conditions with 0- and 20-dB attenuation [$F(2,32)=0.31$, $p=0.74$]. Thus the intelligibility level difference for 90° dL noise is not affected by binaural unmasking, but relies on the ear receiving the signal with the best S/N ratio. However, the data show that a 20-dB attenuation of this signal with respect to the signal for the other ear increases the SRT by 3.5 dB. This cannot be caused by a floor effect, since the speech level on the attenuated side was still 35 dB(A) in this condition, well above the threshold in quiet of approximately 15 dB(A) (Duquesnoy, 1982), nor by cross-hearing, since the maximum ILD plus 20-dB attenuation is still about 20 dB lower than the acoustic cross-talk between both ears (Zwislocki, 1953). Therefore, it seems that the shift in SRT is due to a "disturbing" effect of the relatively loud noise presented to the other ear. In the limiting monaural case with the ipsilateral noise the mean SRT is only -3.1 dB. The difference of 13.2 dB between this value and the result for FF noise and 0-dB attenuation represents the maximum advantage of binaural over monaural listening for the chosen configuration.

Comparison of the results obtained for dL noise and FF noise shows that the binaural unmasking due to ITD in the latter case yields a gain varying between 2.1 dB and 2.9 dB. Thus we find here, just as for dT noise, that the overall attenuation of one channel has but a limited effect on binaural unmasking.

2.5 DISCUSSION

Evidently, an important question relating to the interpretation of our results is whether our simulation of the free-field condition is accurate enough. The analysis of the noise samples (see Sec. 2.2) already indicated that the ITD and headshadow present in the recorded noise agree well with measurements on individuals, reported in the literature. This is confirmed by our experimental results for FF noise, which show a close correspondence with data obtained by Plomp and Mimpen (1981) in a free field using the same sentences and noise. Only at backward angles is there a 1- to 2-dB discrepancy, which is slightly greater than the intersubject variability. As already mentioned in Sec. 2.2, the headshadow introduced by KEMAR at these angles is small compared with measurements of Blauert (1974). This must be due to a difference in dimensions of head and pinna between KEMAR and the 25 subjects used by Blauert. Most of KEMAR's dimensions were chosen between the average values for males and females. Interestingly, the "tragion to wall" length and the ear protusion, which both affect the response at backward angles, are relatively small (Burkhard and Sachs, 1975). Since lesser shadowing causes greater SRT values, this might well explain the discrepancy between Plomp and Mimpen's and our results. In view of the magnitude of the discrepancy, it seems safe to conclude that the free-field condition is indeed simulated accurately enough. The question remains however, to which degree the headshadow introduced by KEMAR at backward angles deviates from the response of the average human head.

In the considerations concerning the advantage of binaural over monaural hearing, initially many thought that ITD was the primary cue for the binaural system. The results of Carhart *et al.* (1967) showed that ITD yields, in fact, only a moderate contribution when considering the gain in intelligibility. Larger unmasking is accomplished when only detection, instead of discrimination, is required. This was demonstrated by Levitt and Rabiner (1967a), who found differences in unmasking up to 10 dB between both listening tasks. In addition, it was suggested by Carhart *et al.* that the ILDs caused by headshadow have a degrading effect on binaural unmasking due to ITD. This is supported by the results of our first experiment: when there is no ILD, the BILD due to ITD varies from 3.9-5.1 dB; with ILDs due to headshadow in the noise masker, the ITD effect is limited to values between 2.1 and 3.4 dB. Comparison of these figures with the gain due to ILD alone, which ranges from 3.5-7.8 dB, shows that headshadow is the major determinant of the binaural advantage in a free field (in the case of a single noise source).

In experiment II, it was shown for 90° dL noise that the gain due to ILD relies on the ear presented with the most favorable S/N ratio. We assume that this is also the case for the other noise azimuths, and that the dL curve in Fig. 5 in fact represents the "best ear" curve. Thus our results can be compared with those of other authors, who determined both binaural and "best ear" performance. The difference is sometimes referred to as the "squelch" effect, a term first used by Koenig (1950). Measurements of binaural versus monaural speech intelligibility against a background of noise or other masking sound have been reported by Hirsh (1950), Lochner and Burger (1961), Nordlund and Fritzell (1963), Carhart (1965), Dirks and Wilson (1969a), McKeith and Coles (1971) and Platte (1980). Unfortunately, these studies show a great variety in material, method and experimental configuration. An important aspect is the method used to create the monaural conditions. As shown by McKeith and Coles, the results for these conditions can be affected considerably by incomplete occlusion of the other ear. Some of the authors used artificial heads, and, therefore, had no difficulty in eliminating one ear. However, in these cases, the results depend on the quality of the artificial head.

Reviewing the data, it appears that the results of Hirsh are difficult to interpret, because he did not specify which ear was used in the monaural conditions. It is remarkable that he reports superiority of monaural over binaural listening in a number of conditions, an effect which is not confirmed in any of the later studies. Nordlund and Fritzell only measured percentages of correct responses at a single S/N ratio, which prohibits a reliable comparison with our BILD values. Lochner and Burger reported dB values, but these were calculated from percentage values using a fixed slope factor. In addition, their data show a consistent difference between binaural advantages obtained with the speech source on the left and right sides of the listener, respectively. The results of the remaining authors indicate that, for those conditions where speech and masking sound originate from different directions, the magnitude of the squelch effect mainly depends on the speech material, and ranges from approximately 2-9 dB. The lowest values are obtained with monosyllabic words (1.9-3.7 dB), the use of spondaic words yields intermediate values (4-7 dB), just as closed sets of synthetic sentences did (4.5-7.5 dB), and the highest values (5-9 dB) were measured using the criterion of 20% correctly identified numbers chosen from a limited set. This corresponds with results of Schubert and Schultz (1962) and Levitt and Rabiner (1967a,b), who showed that binaural unmasking of time-delayed or phase-inverted speech in noise depends on the low-frequency content of the speech signals and increases when the listening task is simplified. In our experiments, we used simple sentences, but required correct reproduction of every word. Since our results are in the same range as when monosyllabic words are used, it seems that the information in the high frequencies is of comparable importance in both tasks.

Our results for CT noise show that the BILD increases as a function of ITD and converges towards a maximum of about 5 dB. The shape of the curve agrees with data of Schubert (1956) and Levitt and Rabiner (1969a), who showed that time delays exceeding 0.5-1 ms yield little or no additional gain. The comparison with the results for dT noise indicates that, in the case of a (head-induced) frequency-dependent ITD, the BILD is determined by the ITD for low frequencies (in Fig. 6, only data from experiment II are plotted, but the results for dT noise from experiment I also support this conclusion). This agrees with the earlier findings, mentioned above, that unmasking of speech in noise mainly depends on the interaural differences in the low-frequency area. Related to this is the dependence of the BILD for time-delayed speech or noise upon speech material, which is similar to that of the squelch effect, discussed above. According to the data given by Levitt and Rabiner and Carhart *et al.* (1967), unmasking for monosyllabic words (2.5-3.3 dB) is less than our maximum BILD of 4.6 dB, which lies in the same range as results for spondees (4.1-6 dB). Much higher values (9-12 dB) are obtained when the listening task changes from discrimination to detection (Kock, 1950; Levitt and Rabiner, 1967a). A possible explanation for the fact that our BILDs now compare favorably with values obtained using monosyllabic words, is the difference in maskers: we used speech noise, the above authors used white noise, which has relatively more energy in the high frequencies.

The results obtained in experiment II show that a 20-dB attenuation of either channel has but a limited effect on the binaural gain due to ITD. This applies to both the conditions with and without headshadow. It appears that hardly any data are available in the literature on this subject. Wilson *et al.* (1985) measured the effect of the overall attenuation of one ear on binaural unmasking of phase-inverted speech in noise. They used groups of subjects with normal hearing or with an asymmetrical sensorineural hearing loss. The results for young normal-hearing subjects indicate a slightly greater decrease of unmasking: 2.8 dB for 20-dB attenuation (by interpolation), compared to maximally 1.7 dB in our case. The MLD values reported by McFadden (1968), obtained with phase-inverted low-frequency tones in noise, show a decrease of approximately 3 dB for 20-dB attenuation.

When considering the binaural gain in everyday listening it has to be kept in mind that our results (and those of a number of others) were obtained in a free-field condition and do not directly apply to situations with reverberation. The indirect sound affects binaural performance and reduces the difference between near-ear and far-ear listening (MacKeith and Coles, 1971). An example is the maximum advantage of binaural over monaural hearing, which is 13.2 dB according to our data. In the limiting case of a listener in the indirect sound field of the noise source, the noise signals at both ears are almost uncorrelated and the binaural gain for homophasic speech is only about 3 dB (Koenig *et al.*, 1977). Another aspect is the number of maskers. Carhart *et al.* (1969a) demonstrated (using signals presented over headphones) that multiple maskers can cause considerable unmasking, and that the BILD increases slightly as a function of the number of maskers. A condition with an interesting analog in a free field situation is that of homophasic speech in two maskers with opposed time delays. In this condition, BILDs were obtained ranging from 2.5-6.4 dB, depending on the noise type. Thus, with two masker sources located at opposite sides of the head, a binaural gain is to be expected, which is less affected by headshadow than in the case of a single masker.

ACKNOWLEDGEMENTS

We wish to thank the Application Engineering group of Philips Hearing Instuments, Eindhoven, the Netherlands for the use of their KEMAR manikin and other equipment. We further thank Dr. J.M. Festen and Dr. T.S. Kapteyn for their support and cooperation.

3. BINAURAL SPEECH INTELLIGIBILITY IN NOISE FOR HEARING-IMPAIRED LISTENERS

A.W. Bronkhorst and R. Plomp
Journal of the Acoustical Society of America 86, 1374-1383 (1989).

ABSTRACT

The effect of head-induced interaural time delay (ITD) and interaural level differences (ILD) on binaural speech intelligibility in noise was studied for listeners with symmetrical and asymmetrical sensorineural hearing losses. The material, recorded with a KEMAR manikin in an anechoic room, consisted of speech, presented from the front (0°), and noise, presented at azimuths of 0°, 30° and 90°. Derived noise signals, containing either only ITD or only ILD, were generated using a computer. For both groups of subjects, speech-reception thresholds (SRT) for sentences in noise were determined as a function of (1) noise azimuth, (2) binaural cue, and (3) an interaural difference in overall presentation level, simulating the effect of a monaural hearing aid. Comparison of the mean results with corresponding data obtained previously from normal-hearing listeners shows that the hearing impaired have a 2.5-dB higher SRT in noise when both speech and noise are presented from the front, and 2.6-5.1 dB less binaural gain when the noise azimuth is changed from 0° to 90°. The gain due to ILD varies among the hearing-impaired listeners between 0 dB and normal values of 7 dB or more. It depends on the high-frequency hearing loss at the side presented with the most favorable S/N ratio. The gain due to ITD is nearly normal for the symmetrically impaired (4.2 dB, compared with 4.7 dB for the normal hearing), but only 2.5 dB in the case of asymmetrical impairment. When ITD is introduced in noise already containing ILD, the resulting gain is 2-2.5 dB for all groups. The only marked effect of the interaural difference in overall presentation level is a reduction of the gain due to ILD when the level at the ear with the better S/N ratio is decreased. This implies that an optimal monaural hearing aid (with a moderate gain) will hardly interfere with unmasking through ITD, while it may increase the gain due to ILD by preventing or diminishing threshold effects.

3.1 INTRODUCTION

In everyday listening, there are numerous occasions when speech intelligibility is reduced due to the presence of background noise. It was shown by a number of investigators that, in such situations, there is a superiority of binaural over monaural listening, especially when speech and noise originate from different directions (Carhart, 1965; Dirks and Wilson, 1969a; MacKeith and Coles, 1971). The binaural advantage, which can be expressed as a shift of signal-to-noise (S/N) ratio for a certain constant level of intelligibility, mainly results from interaural differences in time and level (ITD and ILD), caused by the diffraction of incoming sound waves by the head. A remarkable aspect of the above studies is the fact that the authors often used configurations, in which the speech was presented from one side, and the noise from the opposite side. Since the study of binaural speech intelligibility in noise has many practical implications, for instance for the fitting of hearing aids, it seems logical to focus the research on the situation with the speaker in front of the listener, which is most natural in everyday listening. Furthermore, it is difficult to elucidate the dependence of the binaural advantage on ITD and ILD when one uses conditions in which both speech and noise contain binaural cues.

Plomp and Mimpen (1981) simulated a realistic situation by using speech (consisting of short conversational sentences), reproduced in front of the listener, and speech noise, reproduced at various azimuths in the horizontal plane. Their results, obtained from 10 normal-hearing listeners, show that the binaural gain, relative to the (binaural) condition with coinciding speech and noise sources, increases to about 10 dB when the noise is presented from the side. In our previous experiments (Bronkhorst and Plomp, 1988), we have duplicated their measurements using material recorded with a KEMAR manikin (Burkhard and Sachs, 1975) in a free field, and presented through headphones to normal-hearing listeners. By processing these recordings with a computer, we were able to investigate the separate contributions of ITD and ILD to the binaural gain. The results showed that ILD has a greater effect than ITD, and that these two effects are not additive. For instance, for noise presented from the side, we found the same 10-dB binaural gain as Plomp and Mimpen did, while ILD alone yielded about 7 dB, and ITD alone about 5 dB. In addition, the measurements indicated that the ILD effect involves no binaural interaction, since the results for noise with only ILD did not differ from the monaural results at the side with the most favorable S/N ratio. We have further simulated a one-sided hearing loss by including conditions in which the total stimulus was attenuated by 20 dB at one ear. It turned out that the release from masking due to ITD is hardly affected by such an attenuation, but that the gain due to ILD decreases significantly when the level at the side with the better S/N ratio is reduced.

In the present study, we have used the same KEMAR recordings to investigate binaural speech intelligibility in noise for hearing-impaired listeners. We have tested subjects with symmetrical and asymmetrical cochlear hearing losses in two separate experiments. For both groups, we have studied the effects of ILD and

ITD, and we have determined the sensitivity of the binaural advantage to interaural differences in presentation level. The latter condition now simulated the effect of a monaural hearing aid. This is of special interest in the case of asymmetrical impairment, since it might be expected that the binaural gain is enhanced by matching the presentation levels to the degree of asymmetry.

Past research on binaural speech intelligibility in noise for hearing-impaired listeners was mainly focused on clinical application as site-of-lesion test (Olsen *et al.*, 1976; Bocca and Antonelli, 1976; Stubblefield and Goldstein, 1977), on the determination of the benefit of binaural hearing aids (Wright and Carhart, 1960; Dirks and Wilson, 1969b; Zelnick, 1970; Markides, 1977; Festen and Plomp, 1986) or on age-related effects (Tillman *et al.*, 1973; Duquesnoy, 1983; Gelfland *et al.*, 1988). Exceptions to this categorisation are the work by Tonning (reviewed by Lidén, 1976), who investigated the binaural performance of several groups of hearing-impaired listeners with and without hearing aids, and the recent study by Wilson *et al.* (1985), who measured the binaural release from masking as a function of the interaural difference in overall presentation level. The cues used by the above authors to invoke a binaural gain were ITD, phase inversion, or a combination of ITD and ILD, occurring when the stimuli were reproduced in a sound-treated environment. The studies show a great variety in method, material, and degree and type of hearing-impairment of the subjects. It is a pity that most authors presented only average results, and no analysis of the relationship between the binaural gain and the degree of hearing loss. Furthermore, due to the choice of subjects, the results often do not permit separation of the effects of age and hearing loss, and asymmetry and hearing loss (since most asymmetrical losses were in fact unilateral). It is therefore difficult to get from these studies a clear picture of the effect of cochlear hearing-impairment on the binaural gain, occurring when listening to speech in a noisy environment. This is of course partly due to the fact that most studies had other aims.

The relative lack of available data was a further motivation for our study, which was set up as an extension of our work on normal-hearing subjects. We were especially careful in our subject selection, to prevent the ambiguities mentioned above. We used an upper age limit of approximately 50 years, to minimize age-related effects. For our experiment with asymmetrically impaired listeners, we selected subjects who also had a hearing loss at their better ear. The conditions in our experiments were designed to provide answers to the following questions:

(1) How do the SRTs for monotic and diotic stimuli depend on the degree and type of hearing loss? These results can be compared with a large body of data collected by Plomp and co-workers (Plomp, 1986).

(2) How large are the separate contributions of ITD and ILD to the binaural gain and how do they change as a function of the degree and asymmetry of the hearing loss? Results of the studies mentioned above indicate that the hearing impaired benefit less from ILD than the normal hearing, but almost equally from ITD. Furthermore, it seems that binaural unmasking is less affected by symmetrical, than by asymmetrical impairment.

(3) When the binaural gain is determined for listeners with an asymmetrical hearing loss using time-delayed stimuli, do the results depend on whether the good or bad ear is leading? Bocca and Antonelli (1976) found such a difference for listeners with unilateral sensorineural impairment.

(4) What is the effect of an interaural difference in overall presentation level on the binaural gain? Is it relatively small, as in normal-hearing listeners, and can it be beneficial for the asymmetrically impaired? The latter is of interest in view of the results of Wilson *et al.* (1985), who found that matching the presentation levels by lowering the level at the better ear does not improve binaural unmasking. Since these conditions mimick the effect of a monaural hearing aid, the results might have interesting implications for hearing-aid fitting.

3.2 MATERIAL AND METHOD

An extensive description of the preparation of the speech and noise material can be found in in Bronkhorst and Plomp (1988). Here, only a short recapitulation will be given. The speech material consisted of 130 short meaningful sentences, read by a female speaker and adjusted in level for equal intelligibility (Plomp and Mimpen, 1979). The noise had a spectrum shaped according to the long-term average spectrum of the sentences. Speech and noise were reproduced in an anechoic room and recorded using a KEMAR manikin fitted with 2 Brüel & Kjær (B&K) 4157 ear simulators, placed at a distance of 1.5 m from the loudspeaker. The speech was recorded at an azimuth of 0° (frontal), the noise at 7 azimuths ranging from 0° to 180° in steps of 30° (in the present experiments, we only used the noise signals recorded at 0°, 30° and 90°). An equalization filter described by Killion (1979) was used to compensate for the ear-canal resonance introduced by the ear simulator. The ILD and ITD, present in the noise signals, were determined by digitizing the signals and applying FFT and cross-correlation techniques. The results of these calculations are shown in Bronkhorst and Plomp (1988), Figs. 2 and 3. Using these data, noise with only ITD and no ILD (dT noise) was obtained from the recorded noise (hence designated as free-field, or FF noise) by making the Fourier transforms of both channels equal to the 0° spectrum and by subsequently applying an inverse transformation. In a similar manner, noise with only ILD and no ITD (dL noise) was generated from the 0° noise by modifying the spectra according to the ILD data for the lateral angles.

For practical purposes, we used the same division the sentences into 17 lists of 8 sentences (with 6 sentences duplicated), as in the previous experiments. We thus had 17 conditions per experiment, which were presented to 17 subjects according to a Latin-square design. For each subject, a copy of the sentence lists was made on tape and the noise signals corresponding to the different conditions were added on separate tracks in the order prescribed by the Latin square. The experiments were performed in a double-walled soundproof booth. The outputs of the tape deck were fed into two Madsen OB 822 clinical audiometers, connected to a set of Beyer DT-48 headphones. The setting of both audiometers was remotely controlled by a computer. Speech-reception thresholds were determined by varying the presentation level of the sentences using a simple up-down procedure with the correct replication of the entire sentence as criterion (Plomp and Mimpen, 1979). The first sentence of a list was presented initially below threshold, and then raised in level in steps of 4 dB until it was reproduced correctly. The following seven sentences were presented only once, at a level that was lowered by 2 dB after a correct response, and raised by 2 dB after an incorrect response. The SRT was taken as the average presentation level after the second sentence.

The presentation level was calibrated according to the dBA level of the 0° noise, as measured with a B&K 4152 artificial ear. In the previous experiments, we had chosen fixed noise levels of 60 and 65 dBA, respectively. In the present experiments, we would have been forced to accept rather high levels, to prevent threshold effects in the conditions in which one channel is attenuated. For subjects with light hearing losses, such levels would have been unnaturally, and sometimes even uncomfortably loud. Therefore, we decided to vary the noise level according to the degree of hearing loss of the individual subject. However, even with a variable level, threshold effects could not be completely avoided, due to the limited maximum output level of our equipment and the reduced hearing range of some subjects.

Before the presentation of the sentence lists, which took about 50 minutes (including a short break), a pure-tone audiogram was determined using a modified Békésy procedure operating at octave frequencies ranging from 250 to 8000 Hz. Tone bursts with a duration of 200 ms were presented at a rate of approximately one per second and varied in level in steps of 2 dB. The threshold was taken as the average level at eight reversals; the first two reversals were ignored. The stimuli were calibrated according to ISO 389.

3.3 EXPERIMENT I: SYMMETRICAL IMPAIRMENT

3.3 A. Conditions

In order to limit the number of conditions, we only used three noise azimuths: 0° (frontal), 30° (only the dT noise) and 90°. Since the results for the normal-hearing subjects (Bronkhorst and Plomp, 1988, Fig. 5) show a smooth dependence on noise azimuth, which is almost symmetrical around 90°, we decided that inclusion of more azimuths would not add significant information. The result for 30° dT noise will be of interest since the binaural unmasking for the normal hearing increases sharply between 0° and 30°. The 90° FF, dL and dT noises were also presented with a 20-dB attenuation of the ipsilateral (corresponding to the ipsilateral ear of KEMAR during the recordings) or the contralateral channel. In addition, the monaural SRTs in 0° and 90° noise and in quiet were determined. The latter conditions served to check for threshold effects and to indicate whether a subject had abnormally poor speech discrimination.

3.3 B. Subjects

Seventeen subjects with symmetrical sensorineural hearing losses were selected from the files of the Audiology department. In all cases, there was no indication of middle-ear disease and the absence of an air-bone gap had been confirmed by bone-conduction audiometry. Four subjects had a congenital cochlear impairment; the remaining had cochlear hearing losses of unknown etiology. There were eight male and nine female subjects and the age range was 18-45 years, with a median age of 36.5 years. Over all 34 ears, the pure-tone average (PTA, average over 500, 1000 and 2000 Hz) of the hearing levels ranged from 24 to 48.7 dB, the mean PTA being 37.4 dB. The interaural difference in PTA was less than 6.3 dB and had a mean absolute value of 2 dB. The reference noise level used during the presentation of the sentence lists was set for each subject at a fixed level between 75 and 95 dBA. It was chosen according to the measured hearing levels of the subject and the uncomfortable loudness levels, which had been determined previously during routine audiometry. Normal and

Table I. Results obtained in experiment I. For each condition, the mean SRT in noise, expressed as S/N ratio, and the standard deviation are given. The BILDs in the last column were obtained by subtracting the mean SRTs from the mean result for 0° FF noise. In the description of the noise type, "ipsi" refers to the channel originally recorded on the side of KEMAR, turned towards the noise source. "Contra" refers to the other channel.

Noise Type	Noise Azimuth (°)	SRT (dB) mean	SRT (dB) s.d.	BILD (dB) *re:* 0° azimuth
FF	0	-3.9	1.5	
FF monaural ipsi	0	-2.5	1.7	-1.4
FF monaural contra	0	-2.9	1.5	-1.0
FF	90	-11.0	3.0	7.1
FF 20-dB att. ipsi	90	-10.5	2.8	6.6
FF 20-dB att. contra	90	-7.6	2.5	3.7
FF monaural contra	90	-7.9	2.7	4.1
FF monaural ipsi	90	0.5	2.4	-4.4
dL	90	-8.5	2.9	4.6
dL 20-dB att. ipsi	90	-8.2	2.7	4.4
dL 20-dB att. contra	90	-5.3	3.2	1.4
dT	30	-6.4	1.5	2.5
dT	90	-8.1	2.3	4.2
dT 20-dB att. ipsi	90	-7.2	1.8	3.3
dT 20-dB att. contra	90	-6.8	1.9	2.9

reverse earphone positionings were alternated between subjects.

3.3 C. Results

The adaptive procedure for the determination of the SRT converged for all subjects in all conditions. In no case was the SRT in quiet abnormally elevated with respect to the pure-tone hearing levels. This would have indicated a severe discrimination loss and would have been a reason for exclusion of the subject from the experimental group. In Table I, the means and standard deviations of the measured SRTs in noise are presented, as well as the binaural intelligibility level differences (BILD) with respect to the result for 0° noise. In the conditions with a 20-dB attenuation of one ear, two subjects had SRTs in noise close to their speech thresholds in quiet. We decided against exclusion of these data, since this turned out to have only a marginal effect on the outcome of the statistical calculations. An analysis of variance, applied to the SRTs in noise, showed that 62.9 % of the variance is contained in the conditions, and 15.4 % in the subjects. The analysis yielded an error estimate of 1.8 dB. The t tests on paired observations indicated that all BILDs are significant ($p<0.0002$) except for the

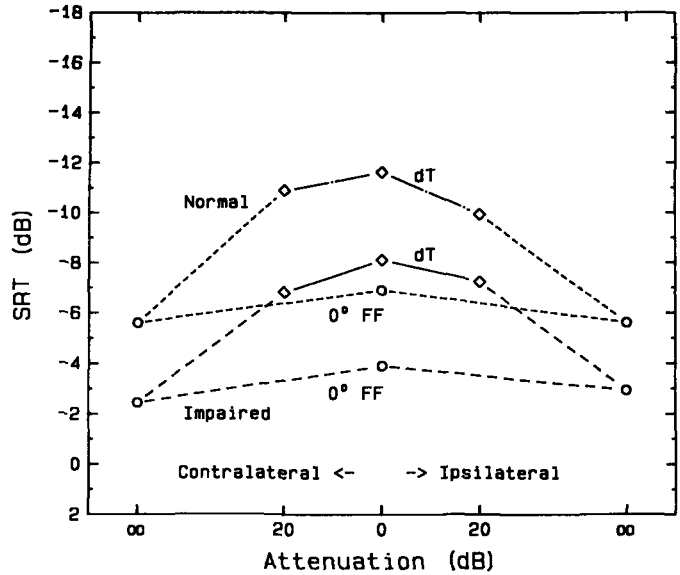

Fig. 1. Mean SRTs for 0° FF noise and for 90° dT noise, obtained in experiment I from listeners with symmetrical hearing-impairment, and the corresponding results for normal-hearing listeners from Bronkhorst and Plomp (1988). The SRTs are expressed as S/N ratio. In the right and left halves of the figure are the results for attenuation of the channel with the ipsilateral and contralateral (leading and lagging) noise, respectively.

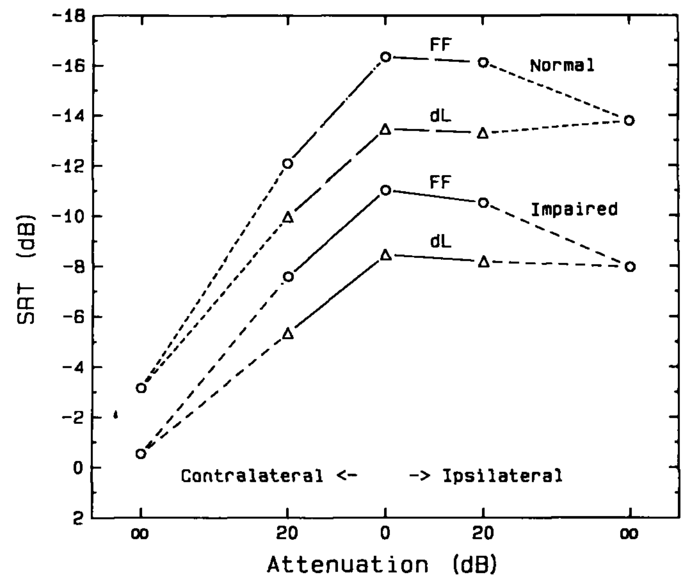

Fig. 2. A similar plot as Fig. 1, but now showing the results for 90° dL and FF noise.

condition with 90° dL noise and attenuation of the contralateral channel (p=0.11).

The mean SRTs for 0° FF and 90° dT noise are plotted in Fig. 1 together with the corresponding results for 17 normal-hearing listeners, obtained previously (Bronkhorst and Plomp, 1988). In the right half of the figure are the results for the conditions where the the channel with the ipsilateral noise was attenuated (the noise recorded originally at the side of KEMAR, closest to the noise source). In the left half are the points, obtained for attenuation of the other channel. The results show that the hearing impaired have clearly higher SRTs in noise than the normal hearing. The mean SRT differences range from 2.7 to 4.1 dB and are all significant (two-sample t tests, p<0.0001). However, when we consider the relative gain due to ITD, it appears that there is hardly any difference in performance between both groups. In the condition with equal presentation levels at both ears, the hearing impaired have a mean BILD of 4.2 dB, which does not differ significantly from the value of 4.7 dB, obtained from the normal hearing (two-sample t test, one-sided, p=0.26). Reducing the overall presentation level on either side has only a limited effect on the SRT: it increases by 0.9 and 1.3 dB for attenuation of the ipsi- and contralateral channel, respectively. Only the latter difference is statistically significant (t test for paired observations, p=0.02).

In Fig. 2, mean SRTs for 90° dL and FF noise are displayed in a similar plot.

The differences with the corresponding results for normal-hearing listeners now range from 3.7 to 5.8 dB. Again, two-sample *t* tests showed strong significance of all differences ($p<0.0001$). Comparison of the mean BILDs for dL noise, obtained from both groups, shows that the hearing impaired benefit less from ILD than the normal hearing. For equal presentation levels at both ears, the BILDs are 4.6 and 6.5 dB, respectively. However, due to the large spread in the data of the hearing impaired, the 1.9-dB difference between these values has only limited significance (two-sample *t* test, one-sided, $p=0.03$). It is demonstrated below that this spread is related to a dependence of the dL and FF results on high-frequency hearing loss. The results further show that the effect of interaural differences in overall presentation level is similar for both groups. The gain due to ILD is not affected when the channel with the worst S/N ratio (the ipsilateral channel) is attenuated or switched off, but it decreases when the level of the other channel is reduced. The difference between the dL and FF results, i.e. the unmasking due to ITD in the presence of ILD, is about 2.5 dB for both groups (2 dB less than the effect of ITD alone) and it is hardly sensitive to a 20-dB attenuation of either channel.

Table I shows that the mean BILDs for 30° and 90° dT noise are 2.5 dB and 4.2 dB. The corresponding results for the normal hearing are 3.7 and 4.7 dB, respectively. The 1.2-dB difference between the results for 30° noise is significant (two-sample *t* test, one-sided, $p=0.02$). Thus, it seems that, for the hearing impaired, there is a less steep dependence of unmasking (due to ITD) on noise azimuth between 0° and 30° even when we take into account that their release

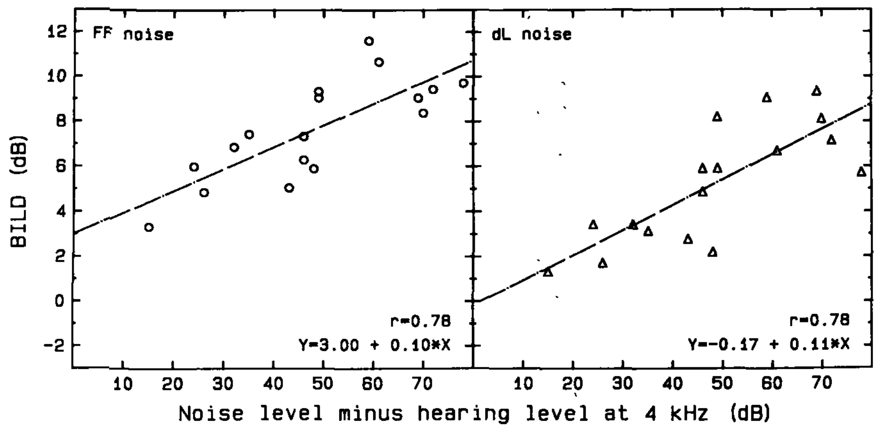

Fig. 3. Binaural gain *re*: 0° SRT for 90° FF noise (left panel) and 90° dL noise (right panel) as a function of the difference between the dBA noise level and the hearing level at 4000 Hz. The plot shows the data points, representing individual results for the 17 subjects of experiment I, and the regression lines.

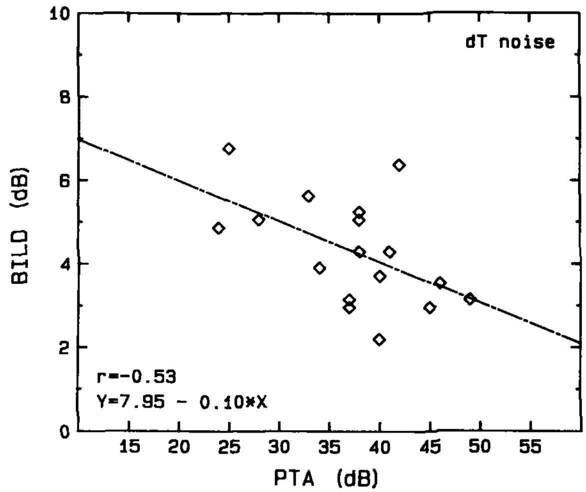

Fig. 4. The BILDs for 90° dT noise, obtained from the subjects of experiment I, and the regression line, both plotted as a function of the PTA.

from masking at 90° is somewhat smaller. Analysis of the individual data shows that the BILDs for 30° and 90° dT noise are strongly correlated in the hearing-impaired group (r=0.79, p<0.001), but not correlated at all in the normal-hearing group (r=0.21, p=0.4). This indicates that the ability to benefit from ITD varies consistently within the hearing-impaired group, which suggests that it is related with the degree or type of hearing impairment.

The relationship between the speech thresholds and binaural gains in noise on one hand, and hearing loss and age on the other, was investigated using multiple regression analysis. Four separate calculations were carried out with as dependent variables: the SRT for 0° FF noise and the BILDs for 90° FF, dL and dT noise. To reduce the spread of the data, the average result of two or three conditions, with similar mean outcome, was taken as input for the calculations. As 0° FF result, we used the average of the binaural and both monaural SRTs. For the 90° FF and dL noise types, the SRTs obtained with 0- and 20-dB ipsilateral attenuation were averaged, and the result was subtracted from the 0° FF result. The 90° dT BILD was obtained by subtracting the average of all 3 SRTs (with and without 20-dB attenuation) from the 0° FF result. As independent variables, we used the hearing levels at 250, 500, 1000, 2000 and 4000 Hz, the PTA, and the age. The PTA and the hearing levels were averaged over both ears. We did not include the 8000 Hz threshold, since it would have affected the statistics due to its large variance, while it can be safely assumed that the speech contains no significant information in that frequency region. The analysis used was a stepwise multiple-regression technique, which isolates the independent variables with the largest regression coefficients by testing the corresponding F ratio. The results showed that the 0° FF result does not depend significantly on any of the independent variables. Furthermore, the analysis yielded no significant dependence of any of the BILDs on age. Significant correlations were found only between the BILDs for FF and dL noise and the hearing level at 4000 Hz ($p<0.001$), and between the BILD for dT noise and the PTA ($p=0.03$). The FF and dL data are shown in Fig. 3, together with the regression lines. Since the dependence is most probably a threshold effect (this is explained in detail in the discussion, Sec. IV), the data are plotted as a function of the difference between the A-weighted reference presentation level and the hearing level at 4000 Hz, to compensate for the interindividual differences in presentation level. In the interpretation of these results, it has to be taken into account that the noise (and speech) level in the 4000 Hz region is about 15 dB lower than the dBA level of the total signal (cf. Bronkhorst and Plomp, 1988, Fig. 1). In Fig. 4, the individual BILDs for dT noise and the regression line are shown as a function of the PTA.

A separate calculation was performed to relate the measured SRTs in quiet to the pure-tone hearing levels. All data were averaged over both ears. The speech thresholds turned out to correlate highly with the PTA ($r=0.81$, $p<0.001$). The following regression equation was found:

$$SRT_q = 8.4 + 1.0\ (PTA).$$

3.4 EXPERIMENT II: ASYMMETRICAL IMPAIRMENT

3.4 A. Conditions

In the second experiment, we again used the 90° FF, dL and dT noise samples, which were now presented four times each: with the ipsilateral (leading and/or loudest) channel at the good or poor ear, and with equal or matched presentation levels. In the latter case, the level was raised at the poor ear and lowered at the good ear by half of the interaural difference in PTA. The remaining five conditions were measurements of the monaural SRT in 0° FF noise and in quiet, and the binaural SRT in 0° FF noise.

3.4 B. Subjects

As in experiment I, 17 subjects were chosen from the patient population of the Audiology department. They had asymmetrical sensorineural hearing losses, confirmed by air- and bone-conduction audiometry, and no indication of middle-ear or retrocochlear pathology. Twelve subjects had a possibly hereditary impairment of unknown etiology, two had Meniere's disease, and three had suffered from sudden deafness. The subjects, 6 male and 11 female, were aged 21-52 years and had a median age of 41.9 years. The PTA of the hearing levels ranged from 9.7 to 45 dB (mean 21.7 dB) for the good ear, and from 26 to 53.3 dB (mean 40.5 dB) for the poor ear. The interaural difference in PTA varied between 4.7 and 31.3 dB. For each subject, the reference noise level was chosen between 70 and 80 dBA, depending on the measured hearing levels and uncomfortable loudness levels.

3.4 C. Results

As in experiment I, the adaptive procedure converged in all cases, and none of the subjects showed abnormally poor speech discrimination at either ear. Means and standard deviations of the obtained SRTs in noise are shown in Table II. In the last column, the BILDs with respect to the binaural result for 0° FF noise are given. For two subjects, some of the measured SRTs in noise were close to the SRT in quiet of the poor ear. Due to the minimal effect on the results of the statistical calculations, these data were not excluded. According

Table II. Results obtained in experiment II. The mean SRTs in noise (expressed as S/N ratio) and the standard deviations are given, as well as the BILDs *re:* 0° azimuth. For the 90° conditions, the orientation of the good ear is given with respect to the position of the noise source during the KEMAR recordings. The signals at both ears were presented either at equal levels, or at levels matched to the asymmetry of the hearing loss.

Noise type	Method of presentation	SRT (dB) mean	SRT (dB) s.d.	BILD (dB) *re:* 0° azimuth
0° FF	Binaural	-4.4	1.5	
0° FF	Good ear	-3.4	1.7	-1.0
0° FF	Poor ear	-1.4	3.0	-3.0
90° FF	Good ear ipsi, equal levels	-9.1	2.4	4.7
90° FF	Good ear ipsi, matched levels	-9.4	3.3	5.0
90° FF	Good ear contra, equal levels	-11.5	2.6	7.2
90° FF	Good ear contra, matched levels	-10.2	2.7	5.9
90° dL	Good ear ipsi, equal levels	-7.2	3.0	2.8
90° dL	Good ear ipsi, matched levels	-7.6	3.1	3.2
90° dL	Good ear contra, equal levels	-9.5	2.6	5.2
90° dL	Good ear contra, matched levels	-8.1	2.8	3.7
90° dT	Good ear ipsi, equal levels	-7.0	3.0	2.6
90° dT	Good ear ipsi, matched levels	-6.8	2.2	2.5
90° dT	Good ear contra, equal levels	-6.9	1.8	2.6
90° dT	Good ear contra, matched levels	-6.6	1.9	2.2

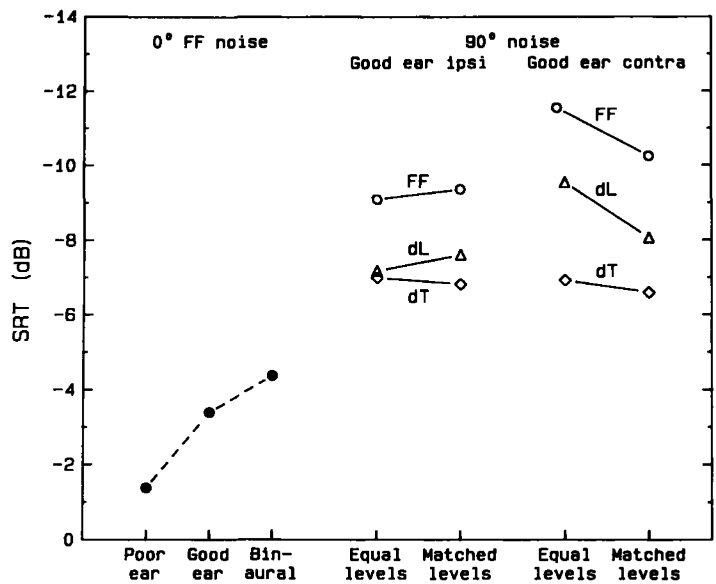

Fig. 5. Mean SRTs in noise, expressed as S/N ratio, for the 0° and 90° conditions, measured in experiment II. For the 90° noise types, the orientation is given relative to the position of the noise source during the original KEMAR recording. The presentation levels were either equal at both ears, or matched to the degree of asymmetry.

to an overall analysis of variance, applied to the SRTs in noise, 51 % of the variance is contained in the conditions, and 23.9 % in the subjects. The error estimate is 1.9 dB. Pairwise comparisons using t tests showed that all BILDs are significant ($p<0.002$), except for the monaural result for the good ear ($p=0.06$).

The mean SRTs in noise, obtained in experiment II, are displayed in Fig. 5. The mode of presentation is denoted along the horizontal axis. In the upper part of the plot, the noise angle and the relative position of the noise source with respect to the good ear are indicated. Comparison with the results for normal-hearing listeners, obtained with identical stimuli, shows that the asymmetrically impaired perform worse in all conditions. The mean SRT differences range from 2.2 to 7.2 dB, and are all significant (two-sample t tests, $p<0.0002$). When we consider first the results for 0° FF noise, it appears that the good ear has, on the average, a 2-dB better SRT in noise than the poor ear, and that the latter ear still yields a 1-dB contribution to the binaural SRT. Both differences are statistically significant (t test for paired observations, one-sided, $p=0.01$ and $p=0.03$, respectively). Next, we can see that the 90° dT noise yields about the same mean SRT in all four conditions. This means that the release from masking for dT noise does not depend on whether the good or poor ear recieves the leading signal, or whether or not the presentation levels are matched to the degree of asymmetry. The BILDs in the four conditions are about 2.5 dB, which is less than the corresponding result for the symmetrically impaired (4.2 dB), though the differences have only limited significance (two-sample t tests, p=0.02-0.09). Pooling the results for equal and matched presentation levels, we compared individual BILDs for "poor ear leading" and "good ear leading". It appeared that the BILDs correlate strongly (r=0.78, p<0.001) and that the slope of the regression line approximates unity, which contradicts the notion, suggested by the results of Bocca and Antonelli (1976), that both stimuli are processed differently by the asymmetrically impaired. In the case of 90° dL and FF noise, the mode of presentation has a clear effect on the SRTs. The previous results for dL noise already showed that the listener relies on the ear presented with the better S/N ratio. It appears that the asymmetrically impaired benefit less from headshadow when they have to depend on the poor ear (with the noise source at the side of the good ear), than in the reverse situation. Matching the presentation levels by raising the level at the poor ear, and lowering the level at the good ear, gives a slight improvement in the first case, and a marked reduction in the second. According to an analysis of variance, applied to the FF and dL results, the effect is only significant in the latter case [F(1,16)=8.9, p<0.01]. For all modes of presentation, there is a difference of about 2 dB between the mean SRTs for dL and FF noise. This value does not differ significantly from the corresponding result for listeners with symmetrical impairment (two-sample t tests, p>0.3).

As in experiment I, the SRTs and BILDs were correlated with hearing levels and age using a stepwise multiple regression analysis. The dependent variables were: the SRT for 0° FF noise and the BILDs for 90° FF, dL and dT noise, obtained with the good ear either at the ipsilateral or at the contralateral side. As 0° FF result, the average of the binaural SRT and the monaural SRT of the good ear was taken. The BILD for dT noise was calculated by subtracting the average of the two SRTs, obtained for equal and matched presentation levels, from the 0° FF result. The BILDs for FF and dL noise were calculated using only the SRT, obtained for equal presentation levels. As independent variables we used the hearing levels for both ears at 250-4000 Hz, and derived variables such as the PTAs for both ears, the interaural differences in hearing level at the individual frequencies and the average interaural difference expressed as PTA. Though these variables are in effect highly correlated, they are treated as independent variables in the stepwise regression analysis. The results were similar as in experiment I. Again, the 0° FF result did not correlate significantly with any of the independent variables, and there were no significant correlations between age and any of the BILDs. The BILDs for FF and dL noise showed a significant dependence (p<0.007) on the hearing level at 4000 Hz of the ear, contralateral to the noise source. The data and regression lines are shown in Fig. 6. As in Fig. 3, they are plotted as a function of the difference between the noise level in dBA, and the hearing level at 4000 Hz, to compensate for the interindividual differences in presentation level. No corroboration could be found of the dependence of the dT BILD on the PTA, which was present in the data of experiment I. The correlations between the dT BILDs and the PTAs of the good and poor ear were now rather low ($r<0.35$, $p>0.18$). Though the results of both experiments show that the mean BILD for dT noise is poorer for asymmetrical, than for symmetrical impairment, the analysis yielded no significant dependence of the BILD on interaural differences in hearing level.

The linear regression between the monaural SRTs in quiet and the PTAs of the respective ears yields a different equation as in experiment I:

$$SRT_q = 19.9 + 0.8\ (PTA).$$

The correlation is quite high: $r=0.926$, $p<0.001$. The difference between both regression equations is probably caused by the fact that the SRT in quiet has a lower limit of approximately 20 dB, which affects the results for the good ear in experiment II.

3.5 DISCUSSION

It is evident from past research and clinical experience that hearing-impaired listeners are often especially handicapped in noisy situations. Experimental work by Plomp, Duquesnoy and others (reviewed by Plomp, 1986) has shown that the

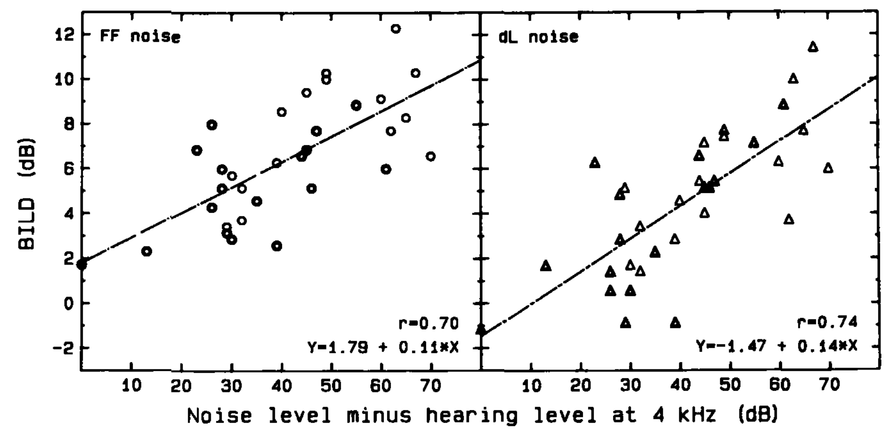

Fig. 6. The BILD for 90° FF noise (left panel) and 90° dL noise (right panel) as a function of the difference between the noise level in dBA and the hearing level at 4000 Hz of the contralateral ear. The individual results for the good ears (normal symbols) and poor ears (boldface symbols) of the 17 subjects of experiment II are shown, as well as the regression lines.

speech-reception threshold of most hearing-impaired listeners is elevated even at suprathreshold noise levels (in the model based on the data this is called a Class D hearing loss), and that hearing aids will only have a beneficial effect at lower noise levels. Since the sentence intelligibility rises steeply as a function of S/N ratio (about 20%/dB near the 50% point), a few dB of Class D hearing loss will substantially decrease the ability to understand speech in critical situations. However, the true handicap of hearing-impaired listeners in noisy situations may well be underestimated when only their Class D hearing loss is taken into account. Their SRT in noise may be elevated further due to the fact that they benefit less from binaural cues than the normal hearing do. In the present experiments, we have investigated in detail not only the performance of the hearing impaired for monotic and diotic presentation of speech and noise, but also their ability to benefit from the binaural cues ILD and ITD in a situation where speech and noise sources are spatially separated. In addition, we have studied to which degree the SRT in noise is affected by an interaural difference in overall presentation level, simulating the effect of a monaural hearing aid.

3.5 A. SRT for monotic and diotic stimulation

The Class D hearing losses for our groups of hearing-impaired subjects can be determined by comparing the results for 0° FF noise with normal values, obtained previously (Bronkhorst and Plomp, 1988). Subtracting the monaural results yields average Class D hearing losses of 2.9 dB for the symmetrically impaired, 2.2 dB for the good ear of the asymmetrically impaired, and 4.2 dB for their poor ear. For both normal-hearing and hearing-impaired listeners, diotic stimulation yields a slight (1- to 1.5-dB) improvement of the SRT with respect to the monaural result (of the good ear in case of asymmetrical impairment). Thus, the Class D hearing loss observed during binaural stimulation does not differ from the monaural result (for the better ear). Our mean values agree well with results reported by Plomp (1986) for groups of listeners with either presbycusis or noise-induced hearing loss. The rather large spread in his data and the weak correlation between the hearing loss in quiet and the hearing loss in noise also correspond with our findings: in both our experiments, the Class D hearing loss varied across subjects over a range of about 6 dB, but no significant dependence on the hearing levels could be demonstrated.

3.5 B. Effect of spatial separation of speech and noise sources

According to our data, the binaural gain (BILD), occurring when the speech source remains in front and the noise source is moved from the front to the side, is on the average 9.8 dB for the normal hearing, 7.1 dB for the symmetrically impaired, and 4.7 or 7.2 dB for the asymmetrically impaired, depending on whether the noise source is moved to the good or bad side, respectively. Thus, the hearing loss in noise of the hearing impaired for binaural listening in our free-field configuration is 2.6-5.1 dB higher than their Class D hearing loss.

These results are in general agreement with free-field data, reported by other authors. Festen and Plomp (1986), using the same material and identical positions of speech and noise sources as we did, found mean binaural gains of 6.5 and 5.5 dB for groups of hearing-impaired listeners with PTAs less than 50 dB and more than 50 dB, respectively. The subjects had nearly symmetrical hearing losses; the noise level was 80-85 dBA. Duquesnoy (1983), also using the same configuration, studied the performance of a group of elderly listeners with a hearing loss which was normal for their age (ranging from 76-88 years). His results, obtained with the same speech list as we used but with a noise masker based on a male voice, show an identical gain for the normal hearing (9.6 dB), but a gain of only 2.5 dB for the elderly. As discussed below, such a low binaural gain is to be expected when, as a result of impaired high-frequency hearing, the listeners experience no headshadow advantage at all. This was probably the case during Duquesnoy's measurements since the noise level was rather low (55 dBA), while the subjects had sloping audiograms with an average hearing level of almost 60 dB at 4000 Hz. Gelfland *et al.* (1988), replicating the measurements of Duquesnoy (except for the environment, which was not anechoic in their case), found a 5- to 6-dB binaural gain for normal-hearing listeners, independent of age, and a gain of about 3 dB for hearing-impaired elderly listeners. These findings support the notion that a peripheral hearing loss is the primary cause of reduced binaural performance of elderly listeners.

1. The ITD cue

It appears from our data that the binaural release from masking due to ITD is not strongly affected by a moderate cochlear hearing-impairment. Though a dependence (with limited significance) was found of the binaural gain on the PTA for the subjects with symmetrical impairment, there was only a small difference between the mean BILD for that group and the normal value. For both groups, the gain due to ITD in noise with ILD (the difference between corresponding dL and FF results) had a fairly constant value of 2-2.5 dB in all conditions, which is only slightly less than the mean normal value of 2.6 dB. A marked reduction of about 2 dB with respect to normal values was present in the BILDs for dT noise, obtained from the subjects with asymmetrical impairment. Thus, the binaural interaction is deteriorated by differences in degree (and possibly type) of impairment of both inner ears. The regression analysis, however, showed no significant relation between the BILDs and the interaural differences in hearing level. A clearly reduced binaural unmasking for listeners with an asymmetrical sensorineural impairment was found also by Olsen *et al.* (1976), Bocca and Antonelli (1976), and Wilson *et al.* (1985). The difference between results with the good or poor ear leading, reported by Bocca and Antonelli (1976), is not confirmed by our data, which show no difference whatsoever. This might be explained by the fact that their subjects had unilateral, instead of bilateral asymmetrical impairment.

2. The ILD cue

Our results indicate that the increased hearing loss in noise of the hearing impaired during binaural listening is mainly due to the fact that they are not able to take full advantage of the ILD cue. According to the regression analysis of the FF and dL data, obtained in both experiments, the gain due to ILD depends on the hearing level at 4000 Hz of the ear, contralateral to the noise source. By comparing binaural and monaural SRTs, we already found for normal-hearing and hearing-impaired listeners that the gain due to ILD in fact represents the monaural performance of the contralateral ear. Since the headshadow at 90° is most prominent between 3 and 5 kHz (cf. Bronkhorst and Plomp, 1988, Fig. 2), it is not surprising that the gain is affected by a hearing loss in that frequency region. Figs. 3 and 6 show that the gain increases from about 0 dB, when the hearing loss completely cancels the headshadow advantage, to 7 dB or more for the subjects with normal high-frequency hearing, which corresponds with the mean value of 7.2 dB for the 34 normal-hearing subjects tested previously. The BILDs for FF noise increase from 2-3 dB to normal values of about 10 dB. The

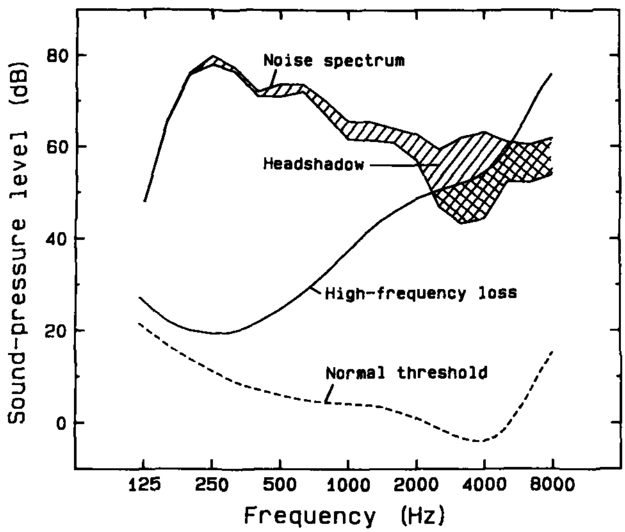

Fig. 7. Plot showing the free-field spectrum of the speech noise (reproduced at 80 dBA), the effect of the headshadow at 90°, and the free-field threshold curve corresponding to a typical high-frequency hearing loss. The figure illustrates that the level of the noise at the shadowed ear is below threshold in the high frequencies.

figures show further that, in order to get normal BILDs, the dBA level of the noise (i.e. the reference level, measured at 0°) must be about 50 dB higher than the hearing level at 4000 Hz. When we subtract 15 dB, to obtain from the dBA level of the noise the level in the 4000 Hz region, and another 20 dB to account for the ILD at 90°, we find that the true sensation level of the noise in the 4000 Hz region at the shadowed ear is only 15 dB. From the dependence of the SRT on noise level (Plomp, 1986), it follows that such a margin is normally required in order to obtain SRTs in noise exclusively determined by the S/N ratio. An illustration of of the effect of a high-frequency hearing loss on the headshadow advantage is given in Fig. 7. It shows the free-field spectrum of our speech noise, reproduced at 80 dBA (these data are taken from Duquesnoy, 1983), and the headshadow at 90°, plotted as a shaded area. In addition, a typical free-field threshold curve is shown, obtained by adding the interpolated hearing levels of one of our subjects to the normal free-field threshold values. The figure demonstrates that there will be a reduced headshadow advantage in this case since the signals in the high frequencies become subthreshold.

3.5 C. Effect of interaural differences in overall presentation level

For both groups of subjects, we have determined the effect of an interaural difference in overall presentation level on the SRTs in noise. The binaural unmasking due to ITD, with or without ILD present as well, turns out to be hardly sensitive to the level differences, applied in our experiments. The largest effect was measured in experiment I: a reduction of 1.3 dB due to the 20-dB attenuation of the contralateral channel. This deviates from the results of Wilson *et al.* (1985), who found that the BILD (for phase-inverted spondees in noise) was unaffected by a 6-dB attenuation of one channel, but decreased beyond that value by about 0.2 dB per dB attenuation. We did find a reduction of the binaural gain due to ILD in those conditions where the presentation level at the contralateral ear was lowered. This is probably due to the threshold dependence of the results for noise with ILD, discussed above.

When we consider the conditions with differences in presentation level as a simulation of a monaural hearing aid (a high-fidelity aid with a flat amplitude-frequency response and a moderate gain), we can deduce from the results that such an aid will hardly affect binaural unmasking due to ITD and that it will increase the gain due to ILD in certain situations by preventing or diminishing threshold effects. This seems to be in contradiction with the results of Festen and Plomp (1986), who found a superiority of unaided over aided listening, and (for their group with PTAs less than 50 dB) no difference between the results for one and two hearing aids. However, their measurements of the headshadow at 90° show that there is significantly less attenuation of the high frequencies at the microphone position of a behind-the-ear hearing aid than in the ear canal. Thus, the aid will have the effect of presenting the signals at a higher level but at a less favorable S/N ratio. The observed superiority of unaided listening suggests that the loss caused by the latter effect is greater than the gain caused by the former. The dependence of ILD on microphone location implies further that speech intelligibility in noise may be improved by using in-the-ear, instead of behind-the-ear hearing aids. In view of this, it is a pity that Jerlvall *et al.* (1983) and Leeuw and Dreschler (1986) compared both types of hearing aid by using a symmetrical configuration of the signal sources, with speech presented from the front, and noise from both left and right sides. Thus, by using stimuli without ILD, a significant advantage of in-the-ear hearing aids may have been overlooked.

3.6 CONCLUSION

Our study of free-field conditions with speech presented from the front, and noise either from the front or from the side has shown that the hearing impaired not only have poorer SRTs in noise than the normal hearing in the former (diotic) condition, but also less binaural advantage when changing from the former to the latter (dichotic) condition. When considering the separate contributions of the binaural cues ILD and ITD, it appears that, in general, the hearing impaired benefit less from ILD, but almost equally from ITD, compared with the normal hearing. Listeners with asymmetrical impairement, however, benefit also significantly less from ITD, when it is presented as single cue. The reduced ability to take advantage of ILD is mainly due to threshold effects. It is a combination of, on one hand, the frequency-dependence of the headshadow, which is greatest in the high frequencies, and, on the other hand, the presence of a high-frequency loss at the ear, contralateral to the noise source.

The introduction of an interaural difference in overall presentation level has only a limited effect on the release from masking through ITD. Specifically, we found for the listeners with asymmetrical impairment that matching the presentation levels to the degree of asymmetry neither degrades nor improves binaural unmasking. In contrast, the contribution of ILD to the binaural gain can be affected considerably by changes in overall presentation level. An explanation for this is the observed threshold dependence of the ILD advantage. Our results imply that an optimal hearing aid with a moderate gain will hardly interfere with binaural unmasking due to ITD, while it may increase the gain due to ILD by preventing or diminishing threshold effects.

4. BINAURAL SPEECH INTELLIGIBILITY FOR SIMULATED COCKTAIL-PARTY CONDITIONS

A.W. Bronkhorst and R. Plomp
Submitted to the Journal of the Acoustical Society of America.

ABSTRACT

Speech-reception thresholds (SRT) were measured for 17 normal-hearing and 17 hearing-impaired listeners in conditions simulating free-field situations with between one and six interfering talkers. The stimuli, speech and noise with identical long-term average spectra, were recorded with a KEMAR manikin in an anechoic room and presented to the subjects through headphones. The noise was modulated using the envelope fluctuations of the speech, in order to obtain a masking signal with similar properties as true interfering speech. A variety of conditions was simulated with the speaker always in front of the listener, and the maskers either also in front, or positioned in a symmetrical or asymmetrical configuration around the listener. Results show that the hearing impaired have significantly poorer performance than the normal hearing in all conditions. The mean SRT differences between the groups range from 4-10 dB. It appears that the modulations in the masker act as an important cue for the normal-hearing listeners, who experience up to 5 dB release from masking, while being hardly beneficial for the hearing impaired listeners. The results for the configurations with spatial separation of speech and noise sources indicate that the unmasking due to interaural time delays is the same for both groups (approximately 2.5 dB), and independent of the locations of the maskers. In contrast, the binaural advantage due to headshadow ranges from 0 to 8 dB for the normal hearing, and from 0 to 4 dB for the hearing impaired, depending on the masker configuration. The results show that the handicap of the hearing impaired in practical listening situations will be underestimated when the effects of fluctuating interference and interaural level differences are not taken into account.

4.1 INTRODUCTION

In 1958, a now classic paper of Pollack and Pickett was published, in which the effect of stereophonic listening on speech intelligibility in noise was studied. Using stimuli presented through headphones, the authors compared dichotic and monotic listening modes. In their dichotic condition, both ears received the same speech signal, but separate babble signals, recorded from different sets of talkers. In the monotic condition, the speech and one of the babble signals were presented to one ear. The results showed a superiority of binaural over monaural listening which they named the "cocktail party effect". Since then, it has become popular to use this term when referring to binaural listening in noisy situations. Though the name of the effect suggests a situation in which the listener is confronted with multiple interfering signals arriving from different directions, such a situation has, to our knowledge, never been the subject of systematic study. In the original experiment, the stimuli did not contain the spectral effects of headshadow, nor the interaural time delays (ITD), which occur during true binaural listening. Other research has been limited either to free-field configurations with a single masker (Carhart, 1965; Dirks and Wilson, 1969a; MacKeith and Coles, 1971; Plomp and Mimpen, 1981), or to headphone experiments with multiple maskers containing only interaural differences in time or phase (Carhart *et al.* 1969a; Tillman *et al.*, 1973). In view of our previous findings (Bronkhorst and Plomp, 1988, 1989), which indicate not only that headshadow has a considerable effect on free-field speech intelligibility in noise, but also that the interaural level differences (ILD) caused by headshadow interfere with unmasking through ITD, it is of interest to investigate the cocktail party effect using free-field stimuli containing both cues. This has been undertaken in the present study.

Whereas the presence of binaural cues, and the resulting improvement of speech intelligibility, is one typical aspect of situations like a cocktail party, another main characteristic is the fact that speech is masked by other speech. There are several ways to incorporate this in an experimental study. The most straightforward way, the use of true speech as masker, has the important disadvantage that the masking efficiency then depends on spectral differences between target voice and interfering voice(s). An alternative would be to use the same voice as both signal and masker, with the latter reversed in time so as to remove all semantic information. This, however, results in a awkward experimental condition, in which the listener is confused by the timbre similarity of both signals (Festen, 1987, 1990). Another possibility is to use steady-state noise with a spectrum equal to the long-term average spectrum of the target speech, as we did in our previous studies. This option has the disadvantage that the effect of level fluctuations in the interfering speech, which may cause a considerable reduction of masking efficiency (Miller, 1947; Pollack and Pickett, 1958; Carhart *et al.*, 1969b; Duquesnoy, 1983; Festen, 1987, 1990), is not taken into account. A solution of this problem was given by Festen (1990), who used the envelope fluctuations of running speech to modulate spectrally-matched noise, and thus generated a masker with virtually the same properties as true interfering speech. We decided to use this masker in the present study, since it enabled us to simulate a multiple-talker environment while preventing confounding effects of spectral differences.

We have demonstrated earlier (Bronkhorst and Plomp, 1988), that a free-field situation can be accurately simulated by making use of artificial-head recordings, presented through headphones. This technique offers several important advantages: it eliminates effects of head movement, it facilitates the realization of a truly monaural condition, and it allows the two binaural cues ITD and ILD to be tested separately. We have therefore also used it in the present study. We have further used the same recordings of speech, presented from the front, and of steady-state noise, presented from several directions in the horizontal plane, which were originally made for our previous experiments. The speech was left unchanged but the noise signals were first modulated according to the algorithm devised by Festen (1990) and subsequently added in different combinations. Between one and six modulated noise signals were added to each other in order to simulate environments with a variable number of interfering talkers. The speech and noise stimuli were used to determine masked speech-reception thresholds (SRT) for two groups of listeners. The first group consisted of young, normal-hearing subjects and the second of young and middle-aged hearing-impaired subjects with symmetrical sensorineural hearing losses. The latter group was included because hearing-impaired subjects often have great difficulty in understanding speech in noisy environments. By testing the performance of the hearing impaired in a situation with multiple fluctuating background sounds, a realistic estimate of their handicap in daily-life circumstances can be obtained.

4.2 METHOD

4.2 A. Material

The material used in the present study was originally developed by Plomp and Mimpen (1979). It consists of 130 short meaningful sentences, read by a female speaker and adjusted for equal intelligibility, and steady-state noise with a spectrum equal to the long-term average spectrum of the sentences. These stimuli were reproduced over high-fidelity equipment in an anechoic room and recorded with a KEMAR manikin (Burkhard and Sachs, 1975), fitted with two Brüel and Kjær (B&K) 4157 ear simulators. The speech was recorded at an azimuth of 0° (frontal), the noise at 7 azimuths ranging from 0° to 180° in steps of 30°. A detailed description of the recording procedure and the analysis of the interaural differences contained in the noise stimuli can be found in Bronkhorst and Plomp (1988). Of each noise recording, a 5-s interval was digitized with a sample frequency of 20 kHz and a resolution of 12 bits. Mirror signals (with azimuths of -30° to -150°) were generated by interchanging left and right channels. Since we wanted to test, in an additional condition, the effect of ITD as single cue, we processed the noise recorded at 90° in such a way that the level differences (ILD) between left and right channels were removed, while the time differences (ITD) were left unchanged (the procedure used to accomplish this is also described in Bronkhorst and Plomp, 1988). This derived noise signal will be designated as dT noise.

All noise signals were modulated according to a procedure devised by Festen (1990). The modulations were derived from a 1-min sample of target speech, consisting of a number of sentences put one after another (without pauses). The speech was first split up in a high- and a low-frequency part using 128-point finite impulse response (FIR) high- and lowpass filters with a slope of 60 dB/oct and a cut-off frequency of 1 kHz. Then, the envelope of the filtered speech was determined by taking the modulus of the signal, following the peaks with an exponential function having a decay time of 12.8 ms, and smoothing the resultant signal using a simple infinite impulse response (IIR) filter with a cut-off frequency of 40 Hz. Subsequently, left and right channels of all digitized noise samples were passed through the same FIR high- and lowpass filters as the speech, and multiplied in both frequency bands by the corresponding speech envelopes. For each noise signal, different envelopes were used, taken from separate segments of the speech sample. Both left and right channels of the noise signals received the same modulation,

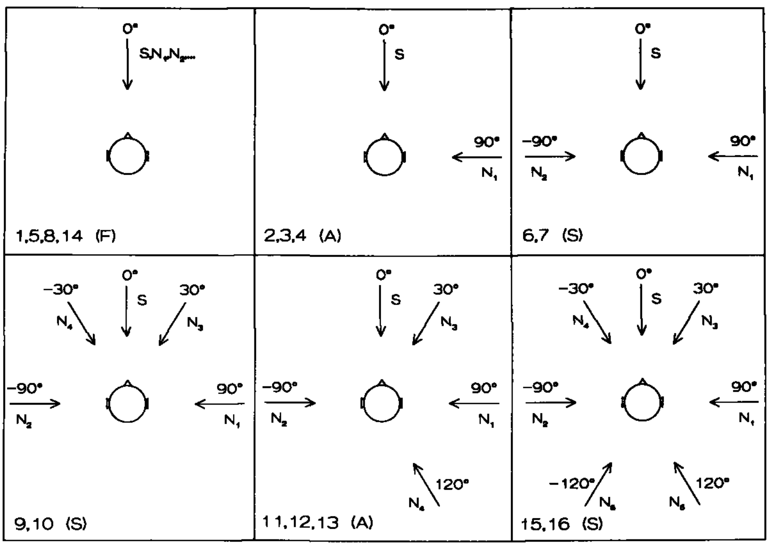

FIG. 1. Schematic plot showing the configurations simulated in the present experiment. The maskers are all in front (F), or distributed either symmetrically (S) or asymmetrically (A). For each configuration, the numbers of the corresponding experimental conditions are given (as listed in the Table).

because the periods of the envelope fluctuations, which range from about 50 ms to 4 s (cf. Houtgast and Steeneken, 1985), are much longer than the ITD, which is less than 1 ms. After the multiplying operation, small level adjustments were made in both frequency bands, to equate the total power contained in the nonmodulated and modulated signals (calculated in both bands and over the full 5-s length of the noise samples). Finally, the bands were added to each other. Using this procedure, 12 separate (non-coherent) noise signals were generated from the noise recorded at 0° azimuth, each modulated according to a different speech segment. For all other azimuths, two separate, independently modulated signals were generated. We made two different signals for each azimuth because we planned to use in our experiment two alternative sets of masking signals, each presented to half of the subjects, in order to reduce the possible bias caused by specific modulations. For the 0° azimuth, a larger number of signals was required since we wanted to include, as reference, configurations with all (up to six) maskers in front of the listener, i.e. without spatial separation of speech and noise sources.

The noise signals were used either as a single masker, or added to each other in order to simulate a multiple-masker environment. After addition, the amplitude of the resultant signal was divided by the square root of the number of constituting signals, to compensate for the increase in signal power. A total of 17 conditions was planned for the experiment. This number was chosen for practical purposes, since we could then use the same division of the sentences (into 17 lists of 8 sentences, with 6 sentences duplicated) and the same Latin-square design as in our previous studies. We had conditions with either one, two, four or six maskers. A schematic diagram of the configurations tested in the experiment is presented in Fig. 1. The speech source was always at 0° azimuth and the noise sources were either all at the same location as the speech source, or distributed in a symmetrical configuration (two maskers at 90° and -90°; four maskers at 30°, 90°, -30° and -90°; six maskers at 30°, 90°, 150°, -30°, -90° and -150°) or, alternatively, distributed asymmetrically (a single masker at 90°; four maskers at 30°, 90°, 150° and -90°). Thus there were four configurations with all signal sources in front, and five configurations with spatial separation of all signal sources. These nine configurations were tested using binaural presentation. The five latter configurations were also tested monaurally. This yields seven extra conditions, as there are two separate monaural conditions for each of the two asymmetrical configurations. The dT noise, derived form the 90° noise signal and modulated like the other noise signals, was used as masker in the last condition. A listing of all conditions is presented in the Table.

4.2 B. Procedure

The 17 conditions were presented to two groups of 17 subjects each. The first group consisted of normal-hearing listeners, the latter of hearing-impaired listeners with symmetrical sensorineural hearing losses. The order of the conditions was balanced over the subjects according to a Latin-square design, while the order of the sentences was kept fixed. Seventeen different tapes were prepared containing, on two tracks, left and right channels of the sentence list and, on two separate tracks, both channels of the noise samples, put in the order prescribed by the Latin square. The experiment was performed in a double-walled soundproof booth. Speech and noise were played back on a reel-to-reel tape deck, mixed and amplified in two Madsen OB 822 clinical audiometers, and reproduced over a set of Beyer DT-48 headphones. The level settings of both audiometers were remotely controlled by computer. The SRTs were determined by varying the presentation level of the sentences according to the following adaptive strategy. A response was designated as correct when the listener reproduced the sentence without any error. Presentation of the first sentence of each list was started below threshold. This sentence was then raised in level in steps of 4 dB until it was reproduced correctly. Next, the remaining seven sentences were presented at a level that was lowered by 2 dB after a correct response, and raised by 2 dB after an

TABLE. Means and standard deviations of the SRTs obtained from the normal-hearing and hearing-impaired subjects. The SRTs are expressed as S/N ratio. The conditions simulated situations with one to six maskers, all located at an azimuth of 0° (frontal), or placed in a symmetrical or asymmetrical configuration. (A schematical plot of all configurations is shown in Fig. 1, together with the numbers of the corresponding conditions.) Subjects listened either binaurally or monaurally. In the asymmetrical configurations, monaural listening was either at the ipsi- or contralateral side, the former side being the one closest to the masker(s).

				SRT (dB)			
	Masker	Listening	Number	Normal hearing		Hearing impaired	
Nr.	orientation	mode	of maskers	mean	s.d.	mean	s.d.
1	Frontal	Binaural	1	-12.0	2.6	-4.9	3.6
2	Asymmetrical	Binaural	1	-20.0	2.0	-11.4	4.4
3	Asymmetrical	Mon. ipsi	1	-9.7	3.2	-2.1	3.5
4	Asymmetrical	Mon. contra	1	-17.6	2.0	-7.6	4.3
5	Frontal	Binaural	2	-9.6	2.0	-4.7	2.8
6	Symmetrical	Binaural	2	-14.2	2.2	-7.4	3.5
7	Symmetrical	Monaural	2	-9.9	2.3	-3.7	4.3
8	Frontal	Binaural	4	-8.1	1.4	-4.0	2.0
9	Symmetrical	Binaural	4	-10.0	1.5	-5.0	2.3
10	Symmetrical	Monaural	4	-7.2	2.0	-2.3	2.8
11	Asymmetrical	Binaural	4	-11.4	1.4	-5.9	2.8
12	Asymmetrical	Mon. ipsi	4	-6.2	1.7	-1.8	3.4
13	Asymmetrical	Mon. contra	4	-8.5	1.2	-2.6	3.2
14	Frontal	Binaural	6	-7.7	1.1	-3.2	2.4
15	Symmetrical	Binaural	6	-9.2	1.6	-4.9	2.6
16	Symmetrical	Monaural	6	-6.4	1.4	-2.2	2.7
17	(dT noise)	Binaural	1	-17.2	2.4	-9.7	4.4

incorrect response (Plomp and Mimpen, 1979). The average presentation level of the third and following sentences was taken as the SRT.

The presentation level of the stimuli was calibrated according to the dBA level of the (nonmodulated) noise, recorded at 0° azimuth. Levels were measured with a B&K 4152 artificial ear connected to a B&K 2118 sound-level meter. For the normal-hearing subjects, the noise level was fixed at 65 dBA. In the hearing-impaired group, a noise level between 75 and 90 dBA was chosen for each subject individually, depending on the hearing loss and uncomfortable loudness levels of that subject. This was done in order to present the stimuli well above threshold, while at the same time preventing unnaturally or uncomfortably loud levels.

The subjects were tested in a single experimental session, which lasted about 70 minutes, including a short break. Before the presentation of the sentence lists, a pure-tone air-conduction audiogram was determined using a modified Békésy procedure operating at octave frequencies ranging from 250 to 8000 Hz. Tone bursts with a duration of 200 ms were presented at a rate of approximately one per second and varied in level in steps of 2 dB. The stimuli were calibrated according to ISO 389. Presentation continued until 10 reversals had occurred. The threshold was taken as the average level at the eight last reversals.

4.2 C. Subjects

The group of normal-hearing listeners consisted of eight male and nine female subjects, aged 23 to 39 with a median age of 25.3. All had pure-tone hearing levels of 20 dB or less at octave frequencies from 250 to 8000 Hz. The hearing-impaired subjects were selected from the files of the audiology department. Seven were male and ten were female. Their ages ranged from 18 to 59 with a median age of 53.5. In no case did audiometric or other diagnostic tests indicate middle-ear or retro-cochlear pathology. All had maximum discrimination scores of at least 85% for monosyllables presented in quiet. Over all 34 ears, the PTA (average of the hearing levels at 500, 1000 and 2000 Hz) varied between 16 and 56 dB, with a mean of 36 dB. The PTAs obtained from the left and right ears of the subjects differed by no more than 9 dB. On the average, the difference was 3.8 dB.

4.3 RESULTS

The means and standard deviations of the SRTs for both the normal hearing and the hearing impaired are presented in the Table. The SRTs are expressed as signal-to-noise (S/N) ratio. For all conditions, there is a significant difference between the mean results obtained from the two groups (two-sample t test, $p<0.0002$). An overall analysis of variance indicated that, for the

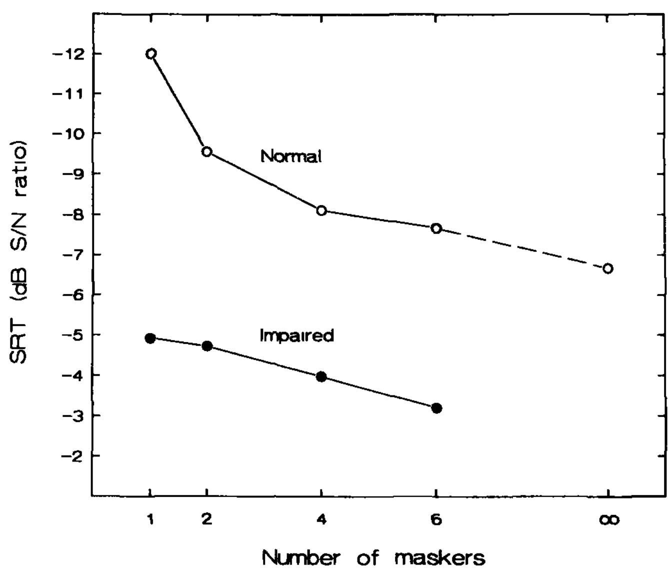

FIG. 2. Mean SRTs for the normal-hearing and hearing-impaired listeners, obtained in the diotic conditions, simulating configurations with all maskers located frontally. The result for steady-state noise (the normalized sum of an infinite number of modulated maskers) is taken from Bronkhorst and Plomp (1988).

normal hearing, 81.3% of the variance is contained in the conditions and only 1.4% in the subjects. For the hearing impaired, only 40% of the variance is accounted for by the conditions and 41% by the subjects. The error estimate, based on the highest-order interaction, is 1.9 dB for both groups.

A plot of the mean SRTs for the diotic conditions, corresponding to configurations with one, two, four or six maskers located in front of the listener, is shown in Fig. 2. Since all noise samples were equalized in terms of total signal power, the dependence of the SRT on the number of maskers is entirely due to differences in signal envelope. Increasing the number of (independently modulated) maskers has the effect that the envelope fluctuations in the resultant signal become smaller. In the limit of an infinite number of maskers, the result will be steady-state noise. The mean SRT for this condition, obtained in an earlier experiment from 34 normal-hearing listeners (Bronkhorst and Plomp, 1988), is also shown in the figure. It appears that when the number of maskers is four or more, the masking effect approaches that of steady-state noise. The figure furthermore shows that the normal hearing benefit more from envelope fluctuations of the masker than the hearing impaired. The difference in SRT between the conditions with one and six maskers is 4.3 dB for the normal hearing and only 1.7 dB for the hearing impaired. These values differ significantly (two-sample t test, one-sided,

$p=0.005$).

In Fig. 3, the mean SRTs for all monaural and binaural conditions with free-field stimuli are displayed, including the data already presented in Fig. 2. The results for one, two, four, and six maskers are shown in separate panels. The letters denoted along the ordinate represent the different masker orientations: all maskers in front (F), or positioned in either a symmetrical (S) or an asymmetrical (A) configuration. The results for the two latter configurations are not only affected by differences in the time envelope of the masking signal, but also by headshadow and ITD.

We will first consider the results for the normal-hearing listeners. Our previous results for steady-state noise have shown that, in the case of a single masker located at 90°, the effect of ITD is about 5 dB when it is presented as a single cue, but only 2.5 dB in combination with ILD due to headshadow (Bronkhorst and Plomp, 1988). It was furthermore found that, for stimuli containing only ILD, there was no difference between the binaural SRT and the monaural SRT at the side presented with the most favorable S/N ratio. This implies that the contribution of ITD can be be obtained by subtracting the binaural SRT from the better monaural SRT. For the present results,

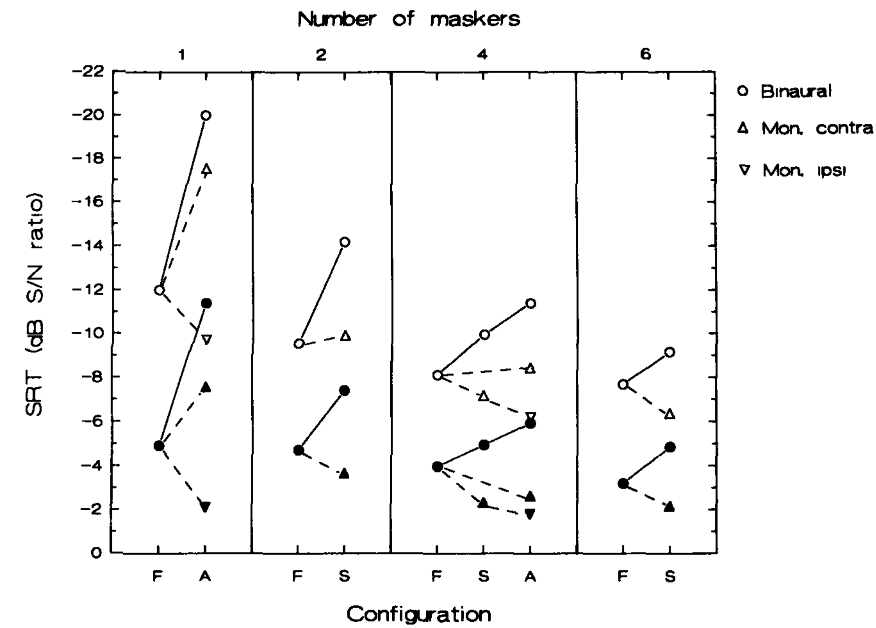

FIG. 3. Mean SRTs for all binaural and monaural conditions with free-field stimuli. The results for both normal-hearing (open symbols) and hearing-impaired listeners (closed symbols) are shown. Configurations were tested with all maskers in front (F), or distributed either symmetrically (S) or asymmetrically (A).

shown in the Table, this yields values of 2.4 and 2.9 dB for the asymmetrical configurations with one and four maskers, respectively, and values of 4.3, 2.8 and 2.8 dB for the symmetrical configurations with two, four and six maskers, respectively. Thus the binaural advantage has a fairly constant value in all conditions except the one with two maskers. As will be discussed below, an additional effect, related to the masker fluctuations, may have affected the results in the two-masker condition.

The effect of headshadow, present when the maskers are located asymmetrically, can be obtained by taking the difference between the SRTs for ipsilateral and contralateral listening. It is 7.9 dB in the case of a single masker, and only 2.3 dB when three maskers are placed at one side, and one at the opposite side.

The SRTs for the hearing impaired, indicated by the closed symbols in Fig. 3, are 4.2 to 10.0 dB higher than corresponding results obtained from the normal hearing. The discrepancy increases as the number of maskers is reduced. As discussed above, this indicates that the hearing-impaired listeners derive less benefit from envelope fluctuations of the maskers than the normal hearing. The mean binaural advantage, obtained by subtracting the binaural SRT from the SRT for monaural (contralateral) listening, is in the asymmetrical conditions with one and four maskers 3.8 and 3.3 dB, respectively, and in the symmetrical conditions with two, four and six maskers 3.7, 2.7 and 2.7 dB, respectively. The mean advantage is in two cases larger than the corresponding value for the normal hearing, but the differences are not significant (two-sample t tests, $p>0.17$). The difference between both monaural SRTs in the asymmetrical conditions is 5.5 and 0.8 dB for one and four maskers, respectively. These values are smaller than those obtained from the normal hearing. This corresponds with results of our previous study (Bronkhorst and Plomp, 1989), showing a reduced ILD effect in the case of impaired hearing.

Our condition with dT noise was meant to check whether the binaural unmasking due to ITD (presented as a single cue) is affected by envelope fluctuations of the masker. The amount of unmasking is the difference between the SRT for dT noise and the SRT for the noise of a single masker located frontally. According to the results presented in the Table, it is on the average 5.2 and 4.8 dB for the normal-hearing and hearing-impaired listeners, respectively. In our previous experiments, where we used steady-state noise (Bronkhorst and Plomp, 1988, 1989), we found that the unmasking caused by ITD is on the average 4.7 dB for the normal-hearing, and 4.2 dB for a group of symmetrically impaired listeners, comparable to the one used in the present experiment. Thus it appears that the binaural process responsible for the unmasking is not sensitive to envelope fluctuations of the masker.

As mentioned above, the overall analysis of variance showed that the percentage of the total variance contained in the subjects is much larger for the hearing impaired than for the normal hearing. The question is, whether this is due to intra- or intersubject variability. In other words, do the hearing impaired perform more randomly than the normal hearing, or do they perform consistently at a certain level determined by their impairment? While we did

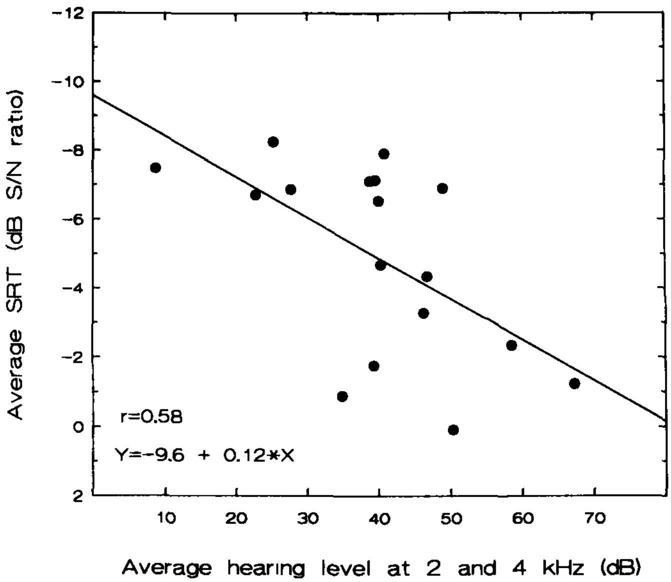

FIG. 4. Plot showing the SRTs, averaged over all conditions, of the individual hearing-impaired subjects as a function of their average hearing loss at 2 and 4 kHz, and the derived regression line.

not measure retests, the similarity of our conditions enables us to estimate the intrasubject variability from the correlations between the different SRTs. Calculation showed that most correlations are high: 118 of the 136 correlations are significant at the 1% level (r>0.60) and 75 are also significant at the 0.1% level (r>0.75). The strong interdependence of the data also manifested itself in a principle components analysis (PCA), applied to the covariance matrix. This analysis showed that 74.4% of the variance can be explained by a single dimension with similar loadings for all conditions. Thus it appears that there is a single factor, linked to interindividual differences, which accounts for most of the variability in the data.

The next step is to relate this factor to the pure-tone audiograms of the subjects. This was done using multiple regression analysis with as independent variables the hearing levels at 250, 500, 1000, 2000 and 4000 Hz (it was presumed that the hearing level at 8000 Hz has no significant effect on speech intelligibility). For simplicity's sake, we used as dependent variable the average SRT over all conditions and not the projection of the data on the single dimension generated by the PCA. The analysis revealed a significant dependence of the average SRT on the hearing levels, especially those at 2000 and 4000 Hz. This is illustrated in Fig. 4, which shows the average SRTs of the individual subjects as a function of the mean hearing level at 2000 and

4000 Hz. The slope of the regression line is 0.12; the correlation coefficient is 0.58 (p=0.014), which means that only about a third of the variance in the average SRTs can be explained from the audiogram.

As demonstrated in our previous study, the SRT for conditions involving headshadow depends strongly on the hearing level at 4000 Hz of the ear contralateral to the noise source (Bronkhorst and Plomp, 1989). Regression analysis applied to the present results for a single masker at 90° reveals a very similar dependence. In contrast, the results obtained in the asymmetrical configuration with four maskers show no dependence on high-frequency hearing loss. This is because the spacing of the maskers, positioned at 30°, 90°, 150° and -90°, results in this case in an ILD which is smaller and less marked in the high frequencies. The dependence of the SRT on high-frequency hearing loss in the case of a single masker at 90° may have enhanced the correlation between the audiogram and the average SRT. We checked this by repeating the analysis excluding the data from this condition. The outcome showed only marginal deviations from the original results: the mean hearing level at 2000 and 4000 Hz was again the best predictor; the slope of the regression line was now 0.11 and the correlation coefficient was 0.54 (p=0.025).

4.4 DISCUSSION

A major problem in the study of binaural speech intelligibility in noise is that there are so many parameters involved and, consequently, so many choices which have to be made for any particular experiment, that it is often difficult to generalize the obtained results. This is further complicated by the fact that the advantage of binaural versus monaural listening depends on multiple interdependent effects, which cannot easily be separated. The risk of producing data with only a limited scope has been an important consideration in the design of our experiments. Our approach has been (1) to simulate a realistic environment, with the talker in front of the listener and the masker(s) either in front or at other azimuth(s), using speech material representative of everyday conversation, (2) to eliminate a number of confounding effects like reverberation, spectral differences between target voice and masker, incomplete occlusion of one ear during monaural listening, and head movement of the listener, and (3) to process the stimuli in such a way that the contribution of several relevant cues, binaural as well as monaural, can be evaluated separately.

In our previous studies, this approach has resulted in estimates of the contributions of ITD and ILD to the binaural advantage experienced by normal-hearing and hearing-impaired listeners in a situation with a single steady-state masker. The present study extends this to a more realistic, but also more complex, multiple-talker situation by employing up to six maskers, modulated like running speech. These modulations act as another important cue since they enable the listener to take advantage of the relatively silent periods in the masking signal. The contributions of both this cue and the cues ITD and ILD have been measured as a function of the number of maskers and the masker configuration. Not only normal performance has been investigated, but also the performance of a group of hearing-impaired listeners with moderate sensorineural hearing losses. We will first dicuss the results for the normal hearing listeners, and then consider the effect of hearing impairment.

4.4 A. Effect of envelope fluctuations on masking efficiency

It was already shown by Miller (1947) that the masking efficiency of voice babble increases as a function of the number of voices, and that it converges to a limit when the number of voices exceeds four. Fig. 2 shows that this effect is also present in our results. However, when we determine from his graphs the signal-to-babble ratios at the 50% intelligibility level, it appears that these ratios cover a range which is about twice as large as the range of our S/N ratios. For example, the increase in masking level when going from one to six voices is more than 10 dB in his case, but only 4.3 dB according to our data. The cause of this discrepancy may lie in factors like speech material, testing procedure, stimulus calibration and spectral differences between masker and target voice. We do not expect it to be caused by inadequacy of the modulated noise masker, as Festen (1990) has demonstrated that such a masker has the same effect as a true interfering voice. Thus the conclusion seems warranted that, when voice babble is used to mask speech (i.c. sentences), the maximum release from masking due to envelope fluctuations of the masker is in the order of 5 dB. In practical situations, the release may be larger due to, for instance, spectral differences between voices. Results obtained by Pollack and Pickett (1958) indicate that the contribution of these additional effects can also be quite small. According to the data presented in their Fig. 5 (for the monaural "control" conditions), they found a difference in masking level of only 4 dB between a single voice and multiple-voice babble. Their results show furthermore that babble signals containing four or seven voices have about the same masking efficiency, which corresponds with both Miller's (1947) and our results.

4.4 B. Effect of the binaural cues ITD and ILD

As already noted above, the results for dT noise show that the release from masking caused by ITD (presented as a single cue) is not affected by the speech-like modulations present in the noise masker: for both steady-state and modulated noise, a value of about 5 dB was obtained. This is also true when the masker contains both ITD and ILD. Comparison of the present data for a single masker at 90° with the results reported in Bronkhorst and Plomp (1988) shows that the SRT difference between binaural and monaural contralateral listening, which in effect represents the contribution of ITD when ILD is present as well, is 2.5 dB for both types of noise. In contrast, the data show that the masker modulations do have an effect on the contribution of ILD to the binaural advantage. A measure of the ILD effect is the SRT difference between monaural ipsi- and contralateral listening, which is 7.9 dB for modulated noise, and, according to the abovementioned study, 10.7 dB for steady-state noise. This discrepancy can be explained from the fact that the performance during contralateral listening depends primarily on the perception of relatively weak high-frequency components (since the headshadow is most prominent in these high frequencies), which implies that there remains in this case less to be gained from level fluctuations of the masker.

In situations with multiple maskers instead of a single one at 90°, the ILDs will be smaller or even absent, depending on the locations of the maskers. In the symmetrical configurations, no ILD will be present, or at least no "mean" ILD. The use of maskers with (independent) envelope fluctuations will in effect cause the ILD to vary in time as well, so that speech masked at one ear may be perceived by the other ear. Thus the total information presented to the listener is augmented and the SRT is lowered. This effect will be largest in the case of two maskers and will diminish as the number of maskers is increased. It most probably explains the relatively large binaural gain of 4.3 dB obtained in our condition with two maskers, as compared with 2.5-3 dB in the other conditions.

While this "varying-ILD" effect will only yield a relatively small contribution in our case, since the varying ILD can not be larger than the maximum ILD due to headshadow, a larger contribution can be expected when different babble signals are presented to the ears through headphones, as during Pollack and Pickett's (1958) experiments. This is supported by their results, which show a difference in binaural gain between the conditions with two and seven maskers of as much as 7 dB. It seems justified to assume that the contribution of the "varying-ILD" effect will be much smaller in practical listening situations, as a result of the limit imposed by headshadow. The difference of approximately 1.5 dB, present in our data, may, on the other hand, be smaller than the true contribution, since the fact that we used maskers with identical long-term average spectra probably caused a further reduction in our case.

Our assumption, made previously (Bronkhorst and Plomp, 1988), that the degrading effect of ILD on the binaural gain due to ITD will be smaller in symmetrical than in asymmetrical masker configurations, is not supported by the present data for the conditions with four and six maskers, which show a binaural gain of only 2.8 dB. The assumption was based on results of Carhart *et al.* (1969a), who found for stimuli presented through earphones that the binaural unmasking remained the same or increased slightly as a function of the number of maskers. However, there remains an important difference between their and our conditions, since, when we consider each masker individually, our stimuli still contain ILDs due to headshadow, while no ILD was present in their case. Thus our finding that the unmasking due to ITD is also reduced in a symmetrical multiple-masker situation suggests that each individual masker is "unmasked" individually.

4.4 C. Performance of the hearing impaired

A remarkable result, also found by Festen (1987, 1990), is that the hearing impaired are hardly able to benefit from envelope fluctuations of the masker. Even more remarkable is, that our subjects with light hearing losses (five subjects with PTAs, averaged over both ears, of less than 30 dB) do not seem to perform better than the subjects with poorer hearing: The mean SRTs for the conditions with one or two maskers are approximately the same for both groups. As fluctuating interference is very common during everyday listening, these results indicate that measurements of speech intelligibility in steady-state noise will underestimate the true handicap of the hearing impaired.

The reduced performance of the hearing impaired may have several causes. First, the advantage will decrease due to threshold effects, since the hearing loss will limit the maximum potential gain during silent periods of the masker. Second, there will be a deterioration due to reduced temporal resolution, which is generally associated with hearing impairment (e.g. Elliot, 1975), and which will have the effect of smoothening the perceived masker modulations. Third, if we may assume that the advantage due to masker modulations is partly caused by comodulation masking release (CMR), results will be affected by a reduction of this release, occurring in impaired hearing (Hall *et al.*, 1988). Quantative assessment of the two former causes falls beyond the scope of this study, and we intend to address it in the future. As to the latter effect, its evaluation awaits data not available at present, for instance on the CMR occurring when the stimulus is a complex signal instead of a pure tone.

The results show that the present selection of subjects had, on the average, somewhat poorer performance than the hearing-impaired subjects used in our previous study (Bronkhorst and Plomp, 1989), even when we consider only the conditions with four or six maskers, where the effect of masker modulations will be limited: In the diotic six-masker condition, the mean SRT of the hearing impaired is 4.3 dB higher than the corresponding normal result, while the above study showed average differences with normal performance of 3.0 and 2.5 dB for subjects with symmetrical and asymmetrical impairment, respectively. The SRT differences between binaural and monaural, and between monaural ipsi- and contralateral listening in the dichotic conditions show that the hearing impaired experience similar unmasking due to ITD as the normal hearing, but less gain due to ILD. The same results were found in the above study, where we also showed that the latter effect can be explained from, on the one hand, the frequency dependence of the headshadow, and, on the other hand, the presence of a high-frequency hearing loss at the ear presented with the most favorable S/N ratio.

Our analysis of the results for the hearing impaired revealed that there is a significant correlation between the individual SRTs, averaged over all conditions, and the hearing levels at 2000 and 4000 Hz. It is remarkable that data of Smoorenburg (1989), obtained from a large number of subjects with noise-induced hearing loss, show exactly the same relationship: He also found that the average hearing level at 2000 and 4000 Hz is the best predictor of the SRTs in noise, and his regression line had the same 0.12-dB/dB slope. In a further analysis of his data (Smoorenburg, 1990), he compares his results with predictions based on the Articulation Index (French and Steinberg, 1947), and concludes that the observed dependence can not be adequately explained by threshold effects only. This indicates that the relationship between the SRT in noise and the audiogram is mostly an indirect one, which is, in essence, the basic assumption underlying Plomp's (1986) signal-to-noise ratio model. Our data also support this assumption, since we found that most of the variance in the data can be explained by a single factor, linked to interindividual differences, while the audiogram proved to have only a limited predictive capability.

4.4 D. Implications for practical situations

The implications of the experimental results for practical multiple-talker situations can be best understood when the SRTs are plotted as S/N ratios relative to the level of a single masker instead of the total level of all maskers. This means that we should add 3 dB, 6 dB, and 7.8 dB to the SRTs for two, four, and six maskers, respectively. The results (for the binaural conditios only) are shown in Fig. 5. In this representation, the S/N ratio of 0 dB has special significance, since it corresponds to the situation where all voices (target as well as interfering) are equally loud. Thus, in situations where the SRT is about 0 dB or more, the intelligibility of speech can only be maintained either when the talker raises his voice, or when the listener employs other (i.e. visual) cues. The plot shows that the normal-hearing listeners can tolerate as many as six interfering talkers, located at different azimuths, but reach the critical region when all talkers are clustered in front, and no binaural cues are present. The hearing impaired, however, already encounter difficulties in situations with four interfering talkers, no matter where these are positioned.

4.5 CONCLUSION

Speech intelligibility in a situation with multiple interfering talkers, like during a cocktail party, is determined by many factors. The listener can, on the one hand, benefit from the binaural cues ITD and ILD, from envelope fluctuations of the interfering sounds, and from spectral differences between these sounds and the voice he is listening to. On the other hand, he can be hindered by reverberation, by an unfavorable position relative to the speakers, or by hearing impairment. In the present study, we have measured monaural and binaural SRTs in noise in a number of simulated multiple-talker conditions in order to determine the effects of ITD, ILD, and masker fluctuations on the intelligibility of speech for both normal-hearing and hearing-impaired listeners. By using free-field stimuli and noise maskers with the same long-term average spectrum as the speech, we have excluded effects of reverberation and spectral differences.

The experimental results show that, over all conditions, the hearing impaired need 4-10 dB better S/N ratios than the normal hearing for equal intelligibility. The largest differences occur in conditions with few maskers, when the normal hearing gain up to 5 dB by taking advantage of envelope fluctuations in the masking signal, while the hearing impaired hardly benefit at all. Possible explanations for the poor performance of the hearing impaired are threshold effects, degraded temporal resolution, and reduced comodulation masking release.

Comparison of the binaural and monaural SRTs (the difference is the actual "cocktail party effect") shows that both normal-hearing and hearing-impaired listeners experience in all conditions about the same release from masking due to ITD (on the average 2.5-3 dB). The contribution of headshadow to the

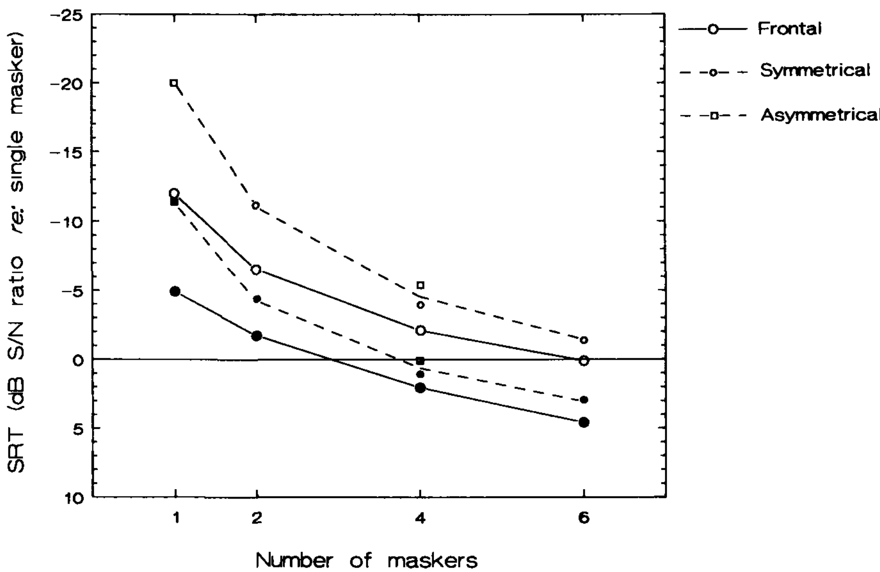

FIG. 5. An alternative representation of the results for the binaural conditions. The data points represent mean SRTs, expressed as S/N ratio relative to the level of a single masker, for both the normal-hearing (open symbols) and hearing-impaired (closed symbols) listeners.

binaural advantage depends on both the number of maskers and their orientation with respect to the listener. It ranges from 0 to 8 dB for the normal-hearing, and from 0 to 4 dB for the hearing-impaired listeners. Especially the listeners with high-frequency hearing losses benefit less from headshadow, since its effect is based on perception of the high-frequency components at the shadowed side. An additional ILD effect is introduced by the use of fluctuating maskers. The fluctuations cause the ILD to vary as well, which reduces the efficiency of the masking signal, since the listener can at any time use the ear with the most favorable S/N ratio. In our conditions, this effect yields only a relatively small advantage (at most 1.5 dB for the normal hearing).

According to an over-all analysis of the results for the hearing-impaired listeners, there is a significant relationship between the SRT, averaged over all conditions, and the average hearing level at 2000 and 4000 Hz. Only about a third of the variance, however, is accounted for. This is not due to intra-individual variability, since a PCA showed that as much as three-quarters of the variance in the SRTs themselves can be explained by a single dimension. These results imply that the factors causing degraded speech perception in noise are only partially represented in the audiogram.

An implication of our findings for "cocktail-party"-like situations is that, when all voices are equally loud, speech remains intelligible for normal-hearing listeners even when there are as many as six interfering talkers, while the hearing impaired can not tolerate more than three talkers. Another implication, relevant for audiological practice, is that the handicap of hearing-impaired listeners in daily-life circumstances will be underestimated when the effects of fluctuating interference and headshadow are not taken into account.

5. A CLINICAL TEST FOR THE ASSESSMENT OF BINAURAL SPEECH PERCEPTION IN NOISE

A. W. Bronkhorst and R. Plomp.
Submitted to Audiology

ABSTRACT

The present paper describes a clinical test for the assessment of speech perception in noise. The test was designed to separate the effects of several relevant monaural and binaural cues. Results show that the performance of individual hearing-impaired listeners deviates significantly from normal for at least two of the following aspects: (1) perception of speech in steady-state noise, (2) relative binaural advantage due to directional cues, (3) relative advantage due to masker fluctuations. In contrast, both the hearing loss for reverberated speech and the relative binaural advantage due to interaural signal decorrelation were essentially normal for almost all hearing impaired.

5.1 INTRODUCTION

In clinical audiology, it is a common experience that hearing impaired patients complain about not being able to understand speech in noisy circumstances. A prerequisite for a proper advise concerning a patient's ability to communicate in noisy situations, or for the selection of the optimal hearing aid, is a reliable clinical test for the assessment of the patient's speech perception in noise. The development of such a test is, however, complicated by the large number of factors involved. Test results will depend not only on the hearing loss of the individual tested and on the hearing aids being worn, but also on the type of test material, the method of presentation and the acoustics of the listening environment.

In developing a suitable test, a reasonable approach is to model the material and test conditions after daily-life circumstances. Thus, in deciding which speech material to use, sentences seem to be a better choice than word lists, as the former simulate everyday speech more closely. It seems furthermore most natural to present the speech only from the front, as this is the most common situation, especially for hearing-impaired listeners who also want to take advantage of visual cues. The importance of the fact that the hearing impaired should be able to partake normally in all kinds of social activities, where speech is the major interfering sound, suggests that a speech-like signal should be used as masker. Shaping the spectrum of this masker according to the long-term average spectrum of the speech itself, provides the additional advantage of a reduced speaker-dependence of the test results. Finally, in selecting suitable listening conditions, a useful approach is to design the test in such a way, that the different effects influencing speech perception in everyday circumstances (such as reverberation or the presence of binaural cues) can be evaluated separately. This will allow generalisation of the test results, so that the listener's performance in many different situations can be predicted. In addition, the information will be of interest when the test is used during hearing-aid fitting.

The aim of the present study was to devise a suitable test of speech perception in noise, using the approach outlined above. As test material, sentences and speech noise were used, developed previously by Plomp and Mimpen (1979). In addition to the original steady-state noise, a derived noise signal was employed, modulated like running speech (Festen, 1989). This masker closely simulates the effect of true interfering speech. By comparing the performance for both types of masker, the gain due to masker fluctuations can be determined. This is of interest because it was shown by a number of investigators (Duquesnoy, 1983; Festen, 1986, 1989; Bronkhorst and Plomp, 1989) that these fluctuations act as an important cue for the normal hearing, while being hardly beneficial for the hearing impaired. The effect of reverberation on speech and interference was investigated by presenting the test material in a reverberant environment, with the listener either in the direct or in the indirect field of the speech and/or noise sources. This simulated situations like meetings or conferences, where a distant speaker may be masked by nearby interfering talkers, as well as one-to-one conversations in, for example, a cafeteria or during a party, where nearby speech is disturbed by background noise. By presenting the noise either from the front or from the side in the conditions where all sound sources were close to the listener, the influence of binaural cues was assessed. The test was presented to a group of ten normal-hearing subjects, so as to obtain normative results, and to a group of 18 sensorineurally impaired subjects with hearing losses ranging from light to moderate. Four of the hearing-impaired subjects wore one or two hearing aids during the test.

5.2 METHOD

5.2 A. Material

The speech material used consisted of a set of 130 short sentences, representative of everyday conversation. They were read by a female speaker and adjusted in level for equal intelligibility (Plomp and Mimpen, 1979). The noise was spectrally shaped according to the long-term average spectrum of the sentences. A fluctuating noise signal was derived from the steady-state noise using a procedure devised by Festen (1989). First, a pseudo-running speech was generated by putting a number of (digitized) sentences one after another, without pauses. This signal was then split up into a high and a low frequency part using digital high- and lowpass filters with a cutoff frequency of 1 kHz. Next, the envelopes of both parts were determined by taking the modulus of the signal, following the peaks with an exponential function having a decay time of 12.8 ms, and smoothing the resultant signal with a simple 40-Hz lowpass filter. The steady-state speech noise was passed through the same 1-kHz high- and lowpass filters, and subsequently modulated using the corresponding high and low frequency speech envelopes. Finally, the bands were adjusted in level, to maintain equal long-term average power, and then added.

5.2 B. Procedure

The testing took place in a large room, 6.6 m long, 6.2 m wide and 4.8 m high, with a reverberation time of about 0.9 s, independent of frequency. Background noise was reduced to a level of approximately 35 dBA. Four small dynamic loudspeakers were used as signal sources. Three of these were located nearby the listener and were positioned in front, to the right and to the left, respectively. The listener was sitting in the centre of the room and was facing one of the corners, where the fourth loudspeaker was located. The near loudspeakers were situated at ear level at a distance of 0.8 m from the listener. The far loudspeaker was located at a distance of 4.2 m, approximately

2 m above the ground, with the baffle facing the listener but tilted upwards. A schematic diagram of the testing configuration is shown in the Figure. Though the near loudspeakers were practically at the calculated critical distance of the room (0.83 m), their directional characteristics insured that the listener was indeed in the direct sound field: measurements showed that the level of speech noise at a distance of 0.8 m from the loudspeaker was 9 dBA above the level of the indirect sound. It was also verified that the listener was in the indirect field of the far loudspeaker. The upward tilt given to this loudspeaker reduced the effect of its directivity.

The test comprised nine different listening conditions, seven of which were presented twice. The order of the conditions was fixed. In all cases, the speech was reproduced in front of the listener. In the first two conditions (Q_N and Q_F), which were presented only once, the speech was reproduced in quiet over either the near or the far loudspeaker. In the conditions N_N, N_R and N_L, both speech and (steady-state) noise were presented from nearby, with the noise coming from the front, the right side, or the left side, respectively. In condition N_F, the noise was reproduced over the far loudspeaker and the speech over the near one in front; the reverse situation was tested in condition S_F. The two last conditions, M_N and M_F, were the same as conditions N_N and N_F, except for the use of modulated instead of steady-state noise.

For each condition, the speech-reception threshold (SRT) was determined by varying the presentation level of the sentences using an simple up-down

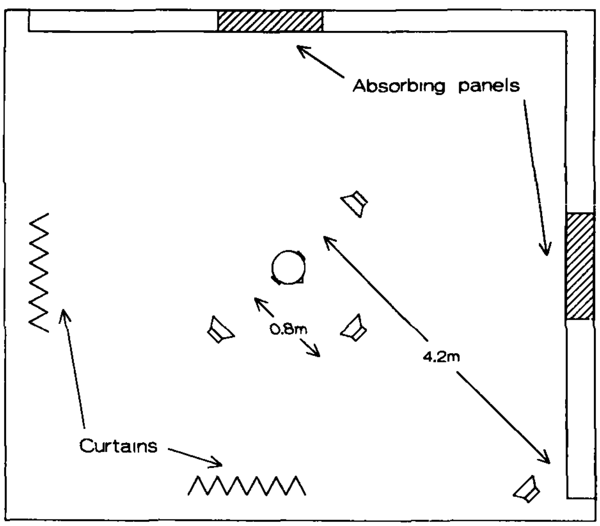

FIG. Schematic diagram of the experimental configuration.

procedure with as criterion the correct replication of the entire sentence (Plomp and Mimpen, 1979). Lists of either nine (only the conditions Q_N and Q_F) or eight sentences were presented per condition. The sentence lists were always used in the same order, as listed in Table I. No balanced order was possible because the test was intended for individual assessment of speech perception. In addition, a fixed order has the advantage that it allows a more accurate comparison between results obtained from different subjects. The step size used during the adaptive procedure was 4 dB for each first sentence of a list, and 2 dB for the remaining sentences. The SRT was taken as the average presentation level of the third and subsequent sentences. In all conditions, the presentation level of the noise was fixed at a level of 60 dBA, as measured with a microphone located at the position of the centre of the listener's head. Thus, an additional amplification of 9 dB had to be applied to the noise signal fed into the far loudspeaker. The noise level of 60 dBA was chosen because the speech level would then vary in most conditions between 50 and 60 dBA, which is the range of levels normally encountered during conversations in quiet and in noise (Plomp, 1978).

Prior to the presentation of the sentence lists, the listener's air-conduction tone audiogram was measured at octave frequencies ranging from 250 to 8000 Hz using a modified Békésy procedure. Tone bursts with a duration of 200 ms were presented at a rate of approximately one per second and varied in level in steps of 2 dB. Levels were calibrated according to ISO 389. Measurement continued until eight reversals had occurred. The threshold was taken as the average presentation level at the last six reversals.

5.2 C. Subjects

Ten normal-hearing and 18 hearing-impaired listeners were used as subjects. The normal-hearing subjects, 6 male and 4 female, with ages ranging from 23 to 29, had hearing levels of 20 dB or less at all octave frequencies. The hearing-impaired group consisted of 8 male and 10 female subjects, aged 22-58, with a median age of 46. They all had a sensorineural hearing-impairment at both ears, confirmed by bone-conduction audiometry (performed during a previous visit to the clinic), and no indication of retrocochlear pathology. The PTAs (average hearing levels at 500, 1000 and 2000 Hz) of these subjects ranged from 11.3 to 55.3 dB (average 31.8 dB) for their better ears, and from 31-57 dB (average 38.3 dB) for their poorer ears. The individual PTAs are shown in Table II. One hearing-impaired subject wore a hearing aid at his right ear; three others wore binaural aids.

5.3 RESULTS AND DISCUSSION

5.3 A. Normal-hearing listeners

The means and standard deviations of the SRTs obtained from the normal-hearing listeners are listed in Table 1. The SRTs for the noise conditions are expressed as S/N ratio, and were obtained by averaging test and retest results. Test-retest differences were relatively small: less than ± 1 dB on the average, with standard deviations in the order of 1.5 dB. The SRTs for both quiet conditions do not represent true thresholds in quiet since they were, in fact, determined by the background noise. (These conditions were actually meant for the hearing-impaired subjects and served to check whether their SRTs in noise were affected by threshold effects.) Comparison of the results for the conditions N_N, N_R, N_L and M_N, where the listener was located in the direct field of the loudspeakers, with free-field results obtained previously (Plomp and Mimpen, 1981; Bronkhorst and Plomp, 1988; Bronkhorst and Plomp, 1990), shows that there is an excellent agreement for the N_N and M_N conditions, but a discrepancy of about 4 dB for the N_L and N_R conditions. The latter can be explained from the fact that, in a free field, the binaural advantage caused by spatial separation of speech and noise sources results primarily from headshadow effects (Bronkhorst and Plomp, 1988). As the interaural attenuation due to headshadow reaches in the high frequencies values of 20 dB or more (Blauert, 1974), a reduction of its effect can be expected in the present reverberant test environment, where the direct sound was only about 9 dB higher in level than the indirect sound.

The contributions of the various cues, presented to the listener, can be derived from the differences between the N_N SRT and the SRTs for the

TABLE 1. Mean results and standard deviations for the 10 normal-hearing subjects, the 14 unaided, and the 4 aided hearing-impaired subjects. The table also lists which sentence lists were used for each condition.

Condition	Sentence lists	Normal hearing 10 ss.		Hearing impaired			
				14 unaided		4 aided	
		mean	s.d.	mean	s.d.	mean	s.d.
Q_N	1	26.8	1.1	38.3	4.7	36.2	4.2
Q_F	2	37.5	1.4	49.4	5.2	46.2	5.8
N_N	3,12	-6.4	0.9	-3.8	1.8	-3.1	0.6
N_R	4,11	-11.8	1.1	-6.8	2.0	-7.6	2.2
N_L	5,10	-12.5	0.9	-8.1	2.0	-6.9	1.6
N_F	6,14	-7.6	1.1	-4.4	1.7	-3.1	1.2
S_F	7,13	3.5	1.0	6.6	1.8	6.3	1.3
M_N	8,16	-13.2	0.9	-5.9	2.6	-5.9	1.0
M_F	9,15	-12.6	0.6	-5.6	3.1	-5.1	2.4

remaining conditions. For the N_L, N_R and M_N conditions, this yields the following simple equations:

$$SRT(N_N) - SRT(N_R) = B_R = 5.4 \text{ dB} \quad (1)$$
$$SRT(N_N) - SRT(N_L) = B_L = 6.1 \text{ dB} \quad (2)$$
$$SRT(N_N) - SRT(M_N) = F = 6.8 \text{ dB}, \quad (3)$$

where B_R and B_L are the binaural advantages (due to both interaural time and level differences) resulting from shifting the noise source to the right and left, respectively, and where F is the advantage due to (speech-like) modulations of the masker.

More complex equations result for the N_F, S_F and M_F conditions, as the relevant effects cannot be directly obtained from the differences with the N_N SRT. Firstly, a level correction factor L has to be included, representing the true level difference (at the ears of the listener) between the direct and indirect sound fields. This factor adds to the SRT differences for the N_F and M_F conditions, but is almost compensated for by the 9 dB amplification given previously to the noise signal. It adds negatively to the SRT difference for the S_F condition. Secondly, there will in all three conditions be positive contributions due to binaural unmasking, resulting from the interaural decorrelation of the noise or speech signals. In the equations shown below, this is accounted for by the factors U_N and U_S, respectively. Thirdly, the reverberation will cause the speech to become less intelligible. The resulting increase in SRT is represented by the factor R. Finally, the advantage due to masker modulations in the M_F condition, which are now smoothed as a result of the reverberation, is represented by the factor F'. Thus, assuming additivity of these factors, the following equations are obtained:

$$SRT(N_N) - SRT(N_F) = U_N + L - 9 = 1.2 \text{ dB} \quad (4)$$
$$SRT(N_N) - SRT(M_F) = F' + U_N + L - 9 = 6.2 \text{ dB} \quad (5)$$
$$SRT(N_N) - SRT(S_F) = U_S - L - R = -9.9 \text{ dB}. \quad (6)$$

Combination of eqs. (4) and (5) directly yields $F' = 5$ dB. Comparison with eq. (3) shows that the smoothing causes a reduction of the advantage of almost 2 dB. The value of R can be estimated from results reported by Duquesnoy (1980), who measured monaural SRTs in noise as a function of reverberation time, using the same test material. For a reverberation time of 0.9 s, his data yield (by interpolation) $R = 4.5$ dB. By entering this value into eq. (6), and assuming that $U_N \approx U_S$, one obtains $U_N \approx U_S = 2.4$ dB and $L = 7.8$ dB (results for both tonal stimuli and speech show that the unmasking will, in fact, be somewhat larger when the signal contains interaural differences and the masker is the same at both ears, than in the reverse case, see eg. Licklider, 1948; Jeffress *et al.*, 1952; Carhart *et al.*, 1967). The derived value for U_N deviates from results of Plomp and Mimpen (1979), who obtained an advantage of only 0.7 dB for diffuse versus diotic noise. This suggests that the above assumptions have introduced inaccuracies in the derivation, so that L may, in fact, have a value closer to 9 dB. A discrepancy will, however, remain, as a result of the difference between the open-ear responses for frontal and random sound incidence (Killion, 1979).

5.3 B. Hearing-impaired listeners

Inspection of the individual data obtained from the hearing-impaired listeners showed that the SRTs for the noise conditions were always higher than the corresponding SRTs obtained in quiet. This insured that the data were not contaminated by simple threshold effects. The thresholds in quiet are not substantially elevated because all subjects had mild to moderate hearing losses, while the subjects with poorest hearing wore one or two hearing aids.

In order to get an impression of the average performance of the hearing-impaired listeners, means and standard deviations were calculated of the SRTs obtained from the 14 subjects without hearing aids and from the 4 subjects with hearing aids. Test and retest results for the noise conditions were averaged. The results, expressed as S/N ratio, are shown in Table 1. For all conditions and for both groups of hearing-impaired listeners, there is a significant difference with the mean SRTs obtained from the normal-hearing listeners (two-sample t tests, $p<0.001$). The mean SRTs for the N_N and M_N conditions agree well with free-field results obtained previously from similar groups of sensorineurally impaired listeners (Bronkhorst and Plomp, 1989, 1990). Analysis of test-retest differences for the noise conditions yielded mean absolute values of less than 1 dB, as for the normal hearing subjects, but slightly larger standard deviations, ranging from 1.5 to 2 dB.

Individual results for all 18 listeners are given in Table 2. The first column lists the SRTs obtained with speech and steady-state noise presented from nearby. The five following columns list SRT differences between the N_N condition and the N_R, N_L, N_F, S_F and M_N conditions, respectively. The seventh column lists the SRT differences between the N_F and M_F conditions. Thus, columns two and three yield the gains B_R and B_L due to directional cues, columns four and five show the effects of interaural decorrelation (U_N and U_S) and of speech distortion due to reverberation (R), and columns six and seven show the gains F and F' resulting from normal and smoothed masker modulations, respectively. No attempt was made to compensate for the factor L in columns four and five, as no direct estimate of its magnitude could be made.

Shown at the bottom of the table are the means and standard deviations for all 18 hearing-impaired and for the 10 normal-hearing subjects, and the significance levels of the differences between corresponding means. Individual results differing more than two or three standard deviations from the mean normal results are indicated by a single or double asterisk, respectively. The data reveal several interesting aspects.

(1) It appears that most hearing-impaired subjects have significantly poorer

TABLE 2. Individual results for the 18 hearing-impaired subjects. Listed are the SRT for the condition where speech and (steady-state) noise were presented from the near loudspeaker, and SRT differences indicating the effects of changing the noise source's azimuth (B_R and B_L), of binaural unmasking (U_N and U_S) and speech distortion (R) due to reverberation, and of envelope fluctuations of the masker (F and F'). In addition, the hearing levels of the right and left ears, averaged over 500, 1000 and 2000 Hz, are given, and the sides where hearing-aids were worn during the test. The bottom rows show mean results and standard deviations for both the hearing-impaired and the normal-hearing subjects, and the significance level of the differences between the corresponding means. Individual results of the hearing impaired, differing more than two or three standard deviations from the mean normal result are indicated by a single or double asteriks.

Subject nr.	SRT(N_N)	B_R	B_L	U_N+L-9	U_S-L-R	F	F'	PTA right	PTA left	Hearing aid
1	-4.0*	2.0**	2.9*	0.3	-11.1	2.6*	2.3*	31.0	11.3	
2	-5.4	1.7**	5.1	0.9	-10.0	3.7	1.7**	16.7	31.0	
3	-6.3	1.7**	2.0**	-2.0*	-10.6	2.3*	0.6**	22.0	25.0	
4	-0.3**	0.6**	3.7*	0.6	-11.4	0.6**	-1.7**	26.3	27.0	
5	-6.3	0.6**	2.0**	0.6	-12.3*	0.6**	-1.4**	26.3	30.3	
6	-3.1**	3.7	4.6	0.9	-8.3	2.3*	1.1**	26.3	38.3	
7	-4.0*	5.1	6.6	0.0	-9.4	4.3	2.9*	27.7	26.7	
8	-5.7	2.6*	6.0	-0.3	-10.6	2.3*	4.3	28.0	34.3	
9	-4.0*	4.0	4.3	1.4	-10.6	3.7	2.3*	28.3	31.7	
10	-3.4**	2.9*	3.4*	-1.1	-11.7	-1.1**	-0.3**	30.7	28.7	
11	-3.4**	5.1	4.9	0.9	-10.9	2.3*	2.9*	33.0	39.7	
12	-3.1**	2.6*	6.0	2.3	-10.6	2.9*	2.0*	33.7	49.7	
13	-0.9**	4.9	5.1	1.1	-8.3	0.6**	-1.7**	37.7	39.3	
14	-3.4**	4.3	3.4*	2.9	-9.4	2.6*	1.4**	40.0	42.0	
15	-4.0*	4.0	1.4**	-0.6	-9.7	2.9*	4.0	48.7	33.3	r
16	-2.6**	7.1	6.3	2.0	-10.6	4.0	1.1**	43.7	50.3	r+l
17	-3.1**	4.9	4.6	-0.6	-8.0	2.6*	2.9*	55.3	54.3	r+l
18	-2.9**	1.7**	2.9*	-0.9	-9.4	1.7**	-0.3**	57.0	55.3	r+l
mean	-3.7	3.3	4.2	0.5	-10.2	2.3	1.3	34.0	36.0	
s.d.	1.6	1.8	1.6	1.3	1.2	1.4	1.9	11.1	11.4	
Normal hearing										
mean	-6.4	5.4	6.1	1.2	-9.9	6.8	5.0			
s.d.	0.9	1.1	1.1	1.4	1.1	1.6	1.0			
p	<0.001	0.003	0.002	0.18	0.60	<0.001	<0.001			

speech perception in steady-state noise than the normal hearing. The mean SRT difference is 2.7 dB, which corresponds, given the slope of 15-20%/dB in the steepest part of the sentence intelligibility function (Plomp and Mimpen, 1979), to a reduction in sentence intelligibility of up to 50%. Comparison of the SRTs and the PTAs indicates that performance tends to decrease as a function of hearing loss. Statistical analysis of the data shows, however, that this dependence is not significant (r=0.4, p=0.1). Both the reduced performance in noise of sensorineurally hearing-impaired listeners and the weak relationship with the audiogram and the hearing loss for speech in quiet correspond with earlier results (eg. Plomp, 1986).

(2) The binaural gain due to spatial separation of the speech and noise sources is, surprisingly, rather small for the listeners with light hearing losses (in terms of PTA), and almost normal for most listeners with poorer hearing. Closer inspection of the tone audiograms reveals that almost all listeners with reduced binaural gain had severely impaired high-frequency hearing on the side contralateral to where the noise source was shifted to. This corresponds with earlier findings (Bronkhorst and Plomp, 1989), indicating that the binaural gain, or, specifically, that part of the gain resulting from head shadow, decreases with increasing hearing loss in the high frequencies, because that is the frequency region where there is maximum attenuation of the noise reproduced at the contralateral side of the head. A hearing aid may have a favourable effect on the binaural gain as it can prevent the high-frequency components from becoming subthreshold. The results seem to support this, as the four subjects using aids had about the same high-frequency hearing loss as subjects 1 to 5, while they attained larger binaural gains.

(3) The relative effects of reverberation on speech intelligibility appear to be essentially the same for both groups of subjects. In other words, the hearing impaired listeners, including those wearing hearing aids, benefit equally from the interaural decorrelation of speech or noise as the normal hearing, and experience also a similar disadvantage due to the distortion of the speech. The former corresponds with previous findings, showing normal or nearly-normal binaural unmasking due to interaural time delay for groups of sensorineurallly hearing-impaired listeners (Bronkhorst and Plomp, 1989, 1990). The latter agrees with results of Duquesnoy and Plomp (1980), who investigated the performance of elderly hearing-impaired subjects and found that these suffered approximately equal relative SRT losses when listening to reverberated speech as did his normal-hearing subjects.

(4) Nearly all hearing-impaired subjects benefit considerably less from envelope fluctuations of the masker than do the normal hearing. On the average, the difference in SRT is in the order of 4 dB. They furthermore appear to be fairly consistent in their behaviour, as comparison of the F and F' values yields a correlation coefficient of 0.67 (p<0.0002). The reduction in performance is partly due to threshold effects, occurring in the "valleys" of the masker envelope. However, this does not seem to be the major factor, as no clear dependence of the F values on hearing levels appears to be present. Other factors that may be involved as well are impaired temporal resolution, resulting in an effective smoothening of the perceived masker modulations, and, possibly, reduced comodulation masking release (Hall *et al.*, 1988).

In general, the data show that either one of the hearing impaired subjects has significantly poorer performance than normal for at least two of the elements tested. This means that they will all from time to time encounter situations where they will be more or less seriously handicapped. Perhaps this is one of the reasons why practically all sensorineurally hearing impaired complain about their ability to communicate in noisy circumstances.

The data furthermore fail to show clear differences in performance between the hearing-impaired listeners with and without hearing aids. This seems to contradict results of previous research, indicating, for sufficiently high signal levels, a superiority of unaided over aided speech perception in noise (Duquesnoy and Plomp, 1983; Festen and Plomp, 1986). This negative contribution is due to effects like signal distortion and internal noise introduced by the hearing aid, and, in case of binaural listening, to the unfavourable microphone position of behind-the-ear hearing aids, which has the effect of reducing the head shadow advantage. At lower signal levels, however, unaided performance will start to be influenced by threshold effects, and the relative benefit of hearing aids will increase. The choice of signal level is therefore crucial in comparing aided and unaided speech perception in noise. For a clinical test, aiming at assessment of the performance in everyday circumstances, it seems best to use moderate signal levels, representative of normal conversations, as was done in the present study. In view of the small number of aided subjects tested, the present results serve only as an indication. It will be of interest to investigate, using a larger group of hearing-aid users, whether hearing-aids, and especially in-the-ear aids, can indeed increase the head shadow advantage, as suggested above, and also whether the aids will allow the listener to benefit more from the relatively silent intervals of a fluctuating masker.

5.3 C. Optimization of the test

Application of a test like the present one in a clinical setting requires optimization for maximum efficiency. It would furthermore be advantageous when all conditions could be tested at least twice, not in order to obtain retest results for increased accuracy, as in the present design, but for comparison of listening with and without hearing aids, or with different types of aids. Reduction of the number of conditions is therefore called for.

The first step would be to eliminate the quiet conditions (Q_N and Q_F). When a signal level equal to that of normal conversational speech is used in the test, the presence of threshold effects need not be detected, and should not be compensated for, as these represent a "real" factor contributing to the handicap of the listener. The consequence is, however, that the hearing loss for the N_N condition cannot be interpreted as pure "class D" hearing loss, in terms of Plomp's (1986) signal-to-noise ratio model, as it may be a mixture of "class A" and "class D" hearing loss.

From the similarity of the F and F' values it can be concluded that the M_F condition provides essentially the same information as the M_N condition, so that it may be omitted as well. As for the conditions N_F and S_F, these seem to be of limited interest, as the results show that the hearing-impaired subjects hardly deviated from normal in their performance for these conditions. Another argument against inclusion of these conditions is that they require availability of a testing room with a relatively long reverberation time. An argument in favour is that it remains to be seen whether (aided) listeners with, for instance, severe losses or large asymmetries will also show normal performance. When a reverberation condition is to be included, the S_F condition is the best candidate, as it tests the effects of both binaural unmasking and speech distortion.

The optimized test thus comprises conditions N_N, N_R, N_L, M_N and, optionally, S_F. This means, given the set of 130 sentences used in the present study, that the conditions can be tested twice or even three times, while maintaining sufficiently long sentence lists. Care should, however, be taken to prevent that the comparisons of results for corresponding conditions are contaminated by differences between sentence lists.

5.4 CONCLUSION

A test for the assessment of speech perception in noise can be highly useful in clinical practice. It provides information for the determination of the handicap of a hearing-impaired individual, and it can be used for the selection of the optimal hearing aid. The test described in the present study employs stimuli and conditions that not only simulate actual everyday circumstances as closely as possible, but also allow separate evaluation of the effects of binaural cues, reverberation and masker fluctuations. The results demonstrate that, on the average, hearing-impaired listeners have (1) higher thresholds for speech in steady-state noise, (2) less benefit from a spatial separation of speech and noise sources, and, (3) less advantage due to envelope fluctuations present in the masker than normal-hearing listeners. In fact, either one of the hearing-impaired subjects tested appeared to have significantly poorer performance than normal for at least two of the above aspects. This finding may explain why so many hearing impaired complain about their ability to communicate in noisy situations. The results show on the other hand, that both the negative and positive effects of reverberation, i.c. the increase in threshold for reverberated speech and the unmasking due to interaural signal decorrelation, are essentially normal for all hearing-impaired listeners tested. Furthermore, the results for four aided hearing-impaired listeners indicate that, at moderate signal levels, the aids may improve speech perception in (fluctuating) noise by preventing or diminishing threshold effects. The results warrant the conclusion that, in order to prevent underestimation of the hearing handicap, a clinical test of speech perception in noise should include conditions with both steady-state and fluctuating interference, and also with and without spatial separation of the signal sources.

GENERAL CONCLUSION

In this thesis, a series of experiments on binaural speech perception in noise have been presented. The focus of study has advanced from rather fundamental aspects, e.g., the contributions of headshadow (ILD) and interaural time delay (ITD) to the advantage of binaural listening in normal and impaired hearing, to more applied aspects, e.g., the performance of individual hearing-impaired listeners, aided or unaided, in a noisy and reverberant environment. Thus the present study may be of value for both the experimental audiologist, interested in the basic mechanisms of binaural hearing, and the clinical audiologist, concerned with the hearing handicap of his hearing-impaired patients.

The experimental results lead to a number of conclusions; the following are based on the results obtained from normal-hearing listeners:

(1) The total advantage, in terms of SRT, due to the binaural cues ILD and ITD is less than the sum of the gains caused by both cues individually. The presence of ILD reduces the effectivity of ITD.

(2) The binaural advantage due to ITD, present during sound field listening, is at most 2-3 dB, whereas the contribution of ILD can reach values of 7 dB or more. The ILD advantage, however, decreases when reverberation is present, and when the number of maskers is increased.

(3) A 20-dB interaural difference in overall presentation level, simulating a one-sided conductive hearing loss, has but a limited effect on the binaural advantage due to ITD.

(4) Modulating a noise masker like speech results in a reduction of masking efficiency of about 5 dB. When four independently modulated maskers are mixed (and adjusted for equal total power), the masking efficiency is only 1 dB lower than that of steady-state noise.

The following conclusions are based on the results obtained from sensorineurally hearing-impaired listeners:

(5) In general, the hearing impaired benefit almost equally from ITD and interaural signal decorrelation as do the normal hearing.

(6) The binaural advantage due to ILD decreases with an increasing high-frequency hearing loss at the ear presented with the most favorable S/N ratio. Listeners with severely impaired high-frequency hearing experience no ILD advantage at all.

(7) An interaural difference in overall presentation level in the order of 20 dB hardly affects binaural unmasking due to ITD.

(8) Optimally, a hearing aid does not interfere with unmasking due to ITD, while it may increase the ILD advantage by preventing or diminishing threshold effects.

(9) The hearing impaired derive almost no benefit from speech-like modulations, present in the noise masker.

(10) In order to prevent underestimation of the hearing handicap, a clinical

test of speech perception in noise should employ both fluctuating and steady-state interfering sound, and should use presentation of speech and interfering sound from both the same and from different directions.

SAMENVATTING

Dat de mens beter hoort met twee oren dan met één oor, lijkt een evident, zelfs triviaal, gegeven. De voordelen van het binauraal horen zijn evenwel veelomvattender en complexer van aard dan men op het eerste gezicht zou vermoeden. Allereerst is er het vermogen tot richtinghoren, dat voornamelijk is gebaseerd op een ingewikkeld verwerkingsproces van tussen beide oren optredende verschillen in aankomsttijd en geluidniveau. Een tweede belangrijk voordeel is de betere waarneming van geluiden in lawaai. Eenieder kan dit effect voor zichzelf eenvoudig vaststellen door bij het volgen van een conversatie in een rumoerige omgeving één oor af te sluiten. Men zal merken dat het verstaan dan onmiddellijk veel moeilijker wordt. De binaurale winst in een dergelijke situatie kan niet worden verklaard uit een eenvoudige optelling van de intensiteiten van de beide monauraal waargenomen geluiden. Aangezien dan zowel spraak als lawaai worden opgeteld, zou de intensiteitsverhouding tussen nuttig signaal en ruis (S/R verhouding) ongewijzigd blijven, zodat er geen winst zou resulteren. Er is integendeel gebleken dat het proces meer lijkt op een soort aftrekking van de linker en rechter signalen, voorafgegaan door intern optredende tijdsvertragingen. Het effect treedt op wanneer de aankomsttijden aan beide oren van spraak en lawaai verschillend zijn; dit is meestal het geval wanneer spraak en lawaai uit verschillende richtingen komen.

Dit opmerkelijke vermogen om stoorlawaai enigermate te onderdrukken is in de 40'er jaren ontdekt en sindsdien uitgebreid wetenschappelijk onderzocht. De aandacht heeft zich hierbij steeds meer verlegd naar per hoofdtelefoon aangeboden artificiële geluidstimuli. Zodoende werd minder aandacht besteed aan de invloed van interaurale geluidniveauverschillen, veroorzaakt door de schaduwwerking van het hoofd, die bij het normaal luisteren eveneens optreden. Voor een goed begrip van de invloed van het binauraal horen op het spraakverstaan in dagelijkse omstandigheden is het evenwel noodzakelijk de werking van beide effecten - en ook hun wisselwerking - te bestuderen. Ook is het van belang hierbij niet alleen normaalhorende, maar ook slechthorende luisteraars te gebruiken, om meer inzicht te krijgen in de problemen die de laatsten ondervinden in lawaaiige omstandigheden. Deze beide overwegingen vormden de leidraad van het onderzoek dat in dit proefschrift is beschreven.

Een bijzonder aspekt van de hier gebruikte experimentele methode is dat er een scheiding tot stand werd gebracht van de beide bovengenoemde effecten. Daartoe werden geluidopnamen gemaakt in een reflectievrije ruimte met een met microfoons uitgerust kunsthoofd; de opnamen werden vervolgens gedigitaliseerd, met de computer bewerkt, en tenslotte via hoofdtelefoons aan luisteraars aangeboden. De computerbewerking was erop gericht, signalen te genereren met ofwel enkel interaurale tijdverschillen (ITD, interaural time differences), ofwel enkel interaurale niveauverschillen (ILD, interaural level differences). Door gebruik te maken van deze signalen en van de oorspronkelijke opnamen konden de bijdragen van ITD en ILD zowel apart als in combinatie worden bepaald. Het opgenomen materiaal bestaat uit lijsten met korte, eenvoudige zinnen, en ruis met hetzelfde gemiddelde frequentiespectrum als de spraak (Plomp en Mimpen, 1979). Er werden zinnen gebruikt in plaats van losse woorden, omdat deze de alledaagse spraak dichter benaderen. De drempel voor de verstaanbaarheid van de zinnen (SRT, speech-reception threshold) werd met een adaptieve methode bepaald. Hierbij wordt het aanbiedingsniveau 2 dB verlaagd wanneer een zin in zijn geheel correct wordt nagezegd, en 2 dB verhoogd, wanneer er een fout wordt gemaakt. Deze methode geeft een efficiënte en nauwkeurige schatting van de drempel. Een nadeel is dat hij minder goed kan worden gebruikt bij ernstig slechthorenden, omdat zij vaak niet aan het criterium zullen kunnen voldoen.

De kunsthoofdopnamen en de met de computer bewerkte signalen zijn in verschillende experimenten, beschreven in de Hoofdstukken 2 en 3, aangeboden aan in totaal 34 normaalhorende en 34 perceptief slechthorende luisteraars. De daarbij gesimuleerde configuratie bestond uit een spraakbron recht voor de luisteraar (een hoek Φ van 0°), en één stationaire ruisbron onder een hoek Φ in het horizontale vlak, met Φ tussen 0° en 180°. Doordat de signalen per hoofdtelefoon aangeboden werden bestond de mogelijkheid bij de normaalhorenden éénzijdige slechthorendheid of doofheid te simuleren door één kanaal te verzwakken of uit te schakelen. Een dergelijke verzwakking (of versterking) kon ook bij de slechthorenden worden toegepast, om de werking van een éénzijdig hoortoestel te imiteren. De resultaten laten allereerst zien dat de spraakdrempels in ruis (SRT) voor de slechthorenden duidelijk hoger zijn dan voor de normaalhorenden. De gemiddelde verschillen voor de diverse luistercondities variëren tussen de 3 en 7 dB. Voor beide groepen geldt dat het naar opzij brengen van de ruisbron de SRT fors doet dalen: gemiddeld is de maximale winst 10 dB bij de normaalhorenden en ongeveer 6 dB bij de slechthorenden. Het blijkt dat deze winst minder is dan de som van de individuele effecten van ITD en ILD; 2 à 3 dB raakt verloren doordat het "aftrekkings"proces ten gevolge van ITD blijkbaar minder efficiënt verloopt wanneer er tegelijkertijd ILD aanwezig is. Verder blijkt uit de vergelijking van monaurale en binaurale drempels dat de door ILD veroorzaakte winst uitsluitend door het oor met de gunstigste S/R verhouding wordt bepaald (in de gemeten situatie is dit het oor aan de van de ruisbron àf gerichte zijde), en dus in feite een monauraal effect is. Doordat de hoofdschaduw toeneemt als functie van de frequentie geldt die gunstige S/R verhouding enkel voor de hoge frequenties. Er is dan ook gevonden dat slechthorenden met een hoge tonen gehoorverlies minder profijt hebben van de hoofdschaduw, en dat hun prestatie afhangt van het aanbiedingsniveau van de signalen aan de "gunstige" zijde. De winst ten gevolge van ITD blijft daarentegen vrijwel gelijk wanneer de signalen aan één zijde worden verzwakt of versterkt. Dat de werking van ITD wèl door ILD en nauwelijks door een éénzijdige verzwakking of versterking wordt beïnvloed komt waarschijnlijk doordat in het laatste geval geen verandering optreedt van de S/R verhoudingen aan beide oren. Dit resultaat geeft aan dat het aan één zijde dragen van een hoortoestel nauwelijks invloed zou moeten hebben op de winst ten gevolge van ITD.

In een volgend tweetal experimenten, beschreven in Hoofdstuk 4, is wederom gebruik gemaakt van de kunsthoofdopnamen, maar deze waren nu bewerkt en in verschillende combinaties bij elkaar opgeteld om het effect van meerdere storende sprekers te kunnen simuleren. De bewerking bestond uit een modulatie van de ruis overeenkomstig de fluctuaties van de omhullende van een doorlopend spraaksignaal, waarbij de gemiddelde intensiteit van het signaal gelijk gehouden werd (Festen, 1990). De gesimuleerde configuratie bestond uit een spraakbron die recht voor de luisteraar is geplaatst, en tussen 1 en 6 ruisbronnen, alle verschillend gemoduleerd, die ofwel ook recht voor de luisteraar, ofwel in een symmetrische of asymmetrische verdeling rond de luisteraar zijn geplaatst. Als proefpersonen werden 17 normaalhorende en 17 perceptief slechthorende luisteraars gebruikt. Het meest opmerkelijke van de gevonden resultaten is dat de slechthorenden nauwelijks van de fluctuaties van het maskeersignaal blijken te kunnen profiteren, zelfs niet diegenen met slechts lichte gehoorverliezen. De normaalhorenden behalen daarentegen tot 5 dB winst doordat zij blijkbaar effectief van de "gaten" in het stoorsignaal gebruik kunnen maken. Deze winst vermindert wanneer het aantal stoorbronnen toeneemt: bij 4 stoorbronnen is de maskerende werking vrijwel gelijk aan die van stationaire ruis. De verdere resultaten zijn vergelijkbaar met die van de vorige experimenten: de SRTs voor de slechthorenden zijn duidelijk verhoogd (gemiddeld 4 tot 10 dB), de door de hoofdschaduw veroorzaakte winst is minder bij diegenen met een gehoorverlies voor de hoge tonen, en de winst ten gevolge van ITD is ongeveer gelijk voor beide groepen luisteraars. Een interessante gevolgtrekking op basis van de gevonden resultaten is dat een normaalhorende een spreker kan blijven verstaan wanneer er tot zes even luide storende sprekers om hem heen staan, terwijl een slechthorende met een matig perceptief gehoorverlies het al moet opgeven als er meer dan drie storende sprekers zijn.

Het laatste experiment, dat in Hoofdstuk 5 beschreven staat, was gericht op de ontwikkeling van een klinisch toepasbare test voor binauraal spraakverstaan in lawaai. Er werd gekozen voor een luisteromgeving met een zekere nagalm, waarbij de luisteraar zich in het direkte of indirekte veld van de spraakbron en/of de ruisbron bevond. Zodoende kon de invloed van nagalm op het spraakverstaan worden bepaald. Verder waren er condities waarbij de ruis fluctuerend was in plaats van stationair, en condities waarbij de ruisbron aan de linker of rechter zijde van de luisteraar geplaatst was. Er werden 10 normaalhorende en 18 perceptief slechthorende proefpersonen gebruikt. Vier van de slechthorenden droegen één of twee hoortoestellen tijdens de test. De resultaten laten zien dat elk van de slechthorenden significant slechter scoort dan de gemiddelde normaalhorende op tenminste twee van de volgende onderdelen: (1) spraakverstaan in stationaire ruis, (2) binaurale winst ten gevolge van het naar opzij brengen van de ruisbron, en (3) winst veroorzaakt door fluctuaties van het maskeersignaal. Deze bevinding is wellicht de verklaring van het feit dat vrijwel alle slechthorenden klagen over slecht spraakverstaan in lawaai. Het blijkt daarentegen dat de slechthorenden van nagalm vrijwel evenveel last (door vervorming van de spraak), en ook evenveel profijt (door binaurale ruisonderdrukking net als bij ITD) hebben als de normaalhorenden. De resultaten laten verder zien dat de (kleine groep) geteste hoortoesteldragers niet afwijkend presteerde vergeleken met de andere slechthorenden. Het experiment toont aan dat het belangrijk is om bij een klinische test van spraakverstaan in rumoer niet alleen stationair maar ook fluctuerend stoorlawaai te gebruiken, en dat het tevens aanbeveling verdient de spraak en het lawaai niet alleen vanuit dezelfde, maar ook vanuit verschillende richtingen aan te bieden.

REFERENCES

Abbagnaro,L.A., Bauer,B., and Torick,E.L. (1975). "Measurements of diffraction and interaural delay of a progressive sound wave caused by the human head-II," J. Acoust. Soc. Am. 58, 693-700.

Blauert,J. (1974). *Räumiches Hören.* Hirzel, Stuttgart.

Bocca,E., and Antonelli,A.R. (1976). "Masking level difference: Another tool for the evaluation of peripheral and cortical defects," Audiology 15, 480-487.

Bronkhorst,A.W., and Plomp,R. (1988). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. 83, 1508-1516.

Bronkhorst,A.W., and Plomp,R. (1989). "Binaural speech intelligibility in noise for hearing-impaired listeners," J. Acoust. Soc. Am. 86, 1374-1383.

Bronkhorst,A.W., and Plomp,R. (1990). "Binaural speech intelligibility for simulated cocktail-party conditions,' Submitted to J. Acoust. Soc. Am.

Burkhard,M.D., and Sachs,R.M. (1975). "Anthropometric manikin for acoustic research," J. Acoust. Soc. Am. 58, 214-222.

Carhart,R. (1965). "Monaural and binaural discrimination against competing sentences," Intern. Audiol. 4, 5-10.

Carhart,R., Tillman,T.W., and Johnson,K.R. (1967). "Release of masking for speech through interaural time delay," J. Acoust. Soc. Am. 42, 124-138.

Carhart,R., Tillman,T.W., and Greetis,E.S. (1969a). "Release from multiple maskers: effects of interaural time disparities," J. Acoust. Soc. Am. 45, 411-418.

Carhart,R., Tillman,T.W., and Greetis,E.S. (1969b). "Perceptual masking in multiple sound background," J. Acoust. Soc. Am. 45, 694-703.

Colburn,H.S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise," J. Acoust. Soc. Am. 61, 525-533.

Dirks,D.D., and Wilson,R.H. (1969a). "The effect of spatially separated sound sources on speech intelligibility," J. Speech. Hear. Res. 12, 5-38.

Dirks,D.D., and Wilson,R.A. (1969b). "Binaural hearing of speech for aided and unaided conditions," J. Speech Hear. Res. 12, 650-664.

Duquesnoy,A.J., and Plomp,R. (1980). "Effect of reverberation and noise on the intelligibility of sentences in cases of presbycusis," J. Acoust. Soc. Am. 68, 537-544.

Duquesnoy,A.J. (1982). "Speech intelligibility of the hearing impaired," Doctoral dissertation, Free University, Amsterdam.

Duquesnoy,A.J., and Plomp,R. (1983) "The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise," J. Acoust. Soc. Am. 73, 2166-2173.

Duquesnoy, A.J. (1983). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. 74, 739-743.

Durlach,N.I., and Colburn,H.S. (1978). "Binaural phenomena," In *Handbook of Perception, Vol IV: Hearing*, edited by E.C. Carterette and M.P. Friedman (Academic, New York).

Durlach,N.I., Thompson,C.L., and Colburn,H.S. (1981). "Binaural interaction in impaired listeners," Audiology 20, 181-211.

Egan,J.P. (1965). "Masking-level differences as a function of interaural disparities in intensity of signal and of noise," J. Acoust. Soc. Am. 38, 1043-1049.

Elliot,L.L. (1975). "Temporal and masking phenomena in persons with sensorineural hearing loss," Audiology 14, 336-353.

Festen,J.M., and Plomp,R. (1986) "Speech-reception threshold in noise with one and two hearing aids," J. Acoust. Soc. Am. 79, 465-471.

Festen,J.M. (1987). "Speech-reception threshold in a fluctuating background sound and its possible relation to temporal auditory resolution," in *The Psychophysics of Speech Perception*, edited by M.E.H. Schouten (NATO ASI Series D, Martinus Nijhoff, The Netherlands).

Festen,J.M. (1990). "Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing," Submitted to J. Acoust. Soc. Am.

French,N.R., and Steinberg,J.C. (1947). "Factors governing the intelligibility of speech sounds," J. Acoust. Soc. Am. 19, 90-119.

Gelfland,S.A., Ross,L., and Miller,S. (1988). "Sentence reception in noise from one versus two sources: Effects of aging and hearing loss," J. Acoust. Soc. Am. 83, 248-256.

Hall,J.W., Davis,A.C., Haggard,M.P., and Pillsbury,H.C. (1988). "Spectro-temporal analysis in normal-hearing and cochlear-impaired listeners," J. Acoust. Soc. Am. 84, 1325-1331.

Hirsh,I.J. (1948). "The influence of interaural phase on interaural summation and inhibition," J. Acoust. Soc. Am. 20, 536-544.

Hirsh,I.J. (1950). "The relation between localization and intelligibility," J. Acoust. Soc. Am. 22, 196-200.

Houtgast,T., and Steeneken,H.J.M. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility," Acustica 28, 66-73.

Houtgast,T., and Steeneken,H.J.M. (1985). " A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acoust. Soc. Am. 77, 1069-1077.

Jeffress,L.A., Blodgett,H.C., and Deatherage,B.H. (1952). "The masking of tones by white noise as a function of the interaural phases of both components," J. Acoust. Soc. Am. 25, 523-527.

Jerlvall,L., Almqvist,B., Ovegard,A., and Arlinger,S. (1983). "Clinical trial of in-the-ear hearing aids," Scand. Audiol. 12, 63-70.

Killion,M.C. (1979). "Equalization filter for eardrum-pressure recording using a KEMAR manikin," J. Audio. Engin. Soc. 27, 13-16.

Kock,W.E. (1950). "Binaural localization and masking," J. Acoust. Soc. Am. 22, 801-804.

Koenig,W. (1950). "Subjective effects in binaural hearing," J. Acoust. Soc. Am. 22, 61-62.

Koenig,A.H., Allen,J.B., and Berkley,D.A. (1977). "Determination of masking-level differences in a reverberant environment," J. Acoust. Soc. Am. 61, 1374-1376.

Kryter,K.D. (1962). "Methods for the calculation and use of the articulation index," J. Acoust. Soc. Am. 34, 1689-1697.

Kuhn,G.F. (1977). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. 62, 157-167.

Leeuw,A.R., and Dreschler,W.A. (1987). "Speech understanding and directional hearing for hearing-impaired subjects with in-the-ear and behind-the-ear hearing aids," Scand. Audiol. 16, 31-36.

Levitt,H., and Rabiner,L.R. (1967a). "Binaural release from masking for speech and gain in intelligibility," J. Acoust. Soc. Am. 42, 601-608.

Levitt,H., and Rabiner,L.R. (1967b). "Predicting binaural gain in intelligibility and release from masking for speech," J. Acoust. Soc. Am. 42, 820-829.

Licklider,J.C.R. (1948). "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. 20, 150-159.

Lidén,G. (1976). "Directional audiometry," Scand. Audiol. 5, 93-96.

Lochner,J.P.A., and Burger,J.F. (1961). "The binaural summation of speech signals," Acustica 11, 313-317.

Ludvigsen,C., Elberling,C., Keidser,G., and Poulsen,T. (1990). "Prediction of intelligibility of non-linearly processed speech," Acta Otolaryngol. (Stockh) Suppl. 469, 190-195.

MacKeith,N.W., and Coles,R.R.A. (1971). "Binaural advantages in hearing of speech," J. Laryngol. Otol. 85, 213-232.

Markides,A. (1977). *Binaural Hearing Aids* (Academic, London).

McFadden,D. (1968). "Masking-level differences determined with and without interaural disparities in masker intensity," J. Acoust. Soc Am. 44, 212-223.

Miller,G.A. (1947). "The masking of speech," Psychol. Bull. 44, 105-129.

Nordlund,B., and Fritzell,B. (1963). "The influence of azimuth on speech signals," Acta Otolaryng. 56, 632-642.

Olsen,W.O., Noffsinger,D., and Carhart,R. (1976). "Masking level differences encountered in clinical populations," Audiology 15, 287-301.

Platte,H.J. and vom Hövel,H. (1980). "Zur Deutung der Ergebnisse von Sprachverständlichkeitsmessungen mit Störschall im Freifeld," Acustica 45, 139-150.

Plomp,R. (1978). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am. 63, 533-549.

Plomp,R., and Mimpen,A.M. (1979). "Improving the reliability of testing the speech reception threshold for sentences," Audiology 18, 43-52.

Plomp,R., and Mimpen,A.M. (1981). "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech reception threshold for sentences," Acustica 48, 325-328.

Plomp,R. (1986). "A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired," J. Speech Hear. Res. 29, 146-154.

Pollack,I., and Pickett,J.M. (1958). "Stereophonic listening and speech intelligibility against voice babble," J. Acoust. Soc. Am. 30, 131-133.

Schubert,E.D. (1956). "Some preliminary experiments on binaural time delay and intelligibility," J. Acoust. Soc. Am. 28, 895-901.

Schubert,E.D., and Schultz,M.C. (1962). "Some aspects of binaural signal selection," J. Acoust. Soc. Am. 34, 844-849.

Shaw,E.A.G. (1974). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. 56, 1848-1861.

Smoorenburg,G.F. (1990). "Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram," Submitted to J. Acoust. Soc. Am.

Smoorenburg,G.F. (1990). "On the limited information transfer in the region of noise-induced hearing loss," Acta Otolaryngol. (Stockh) Suppl. 469, 38-46.

Stubblefield,J.H., and Goldstein,D.P. (1977). "A test-retest reliability study on clinical measurement of masking level differences," Audiology 16, 419-431.

Tillman,T.W., Carhart,R., and Nicholls,S. (1973). "Release from multiple maskers in elderly persons," J. Speech Hear. Res. 16, 152-160.

Tonning,F.M. (1971). "Directional audiometry. II. The influence of azimuth on the perception of speech," Acta Otolaryngol. 72, 352-357.

Wilson,R.H., Civitello,B.A., and Margolis,R.H. (1985). "Influence of interaural level differences on the speech recognition masking level difference," Audiology 24, 15-24.

Wright,H.N., and Carhart,R. (1960). "The efficiency of binaural listening among the hearing-impaired," Arch. Otolaryng. 72, 789-797.

Zelnick,E. (1970). "Comparison of speech perception utilizing monotic and diotic modes of listening," J. Aud. Res. 10, 87-97.

Zurek,P.M. (1987). "A predictive model for binaural advantages and directional effects in speech intelligibility," unpublished paper.

Zwislocki,J. (1953). "Acoustic attenuation between the ears," J. Acoust. Soc. Am. 25, 752-759.