SPEECH INTELLIGIBILITY OF THE HEARING IMPAIRED

A. J. H. M. DUQUESNOY

SPEECH INTELLIGIBILITY OF THE HEARING IMPAIRED

VRIJE UNIVERSITEIT TE AMSTERDAM

SPEECH INTELLIGIBILITY OF THE HEARING IMPAIRED

Psychoacoustical modelling of the effect of noise

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
in de wiskunde en natuurwetenschappen
aan de Vrije Universiteit te Amsterdam,
op gezag van de rector magnificus dr. H. Verheul,
hoogleraar in de faculteit der wiskunde en natuurwetenschappen,
in het openbaar te verdedigen
op woensdag 8 september 1982 te 15.30 uur
in het hoofdgebouw der universiteit, De Boelelaan 1105

door

ANTONIUS JOSEPHUS HERMAN MARIA DUQUESNOY

geboren te Hoensbroek

krips repro meppel

Promotor: Prof. dr. ir. R. Plomp
Coreferent: Dr. G. F. Smoorenburg

ACKNOWLEDGMENTS

The research described in this thesis has been performed during a period of four years in the Section Experimental Audiology of the Oto-Rhino-Laryngology and Audiology Department of the Free University. I am especially grateful to Professor Dr. R.Plomp, Head of the Section, for his advice and help in all aspects of the project. I would like to thank my colleagues Mr. J.M.Festen and Mr. W.A.Dreschler for their cooperation and fruitful discussions, and Mr. C. de Boer for his technical assistance.
I am also indebted to Dr. T.S.Kapteyn, Head of the Section Clinical Audiology of the Department, for his cooperation in finding appropriate test subjects and for lending several pieces of audio equipment. Furthermore, thanks are due to Mr. A.M.Mimpen and Mr. H.J.M.Steeneken, both of the Institute of Perception TNO, Soesterberg, for their assistance in preparing the major part of the sound materials applied. I wish to extend my special thanks to my wife, Mildred, for typing the final version of this thesis.

The project was sponsored by the "Beleidsruimte"-fund of the Free University under No. B3, with project title: "Het spraakverstaan van slechthorenden in praktijksituaties".

Aan mijn ouders
Aan Mildred

CONTENTS

SUMMARY

This thesis presents a systematic study on the speech intelligibility of the hearing impaired under everyday listening conditions. The study was prompted by the frequently noticed complaint of hearing-impaired persons of having difficulties primarily in understanding speech in a noisy environment as in group conversation, public halls and traffic. Under these conditions hearing aids appear to be of little or no benefit, as a result of which many users are rather disappointed about the performance of their aids.

A major part of the study has been focused on the psychoacoustical modelling of the effect of noise and reverberation on the unaided and aided Speech-Reception Threshold (SRT; defined as the speech level at which 50% of speech is correctly understood) for conversational sentences. The models were tested by means of SRT values measured against a background of continuous noise with a spectrum identical to the long-term average spectrum of the sentences. The SRT's were measured both on elderly subjects without hearing aids and on younger subjects (age below 65) with and without behind-the-ear hearing aids. The SRT's of the elderly were measured monaurally over headphones. In addition, part of the elderly also participated in an experiment, where binaural hearing was tested under free-field conditions, using different types of interfering sound sources positioned in front or laterally to the subjects. In all investigations reference values for SRT were measured on young normal-hearing listeners tested under the same conditions.

The main results, given below, are ranked in the order of presentation in the Chapters I to V of this thesis:

(1) by means of the Speech-Transmission-Index (STI) model reverberation can be replaced by an equivalent noise in predicting the effects of room acoustics on the SRT of elderly hearing-impaired persons situated in the diffuse sound field;

(2) in rooms, the most effective way of reducing the handicap of elderly hearing-impaired persons is by decreasing the reverberation time rather than by raising the presentation level of the speech;

(3) for any individual, the hearing loss for speech (SHL) can be accurately described, as a function of ambient noise level, by a model which is based on two independent loss components, viz. SHL in quiet, and SHL in noise at high levels (> 70 dBA);

(4) in binaural hearing, mildly hearing-impaired elderly subjects derive less benefit from separating the locations of primary speech source and interfering sound source (speech or noise) than young persons; also, the elderly do not really benefit from the relatively silent periods in competing speech (single speaker), in contrast to a gain of 7 dB for young subjects relative to the condition of continuous interfering noise;

(5) for any hearing-impaired individual, the effect of a hearing aid on SRT can be sufficiently characterized by two components, viz. the functional gain in quiet, and the functional distortion in noise at higher levels (> 55 dBA); in modern hearing aids an average functional distortion corresponding to an increase of the aided SRT in noise of more than 1 dB has been found; thus, the aids provide no benefit in noise, since they are detrimental to the signal-to-noise ratio.

INTRODUCTION

It is an irrefutable fact that speech is a vital and efficient way to communicate. This explains why people whose speech-hearing ability is reduced are so obviously handicapped in daily life. Despite a vast amount of research on numerous details of hearing impairment (cf. the review by Plomp,1978), the question of how hearing-impaired people are exactly handicapped in terms of disability to understand speech under everyday conditions, has not yet been studied systematically. This lack of knowledge manifests itself, for example, in the routine manner in which hearing-impaired subjects are provided with a hearing aid. The primary aim of the present investigation is to gain a quantitative and systematic knowledge of hearing impairment as a communicative handicap, and to explain in what respect a hearing aid can be of rehabilitative potential.

This thesis is basically a collection of five successive papers in which the results of the investigation have been published (given in chronological order in Chapters I to V). Before introducing these papers, two general remarks should be made. The first concerns the symptoms and type of auditory handicap. It has been known for a long time that the first symptoms of a hearing impairment are that the individual has difficulty in understanding speech in church, at the theatre, or in group conversation, but can hear speech at close range without any artificial assistance (cf. Beasley, 1940). Apparently, even moderately hearing-impaired subjects are easily disturbed by interfering noise and reverberation. Nevertheless, in the past decades most attention has been focused on hearing difficulties in quiet, at the expense of research into impaired speech intelligibility in everyday listening situations. This has had a marked, adverse effect on the design of hearing aids, and it has created immoderate expectations of their effectiveness. Indeed, the aids are satisfactorily effective in quiet, but generally of no use in a noisy environment (Carstairs, 1973; Nielsen, 1976). Especially individuals with sensorineural hearing impairment derive rather limited benefit from a hearing aid, as compared to subjects with conductive or mixed losses.

The second remark concerns the prevalence of this auditory handicap. From a literature survey (Plomp,1978) we may conclude that as much as 7.5% of the population has difficulties in understanding speech in noisy environments. The next stage of a hearing handicap, which implies difficulties both in noise and in close-range conversation in quiet (cf. Beasley,

1940), still involves 3.4% of the population. Slightly more than 1% even has difficulties at all ordinary sound intensities, and thus needs very loud speech or artificial assistance. Other figures given by Plomp (1978) show that at the age of 65 no less than 24% of the population is more or less handicapped, as compared to 1% at the age of 20. Furthermore, the cumulative distribution of hearing handicap as a function of age indicates that almost half of the handicapped is over 65. This means that presbyacusis is a major source of hearing difficulties, and that elderly subjects deserve our special attention.

It is the above-described state of affairs that prompted us to the research being introduced next. In Chapter I it is shown how, in cases of elderly listeners with various degrees of hearing loss, the Speech-Reception Threshold (SRT; level at which 50% of speech is correctly understood) for conversational sentences is affected by noise and reverberation. A systematic concept for evaluating the combined effects of reverberation and noise on the SRT of hearing-impaired subjects is presented, which is based on the Speech Transmission Index (STI) (Houtgast and Steeneken, 1973). The STI is a practical, single measure by means of which reverberation can be replaced by an equivalent noise in predicting the effects of room acoustics on speech intelligibility.

In Chapter II the implications of the reverberation study have been elaborated with regard to room acoustics for the aged. The susceptibility to noise and reverberation found on elderly subjects implies that the acoustical requirements for rooms frequented by the aged have to be more stringent than for normal-hearing persons. Based on results obtained from the STI concept, supplemented with data on the hearing loss for speech in quiet and in noise as a function of age (Plomp and Mimpen, 1979), a quantitative specification of the acoustical requirements of elderly subjects is given, which pertains to speech intelligibility both in the direct sound field and the diffuse field of a speaker.

The study presented in Chapter III is focused exclusively on the effect of noise on the SRT's of the same, aged subjects, who participated in the reverberation study. Five noise levels from 0 dBA up to 73 dBA were applied. This systematic way of measuring SRT in noise enables a test to be made of a model of hearing losses for speech developed by Plomp (1978). This model describes SRT as a function of noise level by means of two parameters specific for hearing loss.

All SRT's referred to in Chapters I to III were measured monaurally over headphones. It is known from studies on normal-hearing listeners (e.g. Carhart et al., 1969; Tonning, 1971) that (1) binaural hearing is crucial to cope with degraded listening conditions, and (2) fluctuating interfering signals, like speech, are less disturbing to speech intelligibility than continuous ones. In the study presented in Chapter IV it is investigated whether the above two observations also hold for slightly to moderately hearing-impaired elderly subjects. Their binaural free-field SRT's for sentences have been measured with either an interfering noise source (continuous signal) or an interfering speech source (fluctuating signal) in frontal and lateral positions, respectively.

As stated before, it is important to gain a better insight into the rehabilitative potentialities of a hearing aid under everyday listening conditions. In particular people with sensorineural hearing losses (like presbyacusis) are seriously handicapped under these noisy conditions, and a hearing aid does not alleviate this aspect of their handicap. A basic requirement for a hearing aid to provide substantial benefit in noise is that it increases the signal-to-noise (S/N) ratio. Plomp (1978) presented a simple model of how hearing aids may influence speech intelligibility of the hearing impaired. In Chapter V of this thesis an experimental basis is provided for Plomp's model, and the validity of the model is demonstrated, irrespective of the frequency responses of the hearing aids involved and the types of hearing impairment considered. Furthermore, data are presented on the performance of hearing aids in relation to their electroacoustic characteristics as measured in a test box.

In the section "Final discussion" the implications of the most essential results, presented in Chapters I to V, will be considered. First, the applicability of STI is briefly reconsidered, with emphasis on listening conditions with fluctuating, instead of continuous, interfering sounds. In this respect, the experimental results presented in Chapter IV are quite relevant. Secondly, on the basis of both the systematic quantification of the everyday auditory handicap of hearing-impaired persons and their requirements with respect to hearing aids, it is discussed what the prospects are, for the near future, of mitigating their handicap by a hearing aid more substantially than is possible at the moment.

Finally, in the Appendix two methods are described for calibrating sound-pressure levels produced under headphones. In addition, some results are given on the reliability of SRT values obtained for conversational sentences.

CHAPTER I

EFFECT OF REVERBERATION AND NOISE ON THE INTELLIGIBILITY OF SENTENCES IN CASES OF PRESBYACUSIS

Co-author: Prof.dr.ir. R.Plomp

Paper published in:
Journal of the Acoustical Society of America
Volume 68, August 1980, pp 537-544.

ABSTRACT

For 80 male subjects (age 60-90) and 30 female subjects (age 71-89) the monaural Speech-Reception Threshold (SRT) for sentences was investigated under five reverberation conditions at a constant noise level. The reverberation times used were between 0 and 2.3 s. The noise, with the long-term average speech spectrum, had a level of 52.5 dBA. Each subject was assigned to one of several subgroups formed on the basis of the maximum reverberation time at which the subject was still able to understand the sentences correctly. The mean SRT's and the standard deviations are given, per subgroup, as a function of reverberation time. It is shown that, for each subgroup, the SRT in different reverberant sound fields can be expressed as a single number, namely the required Speech Transmission Index (STI) as introduced in room acoustics by Houtgast and Steeneken (Acustica 28, 66-73 (1973)). Furthermore, it is shown that a model of SRT as a function of noise level, developed by Plomp (J.Acoust.Soc.Am.63,533-549 (1978)), can be combined with the STI model and can thus include the effect of reverberation.

INTRODUCTION

Hearing-impaired subjects often complain of being unable to understand speech in a reverberant room. This paper presents a systematic approach for evaluating the extent to which, for the hearing-impaired, conversational speech is interfered with by a combination of reverberation and noise. The approach is based on the Speech Transmission Index (STI), introduced by Houtgast and Steeneken (1973)[1] (see also Houtgast, Steeneken, and Plomp[2]), which is an appropriate measure for describing the combined effects of reverberation and noise on speech intelligibility. As subjects suffering from presbyacusis represent a large percentage of the hearing-impaired, our study is focused on that group.

The effect of reverberation on the intelligibility of speech for hearing-impaired subjects has been studied by only a few investigators. Bullock (1967)[3] investigated word intelligibility in quiet for reverberation times (T) up to 2 s. Nábelek and Pickett (1974)[4] tested subjects wearing hearing aids; they studied combinations of noise and reverberation (T=0.3 s and 0.6 s), as did Finitzo-Hieber and Tillman (1978)[5] for T up to 1.2 s. Gelfand and Hochberg (1976)[6] presented their subjects with word material in quiet, to which artificial reverberation (T=1, 2, or 3 s) consisting of a series of

attenuated echoes was added. Nábelek and Robinette (1978)[7] have recently summarized various results published on word intelligibility under different reverberation conditions.

Because of their fragmentary character, it is difficult to draw systematic conclusions from these studies. In general, the results for normal-hearing subjects indicate that the decrease in word intelligibility scores in quiet is moderate for T increasing up to 2 s (scores typically above 80% for T = 1 s); in the presence of masking noise the scores decrease more rapidly. Hearing-impaired subjects prove to be more hindered by increasing reverberation than normals. When noise is added, they are seriously handicapped.

A more systematic approach seems to be necessary in order to gain a clear insight into the effect of reverberation and noise on speech intelligibility. For the case of normal hearing some attempts have been made to find a theoretical framework in order to generalize the results.

Lochner and Burger (1961)[8] presented a method to compute speech intelligibility under reverberation conditions from acoustic measures, taking ambient noise into consideration. In essence, the technique consists of splitting the sound reflections in a room into early components to be regarded as useful signals, and late components acting as masking noise. All echoes arriving within a period of 95 ms are integrated with the direct sound in accordance with a weighting function which defines, as a function of echo intensity and delay time, the fraction of the echo intensity to be integrated with the direct sound. By means of this technique the effective signal-to-noise (S/N) ratio in a room could be computed from the impulse response known as the echogram; from this ratio word intelligibility was predicted. The calculated scores agreed very well with the results from word tests in three rooms with T = 0, 0.8, and 1.8 s, respectively.

Peutz and Klein (1973)[9] introduced the Articulation Loss for consonants (ALc) as a measure for predicting speech intelligibility. It was experimentally found that, without noise, ALc gradually increases as a function of distance from the source up to a limiting distance of approximately $0.2(V/T)^{\frac{1}{2}}$ (V = volume of the room in m^3). At larger distances ALc is almost constant, depending merely on T.

Bolt and MacDonald (1949)[10] developed a statistical theory, in which speech is regarded as a series of discrete impulses, the heights of which are uniformly distributed over a range of 30 dB of sound-pressure level in any given audio-frequency band. By considering the extent to which an impulse is masked by components of previous impulses, they derived an expression for the unmasked portion of the impulses. Applying experimental values of speech-impulse durations and intervals obtained from spectrograms divided into seven frequency bands with widths of 500 Hz, unmasked portions were calculated as a function of reverberation time and masking noise-level. By making use of data reported by French and Steinberg (1947)[11] word intelligibility scores could be predicted from these portions. The predicted values agreed encouragingly well with the experimental data which were, it is true, limited in number at that time.

French and Steinberg (1947)[11] introduced the Articulation Index (AI) as a measure for predicting speech intelligibility. Hitherto, results confirming the merit of AI under reverberation conditions have not been reported.

A promising new measure is the Modulation Transfer Function (MTF), introduced in room acoustics by Houtgast and Steeneken (1972, 1973).[1,12] Essentially, the MTF is the Fourier transform of the impulse response. For many cases, a mathematical expression representing the MTF as a function of room volume V, T, S/N ratio and distance can be given. From the MTF a single number, the Speech Transmission Index (STI), can be derived. This index is a very convenient measure for quantifying systematically the combined effect of noise and reverberation on speech intelligibility. In the next section the derivation of the STI will be explained for a simple case.

I. THEORETICAL ASPECTS OF SPEECH TRANSMISSION IN A ROOM

The influence of the acoustical environment on the transfer of acoustic signals consists, essentially, of a smoothing effect on the temporal envelope of the signal caused by masking noise, reverberation and echoes. Houtgast and Steeneken (1973)[1] introduced a noise signal with a sinusoidally modulated intensity to determine the smoothing properties of a room. These properties can be quantified as the MTF, representing the degree of modulation depth of the temporal envelope as a function of modulation frequency. Recently, the method has been described in full detail (Steeneken and Houtgast[13]). In this paper the method will be discussed as far as needed.

Let us consider a room with volume V and reverberation time T. A speaker and an interfering noise source with equally shaped sound spectra are both located far enough from the listener to make the direct sound negligible compared to the reverberant sound reaching the listener after a large, variable number of reflections (diffuse sound field) (according to Peutz and Klein (1973)[9] the distance between the listener and the sound sources should be

larger than $0.2\,(V/T)^{\frac{1}{2}}$). This condition is attractive because it simplifies the equations and because it is the most sensitive way of testing the effect of reverberation on speech intelligibility. The speaker is replaced by a noise input signal with intensity $I_i(t)$, varying sinusoidally between 0 and $2\overline{I}_i$:

$$I_i(t) = \overline{I}_i(1+\cos 2\pi Ft) = \overline{I}_i\{1+\mathrm{Re}\left[\exp(j2\pi Ft)\right]\}, \tag{1}$$

where $\overline{I}_i$ = long-term average of $I_i(t)$, measured at a distance of 1 m from the source (free field condition), and F = modulation frequency. The interfering noise source has a constant intensity I_n at a distance of 1 m. It is clear from Eq.(1) that the input signal is modulated 100%. Noise and reverberation will reduce the modulation depth near the listener. The output signal at the position of the listener is equal to

$$I_o(t) = \overline{I}_o\{1+m(F)\mathrm{Re}\left[\exp\{j2\pi F(t-\Delta t)\}\right]\}, \tag{2}$$

where m(F) = modulation reduction index depending on F, and Δt = time delay relative to the input signal, reflecting the phase response of the transmission path. Index m(F) constitutes the MTF. For the specific condition considered the MTF can easily be calculated on the assumption of an exponentially decaying sound field. In that case, $I_o(t)$ is the sum of the exponentially weighted contributions from all past moments:

$$I_o(t) = \int_{-\infty}^{t}(I_i(t') + I_n)C\exp\{-a(t-t')\}dt', \tag{3}$$

where C = a constant, depending upon the properties of the room,

I_n = constant intensity of the interfering noise source, measured at a distance of 1 m from the source (free field condition), and

a = a factor defined by T: $e^{-aT}=10^{-6}$, viz. a=13.8/T.

Substitution of $I_i(t)$ from Eq.(1) into Eq.(3) yields successively:

$$\begin{aligned} I_o(t) &= C(\overline{I}_i+I_n)\int_0^{\infty}\exp[-(13.8/T)t']dt' + C\overline{I}_i\,\mathrm{Re}\left\{\exp(j2\pi Ft)\int_0^{\infty}\exp[-(j2\pi F+13.8/T)t']dt'\right\} \\ &= \frac{CT}{13.8}(\overline{I}_i+I_n)\left\{1+\frac{\overline{I}_i}{\overline{I}_i+I_n}\,\frac{(13.8/T)\mathrm{Re}[\exp j2\pi F(t-(1/2\pi F)\arctan(2\pi FT/13.8))]}{[(13.8/T)^2+(2\pi F)^2]^{1/2}}\right\}. \end{aligned} \tag{4}$$

Comparison with the expression for $I_o(t)$ in Eq.(2) shows that the MTF is equal to

$$m(F) = (\overline{I}_i/(\overline{I}_i+I_n))\{(13.8/T)[(13.8/T)^2+(2\pi F)^2]^{-\frac{1}{2}}\}$$

$$= (1+10^{-(S/N)/10})^{-1}(1+0.207F^2T^2)^{-\frac{1}{2}} \quad . \qquad (5)$$

This equation shows that the MTF is the product of two independent factors, one depending on T (and F) and the other on the signal-to-noise ratio $S/N = 10 \log(\overline{I}_i/I_n)$.

With respect to speech, the essential modulation frequencies in the temporal envelope are between 0.4 and 20 Hz, as is illustrated in Fig.1. The solid curve represents the envelope spectrum of connected discourse obtained by performing a $^1/_3$-octave frequency analysis on the envelope originating from the audio-frequency octave band centered at 2 kHz (other octave bands

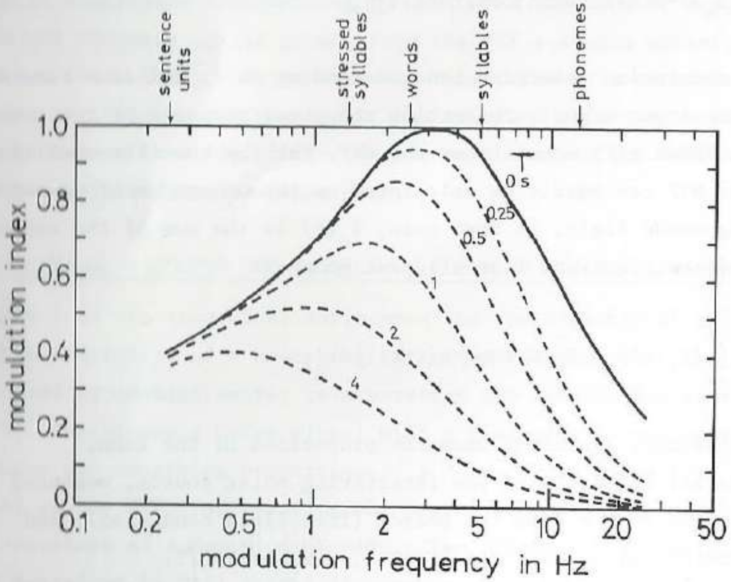


FIG.1. The solid curve is the envelope spectrum (root-mean-square of the fluctuations within $^1/_3$ -octave bands) for the 2-kHz octave band of a 60-s sample of connected discourse of a single speaker. The dashed curves are obtained by multiplying the solid curve with m(F) values calculated in accordance with Eq.(5) with T as the parameter and a noise with intensity I_n=0. (Adopted from Plomp and Duquesnoy, 1980,[14] Fig.2.)

yield similar results). The envelope spectrum has a maximum for F=3.4 Hz situated between the frequency of words (2.5 words/s) and the frequency of syllables (5 syllables/s) of the speech analyzed. The dashed curves are obtained by multiplying the solid curve with m(F) according to Eq.(5) for different reverberation times T (independent of audio frequency). The curves giving the envelope spectrum at the location of the listener situated in the diffuse sound field, show that the speech modulations are transferred rather accurately for T<0.25 s, but that they are strongly reduced for T≥4 s.

In order to achieve an optimal agreement between the MTF and Dutch word-intelligibility scores from normal-hearing listeners, Houtgast, Steeneken, and Plomp[2] developed a weighting function, the Modulation Transfer Index (MTI). The MTI controls the contribution of the MTF's to speech intelligibility for the modulation frequencies involved. It was optimized for a wide variety of rooms by analyzing scores from 80 combinations of noise, single echoes and reverberation times. The resulting procedure consists of the following steps:

(1) express the MTF, for each of the 18 modulation frequencies F (F from 0.4 to 20 Hz in $^1/_3$-octave intervals), in terms of an equivalent S/N ratio:

$$S/N_{eq}(F) = 10 \log\{m(F)/[1-m(F)]\} \text{ in dB.} \qquad (6)$$

(Thus, each modulation reduction index m(F) is interpreted as if it had been caused by interfering noise only);

(2) compute the average $\overline{S/N}_{eq}$ resulting from the above obtained 18 $S/N_{eq}(F)$ values after having been clipped when exceeding the range from −15 dB to +15 dB;

(3) convert the average $\overline{S/N}_{eq}$ into a normalized index MTI according to

$$MTI = (\overline{S/N}_{eq}+15)/30, \text{ where } 0 \le MTI \le 1. \qquad (7)$$

In our case (see Sec. IIA) T and S/N are independent of audio frequency, so that MTI = STI (Speech Transmission Index). In other cases, however, m(F) has to be calculated for seven audio-frequency octave bands in order to obtain STI as a weighted average of the MTI values specific to these bands. Furthermore, in cases of nonexponential reverberation, the simple Eq.(5) does not hold. Houtgast and Steeneken[15] have shown experimentally that such cases are, nevertheless, reliably described by the MTF.

In our simplified conditions STI has been computed for various values of T and S/N inserted into Eq.(5). In Fig.2 the STI is plotted as a function of the indirect sound field of the loudspeaker. The sound field was made as diffuse as possible with sound reflectors hanging from the ceiling. By means of removable absorbers the reverberation time, being almost independent of audio-frequency between 125 Hz and 4 kHz, could be varied. Using a decay-curve averaging method (average of 20 curves, octave bands with center frequencies of 500 Hz and 2000 Hz) the reverberation times were calculated from the slopes of the curves in the range from 0 to -10 dB (early decay time, most relevant in speech transmission, see Houtgast and Steeneken[15]). The reverberation times used were 0.4, 0.5, 0.7, 1.0, 1.3, 1.9, and 2.3 s. Additionally, the MTF was measured for the various conditions. The reverberation times computed from the MTF (see Eq.(5)) agreed with those measured directly within $\pm$ 5%.

B. Procedure

The recordings made for various reverberation times were used for measuring SRT for sentences. The following adaptive procedure was applied. The first sentence of a list was repeatedly presented at, successively, 2 dB higher sound-pressure levels until the listener was able to reproduce the sentence correctly. Then, the second sentence was presented at a 2 dB lower level. If this sentence was correctly understood, the level of the next sentence was, again, decreased by 2 dB; if it was not, the level was increased by 2 dB. All remaining sentences were handled in this manner. The average of the levels adjusted after presentation of sentences 4, 5, etc. was accepted as an estimate of the SRT for sentences. The sentences 1 to 3 were discarded to avoid a possible bias of SRT caused by the first sentence being accidently understood at too high or too low a level.

The procedure implies that a correct reproduction of the entire sentence is required for a positive response. If a listener's speech reception is so poor that he is unable to reproduce sentences correctly even under excellent listening conditions, an adaptive procedure as the above cannot be applied. None of our subjects belonged to that category.

In two separate experiments (Duquesnoy, 1977)[18]; (Plomp and Mimpen, 1979)[17] the reliability of SRT values obtained with the lists was investigated. The SRT values found had a standard deviation of approximately 1 dB.

C. Experimental conditions

The sentences were recorded on one track of a tape. On the other track the

standard masking noise with exactly the same intensity and spectrum as the long-term average of the sentences was recorded. Preceding the sentences the same noise was recorded for easy calibration of S/N ratios. The levels of the output signals of both tracks of the tape recorder could be separately adjusted by means of two attenuators. After attenuation the signals were mixed and fed monaurally into an earphone fitted with circumaural cups with liquid-filled cushion, which attenuates ambient noise effectively (earphone Sharpe-Scintrex Mk IV, attenuation >35 dB at 1000 Hz).

The tests, which took 35 to 45 min, were partly carried out in an anechoic room in our laboratory and partly in a reasonably quiet room (noise level <35 dBA) in a home for the aged. First, air-conduction tone audiograms for both ears were determined, which took about 15 min. Then, the ear with the smaller average hearing level (500, 1000, and 2000 Hz) was chosen to measure SRT under five reverberation conditions, using one sentence list per condition. Varying reverberation times including T=0 s and a constant noise level of 52.5 dBA were used.

It was expected that several hearing-impaired subjects would be unable to understand sentences presented with much reverberation. In order to test the subjects for the reverberation range best adapted to their capacities, each subject was assigned to one of several subgroups by using the following selection procedure. Apart from the sentence lists, a supplementary list of six sentences was recorded three times with T=1.3, 1.9, and 2.3 s, respectively. The greatest T for which at least four of the six sentences, presented against a background noise of 52.5 dBA, were understood correctly determined to which subgroup a subject was assigned. When a subject was unable to understand sentences for T=1.3 s, the maximum T to be considered was apparently 1.0 s or less. In this manner we arrived at the following subgroups with corresponding maximum Ts: subgroup A (T_{max}=2.3 s), B(1.9 s), C(1.3 s), and D(1.0 s). Per subgroup, the five reverberation conditions were distributed as evenly as possible over the range of T applicable to this subgroup.

The five lists were always presented in the same order. To eliminate the effects of training and fatigue as much as possible, the test conditions were counterbalanced for every ten subjects tested successively.

D. Subjects

The elderly people were recruited from two different populations. The first group consisted of 50 male employees (age 60-67) of the Free University. The

the indirect sound field of the loudspeaker. The sound field was made as diffuse as possible with sound reflectors hanging from the ceiling. By means of removable absorbers the reverberation time, being almost independent of audio-frequency between 125 Hz and 4 kHz, could be varied. Using a decay-curve averaging method (average of 20 curves, octave bands with center frequencies of 500 Hz and 2000 Hz) the reverberation times were calculated from the slopes of the curves in the range from 0 to -10 dB (early decay time, most relevant in speech transmission, see Houtgast and Steeneken[15]). The reverberation times used were 0.4, 0.5, 0.7, 1.0, 1.3, 1.9, and 2.3 s. Additionally, the MTF was measured for the various conditions. The reverberation times computed from the MTF (see Eq.(5)) agreed with those measured directly within $\pm$ 5%.

B. Procedure

The recordings made for various reverberation times were used for measuring SRT for sentences. The following adaptive procedure was applied. The first sentence of a list was repeatedly presented at, successively, 2 dB higher sound-pressure levels until the listener was able to reproduce the sentence correctly. Then, the second sentence was presented at a 2 dB lower level. If this sentence was correctly understood, the level of the next sentence was, again, decreased by 2 dB; if it was not, the level was increased by 2 dB. All remaining sentences were handled in this manner. The average of the levels adjusted after presentation of sentences 4, 5, etc. was accepted as an estimate of the SRT for sentences. The sentences 1 to 3 were discarded to avoid a possible bias of SRT caused by the first sentence being accidently understood at too high or too low a level.

The procedure implies that a correct reproduction of the entire sentence is required for a positive response. If a listener's speech reception is so poor that he is unable to reproduce sentences correctly even under excellent listening conditions, an adaptive procedure as the above cannot be applied. None of our subjects belonged to that category.

In two separate experiments (Duquesnoy, 1977)[18]; (Plomp and Mimpen, 1979)[17] the reliability of SRT values obtained with the lists was investigated. The SRT values found had a standard deviation of approximately 1 dB.

C. Experimental conditions

The sentences were recorded on one track of a tape. On the other track the standard masking noise with exactly the same intensity and spectrum as the long-term average of the sentences was recorded. Preceding the sentences the same noise was recorded for easy calibration of S/N ratios. The levels of the output signals of both tracks of the tape recorder could be separately adjusted by means of two attenuators. After attenuation the signals were mixed and fed monaurally into an earphone fitted with circumaural cups with liquid-filled cushion, which attenuates ambient noise effectively (earphone Sharpe-Scintrex Mk IV, attenuation >35 dB at 1000 Hz).

The tests, which took 35 to 45 min, were partly carried out in an anechoic room in our laboratory and partly in a reasonably quiet room (noise level <35 dBA) in a home for the aged. First, air-conduction tone audiograms for both ears were determined, which took about 15 min. Then, the ear with the smaller average hearing level (500, 1000, and 2000 Hz) was chosen to measure SRT under five reverberation conditions, using one sentence list per condition. Varying reverberation times including T=0 s and a constant noise level of 52.5 dBA were used.

It was expected that several hearing-impaired subjects would be unable to understand sentences presented with much reverberation. In order to test the subjects for the reverberation range best adapted to their capacities, each subject was assigned to one of several subgroups by using the following selection procedure. Apart from the sentence lists, a supplementary list of six sentences was recorded three times with T=1.3, 1.9, and 2.3 s, respectively. The greatest T for which at least four of the six sentences, presented against a background noise of 52.5 dBA, were understood correctly determined to which subgroup a subject was assigned. When a subject was unable to understand sentences for T=1.3 s, the maximum T to be considered was apparently 1.0 s or less. In this manner we arrived at the following subgroups with corresponding maximum Ts: subgroup A (T_{max}=2.3 s), B(1.9 s), C(1.3 s), and D(1.0 s). Per subgroup, the five reverberation conditions were distributed as evenly as possible over the range of T applicable to this subgroup.

The five lists were always presented in the same order. To eliminate the effects of training and fatigue as much as possible, the test conditions were counterbalanced for every ten subjects tested successively.

D. Subjects

The elderly people were recruited from two different populations. The first group consisted of 50 male employees (age 60-67) of the Free University. The second group consisted of 30 male subjects (age 69-90, mean age 80.6) and 30 female subjects (age 71-89, mean age 81.2), all of them inhabitants of a home for the aged.

The tone audiograms of both groups were studied to find manifestations of presbyacusis. As Schuknecht (1974)[19] pointed out, four types of presbyacusis can currently be identified, which show an enormous variety in audiometric patterns. This makes identification of presbyacusis solely based on the audiogram rather precarious. With proper restrictions, however, it could be concluded from the audiograms that 26 subjects had predominantly conductive hearing losses, whereas only four subjects (3.5%) had losses mainly due to noise injuries. The remaining 80 subjects were supposed to suffer from presbyacusis.

In addition to the elderly people, 20 young normal-hearing subjects (individual hearing losses for the frequencies 500, 1000, 2000, 4000, and 8000 Hz $\leq$15 dB re ISO-389) were tested in order to obtain a standard of hearing performance. Ten subjects (4 female, 6 male, age 19-27, mean age 21.5) were university students and the other ten (7 female, 3 male, age 18-35, mean age 21.0) were employees of the home for the aged. Criteria for inclusion in this reference group were no history of ear pathology and the ability of the subject to manage the most arduous condition in our test, being the correct understanding of sentences recorded for T=2.3 s and presented against a background noise of 52.5 dBA.

III. RESULTS

The Speech-Reception Threshold (SRT) was measured under conditions all including a masking noise with a constant level of 52.5 dBA. For the elderly subjects who had an SRT in quiet of 50 dBA or more, this noise was nearly inaudible so that the SRT values found were partially due to their absolute threshold rather than to the noise and reverberation. Therefore, these subjects were excluded from our further analysis. This applied to one subject from subgroup A, none from subgroup B, five subjects from subgroup C, and five subjects from subgroup D. Furthermore, all 14 subjects, for which $T_{max} \leq 0.7$ s, were left out of consideration, because 12 of them had an SRT in quiet of more than 50 dBA.

In Table I the averages of air-conduction tone audiograms for frequencies from 500 to 8000 HZ are given for the reference group and for subgroups A to D. The data refer to the ear with the smaller average hearing level (500, 1000 and 2000 Hz).

TABLE I. Means and standard deviations (s.d.) of the air-conduction tone audiograms per subgroup. The number of subjects involved and the mean age are also given. The hearing losses refer to the ear with the smaller average loss (at 500, 1000, and 2000 Hz).

Subjects	Number	Mean age		Hearing loss (dB re ISO-389) 500	1000	2000	4000	8000 Hz
Refer.gr	20	22.6	mean	5.5	-0.7	1.5	3.7	5.5
			s.d.	7.2	7.2	6.5	7.9	6.7
Subgr.A	14	65.3	mean	13.2	12.9	18.2	26.4	31.4
			s.d.	7.5	10.3	11.7	14.1	22.7
Subgr.B	29	64.9	mean	7.9	9.3	18.4	34.1	45.3
			s.d.	7.4	8.8	10.4	16.4	22.0
Subgr.C	25	74.6	mean	18.0	16.0	24.0	46.6	69.2
			s.d.	11.4	8.9	12.4	14.5	15.2
Subgr.D	17	76.7	mean	22.6	22.6	35.6	57.1	74.7
			s.d.	9.2	10.2	13.6	19.0	23.6

In Table II the means and standard deviations (s.d.) of the individual SRT values, expressed in S/N ratio, and the means and standard deviations (s.d.1 and s.d.2) of the corresponding STI values are summarized for the reference group and for subgroups A to D. The STI values were derived from the individual S/N ratios by means of the Eqs.(5) to (7), and s.d.1 is the standard deviation of the original STI values. In particular for subgroups C and D, s.d.1 is large compared to the reference group. This is due to considerable differences in individual performance. Since we are interested in the more general aspects rather than in the interindividual differences, we corrected for these differences. This was done by subtracting, for each reverberation time tested, the mean STI for a subject (average of five different reverberation times) from the actual STI value found. The standard deviation of the STI values thus obtained is called s.d.2. For subgroups B, C, and D s.d.2 is substantially smaller than s.d.1.

TABLE II. Means and standard deviations of SRT, expressed in S/N ratio (in dB), and the corresponding STI values from the reference group and the elderly subjects grouped in subgroups A to D for the respective reverberation times applied. For each subgroup the number of subjects involved is given. With respect to the STI values two standard deviations are mentioned: (1) s.d. of the original STI values, and (2) s.d. of the STI values from which the subject means have been removed.

Subjects			Reverberation time (s)					
Refer.gr			0.0	0.4	0.7	1.3	2.3	
(20 Ss)	S/N	mean	-4.4	-1.9	-0.8	2.1	5.5	dB
		s.d.	1.1	1.7	1.0	1.7	2.7	dB
	STI	mean	0.355	0.37	0.36	0.36	0.34	
		s.d.1	0.035	0.05	0.03	0.04	0.035	
		s.d.2	0.03	0.035	0.025	0.03	0.025	
Subgr.A			0.0	0.4	0.7	1.3	2.3	
(14 Ss)	S/N	mean	-3.1	-0.1	0.9	2.9	12.6	dB
		s.d.	1.4	2.1	1.3	1.8	6.5	dB
	STI	mean	0.395	0.415	0.405	0.375	0.385	
		s.d.1	0.045	0.04	0.03	0.035	0.025	
		s.d.2	0.035	0.03	0.025	0.04	0.025	
Subgr.B			0.0	0.4	0.7	1.3	1.9	
(29 Ss)	S/N	mean	-2.3	0.8	1.8	5.5	10.5	dB
		s.d.	1.5	1.9	1.3	5.0	5.1	dB
	STI	mean	0.425	0.45	0.425	0.415	0.41	
		s.d.1	0.05	0.05	0.03	0.05	0.025	
		s.d.2	0.025	0.035	0.025	0.035	0.03	
Subgr.C			0.0	0.4	0.7	1.0	1.3	
(25 Ss)	S/N	mean	-0.6	3.3	5.2	6.6	13.8	dB
		s.d.	2.2	3.1	4.0	4.4	7.2	dB
	STI	mean	0.48	0.51	0.49	0.46	0.49	
		s.d.1	0.075	0.075	0.07	0.06	0.04	
		s.d.2	0.03	0.03	0.03	0.03	0.04	
Subgr.D			0.0	0.4	0.5	0.7	1.0	
(17 Ss)	S/N	mean	1.9	8.2	8.9	13.1	20.1	dB
		s.d.	2.9	6.1	6.5	9.0	11.5	dB
	STI	mean	0.56	0.605	0.58	0.575	0.56	
		s.d.1	0.095	0.09	0.08	0.085	0.045	
		s.d.2	0.045	0.035	0.03	0.04	0.05	

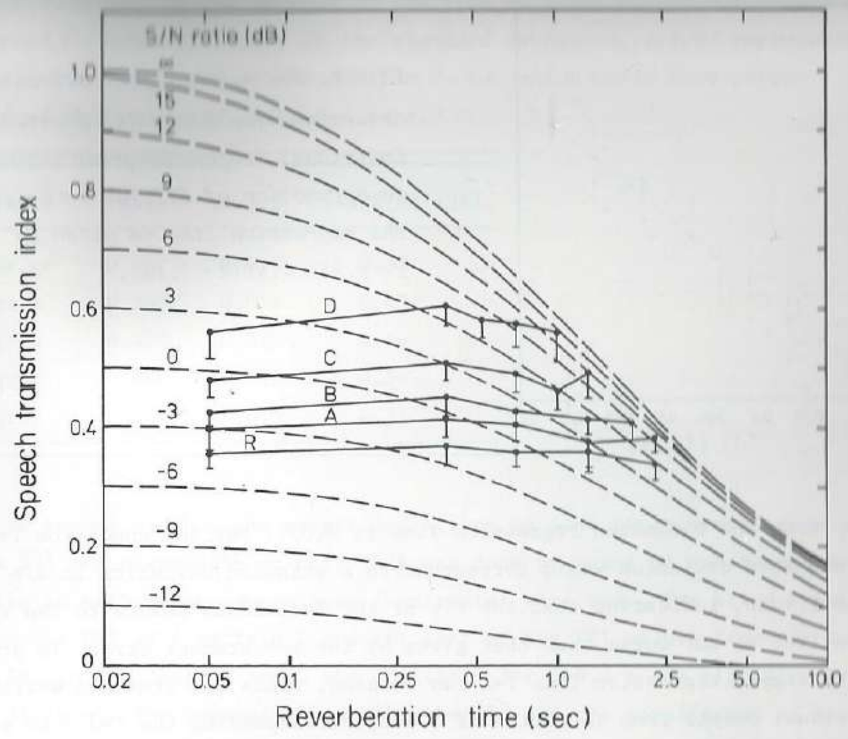


FIG.3. The dashed curves show STI as a function of reverberation time T with S/N ratio as parameter. For each subgroup the mean STI values for various reverberation times, as reported in Table II, are plotted and connected with solid lines. The vertical bars represent the reduced standard deviation s.d.2 (shown one-sided).

In Fig.3 STI has been plotted as a function of T, with S/N ratio as the parameter. In this figure the mean STI values with corresponding s.d.2, as summarized in Table II, are shown for the reference group and four subgroups.

All results given thus far refer to groups of subjects. In addition, the applicability of the STI approach will be shown for the individual case. Each subject was presented with five reverberation conditions, including T=0 s. The predictive power of the single STI, measured for T=0 s, is shown in Fig.4. In this scatter diagram the mean (Y) of four STI values corresponding to the four SRT values for T>0 s is plotted, per individual, as a function of STI for T=0 s (X). The linear regression of Y on X is given by $\overline{Y} = 0.45+0.82(X-0.44)$ and the correlation coefficient is 0.90. The standard deviation in STI of individual

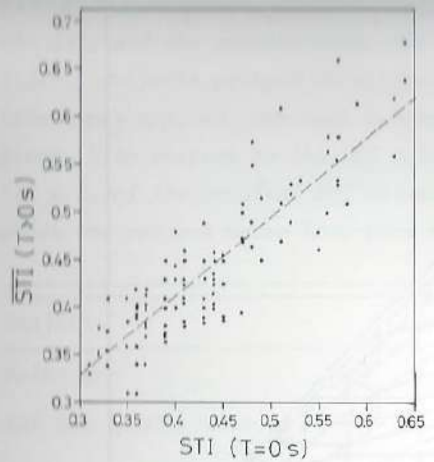


FIG.4. The mean of four STI values determined for T > 0 s (Y) plotted, per individual (reference group included), as a function of STI for T = 0 s (X). The regression line is given by $\overline{Y} = 0.45 + 0.82(X - 0.44)$.

points from the estimated regression line is 0.037. For the condition T=0 s such a standard deviation would correspond to a standard deviation in S/N ratio of 1.1 dB, indicating that the fit of the individual points to the regression line is not worse than that given by the measurement error. In practice, the simple regression line $\overline{Y}$=X can be used, since the standard deviation of individual points from this line is 0.042, corresponding for T=0 s to a standard deviation in S/N ratio of 1.3 dB.

IV. DISCUSSION

As was explained in Sec.I, SRT values of normal-hearing subjects obtained for different T, should be ordered along a horizontal line in Fig.3. From visual inspection of the data points for the reference group it appears that they agree well with such a line. This also holds rather well for the data points for each of the four subgroups of hearing-impaired subjects. The degree of SRT elevation is reflected in a parallel shift of the subgroup curves to larger STI values.

In Table III the mean STI values for the five data points per subgroup as well as their standard deviations, s.d.3, are given. For a better evaluation of s.d.3 the last column, Δ(S/N), represents standard deviations of SRT for T=0 s, which would have given standard deviations of STI equal to s.d.3. The small Δ(S/N) values, expressed in dB, suggest that it is justifiable to represent the subgroups shown in Fig.3 by straight horizontal lines with $\overline{STI}$ as the ordinate. This assumption was investigated in the following way.

TABLE III. $\overline{STI}$ is the average of the five mean STI values per group, as reported in Table II. S.d.3 is the standard deviation; Δ(S/N) represents the standard deviation of the SRT for T = 0 s which would have given standard deviations of STI equal to s.d.3.

Subjects	$\overline{STI}$	S.d.3	Δ(S/N) (dB)
Refer.gr	0.357	0.011	0.33
Subgr.A	0.395	0.016	0.47
Subgr.B	0.425	0.015	0.46
Subgr.C	0.486	0.018	0.54
Subgr.D	0.576	0.019	0.55

We tested whether the slight deviations of some experimental data points from $\overline{STI}$ are systematic or not. This was done by means of the Newman-Keuls Q-statistic with which comparisons (contrasts) were made, for all groups, between the STI at a certain T and the mean of the STI values from the remaining Ts. In subgroups B and C the contrasts for T=0.4 s are significant at the 5% level. The contrasts for all other Ts are not significant at this level, except for T=1.0 s in subgroup C. Typically, the somewhat inferior performance (STI consistently larger) for sentences recorded with T=0.4 s may be related to a slightly deviating frequency response-curve of the recording room for this condition, as compared to the other Ts.

The general conclusion from the foregoing discussion may be that the effect of T on STI does not seem to be systematic and that the STI is a convenient single number for systematically quantifying the combined effect of noise and reverberation on speech intelligibility both in normal-hearing and in presbyacusis populations.

The validity of STI as a single number for determining the combined effect of noise and reverberation on SRT in case of presbyacusis implies that a model, developed by Plomp (1978)[20] for describing the hearing loss for speech as a function of noise level can be easily extended in order to include the effect of reverberation. According to this model, every hearing loss for speech (SHL) is interpreted as the sum of a loss class A (attenuation), characterized by a reduction of the levels of both speech and noise, and a loss class D (distortion), comparable with a decrease in S/N ratio. From measurements of the SRT in quiet and in the presence of defined noise levels, the hearing loss for speech in quiet, $SHL_{A+D}(=A+D)$, and in noise, $SHL_D(=D)$, can be calculated. The value D can be expressed as a shift in STI by considering that every 3 dB of D corresponds with an increase of 0.1 in STI for T=0 s (see Fig.3, graduated scale along ordinate). This means that, on the basis of SHL_D, it is also known which combinations of noise and reverberation will represent acceptable listening conditions for the hearing-impaired. In another publication (Plomp and Duquesnoy[21]) the implications of the above have been elaborated with regard to room acoustics for the aged.

ACKNOWLEDGMENTS

The authors wish to thank Mr.A.M.Mimpen and Mr.H.J.M.Steeneken for their willing assistance in preparing the sentence lists with variable reverberation. We are indebted to Mr.H.Slettenhaar and Mrs.R.Dalmeyer, both of the home for the aged "Menno Simonsz" in Amsterdam, for their cooperation in performing the experiments.

REFERENCES

[1] T.Houtgast and H.J.M.Steeneken, "The Modulation Transfer Function in Room Acoustics as a Predictor of Speech Intelligibility," Acustica 28, 66-73 (1973).
[2] T.Houtgast, H.J.M.Steeneken, and R.Plomp, "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I: General Acoustics," submitted to Acustica.
[3] M.D.Bullock, "The effects of different reverberation times upon the intelligibility of PB words as perceived by subjects with normal hearing and subjects with sensori-neural impairments and concomitant discrimination losses," Master's thesis, Ohio State University (1967).
[4] A.K.Nábelek and J.M.Pickett, "Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners," J.Speech Hear. Res. 17, 724-739 (1974).
[5] T.Finitzo-Hieber and T.W.Tillman, "Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children," J.Speech Hear. Res. 21, 440-458 (1978).
[6] S.A. Gelfand and I.Hochberg, "Binaural and Monaural Speech Discrimination under Reverberation," Audiology 15, 72-84 (1976).
[7] A.K.Nábelek and L.Robinette, "Reverberation as a Parameter in Clinical Testing," Audiology 17, 239-259 (1978).
[8] J.P.A.Lochner and J.F.Burger, "The Intelligibility of Speech under Reverberant Conditions," Acustica 11, 195-200 (1961).
[9] V.M.A.Peutz and W.Klein, "Articulation Loss of Consonants Influenced by Noise, Reverberation and Echo," Proc. 1973 FASE Symposium on Speech Intelligibility, Liège, Belgium, 89-97 (1973).
[10] R.H.Bolt and A.D.MacDonald, "Theory of Speech Masking by Reverberation," J.Acoust. Soc. Am. 21, 577-580 (1949).

[11] N.R.French and J.C.Steinberg, "Factors Governing the Intelligibility of Speech Sounds," J. Acoust. Soc. Am. 19, 90-119 (1947).
[12] T.Houtgast and H.J.M.Steeneken, "Envelope spectrum and intelligibility of speech in enclosures," Proceedings of the 1972 Conference on Speech Communication and Processing IEEE, New York, (1972), pp. 392-395.
[13] H.J.M.Steeneken and T.Houtgast, "A physical method for measuring speech-transmission quality," J.Acoust. Soc. Am. 67, 318-326 (1980).
[14] R.Plomp and A.J.Duquesnoy, "The Speech-Reception Threshold of hearing-impaired subjects in noise as a function of reverberation time," in Acoustical Factors Affecting Hearing Aid Performance, edited by G.A.Studebaker and I.Hochberg (University Park Press, Baltimore, 1980), pp. 47-53.
[15] T.Houtgast and H.J.M.Steeneken, "Predominance of Early Decay Time for Speech Intelligibility in Rooms" (in preparation).
[16] K.D.Kryter, "Methods for the Calculation and Use of the Articulation Index," J.Acoust. Soc. Am. 34, 1689-1697 (1962).
[17] R.Plomp and A.M.Mimpen, "Improving the Reliability of Testing the Speech-Reception Threshold for sentences," Audiology 18, 43-52 (1979).
[18] A.J.Duquesnoy, "Speech-level measurements for sentences using an artificial ear," Report 1977-AD01, Experimental Audiology, Free University, Amsterdam (1977).+)
[19] H.F.Schuknecht, Pathology of the Ear (Harvard U.P., Cambridge, MA, 1974).
[20] R.Plomp, "Auditory Handicap of Hearing Impairment and the Limited Benefit of Hearing Aids," J.Acoust. Soc. Am. 63, 533-549 (1978).
[21] R.Plomp and A.J.Duquesnoy, "Room Acoustics for the Aged," submitted to J.Acoust. Soc. Am.

+) Reproduced in the Appendix of this thesis

CHAPTER II

EFFECT OF REVERBERATION ON THE SPEECH INTELLIGIBILITY OF AGED PERSONS; IMPLICATIONS WITH REGARD TO ROOM ACOUSTICS

Co-author: Prof.dr.ir. R.Plomp

Paper published in:
NEDERLANDS AKOESTISCH GENOOTSCHAP
Publikatie nr.56, oktober 1980

Summary

For 110 subjects (age 60-90, 80 male, 30 female) the monaural Speech-Reception Threshold (SRT) for sentences was investigated under five reverberation conditions at a constant noise level. The reverberation times used were between 0.05 s and 2.3 s. The noise, with the long-term average speech spectrum, had a level of 52.5 dBA. It is shown that the Speech Transmission Index (STI), as introduced by Houtgast and Steeneken (Acustica 28, 66-73, 1973) for normal-hearing subjects, is also in cases of presbyacusis an appropriate measure for describing the combined effect of reverberation and noise on speech intelligibility. The SRT can be easily expressed in the STI value required at the listener's position. Furthermore, it is shown that, in order to compensate for a hearing loss for speech in noise of 1 dB, the reverberation time of a room has to be reduced by 18% if the listener is situated in the direct sound field, and by 25% in the case of a diffuse field.

---

1. Introduction

Hearing-impaired subjects often complain of being unable to understand speech in a noisy or reverberant environment. This is particularly the case with elderly subjects suffering from presbyacusis. As these subjects represent a large percentage of the hearing impaired, our attention has been focused on the consequences of presbyacusis for the intelligibility of speech. This paper presents a systematic approach for evaluating the extent to which conversational speech is interfered with by a combination of reverberation and noise. The approach is based on the Speech Transmission Index (STI) (Houtgast and Steeneken, 1973), which is a practical measure for the prediction of speech intelligibility in rooms.

The susceptibility of elderly subjects to noise and reverberation implies that the acoustical requirements for rooms frequented by the aged have to be more stringent than for normal-hearing subjects. In practice, the design of the acoustical environment is attuned to the large majority of normal-hearing listeners. This means that in most situations the listening conditions are marginal or insufficient for the hearing impaired. Based on the results of the above-mentioned study, supplemented with data on the hearing loss for speech in quiet and in noise as a function of age (Plomp and Mimpen, 1979a), this paper gives a very applicable quantitative specification of the acoustical requirements of hearing-impaired subjects.

2. Prediction of speech intelligibility in a room

The influence of the acoustical environment on the transfer of speech signals consists of a smoothing effect on the temporal envelope of the signal caused by ambient noise and reverberation. In order to quantify these smoothing effects, Houtgast and Steeneken (1973) calculated the degree to which sinusoidal intensity variations at the speaker's position are preserved at the listener's position. For an input signal $I_i(t) = \overline{I}_i(1 + \cos 2\pi Ft)$, where F = modulation frequency, the output signal is:

$$I_o(t) = \overline{I}_o \left\{1 + m(F).\cos 2\pi F(t-\Delta t)\right\} \quad , \qquad (1)$$

where $m(F)$ = modulation reduction index depending on F, and

Δt = time delay relative to the input signal, reflecting the phase response of the transmission path.

In this paper we will restrict ourselves to the case of ambient noise due to interfering speakers, since they form a very frequent ambient-noise condition. In that case the noise source has roughly the same spectrum as the speech signal. From the equations presented by Houtgast and Steeneken (1973) the index $m(F)$ at the listener's position can be expressed in the parameters characteristic of the listening situation considered:

$$m(F) = \left[\left(\frac{q_s q_1}{r^2/r_c^2}\right)^2 + \frac{1 + \dfrac{2q_s q_1}{r^2/r_c^2}}{1 + 0.207F^2T^2}\right]^{\frac{1}{2}} \times \left[\frac{q_s q_1}{r^2/r_c^2} + 1 + \frac{\overline{I}_n}{\overline{I}_i}.\left(\frac{q_n}{r_n^2/r_c^2} + 1\right)\right]^{-1} \qquad (2)$$

where r = speaker-to-listener distance,

r_c = critical radius of the room (distance from the speaker at which the intensity of the indirect sound, reflected one or more times, is equal to the intensity of the direct sound),

r_n = distance from listener to interfering noise or speech source(s),

F = modulation frequency,

T = reverberation time,

$\overline{I}_n$ = long-term average vocal intensity of the interfering source, measured at a distance of 1 m (free-field condition),

$\overline{I}_i$ = long-term average vocal intensity of the primary speaker, measured at 1 m in front of the speaker (free-field condition),

q_s = directivity index of the speaker's voice,

q_1 = gain factor due to binaural release from masking, and

q_n = directivity index of the interfering source.

It should be noted that Eq.(2) does not hold for rooms with very divergent sound-absorption coefficients at the various boundary surfaces (see Plomp et al., 1980b).

For the rest of this paper it is implicitly assumed that q_s=2 (primary speaker facing the listener), q_1=1.5, and q_n=1 (randomly oriented competing speakers). For the critical radius a rule-of-thumb may be

$$r_c \sim \frac{1}{17} \cdot \sqrt{V/T} \quad , \tag{3}$$

where V = room volume in m^3.

An attractive condition for testing the effect of reverberation and noise on speech intelligibility is to position the listener in the diffuse sound field of both the speaker and the noise source. According to Peutz and Klein (1973) the distance between the listener and the sound sources should be larger than $0.2\sqrt{V/T}$ (i.e. $r > 4r_c$ and $r_n > 4r_c$). For this condition it holds: $r^2 \gg r_c^{\,2}$ and $r_n^{\,2} \gg r_c^{\,2}$, so that Eq.(2) is reduced to:

$$m(F) = \frac{1}{1 + \bar{I}_n/\bar{I}_i} \cdot \frac{1}{(1 + 0.207F^2T^2)^{\frac{1}{2}}} \quad . \tag{4}$$

This equation shows that m(F) is the product of two independent factors, one depending on T (and F) and the other one on the signal-to-noise ratio $S/N = 10 \log(\bar{I}_i/\bar{I}_n)$.

According to Houtgast and Steeneken (1978) the Speech Transmission Index (STI) is obtained from 18 values of m(F), measured for F from 0.4 Hz to 20 Hz in 1/3-octave intervals (these 18 frequencies cover the envelope spectrum of speech), by first expressing these values into equivalent S/N ratios in dB:

$$S/N_{eq}(F) = 10 \log \{m(F)/(1-m(F))\} \quad . \tag{5}$$

Then, after each $S/N_{eq}(F)$ value that exceeds the range from -15 dB to +15 dB has been replaced by the limit, the average value $\overline{S/N}_{eq}$ of the 18

values determines the STI:

$$STI = \frac{\overline{S/N}_{eq} + 15}{30} \quad , \text{ where } 0 \leq STI \leq 1 \quad . \tag{6}$$

With the above procedure an optimum relationship between STI and intelligibility scores for Dutch nonsense syllables could be achieved. This enables the interpretation of a given STI value in terms of listening quality. For normal-hearing subjects an STI below 0.4 is considered a poor listening condition, values between 0.4 and 0.6 are fair, values between 0.6 and 0.8 are good, and values greater than 0.8 are excellent.

In Fig.1 it is shown how STI depends on the various variables included in Eq.(2). In this figure STI is plotted as a function of the normalized speaker-to-listener distance r/r_c, with T and S/N ratio as the parameters

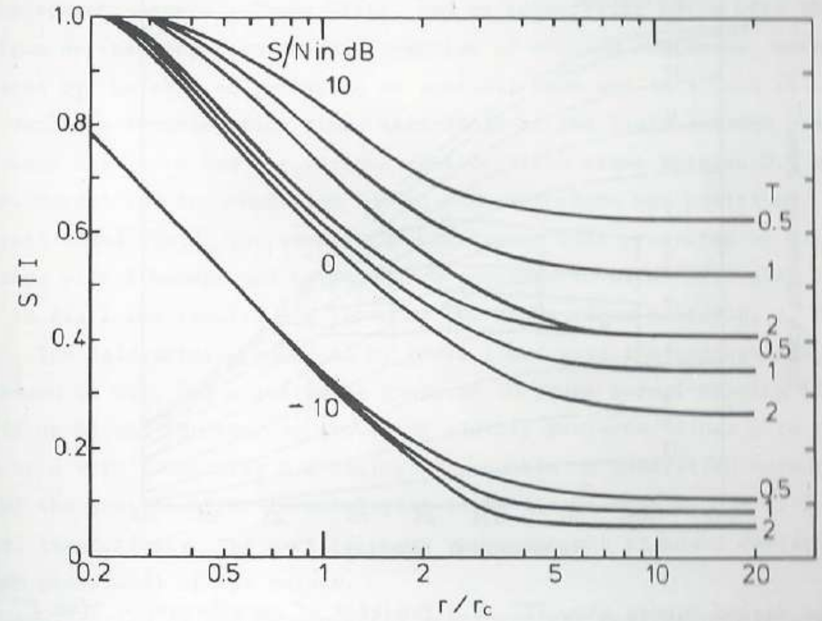


FIG. 1. STI as a function of the speaker-to-listener distance, relative to the critical radius, with T and S/N ratio (at 1 m distance, free-field condition) as the parameters and with the listener situated in the indirect field of the noise source. (Adopted from Plomp and Duquesnoy, 1980).

and $r_n \gg r_c$ (listener in the indirect field of the noise source). The condition S/N=0 dB is relevant to class rooms or conference rooms in which an audience may have to listen to a single speaker, who is interrupted by a second one. Even for T=0.5 s, intelligibility is only good for small distances from the speaker ($r < r_c$), and it deteriorates rapidly for increasing distances. For fair to good intelligibility at all positions in the room, an S/N ratio of 10 dB or more is demanded. The condition S/N=-10 dB fits particularly to rooms like lounges and restaurants in which many local conversations take place. In that case, fair to good intelligibility is only found in the direct sound field ($r < 0.8r_c$).

In Fig.2 it is shown how STI depends on T and S/N ratio for the situation in which Eq.(4) holds. In this figure STI is plotted (dashed curves) as a function of T, with S/N ratio as the parameter. As all conditions on

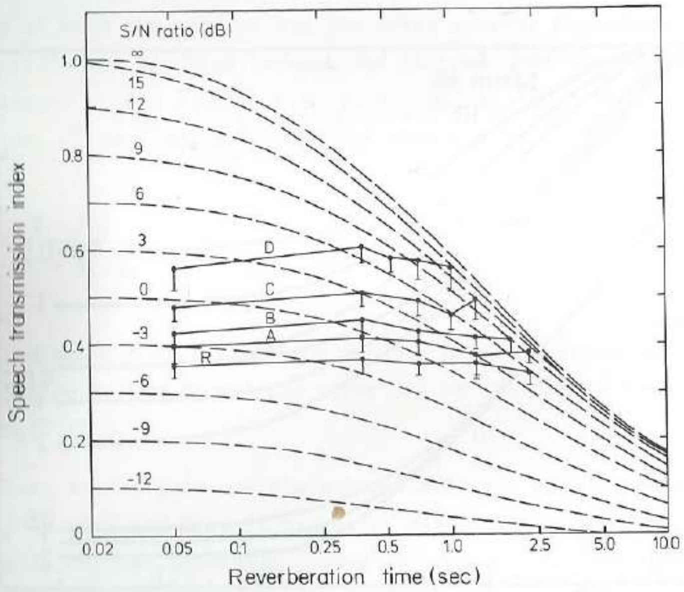


FIG. 2. The dashed curves show STI as a function of reverberation time T with S/N ratio as the parameter (listener situated in the diffuse sound field of both speaker and noise source). The data points and solid lines refer to the average speech intelligibility scores for normal-hearing listeners (line R) and for four groups of elderly subjects (lines A to D). The vertical bars represent the standard deviation of the points, shown one-sided. (Adapted from Duquesnoy and Plomp, 1980).

a horizontal line represent equal intelligibility, this diagram shows, for example, that the condition T=1 s in quiet is comparable with an S/N ratio of 3 dB without reverberation. Note that the curves only hold for distances to the speaker and noise source larger than about $4r_c$.

3. Applicability of STI in cases of presbyacusis

It is evident that for normal-hearing subjects the Speech-Reception Threshold (SRT), defined as the sound-pressure level at which 50% of the speech material is correctly understood, can be represented by a defined horizontal line in Fig.2. Recently, Duquesnoy and Plomp (1980) demonstrated that this also holds in cases of presbyacusis. They investigated the monaural SRT of elderly subjects for sentences with added reverberation. Ten lists of 13 sentences each, developed by Plomp and Mimpen (1979b), which give reliable SRT values (standard deviation about 1 dB) by using an adaptive procedure for the presentation levels of successive sentences, were taken as the speech material. These lists, and an interfering noise with the same spectrum as the long-term average spectrum of the 130 sentences, were reproduced by the same equipment in an anechoic room and in a room (65 m^3) with variable reverberation time. Recordings of the lists and the noise were made for T=0 s and for various reverberation times between 0.4 s and 2.3 s. Except for the condition T=0 s, the microphone was positioned in the indirect sound field. The reverberant sentences were presented by earphone together with a background noise with a constant level of 52.5 dBA.

In Fig.2 the results are shown by the drawn lines marked R, A, B, C, and D. The data points connected by these lines give the average SRT values, expressed in STI, for a reference group of 20 young normal-hearing listeners (line R) and for four subgroups of elderly subjects (lines A to D) who, even at a very favourable S/N ratio, were unable to understand more than 50% of the sentences for reverberation times larger than 2.3, 1.9, 1.3, and 1.0 s, respectively. The vertical bars represent the standard deviation (shown one-sided) of the points.

Application of the Newman-Keuls statistic for testing on differences between means (5%-level of significance) revealed that it is justified to fit the data points of both the reference group and the four subgroups to horizontal lines. We may conclude, therefore, that the combined effect of noise and reverberation on the intelligibility of sentences can be systematically quantified by a single STI value both in normal-hearing and in presbycusic populations. This means that we are able to predict the effect of acoustic environments on speech intelligibility of the hearing impaired by comparing the STI values characteristic of these environments with the STI required in accordance with the SRT (measured in noise for T=0 s) of the hearing impaired. Although the above conclusion is based on experiments in the diffuse sound field, it reasonably holds for any listening situation covered by Eq.(2). All conclusions refer to groups of subjects. The applicability of the STI approach for the individual case has not been definitely established to date.

In a study by Plomp and Mimpen (1979a) data have been published on SRT in noise (for T=0 s) as a function of age. The data were described in terms of a model (see Plomp, 1978) in which every hearing loss for speech (SHL) is interpreted as the sum of a loss of class A (attenuation of the levels of both speech and noise) and a loss of class D (distortion, comparable with a decrease in S/N ratio). In Fig.3 the average percentages of male and female subjects with speech hearing losses in noise (SHL_D) exceeding a value of 1 dB, 2 dB, etc., are given as a function of age.

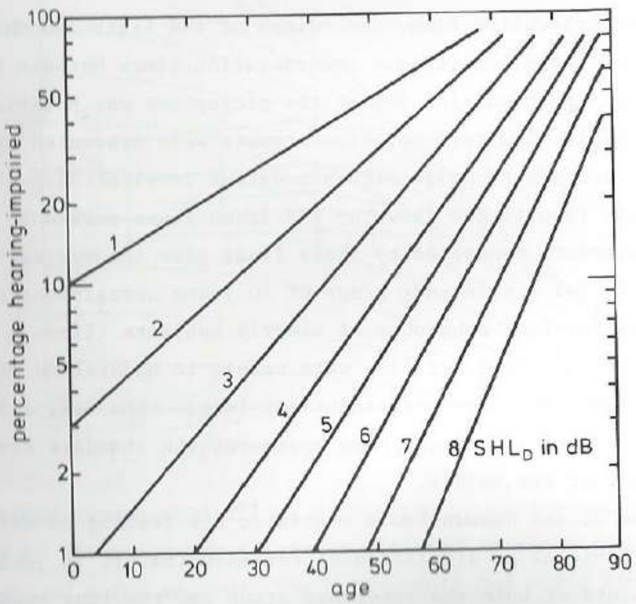


FIG.3. Average percentage of male and female subjects with hearing losses for sentences in noise exceeding 1 dB, 2 dB, etc. up to 8 dB, as a function of age. (Adopted from Plomp and Duquesnoy, 1980).

The losses SHL_D can be expressed as shifts in STI relative to the STI needed for normal-hearing subjects. As the total range of STI covers 30 dB (see Eq.(6)), every 3 dB of SHL_D corresponds to an increase of 0.1 in STI for T=0 s. This means that, on the basis of SHL_D, it is also known which combinations of noise and reverberation will represent acceptable listening conditions for the aged. From Fig.3 we may conclude, for example, that subjects aged between 80 and 90 need an STI which is 0.2 to 0.3 higher than required by normal-hearing subjects.

4. Implications with regard to room acoustics for the aged

The higher STI values needed to compensate for an increased SRT in noise can be achieved by improving the S/N ratio and by reducing T. The S/N ratio increases when the speaker raises his voice level and when the handicapped subject shortens his distance to the speaker. The best way, however, of helping the hearing impaired is by reducing the reverberation time as is discussed by Plomp and Duquesnoy (1980a). They studied the effect of reducing T for two rather diverging rooms:

(1) Auditorium, class room, etc. In these rooms most listeners are situated in the indirect field of the speaker. The interfering sound may originate from an interrupting speaker or from local conversations in a low voice. For conditions where the listener is situated in the diffuse sound field of both the speaker and the noise source, the slopes of the dashed curves in Fig.2 indicate how a given loss SHL_D can be compensated for by reducing T. Let us consider a room with T=1 s and S/N=9 dB, resulting in an STI of 0.51. Then, a listener with SHL_D=3 dB (ΔSTI=0.1) needs an STI of 0.61 to obtain the same intelligibility as normal-hearing listeners for an STI of 0.51. Following the curve of S/N=9 dB, T has to be reduced to 0.48 s for the STI to be equal to 0.61. For the most relevant area of slopes (0.5 s $\leq$ T $\leq$ 2.5 s) an average reduction factor of 0.75 for T per dB of SHL_D has been derived (e.g. for SHL_D=3 dB a factor of $0.75^3 \sim 0.42$ applies).

(2) Lounge, restaurant, etc. In rooms of this kind only the direct sound of the speaker has to be considered. It is assumed, however, that the noise sources are at such a distance that their direct sound can be neglected. Under these conditions a variation of T affects only the level of the indirect sound of the noise source. By means of the Eyring-Norris reverberation formula a reduction factor of 0.82 for T per dB of SHL_D was found.

A better insight into the benefits of reducing T can be obtained from

Fig.3. From the almost equidistant curves in the upper right corner (age $\geq$ 60, percentage $\geq$ 20) it can be concluded that the number of subjects that are handicapped by noise and reverberation to more than a given degree is reduced, on an average, by a factor of 0.7 for every dB in SHL_D compensated for by reducing T by a factor of 0.75 in auditoria, etc., and of 0.82 in restaurants, etc.. For example, suppose that subjects aged 75 are handicapped if $SHL_D > 5$ dB; at this age 35% has $SHL_D > 5$ dB and 25% has $SHL_D > 6$ dB; thus, compensation for 1 dB in SHL_D reduces the number of handicapped subjects from 35% to 25%. In view of the foregoing the authors consider it highly recommendable to give more attention to the acoustics of rooms frequented by elderly people.

References

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

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. (1978). "Application of the Modulation Transfer Function (MTF) in Room Acoustics," Report IZF 1978-20, Institute for Perception TNO, Soesterberg, The Netherlands.

Peutz, V.M.A. and Klein, W. (1973). "Articulation Loss of Consonants Influenced by Noise, Reverberation and Echo," Proc. 1973 FASE Symposium on Speech Intelligibility, Liège, Belgium, 89-97.

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. (1979a). "Speech-Reception Threshold for sentences as a function of age and noise level," J.Acoust.Soc.Am.66, 1333-1342.

Plomp, R. and Mimpen, A.M. (1979b). "Improving the Reliability of Testing the Speech-Reception Threshold for Sentences," Audiology 18, 43-52.

Plomp, R. and Duquesnoy, A.J. (1980a). "Room acoustics for the aged," J. Acoust.Soc.Am.68, 1616-1621.

Plomp, R., Steeneken, H.J.M., and Houtgast, T. (1980b). "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. II. Mirror Image Computer Model Applied to Rectangular Rooms," Acustica 46, 73-81.

+) Reproduced in Chapter I of this thesis

THE SENTENCE INTELLIGIBILITY OF AGED PERSONS IN QUIET AND IN NOISE - TEST OF A MODEL ON SPEECH HEARING

Paper submitted for publication in:
Journal of the Acoustical Society of America.

ABSTRACT

The monaural Speech-Reception Threshold (SRT) for sentences was investigated in quiet and under four noise conditions for 80 male subjects (age 60-90) and 30 female subjects (age 71-89). The noise levels used were 28, 43, 58, and 73 dBA. The noise had the long-term average spectrum of speech. It is shown that a model developed by Plomp (J.Acoust.Soc.Am.63, 533-549, 1978), which interprets any hearing loss for speech (SHL) as a combination of a loss of class A (attenuation of both speech and noise) and a loss of class D (distortion of the sound signals), enables an accurate description to be made of the SRT values measured. Fitting this model, for each individual, to the SRT data yielded values for SHL in quiet (=A+D) and in noise (=D). These individual SHL values were studied in relation to one another and in relation to such parameters as the pure-tone average (PTA), the Fletcher Index (FI), and the increment in intelligibility score per dB near SRT. It was found that (1) subjects with the same SHL in quiet may differ considerably in their SHL in noise; (2) for an individual, PTA and FI are inaccurate measures for predicting SHL in quiet and in noise, and (3) the higher an individual's SHL in noise, the lower is the increase per dB of the intelligibility score for sentences in noise.

INTRODUCTION

Elderly subjects often complain of being unable to understand speech in noisy or reverberant environments. It has been shown previously (Duquesnoy and Plomp, 1980) that reverberation can be replaced by an equivalent noise in predicting the effects of room acoustics upon the Speech-Reception Threshold (SRT) for conversational sentences. Therefore, the present study has been focused exclusively on the effect of noise on the SRT's of the same subjects, who participated in the just-mentioned reverberation study. Their hearing losses are mainly attributable to presbyacusis.

Several authors have published results on speech intelligibility in noise for elderly listeners. In the majority of cases clinical populations were involved (see e.g. Olsen and Carhart, 1967; Groen, 1969; Tillman, Carhart and Olsen, 1970; Jerger, 1973; Surr, 1977; Hayes and Jerger, 1979), consisting of patients with substantial hearing losses attributable to multiple pathological processes rather than to genuine presbyacusis. Van der Waal (1962) reported that only 3% of 604 elderly patients (age 65-74) visiting an audiological center had hearing losses meeting the criterion of

Hinchcliffe (1959) for pure presbyacusis. In addition, most of the clinical studies mentioned above are so fragmentary that it is impossible (1) to draw up representative performance-intensity functions, or (2) to determine how SRT (50%-correct score) of the patients changes as a function of sound-pressure level of interfering noise.

Hitherto, only a few investigators have studied in more detail the effect of noise on the speech intelligibility of aged persons. Kell et al. (1971) presented, as a reference, data on SRT in quiet and in noise for 96 subjects (mean age 64.5) free from aural diseases. The SRT in noise was measured in a sparsely furnished room, as a result of which the data are biased with an unspecified reverberation. As the detrimental effect of reverberation depends on both the amount of reverberation and the size of a subject's hearing loss in noise without reverberation (see Duquesnoy and Plomp, 1980), the data are less suitable, unfortunately, for studying hearing losses in noise.

Jokinen (1973) investigated 100 subjects (age 30-87, grouped by decades) free from aural diseases other than presbyacusis. For disyllables he measured the monaural SRT in quiet and four discrimination scores in white noise. The words in noise were presented at a level 30 dB above an individual's SRT in quiet. Four different signal-to-noise (S/N) ratios were applied. Reference values from 20 normal-hearing listeners (age 20-29) were also included. Jokinen's data will be presented and discussed in Sec.IV.

Orchik and Burgess (1977) published mean scores on synthetic sentence identification for ten subjects (age > 60, individual hearing losses < 30 dB up to 4000 Hz, re ANSI-1969). The sentences were presented monaurally at a fixed level of 40 dB above an individual's SRT in quiet, against a background of continuous discourse. Five different S/N ratios were used. Relative to a reference group of ten subjects (age 20-29, individual hearing losses < 25 dB up to 4000 Hz, re ANSI-1969), the elderly subjects had a mean hearing loss of 10 dB in terms of S/N ratio.

Kalikow et al. (1977) presented ten elderly subjects (age 60-75, individual hearing losses <20 dB up to 4000 Hz, re ANSI-1969) with high-predictability (HP) and low-predictability (LP) sentences against babble-type noise at various S/N ratios. Relative to ten normal-hearing subjects (age 18-25), the elderly subjects had a score reduced by about 14%, corresponding to nearly 1 dB for the HP sentences and 2 dB for the LP sentences in terms of S/N ratio.

The experimental conditions in the three last-mentioned studies are quite different and the results presented are still insufficient to gain a good insight into what noise levels are detrimental to the everyday speech intelligibility of elderly (presbyacusic) subjects. Therefore, in the present study measurements of SRT for sentences were performed in quiet and against masking noise at four different sound-pressure levels up to 73 dBA. Measuring SRT in this systematic way makes it possible to test a model of hearing losses for speech developed by Plomp (1978). This model describes SRT as a function of noise level by means of two parameters specific to hearing loss. Plomp and Mimpen (1979a) applied the model in a study of SRT as a function of age. In the present paper its applicability is tested in more detail by means of a parameter-estimation procedure presented in Sec.I.

I. VERIFICATION OF A NEW MODEL OF HEARING LOSSES FOR SENTENCES

A. A model of hearing losses for speech

It was experimentally verified by Hawkins and Stevens (1950) that for normal-hearing listeners SRT is governed by the S/N ratio. According to Plomp (1978), the data of Hawkins and Stevens agree excellently with

$$SRT = 10 \log \left[10^{L_0/10} + 10^{(L_N - \Delta L_{SN})/10}\right] \text{ in dBA,} \qquad (1)$$

where: L_0 = SRT in quiet for the normal-hearing (in dBA),
L_N = sound-pressure level of the masking noise (in dBA), and
ΔL_{SN} = the number of decibels that SRT in noise for the normal-hearing is below L_N, thus $L_N - \Delta L_{SN}$ represents SRT in noise for the normal-hearing.

If we assume that $L_0 = (L_i - \Delta L_{SN})$, where L_i is the internal noise level of the ear, it is evident that SRT is determined by the combined masking effects of the internal noise L_i and the external noise L_N. In Fig.1 the lower curve, marked R, shows the SRT for sentences as a function of noise level L_N for normal-hearing listeners tested in the present experiments (details in Sec.III).

Plomp extended the above model in order to describe the effect of hearing impairment on SRT in quiet and in noise. To this end, any Speech-Hearing Loss (SHL) was interpreted as caused by a combination of two formalistic impairments: (1) attenuation of all sounds entering the inner ear, resulting in the definition of a class A loss, and (2) distortion of the sounds, resulting in a class D loss. A loss of class A manifests itself in a threshold shift for speech in quiet and it can be fully compensated for by increasing the level of the sounds entering the ear. A loss of class D affects speech intelligibility both in quiet and in noise. The distortion hampers discrimination between signals presented simultaneously. Therefore, a loss of class D can be compensated for by improving the S/N ratio, provided that its effect on intelligibility is not too detrimental. Assuming that the required increase in S/N ratio is independent of noise level, a loss of class D manifests itself in a threshold shift, both in quiet and in noise. Hence, for the hearing impaired the SRT as a function of noise level L_N can be described by (see Plomp, 1978):

$$SRT = 10 \log \left[10^{(L_0+A+D)/10} + 10^{(L_N - \Delta L_{SN} + D)/10}\right] \text{ in dBA,} \qquad (2)$$

where: L_0, L_N, and ΔL_{SN} as defined in Eq.(1),
A+D = Speech-Hearing Loss in quiet (=SHL_{A+D}) re L_0, and
D = Speech-Hearing Loss in noise (=SHL_D) re $L_N - \Delta L_{SN}$.

In Fig.1 it is illustrated how SRT changes, according to Eq.(2), as a function of masking noise level L_N, with the losses A and D as parameters. As can be seen, Eq.(2) represents curves similar to Eq.(1), with slopes of 0° in quiet (L_N=0 dBA) and slopes of 45° at high noise levels (L_N>60 dBA). In other words, for low noise levels SRT is fairly constant, as it is largely determined by SHL in quiet (=A+D). At higher noise levels SRT increases proportionally to the increment of the noise level, independently of the given value for SHL in noise (=D). This theoretical approach suggests that, in principle, any hearing loss for speech can be reliably characterized by measuring just two thresholds, namely SRT in quiet and SRT at a single (high) noise level.

B. Test procedure for the model

The above theoretical considerations have to be verified experimentally. Two aspects are of particular interest. First, does any given hearing loss of class A lose its influence on the SRT of hearing-impaired listeners, for increasing noise level L_N, in the same way as the internal noise of the ear loses its influence on SRT for normal-hearing listeners? In other words, has any experimental SRT curve a turning point as described by Eq.(2)? Secondly, does a class D hearing loss independent of the level L_N exist? In other words, does SRT at higher noise levels increase in direct

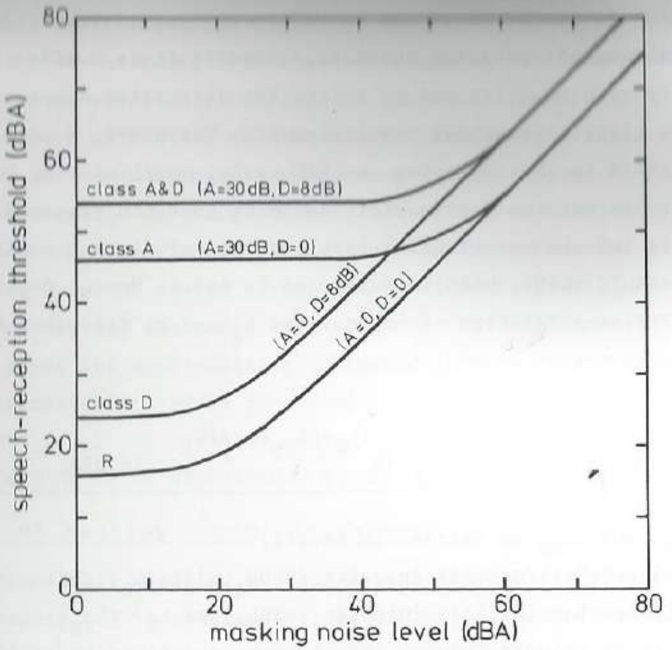


FIG.1. Speech-Reception Threshold (SRT) for sentences as a function of masking noise level. The lower curve, marked R, was determined according to Eq.(1) for normal-hearing listeners, for which it was experimentally found that L_0=16.2 dBA and ΔL_{SN}=4.8 dB. The other curves were determined according to Eq.(2). They hold for theoretical speech-hearing losses in quiet (A+D) and in noise (D) as indicated in the figure.

proportion to the increment of the noise for any given loss D? For verification of these aspects, monaural SRT's of hearing-impaired subjects were measured under five conditions (details in Sec.II) and an iterative procedure was developed for fitting a curve defined by Eq.(2) to the five data points collected per subject.

The fitting procedure applies the sum of the five quadratic differences between the SRT's measured and the SRT's predicted according to Eq.(2) as error criterion. This criterion is minimized by adjusting estimates for A and D (written as $\hat{A}$ and $\hat{D}$) by means of the steepest descent method. The initial estimates $\hat{A}_0+\hat{D}_0$ and $\hat{D}_0$ are chosen as follows: $\hat{A}_0+\hat{D}_0 = SRT_0-L_0$, where SRT_0 is the threshold of the subject measured in quiet, and $\hat{D}_0 = 0.3(\hat{A}_0+\hat{D}_0)$ (ratio adopted from Plomp, 1978, Fig.8). The iterative procedure is stopped when the partial derivatives of the error criterion with respect to A and D become smaller than 0.02 dB. This corresponds to adjustments of $\hat{A}$ and $\hat{D}$ smaller than 0.002 dB. The resulting $\hat{A}$ and $\hat{D}$ values define the best-fitting curve. Next, the standard deviation of the five data points from the best-fitting curve is calculated as a measure of goodness of fit.

Some caution should be exercised when applying this fitting procedure. It is inherent in Eq.(2) that a reliable estimate of D is only guaranteed if, for one noise level L_N at least, the contribution to SRT of the first power term of ten, the exponent of which depends on both A and D, is small relative to the second term. In practice, this means that it should be checked that, at least, the SRT measured at the highest noise level L_{Nmax} is unequivocally positioned on the rising flank of the SRT curve. Assuming a ratio of 1 : 10 between the terms as the desirable minimum, this will be the case if $(L_0+A+D)/10 < ((L_{Nmax} - \Delta L_{SN}+D)-1$, or:

$$A < L_{Nmax}-L_0-\Delta L_{SN}-10 \quad . \qquad (3)$$

As will be demonstrated by the experiments to be considered next, the model describes the SRT for sentences quite well. Consequently, the speech hearing capacities of a subject can be characterized adequately by just two numbers, viz. SHL_{A+D} and SHL_D.

II. EXPERIMENTS

A. Speech material

As everyday listening situations were of interest, conversational sentences were preferred as speech material. An accurate test was required for measuring SRT in noise in such a way that even a change of 1 dB in S/N ratio should have considerable impact on intelligibility. Plomp and Mimpen (1979b) developed such a test for the Dutch language. It consisted of ten carefully selected lists of 13 sentences each and a special masking noise having the same spectrum as the long-term average spectrum of the 130 sentences.

By using a simple up-and-down procedure for the presentation level of the sentences, an estimate of SRT (50% threshold) can be obtained for each list. The procedure requires a correct repetition of the entire sentence for a correct response. If a listener is unable to repeat sentences correctly even under excellent listening conditions, the above test cannot be applied. None of our subjects belonged to that category.

The test-retest reliability of SRT values was investigated in two separate experiments in which each listening condition was tested twice with different lists (Duquesnoy, 1977; Plomp and Mimpen, 1979b). The individual SRT values found had a standard error of estimate of only about 1 dB, including systematic differences among lists.

B. Experimental conditions

The sentence lists were recorded on one track of a tape. On the other track the standard noise with exactly the same intensity and spectrum as the long-term average of the sentences was recorded. The output signals of both channels of the tape recorder could be attenuated separately. After proper attenuation the signals were mixed and presented monaurally over earphones with circumaural cups (Sharp Scintrex Mark IV, ambient noise attenuation >35 dB at 1000 Hz). The earphone sound levels were calibrated against free-field conditions by measuring the monaural SRT in quiet with earphones and with a loudspeaker in front of the listener. The sound level corresponding to SRT was determined by an additional sound level measurement (in dBA) at the position of the listener's ear, with the listener removed (details in Duquesnoy, 1977).

The test, which took 35 to 40 min, was carried out either in an anechoic room in our laboratory or in a reasonably quiet room (noise level <35 dBA) in a home for the aged. First, air-conduction tone audiograms were determined for frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz, which took about 15 min. The pure tones were presented over Beyer DT-48 earphones with flat cushions. Then, the ear with the smaller pure-tone average (PTA, average hearing loss for 500, 1000 and 2000 Hz) was chosen to measure SRT in quiet and against four noise levels of 28, 43, 58, and 73 dBA. The five sentence lists involved were always presented in the same order. To eliminate the effects of training and fatigue as much as possible, the test conditions were counterbalanced for every ten subjects tested.

C. Subjects

The elderly subjects were recruited from two different populations. They consisted of 50 male employees (age 60-67) of the Free University, and of 60 inhabitants of a home for the aged (30 females, age 71-89; 30 males, age 69-90).

The tone audiograms of the subjects were studied to find manifestations

of presbyacusis. As Schuknecht (1974) pointed out, four types of presbyacusis can currently be identified which show an enormous variety in audiometric configurations. This makes identification of presbyacusis solely based on the audiogram rather precarious. In order to aid the interpretation of audiograms, Robinson and Sutton (1979) presented, as a function of age, normative pure-tone hearing levels with a defined probability of being exclusively associated with presbyacusis. In the present case, normative levels representing about 80% probability (viz. the levels given by ΔH+SD, see Robinson and Sutton, 1979, p.333) were chosen and 2 dB was added to these levels in order to make allowance for inaccuracies in the audiograms measured. Interpretation of the 110 audiograms resulted in 77 audiograms for which, at least for the better ear, the hearing losses at all six frequencies were below the normative levels. Thus, the hearing losses of the 77 subjects in question (70%) were taken as being mainly attributable to presbyacusis. From a pattern study of the 33 audiograms not satisfying the criteria it was concluded, with proper restrictions, that four subjects had hearing losses mainly due to noise injuries (noise notch in the audiogram), and that the others had conductive or mixed-type hearing losses (predominantly flat audiograms).

In addition to the elderly people, 20 young normal-hearing subjects (mean age 21; individual better ear hearing-levels at the six above-mentioned frequencies <20 dB re ISO-389) were tested in order to obtain a reference for our speech material.

III. RESULTS

A. Verification of the SRT model

First, the fitting procedure was applied to the individual SRT values of the 20 normal-hearing subjects. By definition, this group as a whole was free from hearing losses for speech. Indeed, by introducing in Eq.(2) as reference values: L_0=16.2 dBA and ΔL_{SN}=4.8 dB, the medians of the individual SHL values, A+D (quiet) and D (noise), were made zero. The median of the individual standard deviations (s.d.) of the SRT values from the best-fitting curves was 0.84 dB, which was taken as the reference for goodness of fit.

After the above standardization, the fitting procedure was applied to the individual SRT values of all 110 elderly subjects, irrespective of type of hearing impairment. For a reliable operation of the estimation procedure, it is necessary that the SRT measured at the highest noise level is clearly situated on the rising flank of the curve given by Eq.(2). For the present study, in which L_{Nmax} is 73 dBA, this means that the allowable maximum value of class A hearing losses is 42 dB, according to Eq.(3). This restriction led to the exclusion of eight elderly subjects from further analysis. All eight subjects had predominantly conductive hearing losses. For the remaining 102 subjects the curve fitting resulted, per subject, in reliable estimates for A+D and D, together with the standard deviation of fit (s.d.).

The applicability of Plomp's model, irrespective of the gravity of hearing impairment, was a matter of prime importance. In view of this, the range of frequently occurring D values (from −1.0 up to 7.0 dB) was divided into four intervals of 2 dB. Thus, four groups of elderly subjects were formed on the basis of SHL_D. Six subjects with $D \geq 7.0$ dB were not grouped

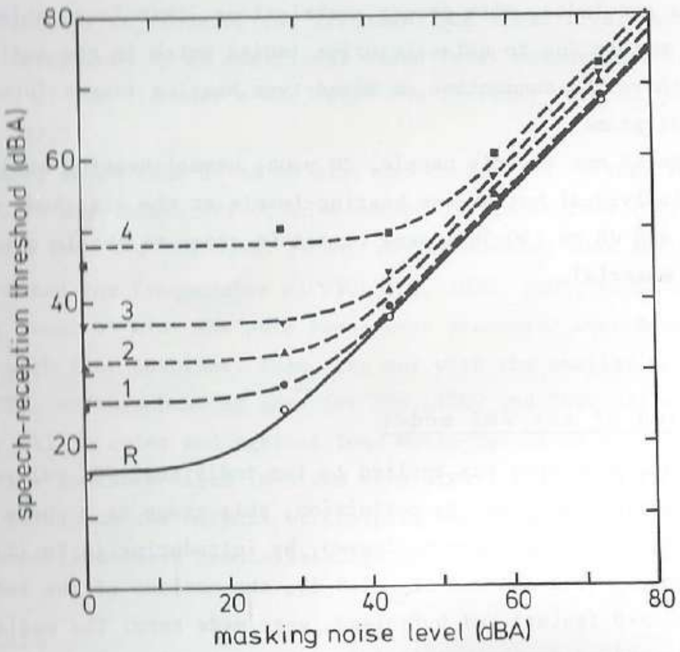


FIG. 2. SRT for sentences as a function of noise level. The data points shown are the medians of SRT as given in Table I for five groups of subjects. The curves marked R, 1, 2, 3, and 4, represent the curves best fitting the experimental data, according to Eq.(2). The fitted SRT values and the standard deviations of best-fit are given in Table II.

because of their small number.

In Table I the medians and upper and lower quartiles of the five SRT values measured per subject, and of the individual values for A+D, D, and s.d. are shown separately for each group. The results for the normal-hearing group are included as a reference.

In Fig.2 the medians for SRT are plotted, per group, as a function of noise level. The dotted curves also shown are defined by Eq.(2) and represent the curves best fitting the SRT medians plotted.

TABLE I. Medians and upper and lower quartiles of the SRT values measured at five noise levels (L_N, in dBA), as shown at the top of the table, and of the estimated individual hearing losses A+D and D together with the standard deviation of fit (s.d.). The groups 1 to 4, consisting of the elderly subjects, were formed on the basis of SHL_D values according to the intervals indicated. The number of subjects (Ss) in each group is also given.

	SRT (dBA)					(dB)		
	L_N=0	28	43	58	73 dBA	SHL_{A+D}	SHL_D	s.d.
Ref.group (20 Ss)								
lower quartile	13.6	24.4	37.6	52.0	67.2	−2.4	−0.5	0.60
median	16.0	25.0	38.0	53.2	68.0	0.0	0.0	0.84
upper quartile	18.8	25.6	39.2	53.6	68.8	2.8	0.2	1.34
Group 1 (23 Ss) →(−1.0 dB ≤ SHL_D < 1.0 dB)								
lower quartile	21.6	26.4	38.0	52.4	68.0	5.1	−0.1	0.74
median	25.6	28.4	39.6	52.8	68.4	8.9	0.3	0.88
upper quartile	31.6	33.6	40.4	53.6	69.6	15.9	0.6	1.26
Group 2 (35 Ss) →(1.0 dB ≤ SHL_D < 3.0 dB)								
lower quartile	22.8	28.8	40.4	54.0	69.2	6.0	1.4	0.87
median	30.8	33.2	42.0	54.4	70.4	14.7	2.2	1.23
upper quartile	42.0	40.0	44.0	55.6	70.8	24.3	2.6	1.79
Group 3 (29 Ss) →(3.0 dB ≤ SHL_D < 5.0 dB)								
lower quartile	31.2	34.0	41.6	56.4	70.8	15.6	3.3	1.12
median	36.8	36.8	44.4	56.8	71.6	19.8	3.7	1.50
upper quartile	46.0	45.6	49.2	58.4	72.4	30.2	4.2	1.72
Group 4 (9 Ss) →(5.0 dB ≤ SHL_D < 7.0 dB)								
lower quartile	39.2	42.0	46.4	58.4	72.6	24.2	5.2	1.28
median	44.8	51.2	49.6	60.4	73.2	31.5	5.4	1.87
upper quartile	58.6	56.0	57.4	62.8	74.2	40.2	5.7	2.60

In Table II the differences between the SRT medians and the corresponding SRT values on the curves best fitting these medians are shown. In addition, the medians of the A+D, D, and s.d. values specific for the individuals of a group are compared with the corresponding values of the best-fitting curve specific for the group as a whole.

The validity of the SRT model can be demonstrated by means of the results presented above on individual and group fitting (see Sec.IIIA).

TABLE II. Comparison of individual data with group data. The rows indicated by "exp. median" give the median of the individual values for SRT, SHL_{A+D}, SHL_D, and s.d. (adopted from Table I, row "median"). The next rows give the SRT values, defined by the curve best fitting the five median SRT values of each group, with the corresponding best estimates of SHL_{A+D}, SHL_D, and s.d.. The rows indicated by Δ(exp-fit) show the differences between the rows of individual data and data for the group as a whole.

	SRT (dBA)					(dB)		
	L_N=0	28	43	58	73 dBA	SHL_{A+D}	SHL_D	s.d.
Ref.group (20 Ss)								
exp.median	16.0	25.0	38.0	53.2	68.0	-0.01	-0.02	0.84
best fitting	16.1	24.1	38.3	53.3	68.3	-0.05	-0.08	0.52
Δ(exp-fit)	-0.1	0.9	-0.3	-0.1	-0.3	0.04	0.06	
Group 1 (23 Ss)								
exp.median	25.6	28.4	39.6	52.8	68.4	8.9	0.3	0.88
best fitting	26.0	27.9	38.8	53.6	68.6	9.6	0.3	0.64
Δ(exp-fit)	-0.4	0.5	0.8	-0.8	-0.2	-0.7	0.0	
Group 2 (35 Ss)								
exp.median	30.8	33.2	42.0	54.4	70.4	14.7	2.2	1.23
best fitting	31.5	32.5	40.9	55.4	70.4	15.1	2.2	0.90
Δ(exp-fit)	-0.7	0.7	1.1	-1.0	0.0	-0.4	0.0	
Group 3 (29 Ss)								
exp.median	36.8	36.8	44.4	56.8	71.6	19.8	3.7	1.50
best fitting	36.7	37.2	43.2	57.1	72.1	20.3	3.9	0.70
Δ(exp-fit)	0.1	-0.4	1.2	-0.3	-0.5	-0.5	-0.2	
Group 4 (9 Ss)								
exp.median	44.8	51.2	49.6	60.4	73.2	31.5	5.4	1.87
best fitting	48.0	48.1	49.5	59.5	74.2	31.6	6.0	2.34
Δ(exp-fit)	-3.2	3.1	0.1	0.9	-1.0	-0.1	-0.6	

B. Relations between SHL_{A+D}, SHL_D, and PTA

It is of practical importance to know how strongly an individual's SHL in noise correlates with his SHL in quiet. Therefore, the next results concern relations between SHL_{A+D} and SHL_D, split for presbyacusis and other types of hearing impairment. Figure 3 is a scatter diagram of SHL_D(=D) versus SHL_{A+D}(=A+D) for the better ear of the 77 subjects constituting the presbyacusis group. The regression of SHL_D on SHL_{A+D} is given by: $\overline{D} = 0.13(A+D)+0.43$, and the correlation coefficient is 0.62. The standard deviation of individual points from the regression line is 1.53 dB.

Figure 4 shows SHL_D as a function of SHL_{A+D} for the better ear of the remaining 25 elderly subjects, who had losses not meeting the criteria for presbyacusis as given by Robinson and Sutton (see Sec.IIC). For this heterogeneous group the regression line is given by: $\overline{D} = 0.14(A+D)-0.17$ and the correlation coefficient is 0.62. The standard deviation of individual points from the regression line is 2.37 dB. The slopes of the two regression lines do not differ significantly.

It is interesting to investigate how well the SHL values A+D (quiet) and D (noise), resulting from our hearing model, can be predicted from the PTA (average hearing level for 500, 1000 and 2000 Hz). The PTA is a very common audiometric measure for assessing hearing impairment, although its usefulness for predicting speech intelligibility seems questionable (Noble,

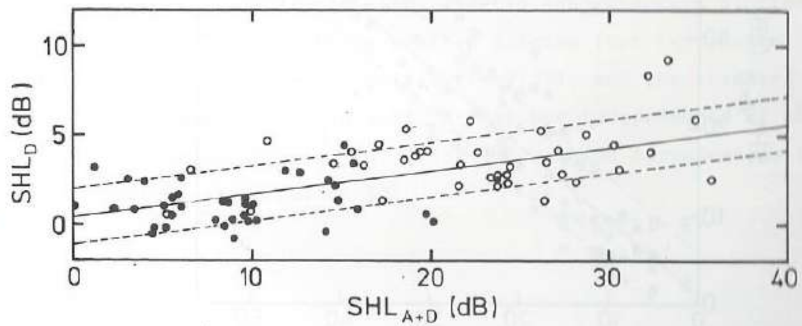


FIG. 3. SHL_D as a function of SHL_{A+D} for the better ear of 77 subjects with predominantly presbyacusic losses. The regression line is given by: $\overline{D} = 0.13(A+D)+0.43$. The standard deviation of individual points from this line, indicated by the two dashed lines, is 1.53 dB. Filled circles represent subjects aged 60-67; open circles subjects aged 69-90.

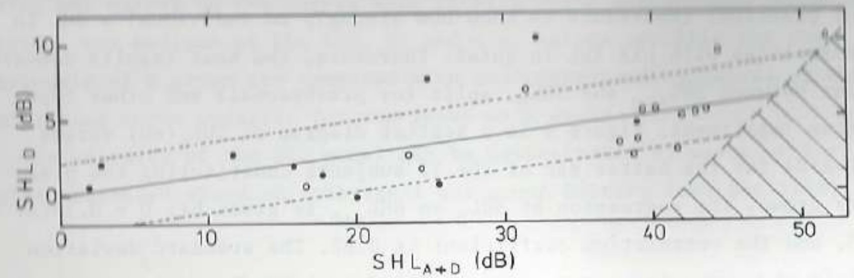


FIG.4. SHL_D as a function of SHL_{A+D} for the better ear of 25 subjects with hearing losses of heterogeneous origin. The regression line is: $\overline{D} = 0.14(A+D)-0.17$. The standard deviation from this line, as indicated by the dashed lines, is 2.37 dB. The dashed area is excluded because of restrictions imposed on values for A, according to Eq.(3). Filled circles represent subjects aged 60-67, open circles subjects aged 69-90.

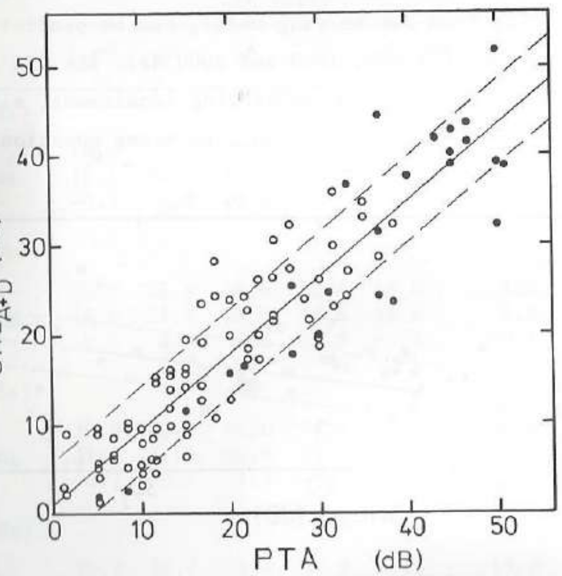


FIG.5. SHL_{A+D} versus PTA for the better ear of all 102 elderly subjects. The regression line is: $\overline{A+D} = 0.85PTA+1.0$. The standard deviation from this line, as indicated by the dashed lines, is 4.75 dB. Open circles represent subjects with presbyacusic losses, filled circles represent the heterogeneous losses.

1973). In Fig.5 the 102 individual SHL_{A+D} values are plotted versus PTA. The regression of SHL_{A+D} on PTA is given by: $\overline{A+D} = 0.85PTA+1.0$, and the correlation coefficient is 0.92. The standard deviation from the regression line is 4.75 dB. Similarly, for the scatter diagram in which SHL_D was plotted versus PTA (not reproduced), the regression was found to be: $\overline{D} = 0.11 PTA+0.5$ ($r=0.61$). The standard deviation from the regression line is 1.85 dB.

C. Increment in intelligibility score near SRT

Generally, the chance of understanding a sentence correctly increases with better S/N ratios. The present data make it possible to investigate if, for the elderly subjects, this increment in intelligibility is affected by the extent of their SHL in noise.

The slopes near SRT (50%-correct score) of the individual intelligibility curves for sentences were calculated as follows. In all five lists presented per subject, the first three sentences were ignored. The presentation levels of the remaining 50 sentences, ten of which represent a separate listening condition, were shifted so that the corresponding five SRT values coincided at a level of zero. Then, the number of sentences repeated correctly at levels of 1, 2, 3 dB, etc. below and above zero was determined. Over the middle range (4 dB width) of the resulting cumulative probability distribution curve, the slope ΔS_{50} of the curve was estimated by means of linear regression (ΔS_{50} expressed in %-score increment per dB).

The individual slopes for the 102 elderly subjects were plotted as a function of SHL_D. In the resulting scatter diagram (not reproduced) the linear regression is: $\overline{\Delta S_{50}} = -0.7D+19.9$ ($r=-0.32$), and the standard deviation from the regression line is 4.9%. A test for independence of ΔS_{50} and D was rejected at the 1% level ($t=-3.35$, $df=100$), which corroborates the regression of individual slopes on SHL_D.

IV. DISCUSSION

A. Applicability of the model

First, the results for individuals will be considered. As shown in Table I, in the column marked s.d., the median of the individual standard deviations of fit is only 0.84 dB for young subjects (reference group) and increases from 0.88 to 1.87 dB for elderly subjects with increasing hearing losses. These values are similar to the standard deviations of individual SRT's.

Furthermore, it was examined whether the differences between the measured and predicted SRT's were positive or negative in a sufficiently random way over the conditions tested. According to the theory of runs (Dixon and Massey, 1969), this was the case for all subjects and conditions. The small standard deviations for individuals and the random distribution of the sign of the above-mentioned differences are first indications of the validity of the model.

Next, the results for groups of subjects will be considered. As stated in Sec.IB, two aspects of the model SRT-curve should be considered especially for any given SHL in quiet and in noise, namely (1) the shape of the turning point at moderate noise levels, and (2) the slope at higher noise levels. From the best-fitting curves presented in Fig.2 it can be concluded both that the medians at moderate noise levels are described well by the turning points of the curves, and that the medians at higher noise levels increase in direct proportion to the increment of the noise, which is also in accordance with the model. Furthermore, from a comparison of individual and group s.d.'s (Tables I and II), it can be concluded that there is a considerable gain in goodness of fit for the group curves as compared to the individual curves. Group 4 is an exception to this, perhaps because of the small number of subjects, in addition to a large spread in hearing losses for low noise levels (L_N<40 dBA). The gain in goodness of fit demonstrates that the description of SRT data by means of Eq.(2) is even more valid, because the spread of the data, due to experimental and individual inaccuracies, is reduced by considering homogeneous groups of subjects. In summary, the model gives a valid description of the experimental individual and group data.

Consequently, a subject's hearing loss for speech is characterized by only two numbers, viz. SHL_{A+D} and SHL_D. The estimates presented in Sec.III resulted from fitting the model to five SRT values, but, since the validity of the model has been demonstrated, any hearing loss for speech can be defined, essentially, by measuring not more than two thresholds, namely SRT in quiet and SRT at a single high noise level (e.g. 73 dBA). The fitting procedure may be left out. It is shown in Table III that this approach results in acceptable, though less accurate, estimates of SHL_{A+D} and SHL_D. In the table the means and standard deviations of the differences ΔSHL_{A+D} and ΔSHL_D are given separately for the reference group (20 ears) and the elderly subjects (102 ears). ΔSHL_{A+D} is the difference for each ear between SHL_{A+D}, determined solely on the basis of the SRT measured in quiet, and

TABLE III. Means and standard deviations (s.d.) of the individual differences ΔSHL_{A+D} and ΔSHL_D, separately given for the reference group and the elderly subjects.

Subjects	ΔSHL_{A+D} (dB)		ΔSHL_D (dB)	
	mean	s.d.	mean	s.d.
Reference (20 ears)	0.1	0.3	-0.3	0.9
Elderly (102 ears)	-0.2	1.4	-0.2	1.3

the value $\hat{A}+\hat{D}$ resulting from the fitting procedure applied to five SRT values. ΔSHL_D is the difference for each ear between SHL_D, determined solely on the basis of the SRT measured in noise at a level of 73 dBA, and the value $\hat{D}$ resulting from the fitting procedure. As can be seen, all standard deviations found are small enough to guarantee reasonably accurate individual estimates of SHL_{A+D} and SHL_D with only two SRT measurements.

B. Relations between SHL_{A+D}, SHL_D, and the PTA

From Table I, columns marked SHL_{A+D} and SHL_D, it can be concluded that, for all groups, the interquartile ranges of SHL_{A+D} are considerable (5.2 to 18.3 dB), although the ranges for SHL_D are small (0.5 to 1.2 dB) as a consequence of the specific group division. Apparently, remarkable interindividual differences exist in the relation between hearing capacity in quiet and in noise. This is clearly illustrated by the scatter diagrams of Figs. 3 and 4. Although for both groups (presbyacusis and heterogeneous) the correlation between SHL_{A+D} and SHL_D is fairly high, the standard error of regression is 1.53 dB and 2.37 dB, respectively. This indicates that for the individual ears the relationship is not that close. Subjects with the same hearing loss in quiet may have differences of several dB in their hearing loss in noise. As shown in Sec.IIIC for different groups of subjects, each dB of hearing loss in S/N ratio lowers sentence intelligibility by 17% to 20%. In view of this great impact, hearing loss in noise should always be measured in order to obtain a good insight into the speech hearing ability of a subject.

The average ratio of 0.13 between SHL_D and SHL_{A+D} for our presbyacusis population is substantially smaller than the value of 0.22 arrived at by Plomp and Mimpen (1979a) with the same sentence material and the same hearing model. Therefore, their data were reanalyzed by means of our approach. This means that (1) only the SRT data for the better ear (smaller PTA) were selected for analysis, (2) the population of, in total, 132 subjects was split into a presbyacusis group, according to the Robinson and Sutton criteria (see Sec.IIC), and a group with hearing losses of various origin, and (3) in order to guarantee a proper operation of the fitting procedure, the allowable maximum value of A, according to Eq.(3), was 33 dB, as in Plomp and Mimpen's data: L_0=19.0 dBA, ΔL_{SN}=5.4 dB, and L_{Nmax}=67.5 dBA. This resulted in a presbyacusis group consisting of 56 subjects (or ears) and a heterogeneous group of 59 subjects. Two scatter diagrams of SHL_D versus SHL_{A+D} were made. The regression of SHL_D on SHL_{A+D} yielded the following results: for the presbyacusis group regression is given by $\overline{D} = 0.15(A+D)+1.2$ with $r = 0.53$ and s.d.=1.75 dB, and for the heterogeneous group $\overline{D} = 0.15(A+D)+1.55$ with $r = 0.57$ and s.d.=2.37 dB. These values agree quite well with ours. The same holds for data on SHL_{A+D} and SHL_D as derived from results published by Jokinen (1973). Interpolation of his mean scores plotted, per age decade, as a function of S/N ratio yields values for SRT in noise. From these, speech hearing losses in noise can be determined relative to reference values also published. The losses are given in Table IV. The regression of SHL_D on SHL_{A+D} for the decades is given by: $\overline{D} = 0.15(A+D)+1.8$ with $r = 0.91$. On the basis of these uniform results, the following conclusion is drawn: both for populations of elderly subjects with

TABLE IV. Mean values for SHL_{A+D} and SHL_D in cases of presbyacusis, derived from data published by Jokinen (1973). The losses, expressed in dB relative to reference values from 20 normal-hearing subjects (age 20-29), are given as a function of age decade (20 subjects per decade).

Age group	Mean age	SHL_{A+D}	SHL_D
30 - 39	35	1.5	1.3
40 - 49	44	3.3	2.7
50 - 59	53	9.5	4.0
60 - 69	66	18.2	3.7
70 - 87	78	26.6	6.0

mainly presbyacusis and for populations with losses of various origin the ratio between a given SHL_{A+D} value and the expected SHL_D value is approximately 7 : 1.

It should be noted that the mean hearing loss of 10 dB in noise reported by Orchik and Burgess (1977) is substantially larger than the average value of 2.3 dB for the elderly subjects in the present study. Orchik and Burgess applied synthetic sentence identification (SSI) as test method. Results from SSI-tests are not comparable with our results, since SSI constitutes an artificial test situation lacking any links with reality. The subject tested cannot make use of linguistic redundancy or contextual constraints, and continuously gets false speech cues, because the primary and the interfering speech originate from the same male speaker. The SSI-test seems to be an intelligence rather than an intelligibility test.

Concerning the regression of SHL_{A+D} on PTA, the slope of the regression line shown in Fig.5 is 0.85. Thus, for elderly subjects the observed SHL_{A+D} (quiet condition) will generally be smaller than predicted from pure-tone hearing losses. This finding agrees with the regression found by Plomp and Mimpen (1979a, Fig.7), but contradicts the conclusion of Gjaevenes (1969), based on results for 100 presbyacusic subjects, that at advanced age the observed SRT will generally be higher than predicted from pure-tone hearing loss alone. The large standard deviation from the regression line of 4.75 dB is in agreement with Plomp and Mimpen (1979a), who found 7.7 dB, and Gjaevenes (1969), who found an interquartile range of 9.3 dB. It must be concluded that for elderly subjects the PTA is an unsuitable measure for reliably predicting SHL in quiet. This conclusion should be extended to the noise condition (prediction of SHL_D) for which a standard deviation of 1.85 dB was calculated. This is a large value, since 1 dB in S/N ratio corresponds to 17 to 20% in speech intelligibility scores.

One reason why SHL_{A+D} is smaller than predicted from PTA might be that the sentences used consist for about 40% of disyllables. According to Young and Gibbons (1962), in hard-of-hearing persons with predominantly high-frequency hearing losses the intelligibility of spondees in quiet correlates better with their hearing levels at 500 and 1000 Hz, than at 2000 Hz. Smoorenburg et al. (1981), who measured SRT's in quiet for the same sentences as used in the present study on subjects with noise-induced hearing losses, also found better correlations for 500 and 1000 Hz, than for 2000 Hz. Therefore, the Fletcher Index (FI), defined as the average hearing loss at the best two frequencies from 500, 1000, and 2000 Hz, was also tested as a predictor of SHL_{A+D}. In cases of sloping audiograms FI will add more weight to the low frequencies. The same 102 individual SHL_{A+D} values as used in Fig.5 were plotted versus FI. Regression of SHL_{A+D} on FI yielded: $\overline{A+D} = 0.87\,FI + 4.1$, with $r = 0.92$ and s.d.=4.75 dB. Similarly, the regression of SHL_D on FI is given by: $\overline{D} = 0.10\,FI + 1.1$, with $r = 0.55$ and s.d.=1.96 dB. Thus, it can be concluded that FI is as unsuitable a measure as PTA for reliably predicting an individual's SHL in quiet and in noise, and, although the correlation coefficients are almost similar, the regression on FI is worse than on PTA in view of the greater offsets (4.1 dB versus 1.0 dB, and 1.1 dB versus 0.5 dB).

C. Slopes of intelligibility curves

The lists of sentences yield very distinct SRT values. This can be explained on the basis of the steep slopes of the intelligibility curves near SRT (median value of the individual slopes 18.6 %/dB for the 110 elderly subjects, and 21.6 %/dB for the young subjects). The individual slopes diminish significantly for larger SHL_D values. Nevertheless, the regression formula presented earlier shows that even for a severe hearing loss of, for example 8 dB, the $\overline{\Delta S_{50}}$ to be expected is still 14.3 ± 4.9 %/dB. The steep slopes indicate that with our speech material the subjects can profit maximally from the sentence context, as is usual in everyday listening situations. Kalikow et al. (1977), who found slopes of 14 %/dB with their high-predictability sentences, also reported that older subjects are as adept as younger ones at taking advantage of sentence context.

V. CONCLUSIONS

(1) Plomp's hearing model (see Eq.(2)) gives an accurate description of any hearing loss for speech, as a function of ambient noise level, on the basis of two loss components, viz. SHL_{A+D}(=A+D), the hearing loss in quiet, and SHL_D(=D), the hearing loss at high noise levels.

(2) In practice, an individual's hearing loss for speech can be specified sufficiently by measuring only two thresholds: (1) SRT in quiet, and (2) SRT at a single high noise level (> 50 dBA), resulting in accurate estimates of SHL_{A+D} and SHL_D.

(3) Especially for the individual listener, SHL_D should always be determined in order to obtain a good measure of his speech-hearing ability.

(4) The PTA and the Fletcher Index are unsuitable measures for estimating an individual's SHL both in quiet and in noise.

(5) For populations of elderly subjects every 7 dB of hearing loss in quiet (SHL_{A+D}) is accompanied, on an average, by 1 dB of hearing loss in noise (SHL_D).

(6) For the lists of sentences used in this study, an SHL_D value of 1 dB corresponds to a decrease of 18% in the intelligibility score of elderly subjects.

REFERENCES

Dixon, W.J. and Massey Jr., F.J., Introduction to Statistical Analysis (International Student Edition, McGraw-Hill Kogakusha, Ltd., 1969)

+) Duquesnoy, A.J. (1977). "Speech-level measurements for sentences using an artificial ear," Report 1977-AD01, Experimental Audiology, Free University, Amsterdam, The Netherlands.

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

Gjaevenes, K. (1969). "Estimating Speech Reception Threshold from Pure Tone Hearing Loss," J.Aud.Res.9, 139-144.

Groen, J.J. (1969). "Social Hearing Handicap; its Measurement by Speech-Audiometry in Noise," Int.Audiol.8, 182-183.

Hawkins, J.E. and Stevens, S.S. (1950). "The Masking of Pure Tones and of Speech by White Noise," J.Acoust.Soc.Am.22, 6-13.

Hayes, D. and Jerger, J. (1979). "Low-Frequency Hearing Loss in Presbyacusis," Arch.Otolaryngol.105, 9-12.

Hinchcliffe, R. (1959). "The threshold of hearing as a function of age," Acustica 9, 303-308.

Jerger, J. (1973). "Audiological Findings in Aging," Adv.Oto-Rhino-Laryng.20, 115-124.

Jokinen, K. (1973). "Presbyacusis. VI. Masking of Speech," Acta Oto-Laryngol. 76, 426-430.

Kalikow, D.N., Stevens, K.N., and Elliott, L.L. (1977). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J.Acoust.Soc.Am.61, 1337-1351.

Kell, R.L., Pearson, J.C.G., Acton, W.I., and Taylor, W. (1971). "Social Effects of Hearing Loss due to Weaving Noise," in Occupational Hearing Loss, edited by D.W. Robinson (Academic Press, London, 1971), 179-191.

Noble, W.G. (1973). "Pure-Tone Acuity, Speech-Hearing Ability and Deafness in Acoustic Trauma," Audiol.12, 291-315.

Olsen, W. and Carhart, R. (1967). "Development of test procedures for evaluation of binaural hearing aids," Bull.Prosthetics Res.10(7), 22-49.

Orchik, D.J. and Burgess, J. (1977). "Synthetic Sentence Identification as a function of the age of the listener," J.Am.Audiol.Soc.3(1), 42-45.

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. (1979a). "Speech-reception threshold for sentences as a function of age and noise level," J.Acoust.Soc.Am.66, 1333-1342.

Plomp, R. and Mimpen, A.M. (1979b). "Improving the Reliability of Testing the Speech Reception Threshold for Sentences," Audiol.18, 43-52.

Robinson, D.W. and Sutton, G.J. (1979). "Age Effect in Hearing - A Comparative Analysis of Published Threshold Data," Audiol.18, 320-334.

Schuknecht, H.F., Pathology of the Ear (Harvard University Press, Cambridge, MA, 1974).
Smoorenburg, G.F., de Laat, J.A.P.M., and Plomp, R. (1981). "The effect of noise-induced hearing loss on the intelligibility of speech in noise," in: Aural Communication in Aviation, AGARD Conference Proceedings No. 311, North Atlantic Treaty Organization, Paper 11.
Surr, R.K. (1977). "Effect of age on clinical hearing aid evaluation results," J.Am.Audiol.Soc. 3(1), 1-5.
Tillman, T.W., Carhart, R., and Olsen, W.O. (1970). "Hearing aid efficiency in a competing speech situation," J.Speech Hear.Res.13, 789-811.
Waal, J. van der (1962). Het horen van bejaarden. Doctoral thesis, State University of Leiden, The Netherlands (English summary at pp.93-97).
Young, M.A. and Gibbons, E.W. (1962). "Speech discrimination scores and threshold measurements in a non-normal hearing population," J.Aud.Res. 2, 21-33.

+) Reproduced in the Appendix of this thesis

++) Reproduced in Chapter I.

CHAPTER IV

EFFECT OF A SINGLE INTERFERING NOISE OR SPEECH SOURCE ON THE BINAURAL SENTENCE INTELLIGIBILITY OF AGED PERSONS

Paper submitted for publication in:
Journal of the Acoustical Society of America

ABSTRACT

The free-field Speech-Reception Threshold (SRT) for sentences was investigated in quiet and under nine conditions involving noise or competing speech for a group of 20 elderly subjects (10 male, age 75-85; 10 female, age 76-88) and a reference group of 10 young normal-hearing subjects. The noise source had the same long-term average spectrum as the competing speech. The interfering signals were presented at a constant level of 55 dBA. All elderly subjects had moderate, nearly symmetrical pure-tone hearing losses with an average loss at 500, 1000, and 2000 Hz of between 9 dB and 40 dB re ISO-389. The main results are: (1) the SRT values in noise and competing speech are about equal, whereas the normal-hearing subjects showed a lower SRT (7 dB lower for the condition where both sound sources are in front of the listener) in competing speech than in noise; apparently, the elderly subjects do not benefit from the relatively silent periods in competing speech; (2) the gain obtained by moving the interfering noise source from the front to the lateral position is 2.5 dB, in contrast to a gain of 9.6 dB for the young subjects; apparently, the elderly are unable to make full use of the spatial divergence between primary speaker and noise source.

INTRODUCTION

Speech in everyday listening situations is very often disturbed by interfering noise and competing speech from one or more talkers. Under such conditions speech intelligibility for normal-hearing listeners is usually hardly hampered, whereas hearing-impaired subjects often complain of being unable to understand speech. In two earlier studies (Duquesnoy and Plomp, 1980; Duquesnoy, 1982) the effects of room reverberation and continuous noise (with the spectrum of speech) on the Speech-Reception Threshold (SRT) for conversational sentences were investigated on elderly subjects. All SRT's were measured monaurally by headphones. As we know from studies in which normal-hearing listeners were tested in the free field (e.g. Carhart, 1965; Tonning, 1971), binaural hearing is crucial to cope with degraded listening conditions. Furthermore, Carhart et al. (1969) found that speech intelligibility for normal-hearing listeners is affected less by fluctuating interfering signals, like speech, than by continuous ones.

In view of the many complaints from hearing-impaired listeners in everyday situations, it is important to know whether they still benefit fully from binaural hearing, and whether they are hindered more by a given type of interference than normal-hearing listeners. Therefore, the present study deals with the effect of a single interfering noise source (continuous signal) or speech source (fluctuating signal) on the binaural free-field SRT for sentences. Since more than half of the hearing-handicapped are over the age of 65 (cf. Plomp, 1978), elderly subjects were used in this study.

Studies on the gain of binaural listening in elderly subjects, tested under conditions with interfering noise, are scarce (see e.g. a review by Pickett et al., 1977). Furthermore, in almost all pertinent studies hitherto published (Dirks and Wilson, 1969; Carhart and Tillman, 1970; Tillman et al., 1973; Tonning, 1973) data were presented which are not comparable, for one reason or another, with the data presented below. Therefore, these studies will not be considered in more detail. An exception to this is the study by Bocca and Antonelli (1976), the results of which are discussed in Sec.III.

I. EXPERIMENTS

A. Sound stimuli

The primary speech material used was the same as in previous investigations (Duquesnoy and Plomp, 1980; Duquesnoy, 1982). It consists of ten lists of 13 sentences developed by Plomp and Mimpen (1979). These sentences were pronounced by a female speaker.

As interfering sounds, the present author made two other lists of 13 sentences, pronounced by a male speaker at a rate of 4.3 syllables per second, and noise with the same spectrum as the long-term average spectrum of these sentences. The sentences were about twice as long as the primary ones in order to mask the primary speech entirely. In Fig.1 the long-term average spectra of both the 130 primary sentences and the 26 competing sentences are shown relative to equal overall intensity.

B. Method

Since ten lists of primary sentences were available, ten listening conditions were selected. They are:

(1) Binaural SRT in quiet; this condition was required in order to verify that SRT's found in the presence of an interfering sound were determined by this sound rather than by the subject's absolute threshold.

(2) Monaural SRT in noise for the right ear; this and the following condition were included in order to check on hearing symmetry for speech in noise.

(3) As (2), but for the left ear.

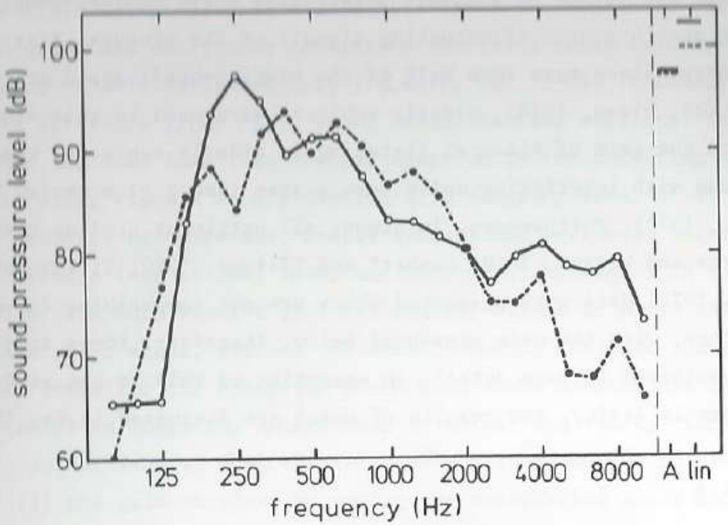


FIG. 1. Long-term average spectra of both the 130 primary sentences (continuous curve, female voice) and the 26 competing sentences (dotted curve, male voice) shown for equal overall intensity in dBA.

(4) Binaural SRT in noise, with the primary sentences as well as the noise reproduced in front of the listener.

(5) Binaural SRT with sentences as the interfering sound, with the primary sentences as well as the interfering sentences reproduced in front of the listener.

(6) As (5), but with the interfering sentences played backwards; this condition was necessary to decide whether possible SRT differences between conditions (4) and (5) are due to the different temporal structure or to the distracting contents of the interfering sentences.

(7) Binaural SRT in noise, with the primary sentences reproduced in front of the listener and the noise reproduced under 90°, either right or left.

(8) Binaural SRT with sentences as the interfering sound, with the primary sentences reproduced in front of the listener and the interfering sentences reproduced under 90°, either right or left.

(9) As (8), but with the interfering sentences played backwards.

(10) Binaural SRT in noise, with the primary sentences reproduced under 90°, either right or left, and the noise reproduced in front of the listener.

This condition was included to check whether the spatial positions of the primary speech source and the interfering noise source can be interchanged without influencing SRT.

In Table I these ten conditions, and their codes, are summarized.

The ten sentence lists were always presented in the same order. To eliminate the effects of training and fatigue as much as possible, the ten conditions were counterbalanced for the ten male as well as the ten female subjects tested, according to a 10x10 digram balanced Latin square (Wagenaar, 1969). The primary sentence lists were recorded on one track of a tape, the interfering sound signals on the other track. The levels of both output channels of the tape recorder could be adjusted separately by means of two attenuators. After attenuation both signals were either mixed and amplified (FF conditions) or amplified separately (other conditions). The amplified signals were fed into one or two Quad electrostatic loudspeakers, the centers of which were placed 1 m above floor level. The distance between the loudspeakers and the entrance of

TABLE I. Survey of the ten listening conditions. Five spatial configurations marked FF, FR, FL, RF, and LF were tested; the first character denotes the position of the primary speech source and the second one that of the interfering source (F = front, R = right (90^{o}), and L = left (-90^{o})). The interfering signals are denoted by: Q = quiet (no interference), N = noise, S = sentences, and invS = sentences played backward. The hearing modes are denoted by: Bin = binaural hearing, Mon = monaural hearing, R = right ear, and L = left ear.

Spatial configuration	Interfering signal	Hearing mode
1) FF	Q	Bin
2) FF	N	Mon R
3) FF	N	Mon L
4) FF	N	Bin
5) FF	S	Bin
6) FF	invS	Bin
7) FR or FL	N	Bin
8) FR or FL	S	Bin
9) FR or FL	invS	Bin
10) RF or LF	N	Bin

the ear canals was 2 m. Monaural hearing was achieved by occluding one ear with a combination of a hearing protector of ear-down (Bilsom Propp-o-Plast) and a high-quality circumaural fluid-seal ear-muff (Eargard, Ferranti Meter Ltd).

The test, which took about 60 min, was carried out in an anechoic room in which the maximum ambient noise level, including instrument hum, was below 17 dBA. First, tone audiograms (see next section) were determined. Next, the speech intelligibility tests were run, during which the subject was seated in an armchair fitted with a head-rest. Under all conditions (except 1) the overall level of the interfering sounds was fixed at 55 dBA, measured at the position of the listener's ear, with the listener removed. Measurement of SRT at only this level of the interfering sound is sufficiently representative, because in an earlier study (Duquesnoy, 1982) it was demonstrated that for levels of 55 dBA and higher the SRT is exclusively determined by the S/N ratio. Throughout this paper all speech levels specified are the long-term average intensity levels of connected discourse, measured by computing the root-mean-square value of a great many samples (sampling rate of 20 Hz).

The SRT values, corresponding to a 50%-correct score, were obtained for each list by using a simple up-and-down procedure for the presentation level of the primary sentences. The procedure requires correct reproduction of an entire sentence for a positive response to be recorded. The test-retest reliability of SRT's, thus obtained against a background of interfering noise, is about 1 dB for normal-hearing subjects (Plomp and Mimpen, 1979).

C. Subjects

Criteria for selection of subjects were (1) the air-conduction tone audiograms of both ears had to satisfy the age-dependent normative threshold criteria for presbyacusis, proposed by Robinson and Sutton (1979), in order to exclude subjects with other ear pathologies, (2) the interaural difference in pure-tone acuity (PTA, average air-conduction hearing level at 500, 1000, and 2000 Hz) must be smaller than 10 dB, and (3) the interaural difference in hearing loss for speech in noise at a level of 55 dBA should not exceed 5 dB. Ten male subjects (age 76-88, mean age 80.5) and ten female subjects (age 76-88, mean age 82.5) participated, meeting these criteria on hearing acuity and symmetry. In Fig.2 the median hearing level as well as the upper and lower quartiles of the 20 subjects are reproduced.

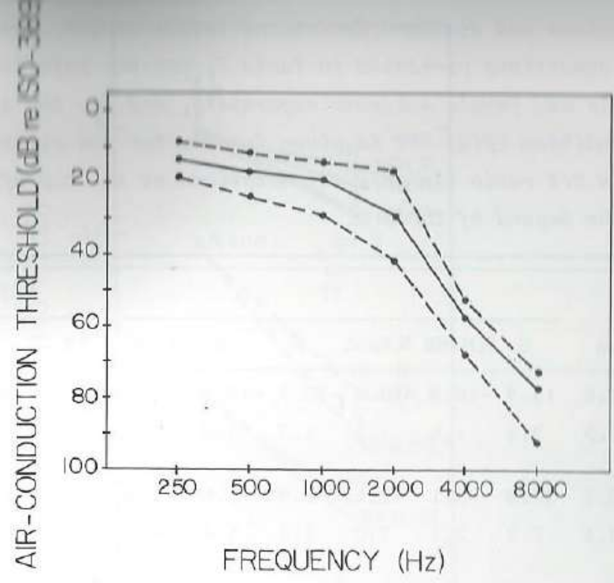


FIG.2. Median values and upper and lower quartiles of the air-conduction tone audiograms of the 20 elderly subjects selected for the tests.

In addition to the elderly people, ten young normal-hearing subjects were tested as a reference group (mean age 22.6; individual hearing losses for the two ears at the frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz <15 dB re ISO-389).

II. RESULTS

In Table II the mean values and standard deviations of SRT for the ten listening conditions are given for the reference group and the elderly subjects. The SRT values are given in terms of S/N ratio, except for the quiet condition, for which SRT is given in dBA.

It was tested whether the differences between the mean SRT values for the various listening conditions were significant. As homogeneity of variance was not met for all SRT's involved, t-tests with variance-dependent degrees of freedom were used (cf. Dixon and Massey, 1969). It was found that, generally, SRT differences exceeding 2 dB are significant at the 2%-level, those exceeding 3 dB significant at the 1%-level.

In Fig.3 the mean SRT values in noise are shown for the reference group and the male and female elderly subjects together. It is evident that the

TABLE II. Mean values and standard deviations (s.d.) of SRT, measured under the ten listening conditions presented in Table I, for the reference group, for the elderly male and female subjects separately, and for the elderly together. For the condition FF(Q) SRT is given in dBA. For the other conditions SRT is expressed in S/N ratio (in dB). All notations at the top of the table are specified in the legend of Table I.

			FF						FR/FL			RF/LF
Subjects		Age	Q	N,MonR	N,MonL	N	S	invS	N	S	invS	N
Refer.gr (10 Ss)	mean	22.6	13.9	-10.5	-10.4	-10.7	-17.6	-17.6	-20.3	-24.3	-23.4	-17.7
	s.d.	2.5	3.4	1.6	1.5	1.3	3.0	3.1	2.4	2.4	3.4	2.2
Elderly male (10 Ss)	mean	80.5	32.3	-3.1	-2.4	-5.1	-5.6	-6.4	-7.4	-10.2	-10.4	-7.0
	s.d.	3.5	7.9	2.1	3.0	2.2	5.4	4.4	4.0	4.8	5.5	3.3
Elderly female (10 Ss)	mean	82.5	37.6	-3.1	-2.8	-5.5	-4.0	-4.4	-8.2	-8.3	-8.4	-6.6
	s.d.	4.2	5.0	3.1	2.5	2.4	3.3	3.5	2.3	4.5	3.7	3.8
Combined elderly (20 Ss)	mean	81.5	35.0	-3.1	-2.6	-5.3	-4.8	-5.4	-7.8	-9.3	-9.4	-6.8
	s.d.	3.8	7.0	2.6	2.7	2.3	4.4	4.0	3.2	4.6	4.7	3.5

SRT differences between the young and elderly subjects are considerable for all conditions; they range from 5.4 dB to 15.0 dB (all significant at the 1%-level). The SRT differences in question are, in fact, the hearing losses for speech (SHL) of the elderly subjects (relative to the reference group).

Upon further consideration of the SRT differences per individual (not shown), it would seem that age effects are present for conditions with competing speech played forward or backward. In order to verify whether these differences actually depend on age, the 20 elderly subjects were subdivided according to their age into four subgroups of five persons each. On these subgroups a three-variable (speech, age, subjects) analysis-of-variance was performed for the conditions FF(S, invS) and FR/FL(S, invS). The factors speech and age were not significant, whereas the factor subjects was highly significant, viz. at levels of 0.2% (contribution of 85.9% to the total variance) and 0.3% (77.4% contribution) for the respective conditions. This implies that the subjects are by far the most relevant

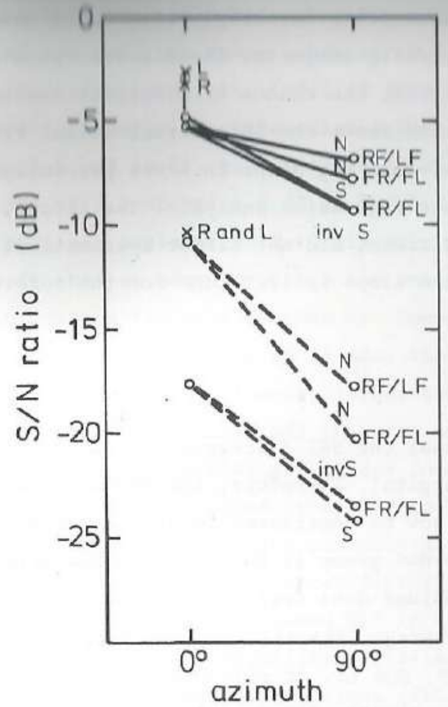


FIG.3. SRT for sentences, expressed in S/N ratio, as a function of (1) the azimuth between primary source and interfering source and (2) the type of interfering signal. Open circles represent binaural SRT's, crosses represent monaural SRT's. Dotted lines connect SRT's of the reference group, and drawn lines those of the elderly subjects. For the notations in the figure see the legend of Table I.

source of variance, whereas age effects are negligible. Apparently, the range of age of the elderly subjects tested was too small to find a significant age effect.

Since the factor speech (i.e. S versus invS) in the analysis-of-variance had no significant effect on SHL, it may be permissible to consider the conditions FF(invS) and FR/FL(invS) as retests of the conditions FF(S) and FR/FL(S), in order to gain insight into the test-retest reliability of SRT's, measured against a background of competing speech. The resulting standard error of individual SRT's is 1.7 dB for the normal-hearing subjects and 1.6 dB for the elderly.

Finally, the increment in sentence-intelligibility score near SRT was determined for the young and elderly subjects. These scores were obtained by calculating, for each condition, the chance of a correct response at levels of 1 dB and 2 dB below and above the 50%-correct level. From these values, cumulated over the subjects, the slope in score percentage per dB was estimated by means of linear regression analysis. The slopes, determined for the various listening conditions, did not differ systematically. For all conditions combined the mean slope is 17.3 %/dB for the reference group and 17.2 %/dB for the elderly.

III. DISCUSSION

It can be seen from Table II that the SRT differences between the male and female elderly subjects are marginal. Therefore, the SRT's of male and female subjects combined will now be considered in greater detail. It should be kept in mind that this combined group is very homogeneous with respect to age effects on SRT. The combined data (see also Fig.3) reveal that:

(1) Condition FF(N) shows an average binaural gain in noise, versus conditions FF(N, Mon R) and FF(N, Mon L), of only 0.25 dB for the young subjects and of 2.5 dB for the elderly. Apparently, for the normal-hearing subjects this diotic listening situation, with identical signals in the two ears, does not contribute appreciably to a better speech intelligibility.

(2) There are substantial SRT differences between the young and the elderly subjects, which range from 5.4 dB (FF, N) to 15.0 dB (FR/FL, S). Since each dB of hearing loss in S/N ratio lowers sentence intelligibility by approximately 17%, the elderly are at a serious disadvantage relative to the normal-hearing young listeners.

(3) Switching the sources of interference (N, S, invS) from F to either R (90^{o}) or L (-90^{o}) yields very significant SRT reductions, viz. 9.6, 6.4, and 5.8 dB for the young subjects, and 2.5, 4.5, and 4.0 dB for the elderly. In addition, these figures indicate that for the elderly the gain of moving the noise source sideways is 7.1 dB smaller than for the young subjects.

(4) Switching the primary speech source from F to either R or L during presentation of noise in front results in a significant SRT reduction (7.0 dB) for the young subjects and scarcely any decrease (1.5 dB) for the elderly.

(5) Forward and backward presented competing sentences give negligible SRT differences for both the young subjects (0.0 dB (FF) and -0.9 dB (FR/FL)) and the elderly subjects (-0.5 dB (FF) and 0.1 dB (FR/FL)).

(6) Young subjects benefit from the relatively silent intervals in competing speech (both S and invS), as demonstrated by SRT differences of 6.9 dB re FF(N) and of 3.1 dB and 4.0 dB re FR/FL(N), whereas the elderly do not. The possible explanation for this gain for young normal-hearing listeners is that at moments with a low overall level of the fluctuating sound, i.e. during the gaps between the competing words and syllables presented at an average level of 55 dBA, the contribution to their speech perception is significantly larger than for the hearing-impaired elderly, who have an enhanced SRT in quiet of 35.0 ± 7.0 dBA (cf. Table II). Furthermore, the masking effect due to moments of high overall level of the fluctuating sound is likely to be stronger for the hearing-impaired elderly subjects, being probably attributable to a reduced frequency-resolving power and higher distortion in the ear.

The adverse effect of an increased SRT in quiet on the SRT in competing speech is corroborated by the significant (1%-level), high correlations found between corrected speech hearing-loss (SHL) values in quiet (corrected for SHL in continuous noise by subtracting this value) and corrected SHL values in competing speech (also compensated for SHL in noise), viz. correlation coefficient, r, equals 0.72, both for the combined listening conditions FF(S) and FF(invS), and for the conditions FR/FL(S) and FR/FL(invS). No significant correlation ($r = 0.26$) was found between corrected SHL values in quiet and SHL in noise. Furthermore, virtually no correlations ($r < 0.10$) were found between SHL values in noise and corrected SHL values pertinent to competing-speech conditions. This indicates that SHL in noise is a poor predictor for SHL to be expected in the case of a (single) competing speaker.

Another result from this investigation is that the accuracy of the SRT determination is independent of age and hearing loss. This is indicated by similar standard errors of individual SRT's for young and elderly subjects against a background of competing sentences (1.7 dB and 1.6 dB, respectively), and by the similar increments in sentence-intelligibility score near SRT (17 %/dB).

The main impression from the above presented results is that even moderately hearing-impaired elderly subjects find it much more difficult to separate source locations and to discriminate between time-dependent sound signals than young normal-hearing listeners. It can hardly be concluded whether the handicap of the elderly is due entirely to the decline of their peripheral hearing system, or that additional deterioration of the central auditory pathways also plays a role. In this respect the study of Bocca and Antonelli (1976) is important. They investigated speech intelligibility for sentences masked by broad-band noise under the conditions FF (Mon R), FF(Mon L), FF, RF, and LF. Among others, ten presbyacusic subjects (age unspecified) were tested who had nearly normal hearing, with an average PTA of 10 dB re ISO-389 and a hearing level at 4000 Hz of 35 dB only. They achieved a negligible binaural gain under condition FF (re FF(Mon R) and FF(Mon L)) and a gain of 6 dB under conditions RF and LF (re FF). These results contrast with the corresponding values found for the elderly in the present study (cf. Fig.3). They achieved a gain of 2.5 dB under condition FF, and a gain of 1.5 dB only under conditions RF and LF. It should be noted, however, that the elderly of Bocca and Antonelli were nearly normal hearing. Their results agree closely with those achieved by the young subjects of the present study (viz. 0.25 dB and 7.0 dB). These results of Bocca and Antonelli suggest that age as such does not necessarily bring with it a decreased performance under conditions where our moderately impaired subjects fail, but that the hearing loss itself (peripheral origin) may be the cause of the increased SRT's. A possible reason for the moderate pure-tone losses of the elderly having such a great impact is that just the frequencies at and above 2000 Hz are almost inaudible (compare Figs.1 and 2), whereas the profitable effects of head shadow on binaural listening are most pronounced at high frequencies.

IV. CONCLUSIONS

(1) Even elderly subjects with moderate pure-tone hearing losses are seriously handicapped under everyday listening conditions, as is demonstrated by the considerable binaural hearing loss for sentences of 5.4 to 15.0 dB in S/N ratio, found for an average PTA of 25 dB.

(2) The gain of moving the primary speech source or the interfering noise or speech source from the front to a lateral position is greater for young listeners than for the elderly. An average gain of 7 dB was found for the young ones versus a gain of 3.1 dB for the elderly. This may be indicative of poorer auditory space perception among the elderly, probably due to their pronounced high-frequency hearing losses.

(3) The negligible effect for both the young and elderly subjects of playing the interfering sentences backwards implies that it is the physical aspects of the interfering signal that are decisive, and not such a psychological factor as distraction due to the simultaneous intelligibility of primary and competing sentences.

(4) The elderly subjects, although having pure-tone hearing losses not more than predicted for their age, do not really benefit from the relatively silent intervals in competing speech (single speaker), whereas the gain for young subjects amounts to 7 dB relative to interfering continuous-noise conditions.

REFERENCES

Bocca, E. and Antonelli, A.R. (1976). "Masking Level Difference: Another Tool for the Evaluation of Peripheral and Cortical Defects," Audiol. 15, 480-487.

Carhart, R. (1965). "Monaural and Binaural Discrimination Against Competing Sentences," Int. Audiol.4, 5-10.

Carhart, R. and Tillman, T.W. (1970). "Interaction of Competing Speech Signals with Hearing Losses," Arch.Otolaryng.91, 273-279.

Carhart, R., Tillman, T.W., and Greetis, E.S. (1969). "Perceptual Masking in Multiple Sound Backgrounds," J.Acoust.Soc.Am.45, 694-703.

Dirks, D.D. and Wilson, R.H. (1969). "Binaural Hearing of Speech for Aided and Unaided Conditions," J.Speech Hear.Res.12, 650-664.

Dixon, W.J., and Massey Jr, F.J., Introduction to Statistical Analysis (International Student Edition, McGraw-Hill Kogakusha, Ltd., 1969).

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

Duquesnoy, A.J. (1982). "The sentence intelligibility of aged persons in quiet and in noise - test of a model on speech hearing," submitted to J.Acoust.Soc.Am. ++)

Pickett, J.N. (1977). "Speech Understanding and Aging," Report of Working Group 75, Committee on Hearing, Bioacoustics and Biomechanics, National Academy of Sciences, Washington DC.

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," Audiol.18, 43-52.

Robinson, D.W. and Sutton, G.J. (1979). "Age Effect in Hearing - A Comparative Analysis of Published Threshold Data," Audiol.18, 320-334.

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.

Tonning, F.-M. (1973). "Directional Audiometry VII. The Influence of Azimuth on the Perception of Speech in Aided and Unaided Patients with Binaural Hearing Loss," Acta Otolaryngol.75, 425-431.

Wagenaar, W.A. (1969). "Note on the Construction of Digram-Balanced Latin Squares," Psych.Bull.72, 384-386.

+) Reproduced in Chapter I of this thesis

++) Reproduced in Chapter III.

CHAPTER V

THE AIDED SENTENCE-INTELLIGIBILITY OF HEARING-IMPAIRED SUBJECTS IN QUIET AND IN NOISE - TEST OF A HEARING-AID MODEL

Paper submitted for publication in:
Journal of the Acoustical Society of America

ABSTRACT

The monaural free-field Speech-Reception Threshold (SRT) for conversational sentences was investigated under unaided and aided conditions for 50 subjects (24 male, 26 female; age below 65), which were grouped according to five different types of pure-tone audiograms. The unaided SRT's were measured in quiet and in noise at levels of 40, 55, 70, and 85 dBA. The aided SRT's were obtained by means of the subjects' own behind-the-ear hearing aids in quiet and in noise at levels of 25, 40, 55, and 70 dBA. The noise had a long-term average spectrum identical with that of the sentences. The five types of audiograms were representative of the following hearing losses: sensorineural high-tone both with and without recruitment, sensorineural flat, mixed, and purely conductive. It is shown that a model on aided hearing developed by Plomp (J.Acoust.Soc.Am.63, 533-549, 1978) enables an accurate description to be made of the unaided and aided SRT values measured, irrespective of amount and type of hearing impairment. The model interprets any hearing loss for speech as a combination of a loss of class A (attenuation) and a loss of class D (distortion), and describes the hearing aid in terms of a gain G, which compensates for a class A loss, and a distortion S, which adds to the class D loss. Values for A, D, G, and S were obtained from fitting this model, per subject, to the unaided and aided SRT data. The four parameters characterize the individual aided speech-hearing very conveniently. The mutual relationships of these four parameters were studied, as were the relations to several electroacoustic hearing-aid properties, such as gain in quiet, internal noise, and second-harmonic and third-harmonic distortions. Among other things, it was found that: (1) generally, the model-derived hearing-aid gain in quiet, G-S, is identical with the functional gain G_f (difference, in dB, between the measured unaided and aided SRT in quiet); (2) modern hearing aids provide no benefit in noise, as they do not improve the signal-to-noise ratio; actually, an average distortion S of slightly more than 1 dB was found; and (3) the aids with third-harmonic distortions are particularly detrimental to SRT in noise; they show an average S value of 2.4 dB in contrast to an average of 0.7 dB for the other aids.

INTRODUCTION

It is well-known from the literature that, particularly, people with sensorineural hearing losses are seriously handicapped in communicating

in noisy environments. This problem is usually not mitigated by wearing a hearing aid, since mere amplification of signals cannot really compensate for these losses. A basic requisite for a hearing aid to provide substantial benefit in noise is that it increases the signal-to-noise (S/N) ratio. For example, Gengel (1971) found that only S/N ratios of +15 dB or more can be considered acceptable for aided speech discrimination by the hearing impaired. It is amazing how much better hearing-aid performances audiologists are still expecting from careful selection and fitting of the currently available hearing aids, although these aids cannot but amplify or attenuate in a more or less satisfactory way. Inevitably, practice continues to prove that most hearing-impaired listeners are disappointed about the benefit of their aids under noisy conditions.

Recent research suggests that subjects with sensorineural hearing losses should derive extra benefit from hearing aids with reduced low-frequency gain and extended high-frequency gain with frequencies up to 6500 Hz (cf. Pascoe, 1975; Nielsen,1976; Harford and Fox, 1978; Schwartz et al., 1979; Skinner, 1980). The next step seems to be the precise matching of the frequency response of a (master) hearing aid to the specific hearing of each single individual (Miller et al., 1980). All these recent efforts of finding the optimum frequency response have two factors in common: (1) there is no purposeful search for a method or device capable of substantially improving the S/N ratio; the approach so far pursued is essentially limited to yielding insufficiently better results in noise (a few decibels at best); (2) there is a lack of a simple concept of how speech intelligibility may depend on some basic characteristics inherent in all hearing aids (cf., for example, Miller et al., 1980).

As distinct from the above mentioned authors, Plomp (1978) discerned the dominating problem of S/N ratio in aided hearing. He rightly observed that the auditory handicap of the hearing impaired can only be lessened by improving the S/N ratio. Some possible ways of achieving this were suggested, partly based on room acoustics, partly on signal processing by a hearing aid; they will not be repeated here. Another important aspect of the paper was the presentation of a simple model of how hearing aids may influence speech intelligibility of the hearing impaired.

The primary aim of the present report is to provide an experimental basis for Plomp's model. It will be shown that the model is generally valid, irrespective of both the frequency response of the (head-worn) hearing aids tested, and of the type of hearing impairment. Furthermore, data will be presented on the benefit of hearing aids. The second aim is to present relations between, on the one hand, several electroacoustic characteristics of hearing aids, as measured in a hearing-aid test box, and, on the other hand, the concomitant psychophysical results, expressed in terms of Speech-Reception Thresholds (SRT) measured in quiet and at various masking-noise levels up to 85 dBA.

I. A MODEL OF THE AIDED SPEECH-RECEPTION THRESHOLD

A. Description of the model

It has been verified previously (Duquesnoy, 1982) that the effect of hearing impairment on the SRT in quiet and in noise is well described by a model proposed by Plomp (1978). In this model any Speech-Hearing Loss (SHL) is interpreted as being caused by a combination of two formalistic impairments: (1) attenuation of all sounds entering the ear, defined as a class A loss, and (2) distortion of the sounds, described as a class D loss. Then, for the hearing impaired, the SRT as a function of noise level can be expressed by

$$SRT = 10 \log \left[10^{(L_0+A+D)/10} + 10^{(L_N-\Delta L_{SN}+D)/10}\right] \text{ in dBA,} \qquad (1)$$

where: L_0 = SRT in quiet for the normal-hearing (in dBA),

L_N = sound-pressure level of the masking noise (in dBA),

ΔL_{SN} = the number of decibels that SRT in noise for the normal-hearing is below L_N, thus $L_N-\Delta L_{SN}$ represents SRT in noise for the normal hearing,

A+D = Speech-Hearing Loss (SHL) in quiet re L_0, and

D = SHL in noise re $L_N-\Delta L_{SN}$.

In Fig.1 it is shown how SRT changes, according to Eq.(1), as a function of L_N, with the losses A and D as parameters.

Duquesnoy (1982) applied an iterative procedure for fitting a curve defined by Eq.(1) to SRT values of hearing-impaired subjects measured in quiet and for, at least, one level of masking noise. Accurate estimates of A+D and D were achieved if the SRT measured at the highest noise level, L_{Nmax}, was unequivocally positioned on the rising flank of the SRT curve. This was shown to be the case if

$$A < L_{Nmax}-L_0-\Delta L_{SN}-10 \quad . \qquad (2)$$

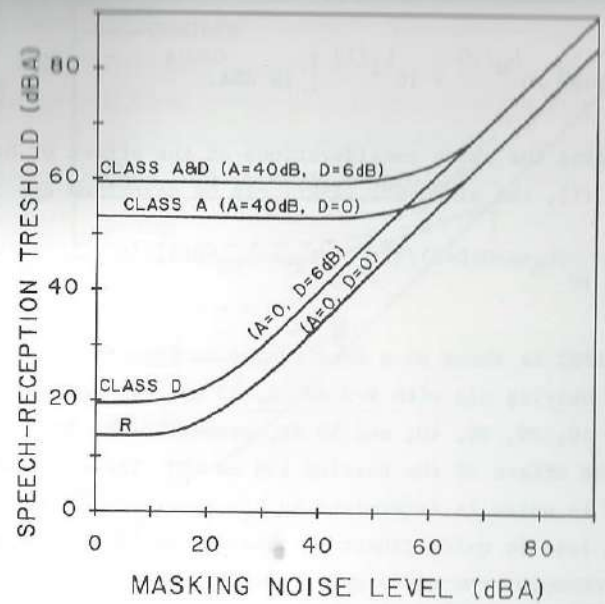


FIG.1. Speech-Reception Threshold (SRT) for sentences as a function of masking noise level L_N. The lower curve, marked R, was determined according to Eq.(1) for normal- hearing listeners, for which monaural tests in the free-field showed that L_0=13.5 dBA and ΔL_{SN}=5.8 dB. The other curves, also according to Eq.(1), hold for hypothetical speech-hearing losses in quiet (A+D) and in noise (D) as indicated in the figure.

Plomp (1978) extended his model in order to include the effect of hearing aids on the impaired SRT. Again, a simple formalistic concept was adopted, in which the following three quantities characterized the hearing aid: (1) the acoustic gain G of the aid (in dB), equal to the increase of sound-pressure level at the ear drum, (2) the increase S (in dB) of SRT in noise, due to signal distortion in the aid (linear and non-linear), and (3) the aid's internal noise L_I (in dBA), interpreted as an additional (external) noise level at the aid's microphone. These three hearing-aid parameters can be easily accounted for in the expression for SRT given by Eq.(1). The gain G compensates for the class A hearing loss, resulting in an apparent class A loss of (A-G) dB. The distortion term S of the aid can simply be added to the class D hearing loss of the ear. The internal noise L_I contributes, together with the background noise level L_N, to the apparent total noise level

L_{NT} at the aid's microphone, given by

$$L_{NT} = 10 \log \left[10^{L_N/10} + 10^{L_I/10} \right] \text{ in dBA.} \quad (3)$$

Then, by including the above considerations on the effect of hearing aids on SRT into Eq.(1), the aided SRT (ASRT) can be described by

$$ASRT = 10 \log \left[10^{(L_0+A-G+D+S)/10} + 10^{(L_{NT}-\Delta L_{SN}+D+S)/10} \right] \text{ in dBA .} \quad (4)$$

In Fig.2 ASRT is shown as a function of background noise level L_N for a hypothetical hearing aid with S=3 dB, L_I=25 dBA (values adopted from Plomp, 1978), and G = 10, 20, 30, 40, and 50 dB, respectively. The figure clearly demonstrates the effect of the hearing aid on SRT. The obtainable minimum value for ASRT in quiet is determined by L_I. Therefore, full compensation of the hearing loss in quiet cannot be obtained by mere amplification; increasing, for example, the aid's gain from 40 dB to 50 dB does not substantially improve the ASRT. The convergence of the ASRT curves at higher noise levels points to a serious limitation of the aid: it amplifies all sounds, but it does not increase the S/N ratio (in the case chosen here, with S = 3 dB , the aid significantly worsens the S/N ratio, but recent commercial aids generally have somewhat lower S values; see Sec. III). Another important aspect related to S/N ratio is the maximum acceptable background noise level for a conversation being possible. Apparently, for the combination of hearing losses and hearing aid shown, this level is decreased by as much as 15 dB relative to normal-hearing (compare the two arrows in Fig.2). For a more elaborate theoretical examination of the benefit of hearing aids in noise the reader is referred to Plomp (1978).

B. Test procedure for the model

The theoretical description of hearing-aid performance presented has to be verified experimentally. In the model given by Eq.(4) the hearing aid has been simplified. Effects of peak clipping (PC) and automatic gain control (AGC) are not included. Furthermore, the model does not explicitly account for the effect of different frequency responses of the aid, nor for the role of recruitment (reduced dynamic span of hearing). At present, most prescribed hearing aids have PC or AGC circuits, and the frequency responses vary between basically wide-range and narrow-band. Also, a considerable

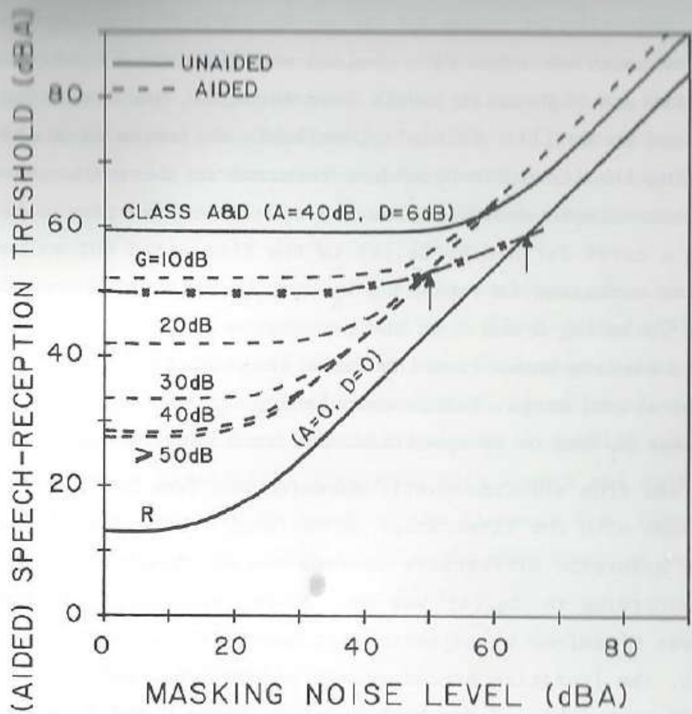


FIG. 2. Unaided and aided SRT curves for sentences as a function of masking noise level L_N. The solid curve, marked R, represents normal-hearing SRT's. The upper solid curve describes SRT for hypothetical hearing losses A and D as indicated. The dashed lines show the effect of increasing the gain G of a hearing aid with S=3 dB and L_I=25 dBA. The curve, marked with the cross symbols, represents the average level in dBA of conversational speech at a distance of 1 m (cf. Heusden et al., 1979). When interfering noise is present, people unconciously raise their voice level. The right vertical arrow indicates the maximum acceptable noise level (65 dBA) during conversation among normal-hearing listeners. The left arrow gives the maximum noise level (50 dBA) for the aided hearing-impaired listener in order to be still able to understand the conversation.

percentage of subjects shows recruitment in the aided ear (e.g. in cases of noise trauma and presbyacusis). Thus, for the model to have any practical value it must apply in all these cases, and if it does not, minor adaptations should be sufficient to make it applicable. In the present investigation all the above-mentioned aspects on aids and ears were studied with regard to the applicability of the model.

For verification of the model, monaural SRT's of hearing-impaired subjects were measured under five unaided and five aided conditions, including both quiet and high noise levels (see Sec.IIB). For the unaided situation Eq.(4) reduces to Eq.(1). Initially, reliable estimates of A and D were obtained using the iterative procedure referred to above (Duquesnoy, 1982). Then, with some simple modifications the same iterative procedure was used for fitting a curve defined by Eq.(4) to the five aided SRT values. The modifications consisted in replacing L_N by L_{NT}, calculated according to Eq.(3), and including G and S as the parameters to be estimated. The values for A and D, already known from the first stage of fitting, were kept constant in the second stage. Before calculating L_{NT} values by Eq.(3), the internal noise L_I had to be specified. For each hearing aid an estimate of L_I was derived from electroacoustic measurements (see Sec.IIC).

In common with the first stage of fitting, in the second stage the sum of the five quadratic differences between the ASRT's measured and the ASRT's predicted according to Eq.(4) was applied as performance criterion. This criterion was minimized by adjusting estimates (via steepest descent method) for G and S. The iterative procedure was stopped when adjustments smaller than 0.002 dB were achieved for both G and S. These G and S values define the best fitting curve. The standard deviation of the five data points from the best fitting curve is considered to be a measure of the goodness of fit. In order to obtain a reliable estimate of S, it must be verified that, at least, the ASRT measured at the highest noise level L_{Nmax} is unequivocally positioned on the rising flank of the ASRT curve. With respect to Eq.(4) this means that, at least for L_{Nmax}, the contribution to ASRT of the first power term of ten, the exponent of which depends on both G and S, is small relative to the second term. Assuming a ratio of 1:10 as the desirable minimum, this holds if $L_0+A+D-G+S < L_{NTmax}-\Delta L_{SN}+D+S-10$, or:

$$G > L_0+\Delta L_{SN}+A+10-L_{NTmax}, \qquad (5)$$

where: $L_{NTmax} = 10 \log \left[10^{L_{Nmax}/10} + 10^{L_I/10} \right]$.

The experimental results presented in Sec.III will demonstrate that the model describes the ASRT for sentences satisfactorily. Consequently, the effect of hearing aids on SRT can be characterized adequately by just two quantities, viz. the acoustic gain G and the distortion S.

Traditionally, hearing-aid performance is evaluated electroacoustically in terms of frequency response, average gain, harmonic and intermodulation distortion, and saturation output levels. It is of interest to know the relationships between these electroacoustic hearing-aid properties and the psychoacoustic hearing-aid parameters G and S. Therefore, all hearing aids involved in the present investigation were also tested electroacoustically (see Sec.IIC).

Attention is drawn to the fact that the present evaluation of the model only deals with monaurally aided speech intelligibility. Plomp (1978) presupposed binaurally fitted aids, in order to have no losses due to head shadow or to lack of binaural release of masking. In practice, however, monaural fitting is more usual, and binaural release of masking under aided free-field conditions is practically zero when both speech and noise signals originate from the same location in front of the listener (cf. Dirks and Wilson, 1969, and Tonning, 1973). This explains why this location was used for testing speech intelligibility (see Sec.IIB).

II. EXPERIMENTS

A. Sound materials

The primary speech material used was the same as in the previous investigation (Duquesnoy, 1982). It consisted of ten lists of 13 sentences each, pronounced by a female speaker, and a masking noise with the same long-term average spectrum as the 130 sentences. By using a simple up-and-down procedure for the presentation level of the sentences, an estimate of SRT (50%-correct threshold) could be obtained for each list. A correct reproduction of the entire sentence is essential for a positive response to be recorded. The individual SRT values thus obtained against a background of noise have a standard error of estimate of only about 1 dB for normal-hearing subjects (Plomp and Mimpen, 1979).

Prior to testing the aided speech intelligibility, the subjects had to adjust the volume control dial of their hearing aids. For this adjustment a standard recording, consisting of sentences pronounced by a male speaker, was available without interfering noise.

B. Psychophysical measurements

The measurements were performed under free-field conditions and took about

60 min. All sound stimuli were presented via a Quad electrostatic loudspeaker situated in front of the listener, with its center 1 m above floor level. The subject was seated in an armchair fitted with a headrest to ensure direct facing of the loudspeaker. The distance between the loudspeaker and the entrance of the ear canals was made exactly 2 m. The anechoic room in which the measurements took place had a maximum ambient noise level, including instrument hum, below 17 dBA. All tests were performed monaurally with the subjects' own hearing aids. Only behind-the-ear aids were considered; none of the earmoulds was vented.

Prior to the test, the subject adjusted the volume control of the hearing aid to achieve a comfortable listening level for the standard recording of sentences. The recording was presented in quiet at a constant level of 62.5 dBA at the position of the subject's head. This is the appropriate level for establishing the volume control setting as normally utilized in every day listening situations (cf. Walden et al., 1977). The adjusted dial setting was fixed with a piece of plaster to prevent the subject from further changing the setting. In cases of binaurally fitted hearing aids, only the hearing aid at the ear with the maximum speech discrimination score, as specified by the clinical audiologist, was mounted and adjusted.

During the test, monaural hearing was achieved by occluding one ear with a combination of a hearing protector made of ear down (Bilsom Propp-o-plast) and a circumaural ear muff (Silenta Super, Exel Oy Ltd, Finland). This combination yielded very high average attenuation values. Cross-over hearing of the occluded ear by bone conduction was negligible, because in all cases the hearing aid was fitted to the better ear.

After the dial setting and ear occlusion, unaided and aided air-conduction tone audiograms were measured for frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz, which took about 15 min. In general, free-field tone audiograms are related to ISO-R226, 1961, which specifies values for the normal binaural minimum audible field. The minimum audible field for monaural listening should be taken 2 dB higher; this correction was included in our reference. From the unaided audiogram the pure-tone acuity (PTA, average hearing level at 500, 1000, and 2000 Hz) was calculated for classifying the subjects into one of five possible classes (see Sec.IID).

Next, the speech-intelligibility test was run, which took 35 to 40 min. The ten sentence lists were recorded on one track of a tape. On the other track the masking noise with exactly the same intensity and spectrum as the long-term average of the sentences was recorded. The output signals of the

two channels of the recorder could be attenuated independently. After attenuation the signals were mixed. The mixed signal was amplified and fed into the loudspeaker. Five sentence lists were used for measuring the unaided SRT in quiet (actually 0 dBA noise level) and against four noise levels of 40, 55, 70, and 85 dBA. The level of 85 dBA was included to satisfy, for all subjects, the boundary condition given by Eq.(2). The other five lists were used for measuring the aided SRT in quiet and against four noise levels of 25, 40, 55, and 70 dBA. It was expected that the noise level of 70 dBA was high enough to activate the aid's PC or AGC circuit, if present. The ten lists were invariably presented in the same order. For half of the subjects the unaided SRT's were measured first, and for the other half the aided SRT's were measured first. To eliminate the effects of training and fatigue as much as possible, the test conditions were counterbalanced over the subjects.

In addition to the hearing-impaired listeners ten young normal-hearing subjects were tested with the same equipment, for unaided conditions only, so as to obtain a reference for our speech material. Six lists were used for measuring the monaural SRT in quiet and against five noise levels of 25, 40, 55, 70, and 85 dBA. The conditions were also counterbalanced over the subjects.

C. Electroacoustical measurements

After the psychophysical measurements the hearing aids involved were electroacoustically tested with the volume control settings unchanged. Each aid was mounted on an artificial ear (2 cm^3 coupler) of a B&K hearing aid test box Type 4212, which provides frequency characteristics representative of free-field conditions. The input sound-pressure level during frequency sweeping was kept constant by the equalizing circuit of a B&K beat frequency oscillator Type 1022. The output levels of the fundamental, second-harmonic, and third-harmonic responses were measured separately in dB SPL using the 1/3-octave filters of a B&K audio-frequency spectrometer Type 2112. All curves were drawn on calibrated paper by means of a B&K level recorder Type 2305. The hearing aids were tested successively at sound levels of 50, 70, and 90 dB SPL. In addition, the fundamental-frequency response of the aid was recorded in the absence of an input signal to obtain an estimate of the aid's internal noise L_I in dBA.

The noise L_I was determined by the following procedure. For reasons of speech intelligibility, the frequency region of interest was restricted to five octaves, between 125 Hz and 4000 Hz. For each hearing aid tested the average amplification in each of the five octave-bands was read in dB from the fundamental-frequency response for the 50 dB SPL input sweep-tone. These amplification values were subtracted from the respective octave-band noise levels recorded with zero input. Since the internal noise L_I is to be specified in dBA, an A-weighting was applied to the noise levels. Notably for high-tone hearing aids with limited bandwidth, L_I was determined on the basis of four octaves by leaving the 125-250 Hz band out of consideration.

Some hearing aids showed, for one or more octaves, very low octave-band noise levels, so that the internal noise of the test box system had to be taken into account. The system noise level between 125 Hz and 4000 Hz was 8.1 dBA.

D. Subjects

All hearing-impaired subjects were recruited from the Audiology Center of the Free University. It was decided to investigate subjects below the age of 65, for which the hearing impairment was likely to come from a pathological origin rather than from presbyacusis. Besides, it was expected that younger persons would generally make the most of the rehabilitative potentialities of a hearing aid.

As known from the literature, the benefit of a hearing aid depends strongly upon the type and gravity of hearing impairment. For instance, different shapes of pure-tone audiograms require different hearing-aid frequency responses. In order to test the model presented in Sec.I in as representative cases as possible, five quite different types of commonly encountered audiograms were defined for selecting hearing-impaired subjects. These five types are specified in Table I. Both the subject's air-conduction and bone-conduction audiograms had to fit one of the five specified shapes for participation in the test.

In order to avoid type of hearing loss being confused with degree of hearing loss, ears with about the same amount of hearing loss were compared. To this end, the free-field PTA was chosen as a measure of overall hearing loss. Generally, conductive hearing impairment is attended by higher PTA values than high-tone hearing impairment with recruitment. In order to match the ranges of PTA values typical of the various types of hearing impairment involved, the lowest possible value of PTA was 30 dB and the highest allowable value (mainly determined by technically imposed limitations of sound levels) was 60 dB, both limits relative to ISO-R226(1961). The PTA range thus resulting was divided into five classes, namely 30.1-36.0 dB, 36.1-42.0 dB,

TABLE I. Specification of five types of hearing impairment with the corresponding audiogram shapes. The discomfort level is defined as the maximum sound level to be tolerated for a short period of time.

Type of pure-tone hearing loss	Air-conduction audiogram			Bone-conduction audiogram
	Audiogram slope (dB/octave)	Discomfort level (250-4000 Hz) (dB re ISO389)	Hearing span (250-4000 Hz) (dB)	(500-2000 Hz)
(1) perceptive high-tone (no recruitment)	>15 above 500 Hz	>115	>50	air-bone interweaving
(2) perceptive high-tone with recruitment	>15 above 500 Hz	<115	<50	air-bone interweaving
(3) perceptive flat	<10 above 250 Hz	>115	>50	air-bone interweaving
(4) mixed flat	<10 above 250 Hz	>115	>50	air-bone gap ≥15 dB
(5) purely conductive	<10 above 250 Hz	>115	>50	a) air-bone gap ≥25 dB b) threshold level <30 dB re ISO-389 (250-4000 Hz)

42.1-48.0 dB, 48.1-54.0 dB, and 54.1-60.0 dB. It was accomplished that, for each audiogram type previously defined, ten subjects, viz. two per PTA class, were investigated. Thus, a 5x5x2 experimental design was brought about. The test conditions were also counterbalanced over PTA classes, which enabled an optimal comparison between types of hearing loss.

In order to find suitable subjects, the data files of the Audiology Center were examined. At this Center, all audiograms are measured over headphones (re ISO-389), and word discrimination scores in quiet are obtained with Dutch phonetically balanced monosyllables. Criteria set for tentative selection of subjects were: a behind-the-ear aid must have been fitted, not more than two years ago, to the ear with the smaller PTA as well as the better word discrimination score in quiet; the audiogram of that ear must correspond to one of the audiogram shapes defined in Table I, and the PTA

TABLE II. Age and sex for the ten subjects in each of the five groups formed on the basis of type of hearing impairment. The type numbers refer to those specified in Table I.

Type of impairment	Age interval	Mean age	Sex male	Sex female
1	26 - 67	50.9	4	6
2	28 - 68	53.2	8	2
3	37 - 63	56.3	4	6
4	26 - 63	46.2	4	6
5	21 - 57	42.5	6	4

(re ISO-389) must lie between the upper and lower limit for the free-field PTA. Whether a subject really fitted in with the experimental design was not clear until he had actually participated in the test, so that a free-field PTA was obtained. Because of both duplications in some PTA classes and free-field audiogram shapes not satisfying the criteria, somewhat more than the 50 subjects needed have participated. In Table II, the finally resulting distribution of age and sex is given for the five groups with different hearing impairment.

For the ten normal-hearing subjects (age 19-24, mean age 21), included in the investigation to obtain reference values for the speech intelligibility test, an average free-field audiogram slightly better than specified in ISO-R226 was observed (mean PTA of -2.4 dB).

III. RESULTS

A. Application of the model to individuals

In order to have a reference for the hearing losses for speech, the curve-fitting procedure was first applied to the individual SRT values of the ten normal-hearing subjects. This group as a whole was, by definition, free from speech-hearing losses. By taking A+D = 0 (i.e. SHL in quiet) and D = 0 (SHL in noise) in Eq.(1), the reference values L_0=13.5 dBA and ΔL_{SN} = 5.8 dB were calculated. The fitting procedure determined, per subject, the standard deviation (s.d.) of the six SRT values from the best-fitting curve. The median of these s.d. values is 0.85 dB. This value will be taken as the reference for the goodness of fit.

After finding L_0 and ΔL_{SN}, the first stage of curve fitting was performed on the individual SRT values of the 50 hearing-impaired subjects. This resulted, per subject, in estimates of A+D and D, together with an s.d. of fit. From Eq.(2) it can be concluded that for obtaining reliable estimates, the allowable maximum value of class-A hearing loss is 56 dB. All subjects satisfy this criterion. The median of the 50 s.d. values of fit is 1.14 dB.

After finding A+D and D, the second stage of curve fitting was applied to the individual ASRT values, both for the test situation including L_N=70 dBA (5 ASRT values) and for the situation without that level (4 ASRT values). This distinction was made for studying the effect of PC or AGC circuit activation on the hearing aid's S value. According to Eq.(5), reliable estimates of G and S for the situation where L_{Nmax}=55 dBA can only be expected if G > A-25.5 dB (median L_I=22 dBA, see Sec.IIIF). For two subjects the observed G value was slightly below the criterion. The fitting resulted, per subject, in estimates of G and S, and a value of s.d.. The median of the 50 s.d. values is 1.15 dB for the 5-point fit, and 1.05 dB for the 4-point fit. The medians of the 50 values of S are 1.25 dB and 1.15 dB, respectively.

The small differences between the 5-point medians and the 4-point medians indicate that, at a noise level of 70 dBA, the output control of the aids does not substantially increase ASRT in noise. Therefore, the next data presented on aided hearing will definitively include the noise level of 70 dBA.

B. Test of the model on the basis of PTA class

The applicability of the model, irrespective of type and gravity of hearing impairment and of hearing-aid performance, has to be verified first of all. To this end, the 50 subjects were grouped into the five PTA classes formerly defined. Thus, each class consisted of five pairs of subjects, each pair representing a specific type of hearing impairment according to Table I.

In Table III, for each class of ten subjects the medians of the measured SRT's are compared with the corresponding SRT values derived from fitting a curve, according to Eq.(1), to these medians. Similarly, the medians of individually computed values of A+D, D, and s.d. are compared with the corresponding values of the curve best fitting the SRT medians.

In Table IV, for each class a comparison is made between the medians of the measured ASRT's and the best-fitting ASRT values situated on curves described by Eq.(4). Furthermore, the medians of the individual values of G, S, and s.d. are compared with the corresponding values of the best-fitting

TABLE III. Comparison of individual data with class data. The classes I to V, each consisting of ten hearing-impaired subjects, were formed on the basis of five PTA intervals as defined in Sec.IID. For each class the first row gives the medians of the individual SRT values measured at five noise levels L_N shown at the top of the table, and of the individual SHL values A+D and D. The next row gives the SRT values defined by the curve best fitting the five median SRT values of the class, with the resulting estimates of A+D and D. The value s.d. is the standard deviation of the five SRT medians from the best-fitting curve.

PTA class		SRT (dBA)					SHL (dB)		(dB)
		L_N=0	40	55	70	85 dBA	A+D	D	s.d.
I	median	44.4	45.8	54.2	66.6	81.4	31.2	2.5	
	best fit	44.9	45.5	52.8	67.1	82.1	31.4	2.7	0.85
II	median	46.6	47.8	56.8	68.2	81.4	33.2	2.6	
	best fit	47.2	47.7	54.2	68.3	83.3	33.7	3.9	1.62
III	median	56.4	55.2	59.2	68.6	82.0	42.7	3.9	
	best fit	56.3	56.3	57.9	68.1	82.8	42.8	3.4	0.99
IV	median	65.4	66.4	67.0	70.2	82.8	52.7	3.8	
	best fit	66.2	66.2	66.4	70.1	82.8	52.7	3.4	0.52
V	median	69.0	68.4	69.0	72.0	83.8	55.7	4.9	
	best fit	68.8	68.8	68.9	71.8	83.9	55.3	4.3	0.26

curves. The table also shows the medians and best-fitting values of the aided hearing losses for speech (ASHL) in quiet (=A+D-G+S) and in noise (=D+S). The individual values of ASHL, for which the medians are shown, were achieved by adding the respective individual values of A+D, D, G, and S. The best-fitting ASHL values were achieved by simply adding the seperate best-fitting values of A+D, D, (given in Table III), G, and S (given in Table IV).

In Fig.3 the medians of the measured SRT and ASRT values are plotted, for each class separately, as a function of noise level. The solid curves have been fitted to the SRT medians, the dashed curves to the ASRT medians.

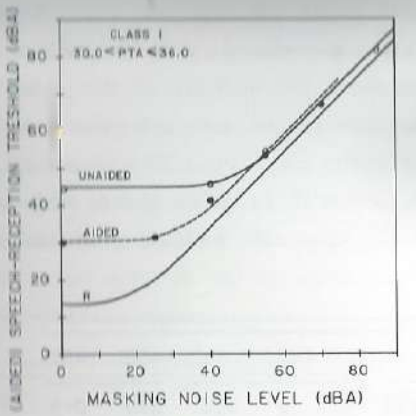
The predominant impression from the results presented above is that the model is valid. The discrepancies, in some cells of Tables III and IV, between measured and fitted results will be discussed and interpreted in Sec.IVA.

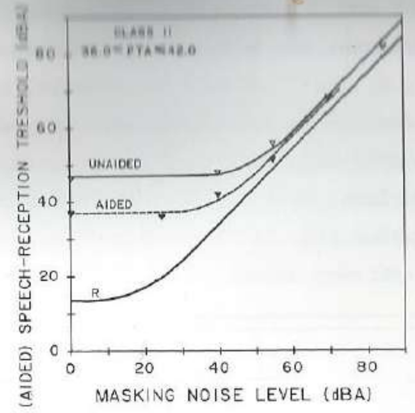
TABLE IV. Comparison of individual data with class data for the aided test conditions. For each class the first row gives the medians of individual ASRT, G, and S values. Furthermore, this row shows the medians of the aided SHL (=ASHL) in quiet (=A+D-G+S) and in noise (=D+S). The next row gives the five ASRT values defined by the curve best fitting the five ASRT medians of each class, with the best estimates of G, S, and ASHL in quiet and in noise. The value s.d. is the standard deviation of the five ASRT medians from the best-fitting curve.

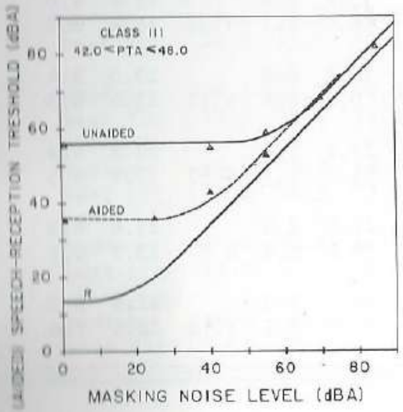
PTA class		ASRT (dBA)					(dB)			ASHL (dB)	
		L_N=0	25	40	55	70 dBA	G	S	s.d.	A+D -G+S	D+S
I	median	29.8	31.2	41.0	52.8	66.8	13.7	0.9		15.4	4.0
	best fit	30.3	31.0	38.9	53.2	68.2	16.1	1.1	1.32	16.4	3.8
II	median	37.6	36.2	42.0	51.8	68.0	10.4	0.0		23.3	3.4
	best fit	37.2	37.4	40.5	52.8	67.7	9.6	-0.6	1.11	23.4	3.3
III	median	35.4	36.6	43.4	52.8	68.4	21.2	1.2		22.3	4.8
	best fit	36.3	36.6	40.8	54.0	68.9	21.2	1.1	1.52	22.7	4.5
IV	median	38.4	39.6	44.6	55.2	70.2	22.0	2.0		25.6	6.3
	best fit	39.2	39.4	43.0	55.7	70.6	29.8	2.8	0.97	25.7	6.2
V	median	45.9	45.4	48.4	55.8	73.1	28.4	3.3		32.6	8.3
	best fit	45.9	46.0	47.4	57.3	72.0	26.2	3.3	1.10	32.4	7.6

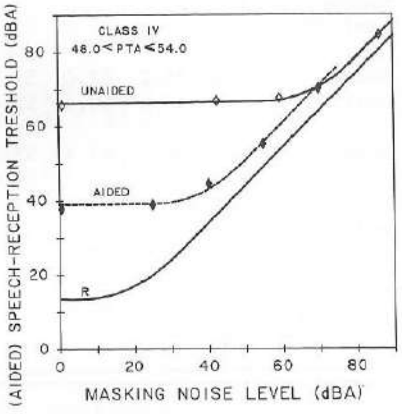
FIG.3. Free-field SRT and ASRT for sentences as a function of noise level, separately shown for five classes of hearing-impaired subjects formed on the basis of PTA. The data points are the medians as given in Tables III and IV. The solid curves, marked R, are reference curves pertinent to unaided normal-hearing listeners. The other solid curves and the dashed curves represent the curves best fitting the experimental unaided and aided data, respectively.

(please turn over, for the figure proper)









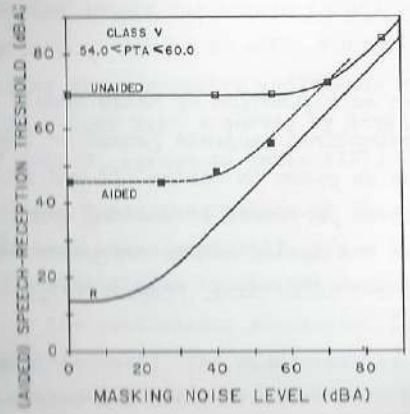


FIG.3.
(for legends see preceding page)

C. Effect of type of hearing impairment on SHL and ASHL

As mentioned before in Sec.IID, a group of ten subjects was investigated for each type of hearing impairment considered. In all groups, two subjects per PTA class participated, as a result of which the distribution of hearing losses is similar in terms of PTA. In Table V the medians of A+D, D, G, S, and ASHL in quiet and in noise are compared, for each group, with the corresponding values of the two curves fitting the medians of SRT and ASRT, respectively.

It is obvious from Table V that the groups differ with respect to the relationship between SHL, and ASHL, in quiet and in noise. This is illustrated in Fig.4, where, for each group, the median of SHL in noise is plotted versus SHL in quiet, as well as the median of ASHL in noise versus ASHL in

TABLE V. For each of the groups of ten subjects, formed on the basis of type of hearing impairment (cf. Table I), the first row shows the medians of the individual SHL values A+D and D yielded by the first stage of curve fitting, and of G and S yielded by the second stage. Furthermore, the medians of ASHL in quiet (=A+D-G+S) and in noise (=D+S) are given. The next row gives the corresponding results from fitting curves, defined by Eqs.(1) and (4) respectively, to the SRT and ASRT medians (not shown) of the groups. The values s.d. are the standard deviations of fit.

Group (type)		Stage 1 (dB)			Stage 2 (dB)				
		SHL						ASHL (dB)	
		A+D	D	s.d.	G	S	s.d.	A+D-G+S	D+S
1	median	44.8	4.6		16.8	0.6		28.4	5.5
	best fit	43.4	5.3	0.85	15.3	0.3	0.96	28.4	5.6
2	median	24.4	4.7		8.3	0.8		18.4	6.2
	best fit	24.6	5.9	1.67	6.0	0.4	1.29	19.0	6.3
3	median	42.5	3.4		15.0	1.2		26.7	5.9
	best fit	43.4	3.8	1.29	18.1	2.0	1.30	27.3	5.9
4	median	43.4	1.9		23.8	1.5		25.0	4.2
	best fit	43.4	2.3	0.43	19.8	1.0	1.48	24.6	3.4
5	median	48.1	1.8		29.4	1.0		22.5	2.4
	best fit	48.2	1.8	0.81	26.0	0.4	0.92	22.6	2.3

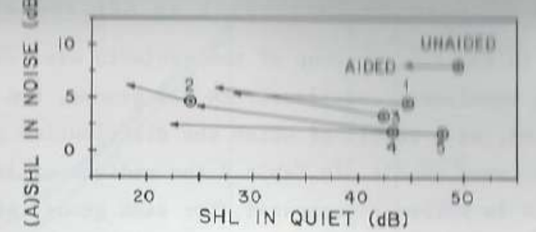


FIG.4. For the same five groups of subjects as presented in Tables V and VI, the median of SHL in noise is plotted versus SHL in quiet. Also, the median of ASHL in noise is plotted versus ASHL in quiet. The arrows, pointing from SHL to ASHL, show the effects of the hearing aids on speech intelligibility both in quiet and in noise.

quiet. It should be realized that the deviating SHL in quiet shown by group 2 is present despite the equal distribution of PTA for the groups. The figure also clearly demonstrates that, in general, benefit of hearing aids is only to be expected under relatively quiet listening conditions.

D. The benefit of hearing aids

For a better insight into the potential benefits of hearing aids for an individual, it is necessary to consider the relationship between SHL and ASHL for the subjects apart. In quiet, all aids are of benefit with reductions of SHL between 2.5 dB and 36.1 dB. In noise, the situation is less positive. In Fig.5 the hearing-aid effect on SHL in noise is represented on the basis of SHL in quiet. It can be seen that most aids intensify an individual's SHL in noise or leave it unchanged. Although the deterioration is small in terms of dB, typically between 1 dB and 5 dB, the aid induces a substantial extra handicap for the hearing-aid wearers in noise, since every dB of hearing loss in noise lowers sentence intelligibility by 17% to 20% (Duquesnoy,1982). Only four subjects derived some benefit from the aid in noise (gain of more than 1 dB; conditional probability of $S < -1$ dB, where mean S is zero, is about 6%).

E. Analyses-of-variance

In order to systematically investigate the effects of relevant variables on unaided and aided speech intelligibility, an analysis-of-variance was performed on the SRT and ASRT values, as well as on the SHL and ASHL values. The variables (factors) considered were in order of hierarchy: (1) type of hearing impairment (T), (2) noise condition (N), (3) unaided/aided situation (UA),

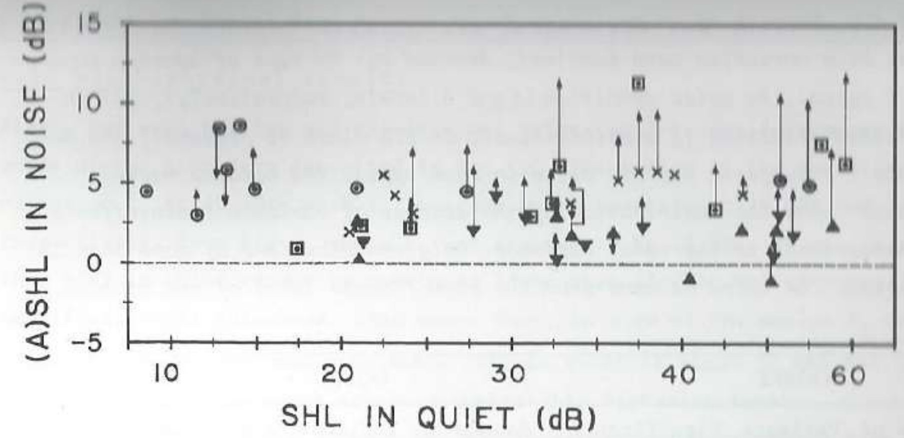


FIG.5. The effect of hearing aid on SHL in noise, shown for each of the 50 subjects investigated. The subjects are represented, according to their type of hearing impairment (cf. Table I), by the following symbols: type 1: X, type 2: O, type 3: □, type 4: ▼, and type 5: ▲. The arrows point to the corresponding ASHL values. The arrows were omitted for individuals for whom the difference between SHL and ASHL does not exceed ±1.0 dB. The SHL (and ASHL) values in noise are plotted versus SHL in quiet, so as to demonstrate the different individual relations between SHL in noise and in quiet, depending on the type of hearing impairment involved.

(4) PTA class (C), and (5) subjects (S). Concerning SRT and ASRT (written below as (A)SRT), four test levels of noise L_N coincided, viz. 0, 40, 55, and 70 dBA. Thus, a 5x4x2x5x2 factorial experiment resulted, involving 200 SRT values and 200 ASRT values. Concerning SHL and ASHL (written below as (A)SHL), only two noise conditions applied, viz. (A)SHL in quiet and (A)SHL in noise, yielding a 5x2x2x5x2 design with 100 SHL values and 100 ASHL values.

In Table VI the results of the analyses-of-variance are listed. Only main effects and interactions accounting for more than 1% of the total variance are given, as far as they are significant. It is surprising that the factor T, although significant, explains less than 1% of the variance of both (A)SRT and (A)SHL. Apparently, the relatively low (A)SRT values for group 2 at noise levels of 0 dBA and 40 dBA are not an important extra source of variance. The factor N accounts for most of the variance. Also, the factors UA and C contribute considerably to the total variance. Concerning (A)SRT, the subjects are a minor though significant source of variance.

TABLE VI. Results of an analysis-of-variance on both (A)SRT and (A)SHL values. Five variables were involved, denoted by: T= type of hearing impairment (5 types), N= noise condition (4 and 2 levels, respectively), UA= unaided/aided situation (2 possibilities), C= PTA class (5 classes), and S= subjects (2 subjects per PTA class in each type). The columns marked "variance" give the contribution of the sources of variance (main effects or interactions) to the total variance. Only contributions of more than 1% are listed. The third columns show the significance level of the sources.

(A)SRT			(A)SHL		
Source of variance	Variance (%)	Significance (%)	Source of variance	Variance (%)	Significance (%)
N	50.2	<0.001	N	69.9	<0.001
UA	13.0	"	N×UA	7.7	"
C	10.9	"	UA	5.4	"
N×UA	6.7	"	C	4.6	"
S	2.6	"	N×C	2.3	"
UA×S	2.4	"	T×N	1.8	"
UA×C	1.9	0.025			
T×N	1.8	<0.001			
N×C	1.6	"			
T×UA	1.3	0.110			

The other relevant sources are all interactions between the factors. Especially, interaction NxUA is very important, which is not surprising in view of the different effects of hearing aids on speech intelligibility (and hearing losses) in quiet and in noise. Two other significant interactions pertinent to both (A)SRT and (A)SHL are NxC and TxN. They imply, with respect to (A)SRT, that the impact of PTA on (A)SRT varies with noise level, and that the influence of noise level on (A)SRT depends on the type of hearing impairment. Similarly, the interactions concerning (A)SHL mean that (A)SHL in quiet and in noise depend in a different way on PTA as well as on the type of hearing impairment. The remaining significant interactions to be dealt with apply to (A)SRT only. They are UAxS, UAxC, and TxUA, and indicate that the effect of a hearing aid on speech intelligibility is determined by the subject using the aid, by the magnitude of PTA, and by the type of hearing impairment.

F. Electroacoustical hearing-aid properties and their relations with psychophysical results

First, for each hearing aid the internal noise level L_I was determined by means of the procedure described in Sec.IIC. The median of the resulting 50 values of L_I is 22 dBA, with a lower and upper quartile of 19 dBA and 26 dBA, respectively. From Fig.2, where L_I was assumed to be 25 dBA, it can be seen that ASRT in quiet cannot be made much lower than 35 dBA, unless excessive amplification is tolerated. This means that, in view of the median L_I of 22 dBA found, the minimum attainable ASRT in quiet is about 32 dBA for the current hearing aids being set at a comfortable listening level.

Next, the relations between the electroacoustical hearing-aid gain in quiet, G_e, and three psychophysical gain measures will be presented. The gain G_e is defined as the average gain between 500 Hz and 4000 Hz, measured at an input sweep-tone level of 50 dB SPL. The psychophysical measures considered are: (1) the average improvement, G_{pt}, in the individual pure-tone thresholds at 500, 1000, 2000, and 4000 Hz, (2) the model-derived acoustic gain G-S, and (3) the functional gain in quiet G_f (=SRT_0-$ASRT_0$; cf. Pascoe, 1975). In Table VII the medians of the four gain measures are given for the five PTA-based classes. Another measure shown in the table is the gap ΔL_N, which is the difference between ASRT in quiet and the apparent total noise level L_{NT} at the

TABLE VII. Comparison between four kinds of hearing-aid gain in quiet for the PTA-based classes of subjects presented in Table III. The measures considered are: (1) the electroacoustic gain G_e, (2) the pure-tone threshold reduction G_{pt}, (3) the model-derived gain G-S, and (4) the functional gain G_f. Furthermore, the last column, marked ΔL_N, gives the gap between ASRT in quiet and L_{NT} defined by Eq.(3). For each class the medians are shown.

Class	G_e	G_{pt}	G-S	G_f	ΔL_N (dB)
I	29.5	16.5	10.6	9.6	11.5
II	25.5	17.7	9.1	10.2	11.2
III	26.5	21.5	17.9	19.6	10.7
IV	36.0	22.0	19.5	18.8	21.0
V	36.0	27.5	25.1	23.8	26.0

aid's microphone, according to Eq.(3). This measure has been calculated in order to check that the internal noise of the aids does not deteriorate the effectiveness of amplification. As stated above, this holds for gaps ≥ 10 dB.

Similarly, in Table VIII the same four gain measures and ΔL_N are given for the five groups of subjects formed on the basis of type of hearing impairment. Concerning group 2, the small medians of G-S and G_f are remarkable in view of the large gain G_e.

In order to gain a more detailed insight into the relationship between psychophysical gain and electroacoustic gain the model-derived gain G-S is plotted, in Fig.6, versus G_e for each individual. Most data points representing individuals suffering from type-2 hearing impairment fall outside the cluster, as would be expected from the results presented in Table VIII. If these ten subjects with recruitment in the ear are left out of consideration, the linear regression of G-S on G_e for the other 40 ears is: $\overline{G-S} = 0.7G_e - 2.0$, with a correlation coefficient, r, of 0.82. The standard deviation of individual points from the regression line is 4.9 dB. For the ears with recruitment no relationship between G-S and G_e seems to be present.

Under noisy conditions the quality rather than the amount of amplification is crucial. In this respect, second-harmonic and third- harmonic distortions are useful measures of electroacoustic quality assessment. For convenience, the following presentation of second and third-harmonic distortion measurements will be confined to the absolute maxima observed in the frequency range between 250 Hz and 4000 Hz. Each peak value is expressed in dB relative to the level of the concomitant fundamental frequency. If a peak value is 24 dB or more below the fundamental level, the distortion is considered negligible (≤6%).

TABLE VIII. Comparison between the medians of the same five measures in quiet as shown in Table VII, for the five types of hearing impairment presented in Table V.

Type	G_e	G_{pt}	G-S	G_f	ΔL_N (dB)
1	26.5	21.2	14.2	12.0	18.5
2	30.5	19.2	7.7	7.4	9.7
3	21.0	16.0	14.4	16.0	21.1
4	30.5	25.7	22.9	21.8	15.8
5	39.0	27.0	27.9	27.2	12.9

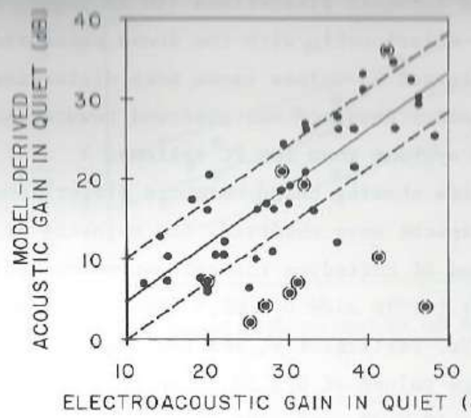


FIG.6. Model-derived gain, G-S, versus G_e for the 50 subjects investigated. The ten dots in parentheses represent subjects suffering from type-2 hearing impairment. The regression line (solid line) holds for the remaining 40 points, and is described by: $\overline{G-S} = 0.7G_e - 2.0$ (r=0.82). The standard deviation from this line, as indicated by the dashed lines, is 4.9 dB.

This is the case with 48 hearing aids tested at an input sweep-tone level of 50 dB SPL. The remaining two aids show slight second-harmonic distortions, only. For the 70-dB test condition, however, 14 aids show second-harmonic distortions up to a level of -10 dB(=32%). Besides, two of these aids show moderate third-harmonic distortions. For the highest test level applied (90 dB SPL) all, but three, hearing aids produce second-harmonic distortions up to 7 dB(=224%), and 21 hearing aids also produce third-harmonic distortions up to a level of 2 dB(=126%).

It has been investigated whether the respective distortions at the 90 dB test level were linked with other electroacoustic hearing aid parameters, such as the gain in quiet G_e, the type and degree of output limitation, and the effective frequency range. The difference, ΔG_e, between G_e and the reduced gain measured for an input level of 90 dB SPL was taken as a measure of output limitation. For 25 aids involved the output limitation was effectuated by an AGC system, for 20 aids by PC circuits, and for three aids by a combination of AGC and PC. The remaining two aids had no separate output control. The effective frequency range, Δf, was determined by using the acoustic gain level of 10 dB; the range between the intersections of this level and the frequency response curve (50 dB input) defines Δf of the aid (cf. Harford and Fox, 1978).

Concerning the second-harmonic distortions (90 dB input, 47 aids involved), the only obvious relationship with the above presented parameters is the one with G_e, i.e. larger G_e values cause more distortions. Also, under comparable conditions a minor tendency was observed towards somewhat lower distortion values for AGC systems than for PC systems.

For the 21 hearing aids showing third-harmonic distortions (also 90 dB input) some interesting aspects were observed. The majority of these aids have values of G_e, ΔG_e, and Δf exceeding the corresponding median values for all 50 aids together, viz: for 16 aids $G_e > 29.5$ dB, for 14 aids $\Delta G_e > 9.0$ dB, and for 16 aids $\Delta f > 4735$ Hz. Particularly, for the nine aids with distortion levels > -12 dB (>25%) large values of G_e, ΔG_e, and Δf were measured, viz: average values of 38.5 dB, 19.9 dB, and 5100 Hz, respectively. Eleven out of the 21 aids considered were equipped with AGC, and nine with PC output control. No differences in distortion level were found between the two systems.

An important question is to what extent the second-harmonic and third-harmonic distortions affect the aided speech hearing in noise. The highest background-noise level applied was 70 dBA, which implies that many peaks in the fluctuating speech signal simultaneously presented with the noise will amply exceed a level of 70 dB SPL. Therefore, it is justified to relate the model-derived speech distortion S of the aids to the harmonic distortions measured at an input level of 90 dB SPL rather than to those measured at 70 dB SPL. The distortion S was plotted versus the second-harmonic peak distortion for each of the 47 aids. The scatter diagram (not depicted) did not reveal any relationship between the two variables. In addition, in Fig.7 distortion S is given as a function of third-harmonic distortion for the 21 aids in question. A certain relationship between S and third-harmonic distortion (HD_3) can be observed. Linear regression yielded: $\overline{S} = 0.16HD_3+4.2$, with r=0.53 and a standard error of regression of 1.9 dB. For these 21 aids the median S value is 2.4 dB, whereas the median S of the remaining 29 aids is only 0.7 dB. The 21 subjects wearing the more distorting aids have a median D value of 3.1 dB, as compared to 3.3 dB for the others. Furthermore, the 21 subjects are evenly spread over the types of hearing impairment involved, but they show the larger hearing losses in terms of PTA.

Finally, some data on the relationship between S and type of output limitation are given. The median S value for the 25 aids with AGC system is 1.0 dB, for the 20 PC systems it is 2.0 dB, and for the other 5 aids (no control or a different one) 0.1 dB. Thus, the advantage of AGC over PC is generally about 1 dB in S/N ratio.

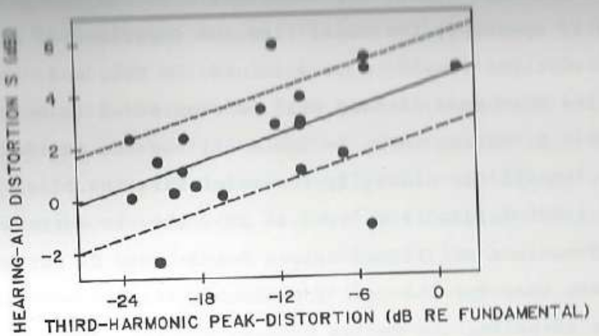


FIG.7. Distortion S plotted versus third-harmonic peak-distortion for the 21 aids involved. The solid line shows the regression of S on third-harmonic distortion (HD_3), described by: $\overline{S} = 0.16HD_3 + 4.2$. The dashed lines show the standard error of regression (1.9 dB).

IV. DISCUSSION

A. Validity of the model

First, the results from applying the model to individuals will be discussed. It is shown in Sec. IIIA that the reference value for the goodness of fit is 0.85 dB. The scarcely larger s.d. values of fit found for the hearing-impaired subjects (a median value of 1.14 dB for the unaided situation, and of 1.15 dB for the aided situation (5-point fit)) are a promising indication of the applicability of the model. These values, however, do not reveal systematic differences, if any, between measured and predicted SRT and ASRT values. Thus, it was tested by means of the Chi square statistic (Dixon and Massey, 1968) whether the differences obtained from the 50 subjects investigated were positive and negative in a sufficiently random way. The statistic was applied, per test condition, to the 50 individual differences under the hypothesis of equal probability of plus and minus signs. The hypothesis was rejected at the 1% level of significance for the following situations: unaided, L_N=85 dBA (10 plus, 40 minus signs); aided, L_N=40 dBA (39+, 11-) and 55 dBA (14+, 36-). In the case of 4-point fitting (aided situation), where the median s.d. value of fit is only 1.05 dB (cf. Sec.IIIA), the Chi square statistic also revealed systematic differences in sign frequencies for L_N = 40 dBA (44+, 6-) and 55 dBA (5+, 45-).

From the fitting results for individuals it can be tentatively concluded that, generally speaking, the model fits the experimental data, although some systematic deviations remain to be examined. To this end, the results pertaining to the five PTA-based classes will be considered in more detail.

From the s.d. values shown in Table III for the unaided situation it is evident that, except for class II, the model fits the class data better than the individual SRT values (s.d.'s < 1.14 dB). This is corroborated by the similarity of medians and fitted values for A+D and D. Furthermore, it is important to note that for the test condition L_N=85 dBA most classes show predicted values exceeding the median values. Nevertheless, only for class II the difference (1.9 dB) exceeds the s.d. value of fit (1.62 dB). However, for classes I to IV another unaided test condition, viz. L_N=55 dBA, yields consistently negative differences between fitted and median values, which all exceed the corresponding s.d. values. This was not at all indicated by the Chi square tests on individual differences. As will be explained later, this deviation at L_N=55 dBA is mainly an arithmetical artefact caused by averaging curves, defined by Eq.(1), with the turning points at different noise levels (cf. the curves in Fig.1, which have turning points at L_N=20 dBA and 60 dBA, respectively).

Summarizing the foregoing discussion, the validity of the model for unaided conditions has clearly been demonstrated, although at very high noise levels (L_N >85 dBA) a trend towards measured thresholds slightly lower than model-derived thresholds (maximally 1 dB) must be taken into account. This trend might be related to an increased bandwidth of the ears at high levels.

Next, the class data from aided situations, as given in Table IV, will be discussed. For classes I and III the s.d. values of fit are larger than the median s.d. for individuals, but for the other three classes a more accurate fit to the class data than to the individual data is observed. These differences between classes in the goodness of fit are not recognized as such in the differences between the medians and fitted values of G and S. On the contrary, for class III (s.d.=1.52 dB) the conformity is striking, whereas for class IV (s.d.=0.97 dB only) the differences are striking. Closer investigation of the latter case reveals that the distribution of the ten G values is so skew as to lead to a misleading median value. This skewness in the distribution of G is also observed in some other classes. It is caused by the diversity of amplification settings between individuals within the same class. The estimation of G and S, with the A and D resulting from the first stage of fitting as

constants, has been accurate enough, in view of the excellent goodness of fit commonly achieved at both fitting stages. The overall accuracy is confirmed by the great conformity of the median and fitted ASHL values (see Table IV), which is attainable only if the results from both fitting stages are correct.

Another aspect to be considered is whether the systematic differences observed between the individually measured and predicted ASRT values at L_N= 40 dBA and 55 dBA also become apparent in the class data. Indeed, for all classes the differences at L_N=40 dBA are consistently positive, and at L_N=55 dBA consistently negative. The differences at L_N=40 dBA exceed, except for class V, the corresponding s.d. values; the differences at L_N=55 dBA show the reverse picture. A complicating factor at L_N=40 dBA is the likely presence of an arithmetical artefact as mentioned before in the unaided case. Therefore, for each class a separate Chi square test was performed on the signs of the ten individual differences per test condition. For the classes I and III significantly too many plus signs (2% level) were found at L_N=40 dBA, but also in the other classes the plus signs prevailed. Besides, the test did not reveal any significant irregularity at L_N=55 dBA, although the minus signs prevailed in all classes.

Summarizing the above, the model is certainly valid for aided listening conditions with background noise levels up to 70 dBA. The minor anomalies in individual and class-based results point to a possibly somewhat smoother turning point than described be Eq.(4), and to an S value not entirely constant with increasing noise level, since at L_N=55 dBA the true S is slightly smaller, and at L_N=70 dBA slightly larger than the model-derived S value. This tendency in S is clearly related to increasing harmonic distortions at higher noise levels (cf. Sec.IIIF). Nevertheless, the model yields accurate and usuful values for G and S for the conditions tested.

Next, the mechanism underlying the arithmetical artefact will be demonstrated. The artefact may appear if SRT values pertinent to various curves, described by either Eq.(1) or (4), are numerically added in order to obtain an average curve. Figure 8 shows two extreme cases of adding up curves, described by Eq.(1). Curves 1 and 2 have their turning points at L_N=55 dBA; curve 3 has its point at 25 dBA. The position of the turning points is determined exclusively by the value of class A hearing loss. The upper dashed curve is characterized by A and D values, which are the arithmetic mean of the corresponding values of curves 1 and 2. The lower dashed curve is described by the mean A and D values of curves 2 and 3. The upper crosses also shown re-

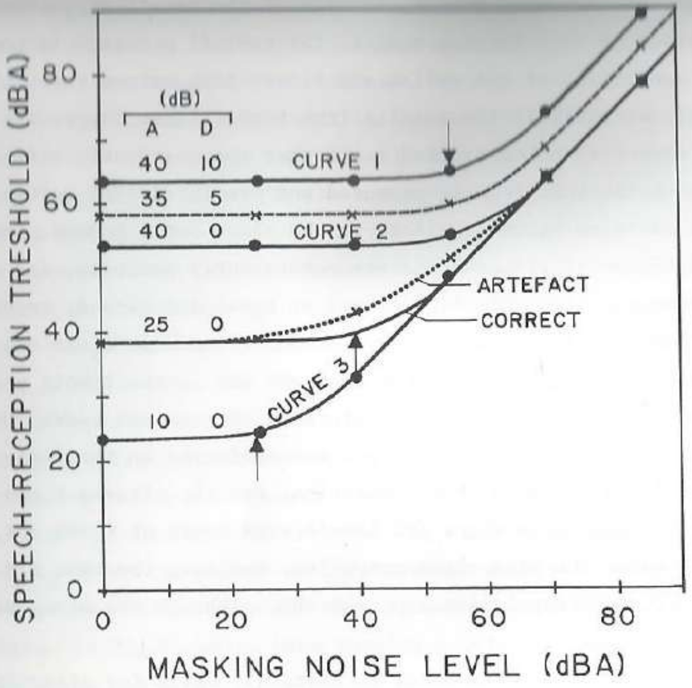


FIG.8. The solid curves 1 to 3 and the two intermediate curves are all described by Eq.(1), with values for A and D as specified. The arrows point to the turning points of the curves, at L_N=25, 40, and 55 dBA, respectively. The upper crosses represent, at six different noise levels, the arithmetic means of SRT's positioned on curves 1 and 2; the lower crosses show the SRT means for curves 2 and 3. The dotted curve, marked 'artefact', connects the lower crosses, as far as they deviate from the dashed curve, marked 'correct'. The arithmetical artefact at L_N=40 dBA amounts to 3.5 dB.

present, at various noise levels, the arithmetic mean of the SRT values positioned on curves 1 and 2, respectively. All these crosses coincide with the upper dashed curve. The lower crosses represent the SRT means for curves 2 and 3. In particular, the crosses at L_N=40 dBA and 55 dBA deviate substantially from the corresponding dashed curve, which defines the correct SRT values. It is evident, therefore, that the addition (and averaging) of SRT values pertinent to curves with differently positioned turning points automatically results in a smoothed average curve, as shown in the figure by the dotted curve, marked 'artefact'.

B. Effects of type of hearing impairment and hearing aids on speech intelligibility

A first aspect to be noted from Table V concerns the generally small s.d. values achieved in both fitting stages, and the great conformity between measured and fitted ASHL values. This agreement confirms the applicability of the model to the divergent types of hearing impairment considered.

Three other aspects of Table V to be noted are: (1) SHL and ASHL in noise are largest for the high-tone and flat sensorineural losses (types 1, 2, and 3); (2) the detrimental effect of the hearing aid in noise (S value) is smallest for subjects with high-tone losses (types 1 and 2); notably, the two subjects who derived some benefit from the aid in noise (S value of -3.2 dB and -2.3 dB, see Fig.5) suffer from high-tone losses with recruitment (type 2); although it is not likely that the aids in question actually improve the S/N ratio, the relative advantage in noise for subjects with steeply sloping audiograms may be attributed to retrieved audibility of speech components in a frequency region not contributing to hearing without the support of an aid; (3) the gain G of the aids in quiet is evident; particularly, the conductive hearing losses (type 5) are easily compensated for by amplification. The gain is less pronounced for the sensorineural losses. Especially the subjects with recruitment in the ear derive only a minor benefit in quiet. Their median SHL in quiet is but small, and the internal noise L_I of the aids hampers greater benefit, despite considerable electroacoustic amplification (cf. Table VIII). Apparently, recruitment is attended by a rather deviating pattern of (aided) speech hearing (cf. Fig.4), even for the present group with average pure-tone hearing losses completely matching those of the other types of impairment considered.

There is not much to add to the results of the analyses-of-variance. They are a systematic affirmation of the aspects of unaided and aided hearing already discussed. Two interesting points to mention are: (1) the PTA is clearly a measure pertinent to speech intelligibility in quiet only, and (2) the benefit attainable with a hearing aid is, under comparable conditions, fairly dependent on the user.

C. Relations between electroacoustic and psychophysical results

First, the gain measures in quiet shown in Table VII, per PTA class, will be considered. Three aspects are to be discussed: (1) the model-derived gain in quiet, G-S, is almost identical with the functional gain in quiet; this is understandable, because for all classes ΔL_N exceeds the limit of 10 dB required for effective amplification (cf. Fig.2 and Sec.IIIF); (2) there is a tendency towards a smaller gain in speech intelligibility than in pure-tone thresholds; in particular class II shows a substantial difference (8.6 dB) between the medians of G_{pt} and G-S, as against 7.4 dB between the corresponding means (not shown); the latter difference is significant at the 2% level (one-sided t-test); the differences in I and II are primarily due to the subjects with high-tone losses (types 1 and 2); (3) the electroacoustic gain is considerably larger than the threshold improvements for both pure tones and speech. Concerning the last point, the internal noise of the aids can hardly be the limiting factor, definitely not for classes IV and V ($\Delta L_N > 20$ dB). Elucidative is the finding of Pascoe (1975) that $2cm^3$-coupler gain measurements greatly overestimate the gain above 2000 Hz; he observed functional gain drops of as much as 24 dB from 1500 Hz to 3000 Hz for aids showing a $2cm^3$-coupler gain up to 4000 Hz. Thus, in terms of real benefit to be expected (in quiet) coupler measurements are misleading. Nevertheless, the linear regression, shown in Fig.6 for subjects without recruitment in the ear, may serve as a rough assessment of the functional gain in quiet from coupler measurements. Hence, a gain G_e of e.g. 30 dB will generally yield a functional gain in quiet between 14 dB and 24 dB.

The deviating values of G-S and G_f for subjects with recruitment (type 2 losses) are clearly revealed by Table VIII. It is important to note that their median ΔL_N value is slightly below 10 dB, which means that the amplification has been partially ineffective due to the internal noise of the aids. This is probably also the case for the subjects with conductive losses (type 5). Anyway, several subjects strived for an ASRT in quiet lower than their aid could manage because of its internal noise. It is desirable, therefore, to construct hearing aids with the maximum L_I value well below 20 dBA, so that an ASRT in quiet of about 25 dBA becomes acoustically realizable. Then, with such an aid revalidation of subjects with moderate hearing losses (PTA between 30 dB and 40 dB) will be, at least in quiet, more substantial than it is now.

A final aspect to be noted in Table VIII is the large difference between G_{pt} and G-S (and G_f) for the subjects with high-tone sensorineural losses (types 1 and 2). The median gap for the 20 subjects is 7.2 dB. This gap is mainly attributable to the higher class-D losses (inner ear distortion) of

these subjects, as shown before in Figs.4 and 5, and to a lesser extent to the distortion term S of the hearing aids. The gap obviously demonstrates the experience of numerous sensorineural-impaired subjects that they can hear what is said, but do not understand it.

Concerning harmonic distortions, all hearing aids tested are satisfactory amplifiers at moderate input sound-pressure levels, apart from the occasionally large L_I values (90th percentile of 28 dBA). However, at higher input levels (> 65 dB SPL) second-harmonic and third-harmonic distortions are readily produced, particularly if a high electroacoustic gain G_e has been adjusted in combination with a progressive output limitation for increasing input levels. A broad power bandwidth ($\Delta f > 5000$ Hz) also causes extra distortions. Furthermore, it is somewhat disappointing that AGC limiting systems produce only slightly less second-harmonic distortions, and no less third-harmonic distortions in comparison with PC output control systems. Third-harmonic distortions are detrimental to speech intelligibility in noise, as shown in Fig.7. Linear regression analysis revealed that every dB of extra third-harmonic distortion adds about 0.16 dB to the psychophysical distortion S. The relation between second-harmonic distortions and S is not that clear. The only small gain in noise of 1 dB of AGC relative to PC systems is possibly due to the almost identical third-harmonic distortions.

An obvious conclusion from the foregoing discussion is that the amplification characteristics of today's hearing aids need further improvement so as to be more equal to the task set by the hearing impaired, especially in noisy environments. This need is also expressed in recent recommendations on larger bandwidths (Pascoe, 1975; Schwartz et al., 1979; Skinner, 1980). Although the extra benefit to be derived from further improving the mere amplification of hearing aids is fundamentally limited to a few decibels, the effort is throughout worthwile because of the great impact of every dB of gain in S/N ratio on speech intelligibility. A real break-through in the field of aided speech hearing is not to be expected, however, until more sophisticated signal processing in the aids will make it possible to increase the aided S/N ratio beyond the S/N ratio commonly prevailing in noisy listening environments.

V. CONCLUSIONS

In the previous sections several conclusions have been drawn which, for ease of survey, are summarized below:

(1) Plomp's model on aided hearing, as described by Eq.(4), gives an accurate

description of the aided SRT (=ASRT), as a function of ambient noise level, on the basis of only five parameters. Two of these parameters (hearing loss classes A and D, respectively) define the subject's SHL in quiet (combination of A and D) and in noise (D only); the remaining three parameters (gain G, distortion S, and internal noise level L_I) describe the hearing aid in a simple, but effective way. The model is valid irrespective of both type of hearing impairment and frequency response of the hearing aid fitted.

(2) In cases where the gap between the aid's L_I and ASRT in quiet is minimally 10 dB, the model-determined gain for speech intelligibility in quiet G-S is identical with the functional gain, which is defined as the difference between the actually measured unaided and aided SRT in quiet.

(3) For a substantial revalidation of hearing-impaired subjects in noisy listening environments improvement of the S/N ratio is a prerequisite. In this respect, the currently available hearing aids, which all provide only aselective amplification of both speech and noise, cannot offer benefit in noise. An average deterioration in S/N ratio of more than 1 dB was found for the hearing aids investigated; only 4% of the aids yielded some benefit in noise.

(4) For a substantial benefit to be derived from an aid in quiet listening environments by subjects with mild to moderate hearing losses, L_I of the aid should not exceed 15 dBA, which is almost the level at which normal-hearing listeners achieve 50% correct intelligibility of conversational sentences. Only 6% of the aids tested satisfied this criterion.

(5) The gain of an aid in quiet can only roughly be estimated from electroacoustic gain measurements in a test box, although for subjects with recruitment in the ear no prediction at all is possible. As a rule-of-thumb, 10 dB of electroacoustic gain yields about 7 dB of functional gain in SRT_0.

(6) The benefit of an aid in noise is adversely affected by third-harmonic distortions. Generally, every 6 dB of extra distortion (relative to the level of the fundamental) worsens the aided SRT in noise by 1 dB. Second-harmonic distortions do not show such a systematic detrimental effect.

(7) In terms of model-derived distortion S, the advantage of AGC output limitation over PC output control is only 1 dB. This means that the need for a less distorting output-control system is not yet sufficiently satisfied.

(8) Subjects with high-tone sensorineural losses, both with and without recruitment in the ear, derive more benefit from an aid with respect to pure-tone detection than to speech intelligibility (in quiet). This effect (7 dB in the present study) is mainly caused by the high class-D losses of these subjects, and also illustrates why they need such favourable S/N ratios when noise comes into the picture.

(9) If there is recruitment in an ear, a different pattern of (aided) speech intelligibility is to be expected in comparison with ears having the same overall pure-tone hearing losses without recruitment. Ears with recruitment show a relatively small SHL in quiet and a maximum of SHL in noise. Especially in cases where only a mild class-A loss has to be compensated for, the benefit of a hearing aid is marginal.

ACKNOWLEDGEMENTS

I wish to thank Dr. T.S.Kapteyn, Head of the Audiology Center of the Free University Hospital, Amsterdam, for his cooperation in finding the appropriate test subjects and for lending the B&K electroacoustic measuring equipment.

REFERENCES

Dirks, D.D. and Wilson, R.A. (1969). "Binaural Hearing of Speech for Aided and Unaided Conditions," J.Speech Hear.Res.12, 650-664.

Dixon, W.J. and Massey Jr., F.J., *Introduction to Statistical Analysis* (International Student Edition, McGraw-Hill Kogakusha, Ltd., 1969).

Duquesnoy, A.J. (1982). "The sentence intelligibility of aged persons in quiet and in noise - test of a model on speech hearing," submitted for publication to J.Acoust.Soc.Am. +)

Gengel, R.W. (1971). "Acceptable Speech-to-Noise Ratios for Aided Speech Discrimination by the Hearing-Impaired," J.Audit.Res.XI, 219-222.

Harford, E.R. and Fox, J. (1978). "The Use of High-Pass Amplification for Broad-Frequency Sensorineural Hearing Loss," Audiology 17, 10-26.

Heusden, E. van, Plomp, R., and Pols, L.C.W. (1979). "Effect of Ambient Noise on the Vocal Output and the Preferred Listening Level of Conversational Speech," Appl. Acoustics 12, 31-43.

Miller, J.D., Niemoeller, A.F., Pascoe, D, and Skinner, M.W., Integration of the electroacoustic Description of Hearing Aids with the Audiologic Description of Clients (Chapter 16, pp.355-377, from *Acoustical Factors Affecting Hearing Aid Performance*, edited by G.A. Studebaker and I. Hochberg, University Park Press, Baltimore, 1980).

Nielsen, T.E. (1976). "Hearing Aid Characteristics and Fitting Techniques for Improving Speech Intelligibility in Noise," Br. J. Audiol. 10, 1-7.

+) Reproduced in Chapter III of this thesis.

Pascoe, D.P. (1975). "Frequency Responses of Hearing Aids and Their Effects on the Speech Perception of Hearing-Impaired Subjects," Ann. Otol. Rhinol. and Laryngol.84, Suppl.23, 1-40.

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," Audiol.18, 43-52.

Schwartz, D.M., Surr, R.K., Montgomery, A.A., Prosek, R.A., and Walden, B.E. (1979). "Performance of High Frequency Impaired Listeners with Conventional and Extended High Frequency Amplification," Audiology 18, 157-174.

Skinner, M.W. (1980). "Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation,"J.Acoust.Soc.Am..67, 306-317.

Tonning, F.-M.,(1973). "Directional Audiometry VII. The Influence of Azimuth on the Perception of Speech in Aided and Unaided Patients with Binaural Hearing Loss," Acta Otolaryngol.75, 425-431.

Walden, B.E., Schuchman, G.I., and Sedge, R.K. (1977). "The Reliability and Validity of the Comfort Level Method of Setting Hearing Aid Gain," J. Speech Hear. Dis.XLII, 455-461.

FINAL DISCUSSION

A. The relevance of STI for the hearing impaired

In Chapter I the validity of the STI concept has been demonstrated for groups of elderly hearing-impaired subjects; only their SRT in noise, with the speech presented without reverberation, needs to be measured to make an accurate prediction of the SRT in a reverberant environment. Furthermore, it has been shown (see Fig.4, Chapter I) that, in general, the predictive power of STI is sufficiently effective also in individual cases. Strictly speaking, the validity of STI in the case of hearing-impaired persons has been tested only in the diffuse sound field of both primary speaker and interfering source. In small to medium-sized rooms this field is found at distances between listener and sound source beyond approximately 2 m. There is no reason to doubt that the STI is equally valid for any listening condition, including direct sound fields, covered by Eq.(2) given in Chapter II. For example, Fig.1 of Chapter II, which is based on Eq.(2), can indicate how closely positioned towards the primary speaker a listener must be to understand a conversation. We know from Chapter I that a normal-hearing listener needs an STI of more than 0.35. Then, if $S/N = 0$ dB and $T < 1$ s, the listener can follow the conversation at any position in the room. On the other hand, a hearing-impaired listener who needs an STI of e.g. 0.50 ($SHL_D = 4.5$ dB) should approach the speaker to $1.75r_c$ or less (only about 1 m in medium-sized rooms).

The STI model has been designed primarily to describe listening situations in rooms where the interfering sounds can be regarded as a continuous noise signal. In view of the speech-hearing results presented in Chapter IV (Table II and Fig.3) for normal-hearing listeners tested under conditions with different types of interference, we must be prudent in plainly applying STI diagrams in the case of fluctuating fields of interfering sounds. Clearly, a complicating factor is that not only the acoustic S/N ratio at the listener's position is decisive for speech intelligibility, but also the sound character (spectral and temporal features) of the interfering source itself.

A practical situation where the above becomes manifest is the cocktail party. In this situation, with a large group of people in a limited area, the listener is positioned both in the diffuse field of many talkers at a distance, and in the (mainly) direct field of competing speakers closer to the listener. The diffuse field can be regarded as continuous, but the direct interfering field is definitely fluctuating. By means of Eq.(2) of Chapter II, an STI value pertinent to this listening condition can be calculated, where the disturbing effect of the interfering speakers in the direct field must be taken into account by the (constant) long-term average intensity of this interference. Since for normal-hearing listeners in the direct field a listening condition with continuous interfering noise is less favourable than a situation where the interfering sound has the same long-term average intensity, but is fluctuating in time (see Chapter IV, Fig.3, maximum difference in S/N ratio 7 dB), this implies that for normal-hearing listeners the calculated STI value represents a less favourable listening condition than actually experienced by them. Because of the random orientation of the competing speakers in the direct field and the considerable contribution of the diffuse noise field, the difference between the SRT of the listener predicted on the basis of the STI value and the actually occurring SRT will be substantially lower than the value of 7 dB, in terms of S/N ratio, found in Chapter IV for conditions without reverberation and with one competing speaker only.

In addition, the results presented in Chapter IV on elderly listeners with perceptive high-tone hearing losses will be considered (see also Table II and Fig.3 of Chapter IV). The SRT's of the elderly are determined not so much by the type of interference as by the acoustic S/N ratio alone. Furthermore, the gain obtained by moving the interfering source to a lateral position is also small in comparison with the gain for normal-hearing listeners. This means that the STI pertinent to cocktail-party conditions may be considered to be representative of the listening situation as it is experienced by hearing-impaired listeners with high-frequency losses.

It is clear that the STI model is a powerful tool for characterizing a broad scala of practical listening situations for both normal-hearing and hearing-impaired listeners. A decisive advantage of the STI model is the functional link it provides between statistical parameters of room acoustics (r/r_c, T, V, S/N ratio) and hearing impairment. By comparing the STI value measured at the position of the listener in a room with the STI value required by hearing-impaired persons (in general: $STI > 0.60$), an assessment of their handicap in the room can be made; and, what is very relevant, the STI model allows a quantitative treatment of how to adapt the acoustics of the room to the enhanced acoustical requirements of the hearing impaired.

B. Prospects for improving hearing aids

The validity of Plomp's hearing model implies that we are able to characterize effectively the unaided and aided speech-hearing ability of the hearing impaired as a function of background noise level by just five parameters, viz. A, D, G, S, and L_I (see Chapter V, Eqs.(1), (3), and (4)). The parameters A and D, pertinent to unaided hearing and, therefore, characteristic of a subject's hearing impairment, can be determined on the basis of two measurements of SRT (for sentences) taking about 8 min. This also holds for the parameters G and S, specific for the hearing aid tested. The internal noise L_I of the aid can be measured electroacoustically, but to save time it is throughout reasonable to take L_I equal to 22 dBA (median value, cf. Chapter V). As a result, it is possible to achieve a satisfactory insight into a subject's unaided and aided speech hearing, representative of his everyday-communication ability, within 20 min of measuring time.

From the results on aided hearing, presented in Chapter V, it is clear that the currently available hearing aids fail to meet the requirements of the hearing impaired in noisy circumstances. A prime requisite for a hearing aid to provide substantial benefit in noise is that it increases the S/N ratio by at least 5 dB. This means that the performance of present-day hearing aids, which have an average S value of 1 dB (cf. Chapter V), has to be improved by 6 dB or more, in terms of S/N ratio.

Can recent developments in hearing aid design meet this requirement? It is unlikely that the present topic of individual fitting of hearing-aid frequency responses, as advocated by Skinner (1980) and Miller et al. (1980), will provide an SRT improvement in noise of more than 3 dB. Another attempt, by Villchur (1973), is to improve the SRT in noise by dynamic range compression of speech. Villchur presumed that loudness recruitment in the ear is the main origin of hearing problems in noise. Although subjects with recruitment in the ear show high type-D losses (cf. Chapter V, Table V, group 2), there are other subjects having similar pure-tone and type-D losses without recruitment in the ear (cf. Table V, group 1). This does not corroborate Villchur's assumption on the origin of a hearing handicap. Apart from this, to date there has not been a promising follow-up on dynamic range compression (see Lippmann et al., 1981).

So, what are the prospects, for the near future, of mitigating a subject's hearing handicap by a hearing aid more substantially than is possible now? Evidently, the short-term possibilities for this are quite limited.

The best realisable points, probably contributing to a (slightly) better performance of the aids, are:

(1) A further reduction of harmonic and intermodulation distortion of the amplifier section.
(2) In aids with AGC (automatic gain control), a reduction of overshoots during the attack and release time.
(3) Extended high frequency amplification as far as tolerated by the user (annoyance should be avoided).
(4) Individual fitting of the aid's frequency response to a subject's hearing by means of a programmable hearing aid.
(5) Improvement of the front-random ratio (directivity) of the directional microphones.
(6) The provision of aids with very low-noise microphones to improve the SRT under quiet listening conditions (cf. Killion, 1976, subminiature microphone XD-985).

It seems unlikely that even a combination of the above-mentioned points of hearing-aid improvement will provide a gain in SRT in noise of more than 5 dB, but every dB of gain in SRT in noise is very valuable for the hearing impaired, as demonstrated in this thesis.

In the future, significant progress in hearing aid performance must be searched for in another direction. In my opinion, the most promising approach is to develop some kind of noise-suppression technique, independent of the type of hearing loss to be compensated for. Many attempts have already been made to increase the S/N ratio in that way, but a real bottleneck in this respect is the continuously changing spatial position and signal structure (in the time and frequency domain) of both the speaker (who must be understood) and the interfering source (the level of which must be sufficiently reduced). Because of this type of signal conditions many, elsewhere successfully applied, correlation techniques for noise suppression become ineffective and useless. Anyhow, it will remain a challenge for engineers in the field of (acoustic) signal processing to search for new and powerful methods of noise suppression.

REFERENCES

For the references made in the Chapters I to V, the reader is referred to the reference list of the chapter in question.

Beasley, W.C. (1940). "Characteristics and Distribution of Impaired Hearing in the Population of the United States," J.Acoust.Soc.Am.12, 114-121.

Carhart, R., Tillman, T.W., and Greetis, E.S. (1969). "Perceptual Masking in Multiple Sound Backgrounds," J.Acoust.Soc.Am.45, 694-703.

Carstairs, V. (1973). "Utilisation of Hearing Aids Issued by the National Health Service," Br.J.Audiol.7, 72-76.

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.

Killion, M.C. (1976). "Noise of ears and microphones," J.Acoust.Soc.Am.59, 424-433.

Lippmann, R.P., Braida, L.D., and Durlach, N.I. (1981). "Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss," J.Acoust.Soc.Am.69, 524-534.

Miller, J.D., Niemoeller, A.F., Pascoe, D., and Skinner, M.W., Integration of the Electroacoustic Description of Hearing Aids with the Audiologic Description of Clients (Chapter 16, pp 355-377, in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G.A. Studebaker and I. Hochberg, University Park Press, Baltimore, 1980).

Nielsen, T.E. (1976). "Hearing Aid Characteristics and Fitting Techniques for Improving Speech Intelligibility in Noise," Br.J.Audiol.10, 1-7.

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). "Speech-reception threshold for sentences as a function of age and noise level," J.Acoust.Soc.Am.66, 1333-1342.

Skinner, M.W. (1980). "Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation," J.Acoust.Soc.Am.67, 306-317.

Tonning, F.-M. (1971). "Directional Audiometry II. The Influence of Azimuth on the Perception of Speech," Acta Otolaryngol.72, 352-357.

Villchur, E. (1973). "Signal processing to improve speech intelligibility in perceptive deafness," J.Acoust.Soc.Am.53, 1646-1657.

SPEECH-LEVEL MEASUREMENTS FOR SENTENCES USING AN ARTIFICIAL EAR

Report 1977-AD01

Experimental Audiology
Faculty of Medicine, Free University
Amsterdam, The Netherlands

1. Introduction

Future research in our department will be directed towards the auditory handicap of hearing-impaired persons with regard to speech understanding in everyday situations. As a measure of speech intelligibility the Speech-Reception Threshold (SRT), defined as the sound-pressure level at which 50% of the speech material is correctly understood, will be used. The speech will be presented monaurally through circumaural earphones, both in quiet and against a background of interfering noise.

According to the definition of SRT, the sound-pressure levels re 2.10^{-5} Pa, produced in the headphones, have to be known. For this purpose we carried out acoustical measurements on circumaural headphones by means of an artificial ear on which a flat-plate coupler was mounted. Generally, the real-ear response of circumaural headphones is larger than the flat-plate coupler response (cf. Shaw and Thiessen, 1962). In the present report two procedures for establishing these response differences (for one type of headphone) are dealt with. In both cases the real-ear responses were measured for ten normal-hearing subjects (7 male, 3 female, age 24-34). The sound materials applied consisted of noise signals and ten lists of everyday sentences.

2. Experimental configuration

The experiments were conducted in an anechoic room. In the longitudinal axis of the room a loudspeaker and an armchair fitted with a head-rest were situated. The distance between the loudspeaker and the entrance of the individual ear canals was made exactly 2 m. The center of the loudspeaker was placed 1 m above floor level. Two electret microphones were stuck just outside the entrance to the narrow part of the left and right ear canals of each subject. The microphones were connected to an amplifier behind the chair. The experimenter and the play-back equipment were outside the anechoic room. Communication took place by means of an intercom. Behind the arm-chair a videocamera was placed to check outside the room whether the subjects correctly carried out the instructions concerning head fixation and alternate replacement of headphones and one-sided ear defender. Both camera and microphone amplifier were direct-current supplied from outside the room to exclude transformer hum.

The playback block-diagram is shown in Fig.1. The speech material and

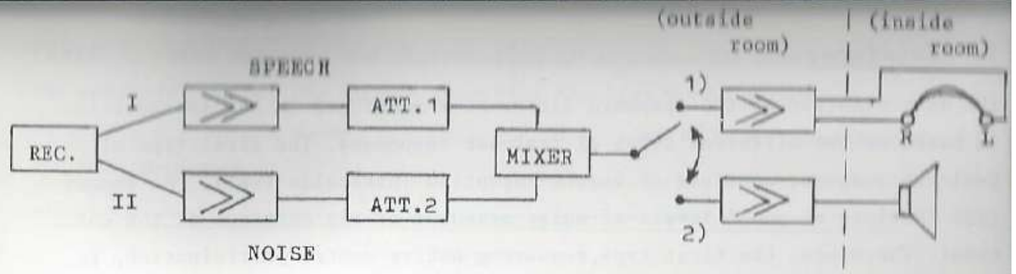


FIG.1. Block diagram showing the play-back equipment inside and outside the anechoic room. The symbols |>>| represent audio amplifiers, and ATT *stands for 'Attenuator'.* REC. *means Recorder.*

the noise were played back on channels I and II, respectively, of the tape recorder. After impedance transformation and separated attenuation to control the signal-to-noise (S/N) ratio, the two signals were mixed. The combined signal was presented to the subject through either the loudspeaker or the earphones (Sharp Scintrex MK IV).

The electret microphones were applied for measuring sound-level differences between various stationary noise signals at the entrance of the ear canal. After a fixed attenuation of 50 dB the microphone signals were electrically coupled to a sound-level meter (B&K Type 2205, provided with a BNC connector) indicating in dB(A). Various noise levels were measured also with the flat-plate coupler. The coupler consists of a rigid metal plate at the center of which a condensor microphone (B&K Type 4144, mounted on B&K artificial ear Type 4152) is positioned with its diaphragm lying in the plane of the plate. The right or left earphone was placed on the coupler in a central position with respect to the microphone and the sealing cushion was held in contact with the plate by a weight of approximately 500 g. The output of the artificial ear was measured with a sound-level meter (B&K Type 2203).

As speech material ten lists of 13 sentences were available (developed by Plomp and Mimpen, Institute for Perception TNO, Soesterberg) with, as far as possible, equal numbers of the various (Dutch) phonemes in each list. The chance of correct recognition of the sentences was equalized for all lists. In addition to the recording of these lists a standard-noise signal having the same level and long-term average sound spectrum as the 130 sentences was recorded both on the second track of the tape and on the sentence track, preceding the sentences, to enable easy calibration of the S/N ratio (for details see Plomp and Mimpen, 1979).

3. Procedures

The determination of the response differences mentioned in the Introduction is based on two different types of real-ear responses. The first type of real-ear response consists of Speech-Reception Thresholds (SRT), the second type consists of sound levels of noise measured at the entrance of the ear canal. Therefore, the first type,requiring active mental participation, is a subjective measure, whereas the second type,which merely utilizes the outer ear of the subject,is objective.

The following adaptive procedure for measuring SRT was applied. The first sentence of a list is repeatedly presented at a higher level until the listener can reproduce the sentence correctly. The level of the second sentence is decreased by 2 dB. If this sentence is correctly repeated, the level of the next sentence is decreased by 2 dB again; if it is not, the level is increased by 2 dB. All remaining sentences are presented in this manner. The average presentation level of   sentences 5 to 14 (the last one is not actually presented, but its level is known from the response to sentence 13) is assumed to represent SRT. Sentences 1 to 4 are for training purposes.

The first type of real-ear response was measured by using all 10 lists. Eight lists were presented under four conditions in quiet, viz. 2 lists to the left ear and 2 lists to the right ear by means of a loudspeaker, and 2 lists to the left ear and 2 lists to the right ear through the earphones. By presenting 2 lists per condition insight can be obtained into the within-subject reliability of the SRT. The remaining lists were presented in noise with a level of 55 dB(A) under two conditions, viz. one list by means of a loudspeaker and one list through earphones, both presented to the left ear only. The SRT measured in noise gives an indication about the possible reproduction-quality difference between the loudspeaker and the earphone, because, if no difference exists, the required S/N ratio should be identical for the two transducers.

Lists 1 to 6 were invariably presented to the left ear, lists 7 to 10 to the right ear. The intelligibility of these split lists does not need to be equal because the responses will be averaged over the left and right ears. In order to compensate for learning processes and fatigue,the list sequence for successive subjects was rotated. In Table I the distribution of the lists and conditions per subject is summarized.

In order to know the SRT levels in dB(A) it suffices to measure the free-field sound level in dB(A) of the noise preceding the sentences, at

TABLE I. List sequence and distribution of experimental conditions over the subjects, where: the number specifies the list to be presented, Q = SRT in quiet, N = SRT in noise, L = left ear, R = right ear, l = loudspeaker, and e = earphones.

Subjects									
1	2	3	4	5	6	7	8	9	10
1-QLe	1-QLl	3-QLl	3-QLe	5-NLe	5-NLl	7-QRl	7-QRe	9-QRe	9-QRl
2-QLl	2-QLe	4-QLe	4-QLl	6-NLl	6-NLe	8-QRe	8-QRl	10-QRl	10-QRe
3-QLl	3-QLe	5-NLe	5-NLl	7-QRl	7-QRe	9-QRe	9-QRl	1-QLl	1-QLe
4-QLe	4-QLl	6-NLl	6-NLe	8-QRe	8-QRl	10-QRl	10-QRe	2-QLe	2-QLl
5-NLe	5-NLl	7-QRl	7-QRe	9-QRe	9-QRl	1-QLl	1-QLe	3-QLe	3-QLl
6-NLl	6-NLe	8-QRe	8-QRl	10-QRl	10-QRe	2-QLe	2-QLl	4-QLl	4-QLe
7-QRl	7-QRe	9-QRe	9-QRl	1-QLl	1-QLe	3-QLe	3-QLl	5-NLl	5-NLe
8-QRe	8-QRl	10-QRl	10-QRe	2-QLe	2-QLl	4-QLl	4-QLe	6-NLe	6-NLl
9-QRe	9-QRl	1-QLl	1-QLe	3-QLe	3-QLl	5-NLl	5-NLe	7-QRe	7-QRl
10-QRl	10-QRe	2-QLe	2-QLl	4-QLl	4-QLe	6-NLe	6-NLl	8-QRl	8-QRe

the position of the listener's ear. On the assumption that the individual SRT's found in quiet and in noise are independent of the type of transducer used to reproduce the sentences, earphone speech-reproduction can be calibrated against free-field speech reproduction. In the case of supra-aural earphones the above assumption is only valid for frequencies below 1500 Hz (Villchur, 1969).

The second type of real-ear response was obtained with the electret microphones to which the standard noise from the speech track of the tape was presented. The responses at the two ears were measured for three different noise levels in the free-field and earphone situation. As a result, earphone noise reproduction can be calibrated against free-field noise reproduction.

In addition to the real-ear responses the flat-plate coupler responses were measured for the same three noise levels. The differences between the two types of real-ear responses and the coupler responses are used for calibrating the artificial ear plus coupler against free-field noise reproduction.

4. Theoretical aspects on calibrations

The above-mentioned calibrations relative to free-field conditions yield output-difference terms C_1, C_2, and C_3 shown in the relation diagram of Fig.2. The noise levels in this figure were achieved as follows:

S_1 = free-field noise level at a distance of 2 m from the loudspeaker, after 50 dB of amplification of the input signal which was supplied to the loudspeaker to obtain a sound signal at SRT level;

S_e = free-field noise level at a distance of 2 m from the loudspeaker, the input of the loudspeaker corresponding to the 50-dB amplified input signal as supplied to the earphone at SRT level;

N_1 = free-field noise level measured with the electret microphones, given that a prefixed input is supplied to the loudspeaker;

N_e = noise level under the headphones measured with the electret microphones, given the same input as in the case of N_1 is supplied to the headphones;

N_1' = identical to S_1;

N_e' = noise level under the headphones measured with the artificial ear, given that the same input as in the case of S_1 is supplied to the headphones.

The output-difference terms can now be expressed as follows:

$$C_1 = S_e - S_1 \tag{1}$$

$$C_2 = N_e - N_1 \tag{2}$$

$$C_3 = N_e' - N_1' \tag{3}$$

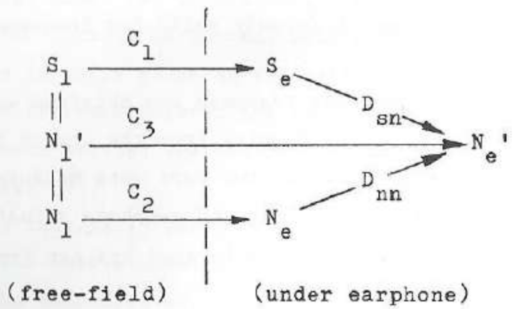


FIG.2. Relation diagram of the noise levels measured (S, N, and N'), where index l means loudspeaker condition and e means earphones. C_1, C_2, and C_3 are the output-difference terms given by Eqs.(1), (2), and (3). D_{sn} and D_{nn} are the response differences (or correction factors) given by Eqs.(5) and (6).

By varying the prefixed input the levels N_1 and N_e can be varied independently of S_1, S_e, and N_1'. Therefore, the current can be chosen so that the following holds:

$$N_1 = N_1' = S_1 \tag{4}$$

Then, the response differences between the coupler responses and the two types of real-ear responses can be calculated by using the expressions (1) to (4), according to

$$D_{sn} = C_3 - C_1 = N_e' - S_e \tag{5}$$

$$D_{nn} = C_3 - C_2 = N_e' - N_e \tag{6}$$

As can be seen from Eqs.(5) and (6), the type of loudspeaker, headphones and electret microphones applied cannot systematically influence D_{sn} and D_{nn} as far as overall energy-output differences of the devices are concerned. Therefore, in theory it should be expected that S_e equals N_e, resulting in D_{sn} being equal to D_{nn}. However, speech intelligibility is not only determined by the overall energy output of a transducer, but also by the shape of its frequency response, whereas electret microphones having an almost flat frequency response are only sensitive to the overall output. In practice, the frequency response of loudspeaker and headphone at the entrance of the ear canal will differ to some extent and, therefore, it can be expected that D_{sn} does not equal D_{nn}.

5. Results

The first type of real-ear response consists of SRT's measured according to Table I. The SRT in quiet was obtained under four conditions; each one was tested twice. The resulting standard deviation of individual SRT's in quiet is 1.3 dB. This number is defined as the root mean square divided by $\sqrt{2}$ of the 40 differences between the two SRT values per subject under each condition. By excluding conditions in which headphones were removed between test and retest, the standard deviation decreases to 1.1 dB. Replacement of the headphones appears to increase the standard deviation to 1.6 dB. The averaging of the free-field SRT's in quiet over all 20 ears resulted in an average value of 15.2 $\pm$ 0.5 dB(A). The output-difference term C_1 found by combining the free-field and headphone SRT's is 10.8 $\pm$ 0.5 dB.

The 20 SRT's measured in noise in order to obtain an indication about the quality difference between the loudspeaker and headphones have, expressed in terms of S/N ratio, an average value of -6.8 dB for the loudspeaker condition, and an average value of -6.2 dB for the headphones. As each condition was tested only once, the reliability of the results cannot be determined. However, in a similar experiment Plomp and Mimpen (1979) found a standard deviation of 0.9 dB. In view of this number the quality difference of 0.6 dB found in the present case is not significant.

In Fig.3 it is shown how the standard deviation of SRT values in quiet depends upon the number of sentences included in calculating SRT. As can be seen, presentation of more than eight sentences hardly increases the reliability, which indicates that a list of 13 sentences, including three training sentences, is sufficient for determining SRT accurately.

Figure 4 shows how the chance of correctly understanding the individual sentences in quiet depends on the presentation level relative to the SRT level (50%-correct score). The curve was obtained as follows. To each subject eight lists were presented in quiet under four conditions. For each list the number of sentences (with omission of the first sentence) correctly repeated after presentation at levels of 1 dB, 2 dB, etc. below or above the average level (rounded off to whole dB's) was determined. The numbers per relative level were almost identical for each condition and, therefore, the results of all 80 lists have been incorporated in the curve shown. Over the middle range of the diagram the intelligibility increment is approximately 15% per dB sound level increase.

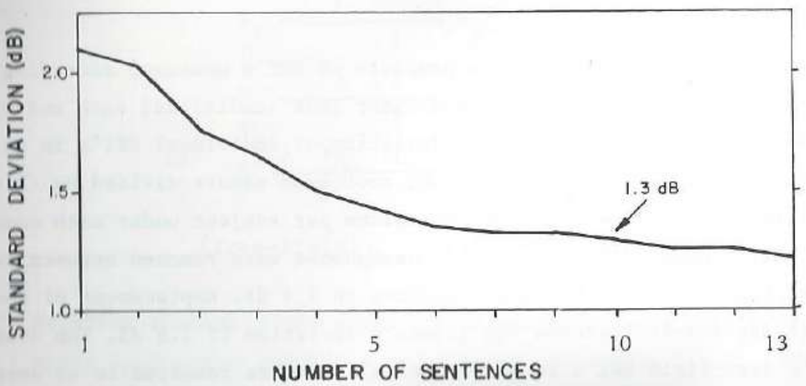


FIG.3. Standard deviation of SRT in quiet as a function of the number of sentences used for determining SRT.

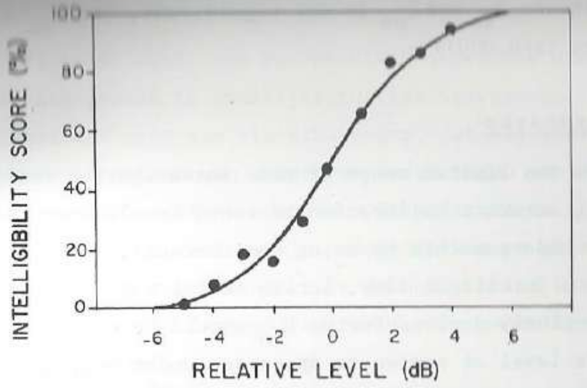


FIG.4. Intelligibility score as a function of presentation level relative to SRT (level of 50%-correct score), averaged over the results from 80 sentence lists presented in quiet (2 lists × 4 conditions × 10 subjects). The curve is the cumulative probability-distribution curve fitted visually to the data points.

The second type of real-ear response consists of noise level measurements using electret microphones. The noise levels were also measured with a sound-level meter under free-field conditions at a distance of 2 m from the loudspeaker. The levels thus resulting were 35, 50, and 65 dB(A), respectively. The lowest noise level may be partially affected by ambient noise ($\leq$15 dBA) and instrument noise ($\leq$17 dBA). The between-subject standard deviation of the electret measurements is 1.5 dB for the free-field condition and 2.1 dB for the headphones condition. The placement of the headphones over the auricle appeared to be rather critical with respect to the electret output, thus causing a larger spread in the results. The output-difference term C_2 found by combining the free-field and headphones electret levels is 9.2 $\pm$ 1.1 dB.

In addition to the real-ear responses the flat-plate coupler responses for the above three noise levels yielded an output-difference term C_3 of 8.8 $\pm$ 0.2 dB.

Based on the above-presented values of C_1, C_2, and C_3 the response differences D_{sn} and D_{nn}, calculated according to Eqs.(5) and (6), become:

$$D_{sn} = -2.1 \pm 0.5 \text{ dB}$$

$$D_{nn} = -0.4 \pm 1.1 \text{ dB} .$$

The difference between D_{sn} and D_{nn} is not significant, but D_{sn} differs significantly from zero ($p \leq 0.0001$).

6. Concluding remarks

With due regard to the limited scope of this investigation (only one speech spectrum applied), accurate calibration of sound levels measured with an artificial ear is made possible by using the correction factors D_{sn} and D_{nn}. Concerning sentence intelligibility, active mental participation is involved, so that the subjectively derived factor D_{sn} should be used. This means that the sound-pressure level of sentences presented under headphones will be known (in terms of free-field levels in dB re 2.10^{-5} Pa) by measuring the level of the equivalent noise, preceding these sentences, with the flat-plate coupler, and by adding afterwards 2.6 dB to this level. If, on the other hand, signals not specifically related to speech perception are to be measured, it is recommended that the objectively derived correction factor D_{nn} be used.

Some secondary results from this experiment based on 10 young normal-hearing subjects are:

(1) the monaural SRT in quiet is 15.2 ± 0.5 dB(A) (free-field condition);
(2) the monaural SRT in noise, expressed in S/N ratio, is -6.5 dB;
(3) the intelligibility-score increment near SRT is about 15 %/dB.

References

Plomp, R. and Mimpen, A.M. (1979). "Improving the Reliability of Testing the Speech-Reception Threshold for Sentences," Audiology 18, 43-52.

Shaw, E.A.G. and Thiessen, G.J. (1962). "Acoustics of Circumaural Earphones," J.Acoust.Soc.Am.34, 1233-1246.

Villchur, E. (1969). "Free-field calibration of earphones," J.Acoust.Soc.Am. 46, 1527-1534.

SAMENVATTING

In dit proefschrift wordt een systematisch onderzoek naar het spraakverstaan van slechthorenden in praktijksituaties beschreven. Uitgangspunt vormde de vaak geuite klacht van slechthorenden, dat zij vooral moeilijkheden hebben met het verstaan van gesproken woord tegen een achtergrond van andere geluiden, zoals geroezemoes, een interfererende stem, verkeerslawaai, etc. In ruimten met veel nagalm, zoals kerken en allerlei openbare gebouwen, ondervinden zij extra hinder. In deze situaties blijkt een hoortoestel weinig profijt op te leveren, waardoor vele gebruikers teleurgesteld zijn over hun toestel.

Een groot deel van het onderzoek was gericht op het testen van twee psychoakoestische modellen. Beide modellen benaderen het begrip spraakverstaan in termen van signaal-ruisverhoudingen. Het ene model, de z.g. Spraak Transmissie Index (STI) van Houtgast en Steeneken (1973), maakt het mogelijk op quantitatieve wijze het effect van nagalm op de spraakverstaanvaardigheid van slechthorenden te vervangen door het gelijkwaardige effect van een additionele ruiscomponent zónder nagalm. Het andere model, ontwikkeld door Plomp (1978), beschrijft de drempel voor het verstaan van spraak als functie van het niveau van storende achtergrondruis (zónder nagalm), met inachtneming van het effect van zowel slechthorendheid als hoortoestel. De drempel voor het spraakverstaan is gedefinieerd als het geluiddrukniveau (in dBA), waarbij 50% van een serie alledaagse korte zinnen correct wordt verstaan.

De modellen werden grotendeels getest aan de hand van drempels verkregen voor het beste oor bij 110 bejaarden (tests inclusief nagalm effecten). De test (zonder nagalm) naar het effect van een hoortoestel geschiedde met behulp van 50 slechthorende toestelgebruikers jonger dan 65 jaar. Bij alle tests werd als storende achtergrond een ruis ten gehore gebracht met hetzelfde geluidsspectrum als dat van de korte zinnen. Onder identieke testomstandigheden werden bij 30 jonge normaalhorenden eveneens de drempels gemeten ter verkrijging van referentiewaarden voor spraakverstaan. Het blijkt, dat de combinatie van beide modellen een solide basis vormt voor quantitatief onderzoek naar slechthorendheid in praktijksituaties.

Ter uitbreiding van bovenstaande beperkte luisterconditie werd bij 20 bejaarden onderzocht in hoeverre de richting en het karakter van de stoorbron de grootte van het gehoorverlies mede bepalen. De stoorbron werd

links of rechts van de luisteraar onder een horizontale hoek van 90° gesitueerd en bestond uit òf een storende spreker òf een ruisbron met een geluidsspectrum gelijk aan dat van deze spreker. De korte, primaire zinnen werden recht vóór de luisteraar ten gehore gebracht. In dit experiment luisterden de proefpersonen met beide oren. Tien jonge normaalhorenden, die aan dezelfde test werden onderworpen, dienden als referentie.

Naast de drempel voor spraak werd bij alle proefpersonen eveneens de drempel voor zuivere tonen in stilte (toonaudiogram) gemeten zonder, en voorzover van toepassing, mêt hoortoestel. De 50 betrokken hoortoestellen werden bovendien electroakoestisch doorgemeten.

Hieronder zijn enige belangrijke resultaten uit het onderzoek naar slechthorendheid samengevat.

(1). Indien het gehoorverlies voor spraak tegen een achtergrond van ruis wordt uitgedrukt in termen van de verbetering van signaal-ruisverhouding (S/R), die voor de slechthorende nodig is om in dezelfde mate spraak te kunnen verstaan als normaalhorenden in dezelfde ruisconditie, kan dit gehoorverlies worden vertaald in een t.o.v. normaalhorenden verhoogde waarde voor de Spraak Transmissie Index (STI). Deze verhoogde STI-waarde voorspelt wat de invloed van een toename van de hoeveelheid nagalm op de spraakverstaanvaardigheid van de slechthorende zal zijn. Bijvoorbeeld, als het gehoorverlies in ruis 6 dB bedraagt, vereist compensatie hiervan (zónder nagalm) een 0,2 grotere STI-waarde. De resulterende verhoogde STI-waarde voorspelt dan, dat de maximale nagalmtijd, waarbij de slechthorende onder zeer storingvrije luisteromstandigheden nog juist de helft van het gesprokene kan volgen, korter dan één seconde behoort te zijn.

(2). De STI kan vergroot worden zowel door de S/R-verhouding te verbeteren als door de nagalmtijd te verkleinen. In de meeste praktijksituaties is het vrijwel niet mogelijk om op directe wijze de S/R-verhouding zodanig te verbeteren, dat slechthorenden er voldoende baat bij vinden. In veel gevallen vormt dan het reduceren van de nagalmtijd een effectieve oplossing voor dit probleem. In het geval, dat de luisteraar zich in het indirecte geluidsveld van de spreker bevindt (zoals in klaslokalen en gehoorzalen), behoort de nagalmtijd voor iedere dB te behalen winst in S/R-verhouding met 25% verkort te worden, ofwel de reductiefactor bedraagt 0,75 per dB (bijv. 3 dB winst vereist reductie tot $0{,}75^3=42\%$ van de oorspronkelijke nagalmtijd). Als de luisteraar zich in het directe geluidsveld van de spreker bevindt (zoals in restaurants, foyers, e.d.) bedraagt de

benodigde reductie 18% per dB (reductiefactor 0,82). Dat een weliswaar kostbare reductie van de nagalmtijd, zéker in ruimten waar vaak bejaarden vertoeven, zinvol is blijkt hieruit, dat het aantal personen met een bepaalde mate van gehoorhandicap met 30% vermindert per dB winst in S/R-verhouding. Bijvoorbeeld: stel men is gehandicapt indien het gehoorverlies in ruis groter is dan 5 dB; op 75-jarige leeftijd heeft 35% der betrokkenen een verlies in ruis van meer dan 5 dB, en 25% een verlies groter dan 6 dB; een winst van 1 dB in S/R-verhouding betekent derhalve, dat het aantal gehandicapten zakt van 35% naar 25% (30% reductie).

(3). Het hoormodel van Plomp kan ieder gehoorverlies voor spraak afdoende karakteriseren door middel van twee verliescomponenten: een verliestype A, dat de geluidsverzwakking in het oor aangeeft, en een type D, dat de geluidsvervorming in het oor vertegenwoordigt. Type A is vooral hinderlijk bij zachte spraak in een rustige omgeving, terwijl type D vooral een handicap vormt voor het verstaan van spraak in geroezemoes. Het is absoluut noodzakelijk om beide typen gehoorverlies te meten om een juiste indruk van iemands gehoorverlies te verkrijgen. Uit regressie-analyse blijkt, dat bij bejaarden de verhouding tussen type A verlies en type D verlies in het algemeen 6:1 bedraagt (correlatie coëfficiënt van 0,62 met N=102). Type D verliezen zijn ondanks hun geringe grootte ernstig, wat onder andere hieruit blijkt, dat elke dB van een dergelijk verlies de zinsverstaanvaardigheid in geroezemoes met ongeveer 18% vermindert.

(4). De experimenten, waarbij zowel de richting als het karakter van de stoorbron gevarieerd werden, wezen uit dat bejaarden in een aantal situaties ernstiger gehandicapt zijn dan het type D verlies aangeeft. Met name als de stoorbron een interfererende spreker is (in plaats van een diffuse ruis met dezelfde intensiteit), kunnen de bejaarden niet langer profiteren van de relatief stille intervallen tussen de woorden, terwijl jonge normaalhorenden hierbij een winst van ongeveer 7 dB behalen. Ook het verplaatsen van de stoorbron van recht vóór naar opzij van de luisteraar levert bij bejaarden 4 dB minder winst op dan bij normaalhorenden.

(5). Het hoormodel van Plomp beschrijft eveneens het effect van een hoortoestel op het verstaan van spraak met behulp van slechts twee variabelen: de versterking G, die een type A verlies kan compenseren, en de toestelvervorming S, die de toename in de drempel voor het spraakverstaan in geroezemoes tengevolge van de vervorming in het hoortoestel aangeeft en derhalve aan het type D verlies kan worden toegevoegd. Deze benadering van een hoortoestel blijkt voor uiteenlopende typen slechthorendheid (hoge tonen -, gemengde -, en geleidingsverliezen) geldig te zijn, onafhankelijk van de versterkingskarakteristieken der toegepaste hoortoestellen. Bij de 50 onderzochte hoortoestellen werd een gemiddelde vervorming S van meer dan 1 dB geconstateerd. Dit betekent, dat slechthorenden in geroezemoes meestal geen baat van hun toestel zullen hebben, omdat dit toestel hun type D verlies in het geheel niet of in verwaarloosbare mate vermindert. De winst van een hoortoestel moet komen van een verkleining van het type A verlies. De trend om ook patiënten met een gering type A verlies en een aanzienlijk type D verlies (vaak personen met slechts matige verliezen voor zuivere tonen en met recruitment in het oor) een hoortoestel aan te passen, zal geen effect kunnen sorteren, gezien de beperkingen van de huidige toestellen ten aanzien van de benodigde S/R-verbetering.

STELLINGEN

1) Het storende effekt van nagalm op de spraakverstaanvaardigheid van slechthorenden kan via de Spraak Transmissie Index (STI) systematisch vervangen worden door een toename aan omgevingslawaai, die hetzelfde storende effekt heeft. (Dit proefschrift)

2) Voor een adequate beoordeling van de auditieve handicap van slechthorenden in praktijksituaties is de bepaling van twee 50%-verstaanbaarheidsdrempels voor korte alledaagse zinnen vereist: de drempel in stilte, en de drempel tegen een achtergrond van ruis met het geluidsspectrum van de spraak.

   (Dit proefschrift)

3) Het effekt van een hoortoestel op het spraakverstaan wordt bepaald door drie toestelgrootheden: 1) de interne ruis, 2) de akoestische versterking bij het trommelvlies, en 3) de "functionele" vervorming in geroezemoes (i.e. de door toestelvervorming verhoogde spraakverstaandrempel, in termen van signaal-ruisverhouding).

   (Dit proefschrift)

4) Indien gehoorklachten van patiënten hoofdzakelijk blijken te berusten op geluidsvervorming (in plaats van verzwakking) in hun gehoor, kan men deze klachten met de huidige hoortoestellen niet verhelpen.

   (Dit proefschrift)

5) Het verdient aanbeveling, dat audiologische poliklinieken, die over een efficiënte spraakverstaantest, representatief voor praktijksituaties, willen beschikken, de op het Instituut voor Zintuigfysiologie TNO (Soesterberg) ontwikkelde zinslijsten met achtergrondruis als testmateriaal gaan toepassen.

6) De door psychofysici veelvuldig toegepaste Two-Alternative Forced-Choice procedure voor het meten van perceptie-drempels is weinig geschikt voor de praktijk van de klinische audiologie.

7) Aangezien ook hedendaagse hoortoestellen geen winst in geroezemoes opleveren, blijft Carhart's "Comfort Level"-methode ter bepaling van de gewenste akoestische versterking van een hoortoestel van belang.

   (R.Carhart, Laryngoscope 56, 1946, 510-526)

8) De konklusie van Nabělek en Robinson omtrent een leeftijdsonafhankelijke winst bij binauraal horen in nagalm wordt door hen experimenteel niet overtuigend onderbouwd.

   (A.K.Nabělek en P.K.Robinson, J.Acoust.Soc.Am.71, 1982, 1242-1248)

9) De opmerking van Nabělek en Mason, dat zij via lineaire regressie (correlatie van 0.76, N=15) het effekt van nagalm op de spraakverstaanvaardigheid van een individu grofweg kunnen voorspellen uit het effekt van ruis, toont aan, dat zij niet begrepen hebben, dat de storende invloed van nagalm progressief toeneemt bij toenemend gehoorverlies.

   (A.K.Nabělek en D.Mason, J.Speech Hear.Res.24, 1981, 375-383)

10) De onderlinge vergelijkbaarheid van resultaten uit spraakverstaanvaardigheidsexperimenten zou sterk bevorderd worden indien minder onderzoekers genoegen zouden nemen met het meten van slechts één scoringspercentage op een arbitrair geluidsniveau.

11) Vanuit het oogpunt van spraakverstaanvaardigheid is de diagnostische waarde van de door Jerger en Speaks als belangwekkend alternatief voor woordlijsten gepresenteerde SSI (i.e. Synthetic Sentence Identification) test uiterst twijfelachtig.

(J.Jerger en Ch.Speaks, J.Speech Hear.Dis.33, 1968, 318-328)

12) Ter vermindering van de handicap van slechthorenden bij door beveiligingsglas afgeschermde loketten is het gewenst de geluidstransmissie van deze loketten te verbeteren en de loketbedienden te doordringen van het belang van gearticuleerd spreken.

13) De door Jansen gevolgde methode van electroëncephalogram (EEG)-analyse, waarbij de EEG's op voorhand onderverdeeld worden in quasi-stationaire segmenten van 1.28 sec., die vervolgens beschreven worden met een autoregressief model, waarvan de coëfficiënten met behulp van het Kalman filter worden geschat, gaat geheel voorbij aan de essentie van het Kalman filter: het adaptief kunnen schatten van in de tijd variërende parameters.

(B.H.Jansen, Academisch proefschrift, Vrije Univ.,1979, Amsterdam)

14) In de late Middeleeuwen was men letterlijk en in vele opzichten ook figuurlijk zéér bij de tijd. Wie deze eeuwen nog als een "duistere" periode beschouwt, moet nodig zijn licht gaan opsteken.

(J.Gimpel, The Medieval Machine, Penguin Books, 1976)

15) De betekenis van schijndissertaties, door de Vrije Universiteit te Amsterdam gedistribueerd bij een wetenschappelijke promotie, kan sterk worden verhoogd door voortaan het opnemen hierin van een samenvatting van het proefschrift verplicht te stellen.

16) De overheidspolitiek van gegarandeerd lage huurlasten en jaarlijkse nieuwbouwplanning dreigt volledig vast te lopen. Het is dringend gewenst om volkshuisvesting weer aan het particulier initiatief over te laten met geleidelijke afschaffing van allerlei subsidieregelingen en knellende bueaucratie, zoals Voorschriften en Wenken voor de Bouw en de Woonruimtebeschikking uit 1974.

17) Uit de benaming "Het Oor van Dionysius" voor een S-vormige grot met krachtige echo bij Siracusa op Sicilië, zou men kunnen afleiden dat de Grieken in de klassieke oudheid reeds op de hoogte waren van het verschijnsel recentelijk populair aangeduid als de Kemp-"echo".

(D.T.Kemp, J.Acoust.Soc.Am.64, 1978, 1386-1391)

8 september 1982 A.J.H.M. Duquesnoy