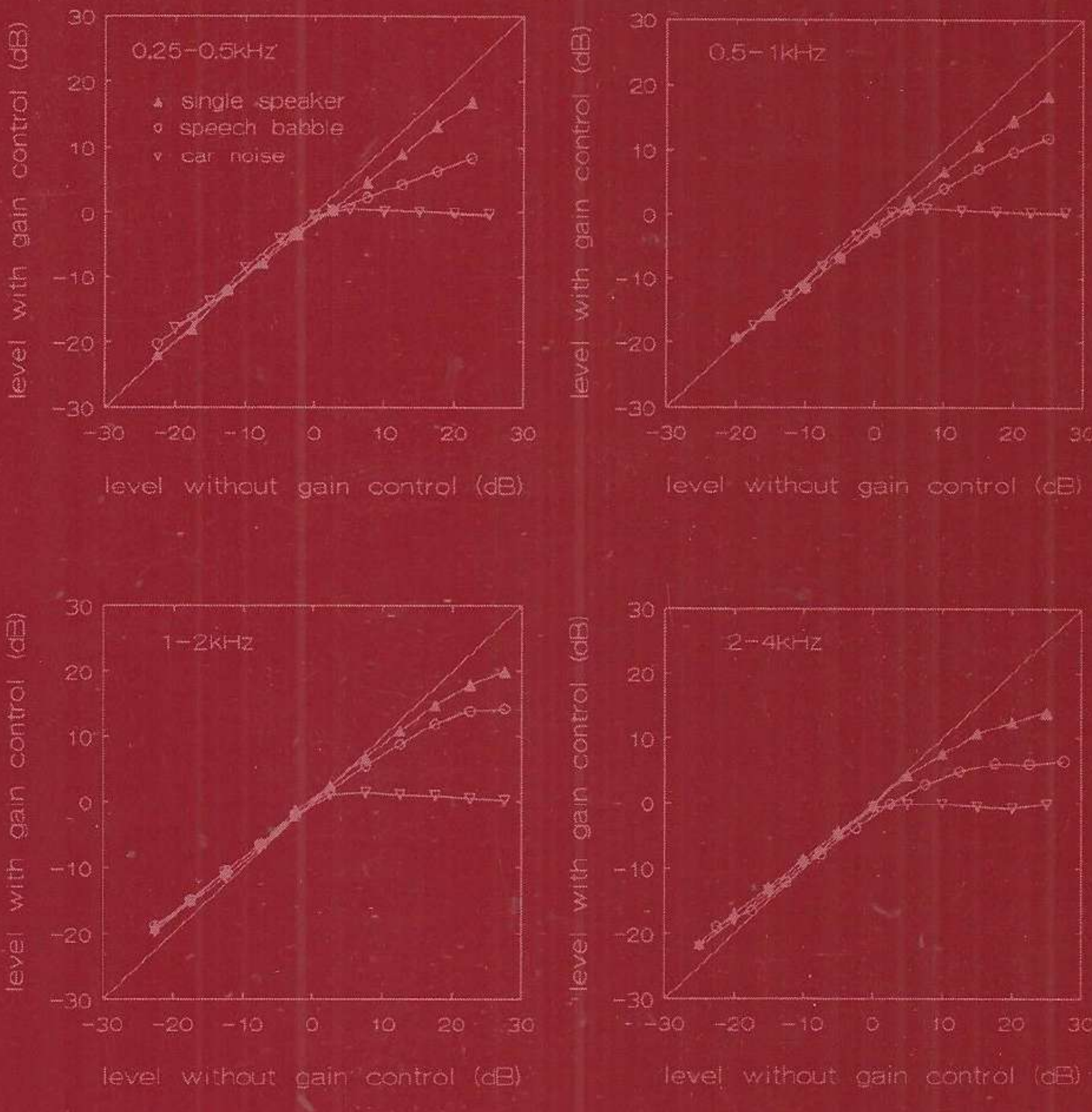
STUDIES ON THE EFFECTIVENESS OF MULTICHANNEL AUTOMATIC GAIN-CONTROL IN HEARING AIDS



Janette N. van Dijkhuizen

STUDIES ON THE EFFECTIVENESS OF

MULTICHANNEL AUTOMATIC GAIN-CONTROL

IN HEARING AIDS

This research was supported by the Innovatiegericht Onderzoek Programma Hulpmiddelen Gehandicapten (IOP-HG) and by Philips Export B.V.

VRIJE UNIVERSITEIT TE AMSTERDAM

STUDIES ON THE EFFECTIVENESS OF MULTICHANNEL AUTOMATIC GAIN-CONTROL IN HEARING AIDS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Vrije Universiteit te Amsterdam,
op gezag van de rector magnificus
dr. C. Datema,
hoogleraar aan de faculteit der letteren,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der geneeskunde
op vrijdag 12 april 1991 te 13.30 uur
in het hoofdgebouw van de universiteit, De Boelelaan 1105

door

JANETTE NICOLE VAN DIJKHUIZEN

geboren te Bussum

Bariet, Ruinen
1991

Promotor : prof.dr.ir. R. Plomp
Copromotor : dr.ir. J.M. Festen
Referent : prof.dr.ir. L.C.W. Pols

VOORWOORD

De in dit proefschrift beschreven experimenten werden uitgevoerd bij de groep Experimentele Audiologie van de vakgroep Keel-, Neus-, en Oorheelkunde van het Academisch Ziekenhuis van de Vrije Universiteit te Amsterdam.

Mijn promotor, prof. dr. ir. Reinier Plomp, dank ik voor zijn zeer inspirerende wetenschappelijke begeleiding. Dr. ir. Joost Festen dank ik voor de altijd energieke wijze waarop hij mij heeft bijgestaan in zowel wetenschappelijke als technische problemen. Verder ben ik hen en mijn andere collega's dankbaar voor een buitengewoon prettige werksfeer. Tenslotte wil ik mijn familie en vrienden bedanken voor de niet aflatende belangstelling voor mijn werk.

Janette van Dijkhuizen
april 1991

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

Many people with a sensorineural hearing impairment complain about having difficulties with understanding speech in the presence of interfering sounds. These difficulties generate the need of a more favorable speech-to-noise ratio compared to normal-hearing people. On the average, the required speech-to-noise ratio is 3 to 4 dB higher than that for normal-hearing listeners (Plomp, 1986); in the case of a fluctuating masker (e.g. a competing speaker) this difference may be even more than 10 dB (Festen and Plomp, 1990). Because in critical listening situations, every decibel in speech-to-noise ratio corresponds with a difference of 15-20% in the intelligibility score of simple sentences, it will be clear that these figures represent a substantial handicap in everyday situations.

The need for a better speech-to-noise ratio has been recognized as a major problem of sensorineural hearing impairment. This has led to a number of attempts in the field of hearing-aid research to improve the speech-to-interference ratio. These attempts have as yet not been successful in the most critical listening situations where the interfering sound has about the same level and spectrum as the desired speech signal, as in the common case of one or more competing speakers. Particularly for conditions where the interfering sound is a single voice, in which case the loss in speech perception relative to normal hearing is largest, the design of a hearing aid that separates the desired speech from the interfering speech appears to be an extremely difficult task.

For the time being, it seems that the hearing-aid user is served best when we concentrate on a more straightforward signal treatment aimed at presenting the total incoming signal of speech in noise as favorably as possible to the listener's ear, independently of signal level. This requires an automatic adjustment of the signal, as a function of time, to such a level that the fluctuations of speech, that carry the information, are audible over a frequency range that is as wide as possible, without exceeding the level of uncomfortable loudness.

This brings us to a frequently encountered complication in impaired hearing: the auditory threshold is elevated without being accompanied by a comparable elevation of the discomfort level. Particularly when the listener has a sloping audiogram, an optimal presentation of the speech signal within the reduced dynamic range is only possible if, for each individual frequency component, the amplification is controlled by the level of the incoming sound. This can be achieved by means of a *multichannel automatic gain control (AGC)* hearing aid with independent control of the amplification in each frequency channel. Because in practice, the spectrum of the acoustical background may vary, the amplitude-frequency response of such a hearing aid will vary along. With wideband rather than frequency-dependent gain control, intense narrowband noise would automatically reduce the amplification for all frequencies. This can easily result in inaudible speech components, even in frequency regions where there is no noise.

Conventional AGC acts on the average level of the sound, irrespective of whether a speech signal is present or not. This means that during periods when no speech communication takes place, background noise is amplified to levels that are experienced as "noisy" by the hearing-impaired listener. This is a major drawback of the application of AGC in hearing aids. This annoyance can be reduced by using the level of the *minima in the temporal envelope* of the sound, rather than the average level, as an AGC criterion. Whereas frequency channels with fluctuating signals such as speech can be amplified to an appropriate level, those channels with signals lacking fluctuations, because speech is absent or masked by noise, can now be suppressed.

A critical parameter of multichannel AGC in hearing aids is the time taken by the frequency-dependent amplification factor to adapt itself to changes of the incoming sound. No decisive data in favor of short attack and release times, as used in syllabic compression, have been reported in the literature (cf. Braida et al., 1979). A strong argument against short time constants in multichannel AGC is that they reduce the fast information-bearing fluctuations of speech within each frequency channel, and this will negatively affect speech reception (Plomp, 1988). *Time constants of roughly 0.5 s* may be a good compromise in the sense that fast intensity fluctuations within syllables are left unaffected whereas variations in the acoustical background, which are usually much slower, can be followed easily.

The aim of the present study was to investigate the merits of a multichannel AGC hearing aid in which, for each individual frequency channel, the rapid intensity fluctuations typical of speech are maximally preserved whereas the level of the more slowly varying background noise is suppressed.

As the major touchstone for hearing-aid effectiveness we adopted the speech-reception threshold (SRT) in noise, defined as the speech-to-noise ratio at which 50% of short conversational sentences are correctly reproduced (Plomp and Mimpen, 1979a). The SRT provides an accurate quantification of speech understanding abilities in critical listening conditions, which, due to its fixed performance criterion, can be easily compared from one experiment to another. In all experimental conditions, signals were presented well within the listener's auditory range, limited by the threshold of hearing and the level of uncomfortable loudness, at all frequencies. This was done to ensure that the threshold for speech reception in quiet would not play a role in the determination of the SRT in noise.

The application of multichannel AGC implies that the amplitude-frequency response depends on the spectrum of the incoming sound. It follows that spectral variations in the background noise will cause the speech spectrum to vary accordingly. We need to know the limits, in terms of shape and rate of change, within which the amplitude-frequency response can be varied before having a detrimental effect on the intelligibility of speech in noise. Therefore, in the *first part* of this study, presented in Chapters 2 and 3, we investigated, for 20 normal-hearing and 20 hearing-impaired subjects, respectively, the effect of varying the slope of the amplitude-frequency response on the SRT for sentences in noise. The noise had a spectrum identical to the long-term average spectrum of the sentences. This spectrum was adopted, first, because the most frequent and prominent interference in everyday listening situations is the human voice itself (e.g. second speaker, speech babble); second, because with this restriction, a change in signal-to-noise ratio has comparable effects for all frequency

regions, which allows accurate measurement through a steep discrimination curve. Results show that, for both subject groups, the SRT in noise is relatively insensitive to variations in the amplitude-frequency response, either when its slope is steady-state or slowly changing in time.

The only way in which frequency-dependent control of the amplification may be expected to weaken the masking effectiveness of the noise, is through a reduction of (upward) spread of masking caused by intense bandlimited interfering sounds. Therefore, in the *second part* of this study, presented in Chapter 4, we measured in conditions of intense low-frequency noise, the beneficial effect of frequency-selective attenuation on the SRT for sentences, for 12 normal and 12 impaired listeners. The noise had a spectrum identical to the long-term average spectrum of the sentences, except that in one octave band (0.25-0.5 or 0.5-1 kHz) its level was increased by 20 dB. Therefore, speech components in that band could not contribute to intelligibility. Results for both subject groups show that frequency-selective attenuation of speech and noise in the octave band containing the extra 20 dB of noise gives a substantial decrease in masked SRT, and is clearly more beneficial than wideband attenuation. These results were obtained both for an increased noise level that is steady-state and for an increased noise level that develops slowly in time.

The promising results reported in the first and second part of this study are prompting to the application of slow-acting frequency-dependent AGC in hearing aids. In these studies, however, the variation of the frequency-dependent amplification was still experimentally controlled. In the *third and final part* of this study, presented in Chapter 5, we studied the ultimate effectiveness of the concept multichannel AGC hearing aid by investigating, for 10 hearing-impaired subjects, the effect of frequency-dependent amplification actually controlled by the minima in the temporal sound envelope. A condition without gain control, but with the amplification in the different frequency bands adjusted to warrant 100% speech intelligibility in quiet, was the reference. Results show that, in the presence of sounds frequently interfering in everyday listening situations, with spectra that are roughly comparable to that of the speech signal, the multichannel AGC acting on the envelope minima does not affect the SRT in noise. However, it substantially reduces the sensation of noisiness when no speech communication takes place. This reduction of noisiness was most obvious for sounds with a more or less continuous character (e.g. stationary noise, music), where the AGC was most active.

Chapter 6 summarizes the experiments carried out in this study and restates the main conclusions. Chapters 2 to 5 are based on papers by van Dijkhuizen, Anema, and Plomp (1987) and by van Dijkhuizen, Festen, and Plomp (1989, 1991a, and 1991b), respectively.

CHAPTER 2

THE EFFECT OF VARYING THE SLOPE OF THE AMPLITUDE-FREQUENCY RESPONSE ON THE MASKED SPEECH-RECEPTION THRESHOLD OF SENTENCES

Janette N. van Dijkhuizen, Peter C. Anema, and Reinier Plomp
published in: Journal of the Acoustical Society of America 81, 465-469 (1987)

ABSTRACT

Within the framework of a study on the merits of a frequency-dependent automatic gain control in hearing aids, the effect of varying the slope of the amplitude-frequency response on the speech-reception threshold (SRT) for sentences in noise was studied for normal-hearing listeners. Speech and noise were both subjected to the same amplitude-frequency response. In the first experiment, the effect of a constant slope was investigated (20 listeners). Over a range of about -7 to +10 dB/oct, the SRT in noise remained constant. In the second experiment, a single change in the slope of the amplitude-frequency response was introduced halfway through the sentence. The effect of varying the transition time over a range down to 0.125 s appeared to be very small. In the third experiment, the slope varied continuously with range and variation frequency (0.25-2 Hz) as the parameters. The masked SRT increased gradually with variation frequency. The results indicate that the masked SRT for sentences is remarkably resistant to dynamic variations in the slope of the amplitude-frequency response.

The introduction of very-large-scale integrated circuits makes it technically possible to design multichannel hearing aids in which the frequency-dependent amplification factor adapts itself automatically to the physical parameters of the sound picked up by the microphone. As explained elsewhere (Plomp et al., 1986), such an automatic gain-control system has several advantages and should preferably take about 0.5 s to adapt itself to a new acoustical condition. For shorter durations, as used in syllabic compression, no conclusive results have been reported (cf. Braida et al., 1979). However, short attack and release times of compression will distort amplitude relations between and within speech sounds, and thus may affect the speech-reception threshold negatively. Longer times may cause discomfort to the user.

It follows that such a hearing aid has an amplitude-frequency response varying in time. Since both the wanted speech sound and the interfering sound (e.g. a second speaker, voice babble, or noise) are subjected to the same amplitude-frequency response, these variations leave the speech-to-noise ratio at the different frequencies untouched. Then, the first question that needs to be answered is: Within which limits of shape and rate may the amplitude-frequency response be varied before having a detrimental effect on the speech-reception threshold (SRT) in noise?

We studied this question by investigating the effect of varying the slope of the amplitude-frequency response on the SRT for meaningful sentences masked by noise with a spectrum identical to the long-term average spectrum of the sentences. Speech and noise were shaped by the same amplitude-frequency response. The present experiments were carried out with groups of normal-hearing listeners in order not to complicate the interpretation with the effects of hearing impairment. In experiment 1, the slope was constant during the presentation of the sentence; in experiment 2, a single change in the slope of the amplitude-frequency response was given; whereas, in experiment 3, the slope varied continuously. As far as we know, no experiments on similar conditions have been reported in the literature.

I. EXPERIMENT 1: CONSTANT SLOPE

In this experiment, the slope of the amplitude-frequency response in each condition was kept constant during presentation. According to the concept of the Articulation Index (French and Steinberg, 1947; Kryter, 1962a), speech intelligibility is a function of the "local" speech-to-noise ratios in a series of filter bands covering the speech frequencies from low to high. Consequently, this concept predicts that, as long as the limited frequency-resolving power of the auditory system does not play a role in the masking pattern of the noise, SRT in noise is independent of the slope of the amplitude-frequency response.

Many experiments on the effect of high-pass and low-pass filtering have been carried out in the past. Speech energy in our experiments was shaped by different slopes of the amplitude-frequency response but remained above hearing level at all frequencies. Because filtering removes part of the acoustical information in speech, experiments on the effect of filtering in noise cannot easily be related to our experiments. No data were gathered on the

effect of systematically varying the slope of the amplitude-frequency response on the SRT in noise.

A. Method

Ten different slopes were adjusted by means of a parallel set of 13 1/3-oct bandpass filters (Brüel & Kjaer spectrum shaper type 5612) with center frequencies of 250, 316, 400, 500, ..., 4000 Hz. These slopes were -15, -12, -9, -6, -3, 0, +3, +6, +9, +12 dB/oct. Extra RC circuits with slopes of -6 and +6 dB/oct were used to effectuate the extreme values of -15, -12, and +12 dB/oct.

The speech material consisted of 10 lists of 13 short (8 or 9 syllables) everyday Dutch sentences spoken by a female speaker (Plomp and Mimpen, 1979a). The masking noise had the same spectrum as the long-term average spectrum of the sentences. This spectrum was adopted because speech is the most frequently and prominently interfering signal in normal listening situations.

For each of the ten slope conditions, magnetic tapes were prepared with the sentences on one track and the noise on the other. Due to the limited dynamic range of the tape recorder (Tandberg Actilinear Studio Type TD 20) in combination with the negative slope of the speech spectrum, it was not possible to cover for the extreme slopes the very large differences in sound level over the entire frequency range. As a result, speech components for the -15 dB/oct and the -12 dB/oct conditions were masked beyond 1300 and 1700 Hz, respectively, and for the +12 dB/oct condition below 450 Hz.

Twenty normal-hearing listeners (11 male and 9 female, age 20-28) participated as paid volunteers. When the subject did not express preference, the right ear was chosen as the test ear. None of the listeners had a pure-tone hearing level in the test ear of more than 15 dB over the range 250-4000 Hz, nor did any listener have a history of ear pathology.

The subjects listened to the stimuli monaurally via a headphone in a soundproof room. The sentence lists were presented to the listeners in a fixed order, but with the ten conditions distributed differently over the lists according to a digram-balanced design per 10 listeners in order to avoid effects of learning and fatigue.

The level of the sentences was adjusted according to an up-down adaptive procedure (Plomp and Mimpen, 1979a). The masked SRT was defined as the speech-to-noise ratio in dB at which 50% of the sentences could be reproduced correctly. In order to use, for all conditions, signals that are near to 80 dB SPL, the signal-to-noise ratio was varied by either changing the noise level or the speech level, whichever was lower. Since the SRT expressed in speech-to-noise ratio is independent of absolute level (cf. Plomp and Mimpen, 1979b), the masked SRT is not influenced by this procedure.

B. Results and discussion

Values of the mean SRT in noise are given in Table I, with their standard deviations, and are plotted in Fig. 1. We see that SRT for the slopes from -6 to +9 dB/oct is rather stable. It was verified that, for these data points, the slope of the regression line does not

Table I. Mean SRT in noise, expressed in speech-to-noise ratio, and standard deviation, for the ten constant slopes of the amplitude-frequency response (20 listeners).

Slope (dB/oct)	Mean SRT in noise (dB)	Standard deviation (dB)
-15	30.3	21.7
-12	1.3	4.5
-9	-2.1	2.1
-6	-4.5	1.5
-3	-4.2	1.8
0	-3.3	1.3
+3	-5.3	1.7
+6	-4.1	1.6
+9	-5.5	1.7
+12	-2.4	1.6

differ significantly from zero.

The following factors can explain why the SRT in noise values for the steep slopes, particularly for the negative ones, are higher than for the middle range of slopes:

(1) *Interference of tape noise*, by which faint frequency components of the speech were masked. Because the SRT in noise for sentences that are low-pass filtered with a cutoff frequency of 1500 Hz is about 4 dB higher than for those with a cutoff frequency of 4000 Hz (Plomp, 1986), the interference of tape noise has to be considered to be partly responsible for the higher SRT in noise values for the slopes of -12 and -15 dB/oct.

(2) *Upward spread of masking*. For a low-frequency masker with an SPL of 80 dB, the slope of the masking pattern can be estimated to be -20 to -30 dB/oct (cf. Ehmer, 1959). Because the slope of the speech spectrum itself is -5 to -10 dB/oct, the substantial increase in masked SRT for the -15 dB/oct condition may be largely due to self-masking of the higher frequency components of the speech signal by the lower frequency components.

(3) *Unfamiliarity with the speech sounds*. Whereas the previous two factors refer to loss of information, we should not exclude the possibility that the listeners had difficulties in understanding the sentences in the steep-slope conditions because they sound so different from the everyday situation. Whether the scores can be improved by training or whether the extreme slopes, in particular -15 dB/oct, exceed the ear's flexibility in spectral adaptation, remains an open question.

We may conclude from this experiment that over a considerable range, from about -7 up to +10 dB/oct, the masked SRT of normal-hearing listeners does not systematically change with the slope of the amplitude-frequency response. This finding confirms the concept of the Articulation Index (French and Steinberg, 1947; Kryter, 1962a) that, as long as self-

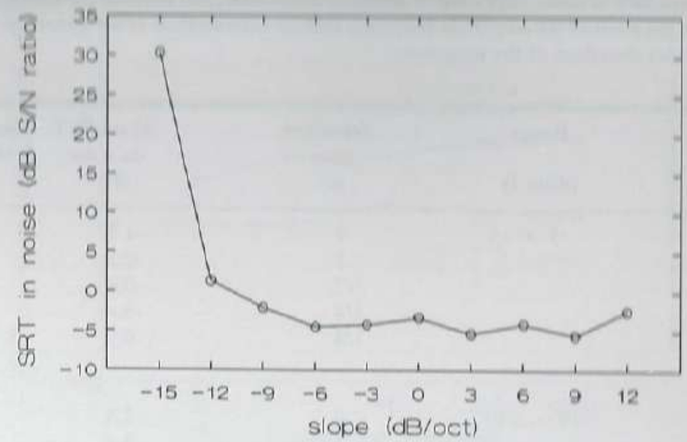


Fig. 1. Mean SRT in noise as a function of the slope of the amplitude-frequency response.

masking is excluded, the SRT in noise depends exclusively on the speech-to-noise ratio.

II. EXPERIMENT 2: SINGLE SLOPE TRANSITION

In this experiment, the slope of the amplitude-frequency response was varied once during the presentation of the sentence from a negative value to an equally large positive value or vice versa.

A. Method

The slope transitions were realised by means of four parallel octave bandpass filters of 250-500, 500-1000, 1000-2000, and 2000-4000 Hz, respectively, programmed digitally in a TMS 32010 signal processor (sampling rate 10 kHz). The amplification factor for each individual filter was controlled by a PDP-11/10 computer and updated at a 50-Hz rate.

The slope varied once during the sentence from -10 to +10 dB/oct or from -5 to +5 dB/oct, or vice versa. In the first condition, for example, the relative amplification factors of the four successive octave filters changed from +15, +5, -5, and -15 dB, respectively, to -15, -5, +5, and +15 dB, respectively. The transition function was according to half a cosine along a dB scale, accomplished in 2, 1, 1/2, 1/4, or 1/8 s, corresponding to 1/4, 1/2, 1, 2, and 4 Hz, respectively. This resulted in 20 conditions (two slopes, two directions, five rates).

The same sentences and noise as in experiment 1 were recorded on the two tracks of a tape recorder. Their levels could be attenuated independently before they were mixed, digitized, and fed into the TMS 32010 signal processor. Since the average duration of the

Table II. Mean SRT in noise, expressed in speech-to-noise ratio, and standard deviation, for single transitions in the slope of the amplitude-frequency response (five values of the transition time, two ranges, and two directions of the transition).

Direction	Range (dB/oct)	Transition time (s)	Mean SRT in noise (dB)	Standard deviation (dB)
Negative to positive (subgroup of ten listeners)	-5 → +5	2	-1.2	2.0
		1	-0.8	1.6
		1/2	0.4	1.8
		1/4	-0.4	1.5
		1/8	-0.2	2.0
	-10 → +10	2	2.8	5.9
		1	3.4	5.2
		1/2	4.9	6.4
		1/4	4.6	3.9
		1/8	5.4	5.7
Positive to negative (subgroup of ten listeners)	+5 → -5	2	-1.2	1.3
		1	-1.3	1.8
		1/2	-0.3	2.2
		1/4	-0.3	1.7
		1/8	-1.0	1.5
	+10 → -10	2	0.0	1.6
		1	1.9	2.8
		1/2	0.9	1.9
		1/4	2.6	2.6
		1/8	2.0	2.2

sentences was 2 s, the starting moment of the slope transition was adjusted so that the slope always passed the 0-dB/oct state exactly 1 s after the beginning of the sentence.

A new group of 20 normal-hearing listeners (10 male and 10 female, age 18-24) participated as paid volunteers. Screening and selection of this group were the same as in experiment 1. Since we had ten lists at our disposal, these listeners were divided in two subgroups, one for the ten conditions with slope transition from negative to positive, and the other subgroup vice versa. For both groups, conditions were presented in a digram-balanced order. The up-down procedure and further details were the same as in experiment 1, except that the noise was always presented at a constant level of 70 dB SPL.

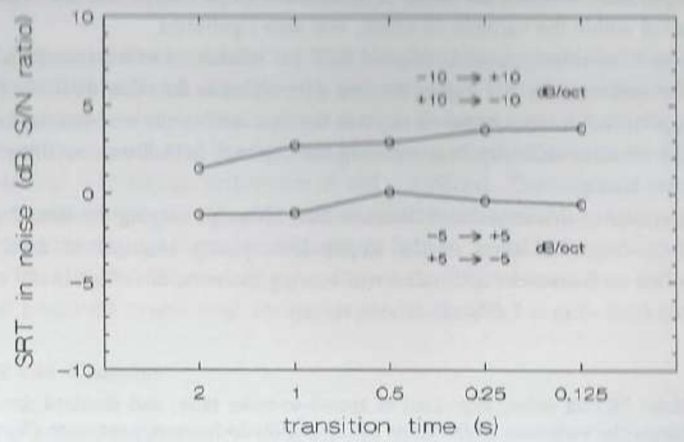


Fig. 2. Mean SRT in noise for single transitions in the slope of the amplitude-frequency response, as a function of the transition time. Parameter is the range of slope transition.

B. Results and discussion

The mean SRT in noise values for the 20 conditions are presented in Table II, with their standard deviations, and are plotted in Fig. 2.

An analysis of variance (see Table III) showed that the effects of range and transition

Table III. Mean effects and first-order interactions from an ANOVA on the results of experiment 2.

Source	Sum of squares	df	Mean squares	F ratio	p (%)
Direction	119.2	1	119.2	2.3	14.53 (n.s.)
Range	608.3	1	608.3	33.0	0.01
Transition time	68.7	4	17.2	5.0	0.16
Listeners within direction	941.3	18	52.3	11.2	0.00
Direction x range	71.0	1	71.0	3.9	6.23 (n.s.)
Direction x tr.time	16.2	4	4.0	1.2	32.84 (n.s.)
Range x tr.time	21.8	4	5.4	1.2	33.39 (n.s.)

time were significant, whereas the effect of direction of slope change was not. The effect of listeners, nested within the variable direction, was also significant.

The interindividual spread in masked SRT for conditions with a transition from -10 to +10 dB/oct was considerably higher than for a transition in the other direction (see Table II). This suggests that a steep negative slope in the first half of the sentence introduced for some listeners an extra difficulty in completing the sentence (which was confirmed by their comments after testing).

The results of this experiment indicate that, although varying the time for a single slope transition from one slope of the amplitude-frequency response to another has a significant effect on the masked SRT of normal-hearing listeners, this effect is still very small for transitions from -5 to +5 dB/oct, or vice versa.

Table IV. Mean SRT in noise, expressed in speech-to-noise ratio, and standard deviation, for continuous sinusoidal variations of the slope of the amplitude-frequency response (four variation frequencies, two ranges, and the average results for two conditions with a flat amplitude-frequency response; ten listeners).

	-5 ⟷ +5 dB/oct		-10 ⟷ +10 dB/oct	
Variation frequency (Hz)	Mean SRT (dB)	Standard deviation (dB)	Mean SRT (dB)	Standard deviation (dB)
1/4	-1.4	0.9	-0.1	1.2
1/2	-0.2	2.6	1.2	1.8
1	-0.4	1.2	2.6	2.5
2	0.8	1.6	4.8	2.1

	Flat response	
	Mean SRT (dB)	Standard deviation (dB)
(0 Hz)	-2.0	1.5

III. EXPERIMENT 3: CONTINUOUSLY VARYING SLOPE

In this experiment the effect of a dynamic slope variation of the amplitude-frequency response on the masked SRT for sentences was further explored by varying the slope continuously.

A. Method

The apparatus and stimulus material in this experiment were the same as in experiment 2. The essential new element was that now the slope varied continuously, according to a sinusoidal function along a dB scale, with frequencies of 0.25, 0.5, 1, or 2 Hz. The phase of the variation at the beginning of the sentence was random. Slopes varied between -10 and +10 dB/oct, or between -5 and +5 dB/oct. Together with two conditions with a flat amplitude-frequency response, this resulted in ten conditions. Again, a new group of ten normal-hearing listeners (4 male and 6 female, age 18-28) participated as paid volunteers in a digram-balanced design. The up-down procedure was the same as in experiment 1 with the overall level always near to 80 dB SPL.

B. Results and discussion

The mean values of SRT in noise are given in Table IV, with their standard deviations, and are plotted in Fig. 3.

An analysis of variance, the results of which are presented in Table V, showed that the effects of range of slope variation and variation frequency were highly significant.

Figure 3 shows that the curves for slope variations between -10 and +10 dB/oct and between -5 and +5 dB/oct diverge for increasing variation frequency. In general, we may

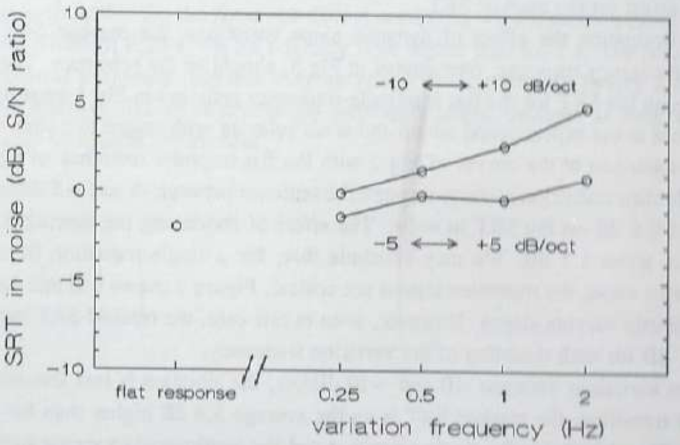


Fig. 3. Mean SRT in noise for continuous sinusoidal variations of the slope of the amplitude-frequency response, as a function of the variation frequency. Parameter is the range of slope variation.

Table V. Mean effects and first-order interactions from an ANOVA on the results of experiment 3.

Source	Sum of squares	df	Mean squares	F ratio	p (%)
Range	120.0	1	120.0	43.4	0.02
Variation frequency	129.5	3	43.2	17.1	0.00
Listeners	70.1	9	7.8	2.6	2.5
Range x var.fr.	23.9	3	7.9	2.7	6.59 (n.s.)
Range x listeners	24.9	9	2.8	0.9	51.62 (n.s.)
Var.fr. x listeners	68.0	27	2.5	0.8	66.57 (n.s.)

conclude that the SRT in noise increases gradually with variation frequency and more rapidly for the larger variations in slope.

IV. GENERAL DISCUSSION AND CONCLUSIONS

The three experiments reported in the previous sections show that the human ear is remarkably resistant to variations in the slope of the amplitude-frequency response. According to experiment 1, constant slopes between about -7 and +10 dB/oct do not show a significant systematic effect on the masked SRT.

In evaluating the effect of dynamic slope variations, the masked SRT for a flat amplitude-frequency response, represented in Fig.3, should be the reference. The difference of 1.3 dB with the SRT for the flat amplitude-frequency response in Fig.1 must be attributed to differences in the experimental set-up and is not relevant with regard to points in question.

Comparison of the curves of Fig.2 with the flat-response reference of Fig.3 reveals that a single slow transition halfway through the sentence between -5 and +5 dB/oct has only an effect of 0.8 dB on the SRT in noise. The effect of shortening the transition time is not greater than about 1.3 dB. We may conclude that, for a single transition from -5 to +5 dB/oct or vice versa, the transition time is not critical. Figure 3 shows that this does not hold for continuously varying slopes. However, even in this case, the masked SRT increases with less than 1 dB for each doubling of the variation frequency.

For variations between -10 and +10 dB/oct, the situation is less favorable. In case of a single transition, the masked SRT is on the average 3.4 dB higher than for the smaller range (Fig.2) and both for the single transition and the continuously varying slope, the SRT increases when speed is increasing (Figs. 2 and 3).

In summary, we may conclude that the masked SRT for sentences is rather stable as long as the slope of the amplitude-frequency response remains within a range of about -7 to +10 dB/oct. For single transitions between -5 and +5 dB/oct, the transition time is not critical, whereas, for continuous variations in this range, the masked SRT is raised by less than 1.8 dB for a variation frequency up to 1 Hz.

CHAPTER 3

THE EFFECT OF VARYING THE AMPLITUDE-FREQUENCY RESPONSE ON THE MASKED SPEECH-RECEPTION THRESHOLD OF SENTENCES FOR HEARING-IMPAIRED LISTENERS

Janette N. van Dijkhuizen, Joost M. Festen, and Reinier Plomp
published in: Journal of the Acoustical Society of America 86, 621-628 (1989)

ABSTRACT

In an evaluation of frequency-dependent automatic gain-control systems in hearing aids, the effect of varying the amplitude-frequency response on the speech-reception threshold (SRT) for sentences in noise is studied for 20 hearing-impaired listeners. The noise has a spectrum identical to the long-term average spectrum of the sentences. Speech and noise are shaped by the same amplitude-frequency response; their spectra are varied relative to the bisector of the individual's dynamic-range of hearing. In four experimental conditions, the effect of a steady-state amplitude-frequency response is studied. Steepening the negative spectral slope of speech and noise appears to cause an increase of masked SRT, possibly due to increased effect of upward spread of masking. The effect of a single transition of the amplitude-frequency response between 10 and -10 dB/oct halfway through the sentence seems to be related to the effect for the fixed -10-dB/oct condition. Two transition times are tested. For a transition time of 0.25 s, the SRT is only little higher than for 1 s. The results suggest that the amplitude-frequency response may be varied in time without having a detrimental effect on the masked SRT of sentences for hearing-impaired listeners as long as strongly negatively sloping spectra are avoided.

INTRODUCTION

A multichannel automatic gain-control (AGC) hearing aid, in which the frequency-dependent amplification adapts itself to the spectral parameters of the incoming sound, may optimally present speech within the dynamic range of the hearing-impaired listener. The amplitude-frequency response of such a hearing aid will vary in time and should, according to Plomp (1988), take about 0.5 s to adapt itself to a new acoustical situation. Such response variations leave speech-to-noise ratios at different frequencies untouched since they affect both speech and noise. However, despite a physically constant signal-to-noise ratio, the speech-reception threshold (SRT) in noise may still be affected by the tilt of the amplitude-frequency response or by the rate of varying this response.

In an earlier study (van Dijkhuizen et al., 1987) it was found that for normal-hearing listeners the SRT for sentences in noise is almost unaffected when the amplitude-frequency response has a steady-state slope within a range between -7 and +10 dB/oct, or when the slope of the amplitude-frequency response is slowly varying in time between +5 and -5 dB/oct. Given the results for normal hearing, a similar study is presently carried out with sensorineurally hearing-impaired listeners by investigating the effect of different steady-state and time-varying amplitude-frequency responses on the SRT of short meaningful sentences masked by noise. The noise has a spectrum identical to the long-term average spectrum of the sentences. Speech and noise are shaped by the same amplitude-frequency response. For optimal speech perception both in quiet and in noise, speech should be presented above threshold level and below the level of uncomfortable loudness over a wide range of frequencies (cf. Skinner et al., 1982). Therefore, we use, as a baseline, speech and noise spectrally shaped for each individual listener according to the bisector of the dynamic-range of hearing (cf. Levitt, 1978). In all conditions with a spectral tilt, the variation is applied relative to this baseline.

According to the concept of the articulation index (AI) (French and Steinberg, 1947; Kryter, 1962a), speech intelligibility is a weighted sum of "local" speech-to-noise ratios in a series of filter bands covering the speech spectrum. For steep masker spectra, corrections have to be applied. Since in our experiment variations of the amplitude-frequency response do not affect speech-to-noise ratio, and also very steep spectral masker slopes are avoided, the change in articulation index over experimental conditions will generally be small. To our knowledge, studies on systematic variations of the amplitude-frequency response with hearing-impaired subjects, with speech and noise presented within the listener's auditory range at all frequencies, have not been reported in the literature.

I. METHOD

Various steady-state and time-varying amplitude-frequency responses are realised by means of four parallel octave filters of 0.25-0.5, 0.5-1, 1-2, and 2-4 kHz, respectively, programmed digitally in a TMS 320-10 signal processor (sampling rate 10 kHz). Six-pole elliptic filters with slopes of approximately 40 dB/oct are used. The gain in the four frequency bands is controlled by a PDP-11/10 computer and, in time-varying conditions, updated at a

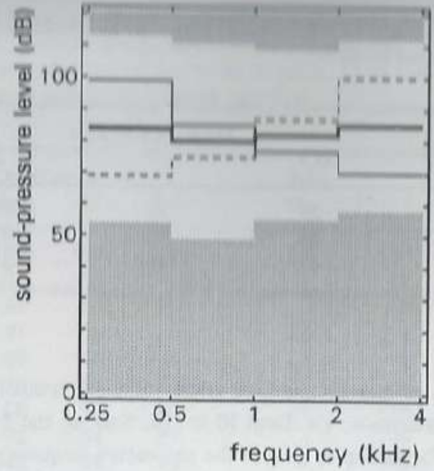


Fig. 1. Examples of spectral slopes in steady-state conditions, expressed in octave-band levels. The non-shaded area represents the listener's dynamic range of hearing. The bold curve gives the dynamic-range bisector. The thin and dashed curves represent -10 and +10 dB/oct relative to the dynamic-range bisector.

30-Hz rate. Maximum output level is 122, 118, 110, and 103 dB for the four octave bands, respectively.

For each listener, the auditory range, limited by the threshold of hearing (THR) and the level of uncomfortable loudness (UCL), is estimated using 300-ms bursts of octave-filtered noise separated by silent intervals of 100 ms. Thresholds are determined for each of the four octave bands according to a Békésy tracking procedure (step size 1 dB). UCL is determined according to a procedure in which the noise level is slowly increased with 1 dB for each new 300-ms burst, and the listener has to push a button when the noise is experienced as uncomfortably loud. Finally, the sound-pressure level corresponding to most-comfortable loudness (MCL) is determined by having the listener adjust the gain of a broadband noise (0.25-4kHz) with a spectrum identical to the long-term average spectrum of speech. The listener is instructed to select a level at which listening would still be comfortable for longer periods of exposure.

Spectral slopes of the experimental conditions will be expressed in dB/oct relative to the bisector of the listener's dynamic-range of hearing. With steady-state amplitude-frequency responses, the spectral slopes are 0, -10, and +10 dB/oct. An example of these slopes is given in Fig.1. In a fourth condition, the amplitude-frequency response is flat, with overall noise level fixed at MCL for the listener. In the time-varying conditions, the slope changes once halfway through the sentence presentation according to half a cosine function along a dB scale. Changes in slope are rising for one half of the listeners and falling for the other

Table I. Median levels with 25 and 75 percentiles of pure-tone air-conduction thresholds expressed in dB HL (20 hearing-impaired listeners).

	Frequency in kHz				
	0.25	0.5	1	2	4
25%	15	29	37	39	39
Median	21	34	42	42	55
75%	30	43	48	51	64

half. Two different rates of change are used (0.25- and 1-s transition times), and three different ranges of slope transition, i.e. from 10 to -10, 6 to -6, and 3 to -3 dB/oct, or vice versa. In one condition, for example, gain in the successive frequency bands changes from +15, +5, -5, and -15 dB to -15, -5, +5, and +15 dB relative to the bisector of the dynamic range. In all conditions, noise levels are dictated by the measurement condition, and the level of the speech is changed relative to the noise in an adaptive procedure in order to determine the speech-reception threshold.

For listeners with a small dynamic range, there are restrictions on the slope of the amplitude-frequency response, in order to avoid signals that are either too loud or too weak in some frequency regions. In the two extreme frequency bands, the gain is limited such that the noise level is always between (UCL - 10 dB) and (THR + 10 dB). In the two center bands, one-third of the range between (UCL - 10 dB) and (THR + 10 dB) for those frequencies is used at most. In time-varying conditions, the shape of the variation always remains a cosine. For an equal dynamic range in all frequency bands, this leads to a similar gain ratio (3:1) between extreme and center bands as for listeners with sufficient dynamic range in each frequency band. The margin of 10 dB was maintained to ensure that also band levels for speech fall within the dynamic range over a sufficiently large range of speech-to-noise ratios.

These restrictions on the amplitude-frequency response only became active in steady-state and time-varying conditions that involve a +10- or -10-dB/oct slope relative to the bisector. For all smaller slope variations, the dynamic range was large enough to allow the variations of speech and noise in all frequency bands. Since in the MCL condition level-preference of the subject is used, this is the only condition where speech and noise may be presented below threshold in one or more frequency bands.

The speech material consists of ten lists of 13 short (8 or 9 syllables) everyday Dutch sentences spoken by a female speaker (Plomp and Mimpen, 1979a). Sentences and noise are recorded on different tracks of a magnetic tape so that their overall levels can be adjusted independently before they are mixed, digitized, and fed into a TMS 320-10 signal processor for adjustment of the amplitude-frequency response. The average duration of the sentences is 2 s. In the time-varying conditions, the onset of the transition from one spectral slope to another is adjusted so that 1 s after the beginning of the sentence the spectral slope crosses

Table II. Octave-band levels (in dB SPL) representing uncomfortable loudness (first row), most comfortable loudness (second row), and auditory threshold (third row), for individual hearing-impaired subjects, determined for bursts of noise. Most comfortable loudness levels are determined for broadband noise with a speech spectrum.

Subj.	Frequency in kHz				Subj.	Frequency in kHz			
nr	0.25-0.5	0.5-1	1-2	2-4	nr	0.25-0.5	0.5-1	1-2	2-4
1	110	105	100	100	11	115	110	106	103
	77	73	65	58		81	77	69	62
	53	49	60	66		43	44	59	72
2	94	91	99	96	12	114	113	105	97
	73	69	61	54		72	68	60	53
	34	36	53	63		64	53	39	34
3	105	98	92	92	13	118	116	105	103
	76	72	64	57		79	75	67	60
	35	48	48	43		59	59	58	56
4	94	87	85	82	14	119	111	110	103
	73	69	61	54		70	66	58	51
	46	41	48	48		71	65	67	61
5	111	107	99	94	15	101	102	110	99
	84	80	72	65		68	64	56	49
	58	55	55	52		51	57	59	56
6	120	116	106	103	16	115	110	106	100
	68	64	56	49		72	68	60	53
	64	51	46	49		45	40	48	55
7	120	115	110	103	17	113	110	106	103
	81	77	69	62		79	75	67	60
	39	64	44	44		53	47	56	61
8	118	110	104	102	18	115	112	104	100
	77	73	65	58		79	75	67	60
	52	42	48	52		50	49	41	41
9	107	105	100	98	19	110	106	96	93
	74	70	62	55		79	75	67	60
	40	41	44	25		50	46	43	47
10	120	114	110	103	20	109	103	106	103
	81	77	69	62		82	78	70	63
	32	46	50	35		56	59	61	73
Mean	111	107	103	99					
(20	76	72	64	57					
subj.)	50	50	51	52					

the shape of the bisector of the listener's dynamic range.

Twenty listeners between 30 and 70 years of age with sensorineural hearing losses participate in the experiment as paid volunteers. For this group, the hearing loss for pure-tones averaged over 0.5, 1, and 2 kHz, is between 30 and 55 dB at their best ear, and

Table III. Octave-band levels for noise (in dB SPL) in steady-state conditions with spectral slopes of -10 (first row), 0 (second row), and +10 dB/oct *re:* bisector (third row), presented for individual hearing-impaired subjects. Since in the MCL-condition the amplitude-frequency response is flat, subtraction of MCL levels (Table II) from the levels in steady-state conditions gives the actual amplitude-frequency responses. For some subjects, amplitude-frequency responses are limited by their dynamic-range of hearing.

Subj. nr	Frequency in kHz 0.25-0.5	0.5-1	1-2	2-4	Subj. nr	Frequency in kHz 0.25-0.5	0.5-1	1-2	2-4
1	96.5	82.0	76.7	76.0	11	94.0	82.0	78.0	82.0
	81.5	77.0	80.0	83.0		79.0	77.0	82.5	87.5
	66.5	72.0	83.3	90.0		64.0	72.0	87.0	93.0
2	79.0	68.5	71.7	73.0	12	104.0	88.0	67.0	50.5
	64.0	63.5	76.0	79.5		89.0	83.0	72.0	65.5
	49.0	58.5	80.3	86.0		74.0	78.0	77.0	80.5
3	85.0	78.0	66.0	53.0	13	103.5	92.5	77.0	66.0
	70.0	73.0	70.0	67.5		88.5	87.5	81.5	79.5
	55.0	68.0	74.0	82.0		73.5	82.5	86.0	93.0
4	84.0	68.3	63.7	58.0	14	109.0	92.3	84.7	71.0
	70.0	64.0	66.5	65.0		95.0	88.0	88.5	82.0
	56.0	59.7	69.3	72.0		81.0	83.7	92.3	93.0
5	99.5	86.0	73.0	62.0	15	91.0	83.7	79.5	66.0
	84.5	81.0	77.0	73.0		76.0	79.5	84.5	77.5
	69.5	76.0	81.0	84.0		61.0	75.3	89.5	89.0
6	107.0	88.5	71.0	61.0	16	95.0	80.0	72.0	65.0
	92.0	83.5	76.0	76.0		80.0	75.0	77.0	77.5
	77.0	78.5	81.0	91.0		65.0	70.0	82.0	90.0
7	94.5	94.5	72.0	58.5	17	98.0	83.5	76.0	71.0
	79.5	89.5	77.0	73.5		83.0	78.5	81.0	82.0
	64.5	84.5	82.0	88.5		68.0	73.5	86.0	93.0
8	100.0	81.0	71.0	62.0	18	97.5	85.5	67.5	55.5
	85.0	76.0	76.0	77.0		82.5	80.5	72.5	70.5
	70.0	71.0	81.0	92.0		67.5	75.5	77.5	85.5
9	88.5	78.0	67.0	46.5	19	95.0	81.0	64.5	57.0
	73.5	73.0	72.0	61.5		80.0	76.0	69.5	70.0
	58.5	68.0	77.0	76.5		65.0	71.0	74.5	83.0
10	91.0	85.0	75.0	54.0	20	97.5	85.0	79.3	83.0
	76.0	80.0	80.0	69.0		82.5	81.0	83.5	88.0
	61.0	75.0	85.0	84.0		67.5	77.0	87.7	93.0
Mean	95.5	83.2	72.6	63.6					
(20	80.6	78.3	77.2	75.3					
subj.)	65.7	73.5	81.7	86.9					

air-bone gap is always less than 10 dB between 0.25 and 4 kHz. A statistical survey of the pure-tone thresholds is given in Table I. Performance functions for monosyllables in quiet

reaches at least 90% intelligibility. Sound-pressure levels of THR, UCL, and MCL for octave-bands of noise are given for each individual listener in Table II. In the 1-2kHz and 2-4kHz bands, UCL is equal to maximum output of the system for four and eight listeners, respectively. The width of the dynamic range, averaged over all listeners, is 61, 57, 52, and 47 dB, for the octave bands from low to high, respectively.

Sentences in noise are presented monaurally to the listener's best ear over a headphone in a sound-proof room. The ten lists of sentences are presented in a fixed order. In the time-varying conditions, the spectral slope varies from negative to positive for subjects 1 to 10, and from positive to negative for subjects 11 to 20. The ten experimental conditions are distributed over the lists according to a digram-balanced design per ten listeners in order to avoid the effects of learning and fatigue in the average results. SRT in noise is defined as the speech-to-noise ratio at which the listener reproduces 50% of the sentences without a single error. Speech-to-noise ratio is varied by adjusting the level of the sentences relative to the level of the noise in an adaptive up-down procedure with a step size of 2 dB (Plomp and Mimpen, 1979a).

Table III gives octave-band levels for steady-state conditions with spectral slopes of -10, 0, and +10 dB/oct relative to the dynamic-range bisector. Since octave-band levels for speech vary with speech-to-noise ratio, only levels for noise are presented. It can be inferred from this table that the dynamic range in the higher frequencies is often too small for the most extreme slope conditions. The average gain relative to the bisector in the octave-band from 2 to 4 kHz is 11.7 dB instead of 15 dB as prescribed for the 10-dB/oct condition. Octave-band levels for the noise in the MCL condition are given in Table II. For time-varying conditions, the limits of octave-band levels can be read from Table III for the +10- to -10-dB/oct slope transitions. For the smaller slope transitions (± 6 and ± 3 dB/oct) these levels can be generated easily by applying the appropriate slopes relative to the spectra in the 0-dB/oct condition.

II. RESULTS AND DISCUSSION

A. Steady-state amplitude-frequency response

The SRT in noise for steady-state amplitude-frequency responses is given in Table IV for each listener. Average thresholds for 20 listeners with standard deviations are plotted in Fig.2. The lowest speech-reception thresholds are found for the bisector condition. In this condition, the average slope of the amplitude-frequency response for our group of listeners is +4.57 dB/oct. This slope agrees well with frequent recommendations in the literature stating that a slightly rising hearing-aid characteristic gives the best performance (cf. Braida et al., 1979). However, because this study was not set up to find the optimal slope for a hearing-aid frequency response, the limited number of slope conditions tested precludes any precise statement about an optimal slope. Lutman and Clark (1986) and Haggard et al. (1986) compared a flat amplitude-frequency response with a response of +9 dB/oct, which corresponds roughly to conditions with a positive and negative slope of 4.5 dB relative to the bisector. They found no difference in masked SRT between the two slopes for hearing-impaired listeners. Although test procedures are not identical, this result suggests a plateau in the SRT curve around the bisector condition.

As stated above, the normal speech spectrum has a slope of about -4.6 dB/oct relative to the bisector. Therefore, the normal slope of the speech spectrum is generally more strongly modified to obtain the +10-dB/oct condition *re:* the bisector than for the -10-dB/oct condition, especially for listeners with increasing thresholds of hearing towards higher frequencies. Nevertheless, the average SRT in noise in the +10-dB/oct condition is raised by only 1 dB relative to the SRT for the bisector condition, whereas, in the -10-dB/oct condition, average SRT in noise is raised by 5.8 dB with a large interindividual spread.

Although the MCL condition is the only one in which the slope of the speech spectrum remains unchanged, the masked SRT for this condition is surprisingly high. Halfway between the 0 and -10 dB/oct relative to the bisector, the amplitude-frequency response would

Table IV. SRT in noise, expressed in speech-to-noise ratio (dB), for steady-state conditions as a function of spectral slope. Results are presented for individual hearing-impaired subjects.

Subj. nr	Spectral slope in dB/oct *re:* dynamic-range bisector -10	0	+10	MCL
1	-0.6	-1.4	-1.4	4.6
2	-0.2	-2.6	1.0	5.4
3	4.6	-1.4	2.2	9.4
4	-1.0	-0.6	3.0	3.0
5	3.8	-1.0	-0.6	4.6
6	7.4	-1.0	1.8	5.0
7	4.6	-1.8	1.8	5.4
8	7.0	3.4	1.8	5.8
9	9.8	0.2	-0.2	2.6
10	5.8	-0.6	2.2	7.0
11	0.2	-2.2	-2.2	4.6
12	10.6	-1.0	-3.4	1.4
13	5.4	-0.6	-2.2	0.2
14	10.6	1.8	1.8	12.6
15	3.4	-1.8	1.8	10.6
16	7.8	-1.4	-0.6	4.2
17	7.0	2.6	1.0	6.2
18	8.2	0.6	-1.4	1.0
19	7.8	0.2	0.6	3.4
20	3.8	-1.8	3.4	5.8
Mean	5.3	-0.5	0.5	5.1
sd	3.6	1.6	1.9	3.1

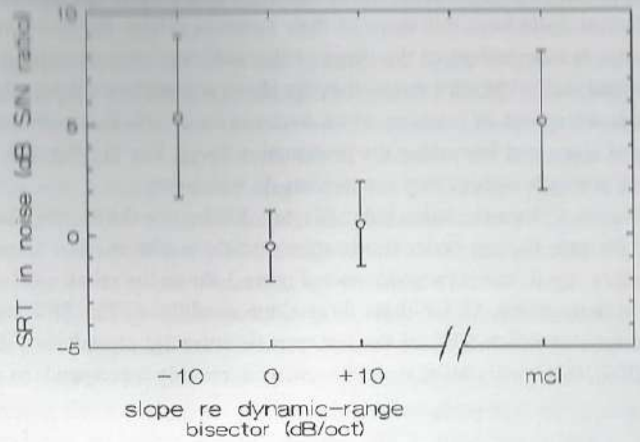


Fig. 2. Mean SRT in noise (20 listeners), with standard deviations, for steady-state amplitude-frequency responses (in dB/oct relative to dynamic-range bisector, except for MCL).

have been approximately flat as in the MCL condition, but the masked SRT for such a condition, obtained from linear interpolation, is about 3 dB lower than for the MCL condition. The relatively high SRT in noise and large spread among listeners in the MCL condition may be partly due to the fact that speech and noise above 2 kHz are presented near or below the threshold of hearing for nine listeners (see Table II). For normal-hearing listeners, the masked SRT for low-pass-filtered speech is about 2.5 dB higher for a cutoff frequency of 2 kHz than for 4 kHz (Plomp, 1986).

A factor that may explain the high SRT in noise values for the -10-dB/oct condition is masking of high-frequency speech components by intense low-frequency components of speech and noise. For a low-frequency masker with an average overall SPL of 87 dB, Jerger et al. (1960) found a slope of the masking pattern of about -6 dB/oct for hearing-impaired listeners. For normal-hearing listeners in similar conditions, this slope can be estimated to be -20 to -30 dB/oct (cf. Ehmer, 1959). It follows that even for the normal speech spectrum with a slope between -5 to -10 dB/oct, like in the MCL condition, masking of high-frequency components by signals in the lowest frequency band may have occurred. From measurements of psychoacoustical tuning curves, Lutman and Clark (1986) concluded that the listeners with more upward spread of masking from frequencies below 500 Hz had better speech intelligibility scores for a rising than for a flat amplitude-frequency response.

Only in the -10-dB/oct *re:* bisector condition, where the average slope of the spectrum is -10.63 dB/oct (see Table III), masked SRT correlates with this spectral slope ($r = -0.80$, $N = 20$). It also appears that the masked SRT in this condition correlates with noise level in the 0.25-0.5kHz band ($r = 0.56$, $N = 20$). Both spectral slope and presentation level of the noise in our experimental conditions depend on the shape of the auditory range of the listener, so both correlations could have this shape as their common origin. However, if we assume that SRT in noise is independent of the shape of the auditory range as long as speech and noise are presented well within this range, then the above correlations support the hypothesis that it is the upward spread of masking which becomes more effective with steepening the negative spectral slope and increasing the presentation level. For the flat and +10-dB/oct conditions, very strong low-frequency components do not occur.

Calculations of the articulation index (Kryter, 1962a; one-third-octave-band method) are performed for each listener under steady-state conditions with spectral slopes of -10, 0, and +10 dB/oct *re:* the dynamic range bisector. Figure 3 shows the relation between masked SRT and the corresponding AI for these three slope conditions. The SRT represents the signal-to-noise ratio at which 50% of the sentences is correctly reproduced. According to ANSI (S3.5-1969), this intelligibility score for sentences roughly corresponds to an AI of 0.2

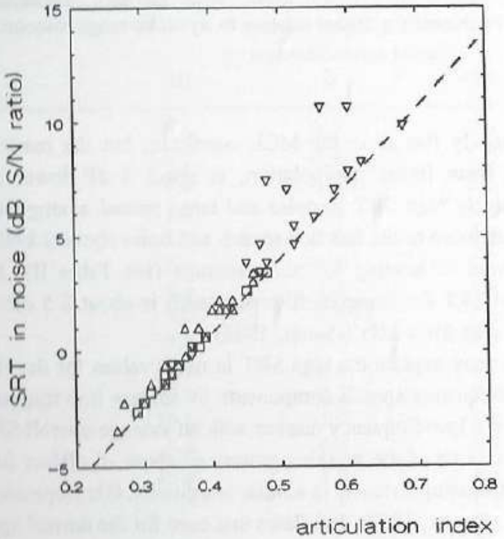


Fig. 3. Relation between masked SRT, representing 50% correct sentence reproduction, and the articulation index (Kryter, 1962a; one-third-octave-band method), for 20 hearing-impaired listeners and three steady-state slope conditions (▿, □, and ▵, for -10, 0, and +10 dB/oct relative to the dynamic-range bisector, respectively). The dashed line represents the linear relation between AI and speech-to-noise ratio for a flat amplitude-frequency response and an overall noise level of 85 dB SPL. Along this line, the correction for spread of masking is constant.

but may vary somewhat with speech material, talker, and listener group. For the talker and speech material used in this experiment and a signal limited between 0.25 and 4 kHz, normal-hearing listeners obtained 50% intelligibility at a signal-to-noise ratio of -2.0 dB with a standard deviation of 1.5 dB (see van Dijkhuizen et al., 1987). The corresponding AI is 0.31 with a standard deviation of 0.05. However, for hearing-impaired listeners and for the slope conditions tested in this experiment, the AI calculated for the signal-to-noise ratio at threshold varies over a much broader range, both for the average results of conditions and for the individual listeners. Since the noise is above threshold in all frequency bands and for all conditions, AI increases linearly with the speech-to-noise ratio between -12 and +18 dB, as long as corrections for spread of masking are constant. The deviations of AI from a linear relation (dashed line) in Fig. 3, all due to the correction for spread of masking, occur mainly for negative spectral slopes but are much too small to account for the observed variability in the SRT data. Whereas in Kryter's method weighting factors are highest for signal-to-noise ratios in frequency bands around 1.6 kHz, a stronger weighing in low-frequency bands may be more appropriate for sentences (cf. Pavlovic, 1987). However, a shift of the weighing towards lower frequencies is a shift to a region where the role of upward spread of masking is less. As a consequence, this would bring the symbols in Fig.3 even closer to the dashed line. It follows that the articulation index grossly overestimates speech intelligibility in noise for the hearing-impaired listeners in our experiment, particularly for the negative spectral slope condition. However, articulation theory was not designed to describe hearing impairment. From the large effect on the masked SRT in the negative slope condition, which is not seen in the AI, it can be concluded that excessive upward spread of masking in impaired ears probably accounts at least for part of the variability in the data.

B. Single transition of the amplitude-frequency response

The SRT in noise for time-varying amplitude-frequency responses is given for individual listeners in Table V, and mean thresholds for 20 listeners with standard deviations are plotted in Fig. 4. An analysis of variance (see Table VI) shows no significant effect of direction of the slope transition. Therefore, the data can be averaged over the two subgroups of ten listeners. The effect of the variable listeners nested within direction, and the effect of range of transition are highly significant, and are the main sources of variance. The masked SRT increases almost linearly with range of transition. Varying the spectra of speech and noise once halfway through the sentence from +3 to -3 dB/oct relative to the auditory range bisector, or vice versa, gives an increase of mean masked SRT of less than about 2 dB. The SRT in noise for the conditions with the largest transition range (+10 to -10 dB/oct, or vice versa) shows a larger interindividual spread than for other time-varying conditions. The effect of reducing transition time from 1 to 0.25 s is also significant, but on the average, less than 1 dB for the lower two transition ranges and about 1.5 dB for the largest transition range. A similar result was found for normal-hearing listeners (van Dijkhuizen et al., 1987).

A comparison of Tables IV and V shows that the mean SRT in noise for the largest slope transitions is much closer to the results for the -10-dB/oct steady-state condition than to the +10-dB/oct condition. The masked SRT for the largest transition range and for both

Table V. SRT in noise, expressed in speech-to-noise ratio (dB), for time-varying conditions, as a function of range limits of spectral slope and transition time. Direction of transition is indicated by - → + or + → -. Results are presented for individual hearing-impaired subjects.

	Range limits of slope transition re bisector (dB/oct)							Range limits of slope transition re bisector (dB/oct)					
	- → +3		- → +6		- → +10			+ → -3		+ → -6		+ → -10	
	trans.time (s)		trans.time (s)		trans.time (s)			trans.time (s)		trans.time (s)		trans.time (s)	
Subj. nr	1/4	1	1/4	1	1/4	1	Subj. nr	1/4	1	1/4	1	1/4	1
1	-0.6	-2.2	-0.6	0.2	2.6	-0.2	11	0.2	-0.2	-0.6	-1.0	2.2	-0.2
2	-0.6	2.2	1.0	1.4	2.2	5.8	12	3.8	0.2	3.4	5.4	9.4	7.8
3	2.2	4.2	5.0	2.2	3.4	3.8	13	-2.2	-0.6	3.4	0.6	1.4	1.8
4	-0.2	-1.4	2.6	4.2	2.6	0.2	14	3.0	3.8	3.8	4.2	9.4	6.6
5	1.4	0.6	0.6	1.0	2.6	1.8	15	2.6	1.8	3.0	2.6	5.4	3.8
6	2.2	3.0	6.2	1.8	8.6	5.4	16	0.2	1.0	0.2	2.2	4.6	5.0
7	1.0	-0.6	2.2	1.4	9.4	3.8	17	3.0	2.6	4.2	5.4	10.2	5.4
8	2.6	1.4	1.8	2.2	6.2	3.8	18	0.2	1.8	0.2	1.8	2.6	2.6
9	2.2	0.6	4.2	1.4	5.0	3.4	19	4.2	-1.8	3.4	1.8	9.8	8.6
10	3.4	0.2	3.8	2.6	5.8	1.8	20	-2.2	-0.2	2.2	2.2	1.0	-0.6
M	1.4	0.8	2.7	1.8	4.8	3.0	M	1.3	0.8	2.3	2.5	5.6	4.1
sd	1.4	2.0	2.1	1.1	2.6	2.0	sd	2.4	1.7	1.7	2.0	3.8	3.2
overall													
M	1.3	0.8	2.5	2.2	5.2	3.5							
sd	1.9	1.8	1.9	1.6	3.2	2.7							

transition times also shows considerably higher correlations with the masked SRT for the steady-state -10-dB/oct slope than for the +10-dB/oct slope [for -10 dB/oct, r=0.62 and r=0.64, and for +10 dB/oct, r=0.11 and r=-0.07, for transition times of 0.25 and 1 s, respectively (N=20)]. It appears that masked SRT for conditions with a single transition of the spectral slope is related to masked SRT for the steady-state negative spectral slope, corresponding to one of the range limits of transition. Probably, the masked SRT for sentences with a single transition of spectral slope halfway through the sentence is determined by the least intelligible half. It should be remembered that the adaptive up-down procedure requires reproduction of the entire sentence without a single error for a correct response.

III. CONCLUSIONS

For the steady-state amplitude-frequency responses tested, the SRT in noise is best when speech and noise are presented at the listener's dynamic-range bisector. From interpolation of the average data, it can be inferred that, for a spectral slope between roughly

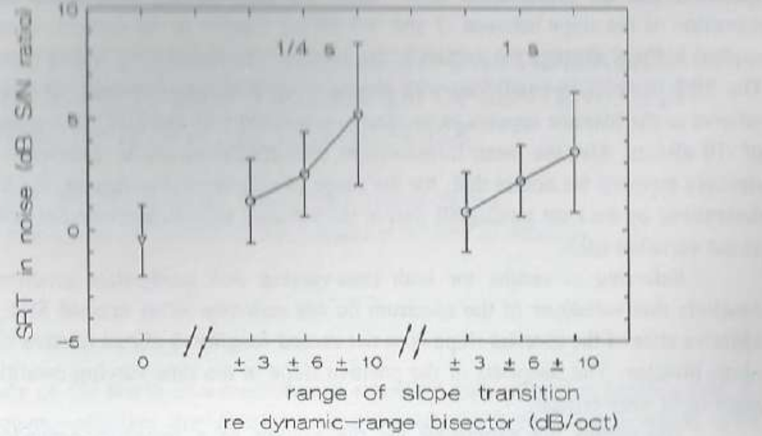


Fig. 4. Mean SRT in noise (20 listeners), with standard deviations, for a single transition of the amplitude-frequency response, as a function of range limits of slope transition (in dB/oct relative to dynamic-range bisector). Parameter is transition time in seconds. The triangle represents masked SRT for the (steady-state) bisector condition (from Fig. 2).

-3 and +10 dB/oct relative to the bisector, masked SRT is not increased by more than 2 dB. Indications are found that steepening the negative spectral slope of speech and noise beyond this range increases the risk of upward spread of masking, depending on steepness of the

Table VI. Mean effects and first-order interactions from an ANOVA on the results for a single transition of the amplitude-frequency response.

Source	Sum of squares	df	Mean squares	F ratio	p
Direction	3.9	1	3.9	0.2	0.65
Listeners within direction	319.6	18	17.8	10.1	<0.005
Transition time	21.2	1	21.2	7.7	0.01
Range	221.7	2	110.8	31.2	<0.005
Direction x tr.time	1.9	1	1.9	0.7	0.58
Direction x range	5.2	2	2.6	0.7	0.51
List. x tr.time	49.8	18	2.8	1.6	0.12
List. x range	127.9	36	3.6	2.0	0.02
Tr.time x range	11.3	2	5.6	3.2	0.05

spectrum and on presentation level. Also, for time-varying conditions, when a single transition of the slope between -3 and +3 dB/oct relative to the dynamic-range bisector is applied halfway through the sentence, the increase of masked SRT is less than about 2 dB. The SRT in noise in conditions with slopes varying in time between -10 and +10 dB/oct relative to the bisector appears to be clearly correlated with the SRT for a steady-state slope of -10 dB/oct. Also the mean thresholds in both conditions are at comparable levels. This strongly supports the notion that, for the range of variations investigated, the SRT is mainly determined by the least intelligible part of the sentence without a substantial contribution due to the variation itself.

Referring to results for both time-varying and steady-state conditions, we may conclude that variations of the spectrum do not seriously affect masked SRT as long as a negative state of the spectral slope does not exceed roughly -3 dB/oct relative to the auditory range bisector. The steepness of the positive slope in the time-varying conditions does not seem to be very critical.

The results are promising for the success of a frequency-dependent automatic gain-control in hearing aids. Disturbing sounds usually have their strongest components at the lower frequencies and, consequently, will frequently call for a rising amplitude-frequency response. Finally, a reduction of the transition time from 1 to 0.25 s gives an increase in the masked SRT of, on the average, about 1 dB. This small increase of threshold may be attributed to loss of intelligibility due to the variation itself and it is an indication that short time constants for transitions of the amplitude-frequency response should be avoided.

CHAPTER 4

THE EFFECT OF FREQUENCY-SELECTIVE ATTENUATION ON THE SPEECH-RECEPTION THRESHOLD OF SENTENCES IN CONDITIONS OF LOW-FREQUENCY NOISE

Janette N. van Dijkhuizen, Joost M. Festen, and Reinier Plomp
accepted for publication in: Journal of the Acoustical Society of America

ABSTRACT

Within a study on the merits of a multichannel automatic gain control in hearing aids, the effect of frequency-selective amplification on the masked speech-reception threshold (SRT) for sentences is measured in conditions of seriously disturbing low-frequency noise, with the effect of wideband amplification as a reference. Speech and noise are both spectrally shaped according to the bisector line of the listener's dynamic-range of hearing, but with the noise in a single octave band (0.25-0.5kHz or 0.5-1kHz) increased by 20 dB relative to this line. The increase of noise level is steady-state in the first experiment, and time-varying in the second experiment. Results for 12 normal-hearing and 12 hearing-impaired listeners indicate that, in both experiments, frequency-selective compression of the signal in the octave band with the 20-dB increase of noise is more beneficial than wideband compression. For the hearing-impaired group, wideband compression does not give any systematic change in intelligibility. Frequency-selective compression in steady-state conditions may, for both groups of listeners, give a decrease of masked SRT (relative to a condition without compression) of up to 4 dB for a compression factor of 100%. Roughly comparable effects are seen for frequency-selective compression in time-varying conditions. The superiority of frequency-selective over wideband compression is attributed to a more effective reduction of upward spread of masking.

INTRODUCTION

A multichannel automatic gain-control (AGC) hearing aid, in which the frequency-dependent amplification adapts itself automatically to the spectro-temporal parameters of the incoming sound, can optimally adapt the level and spectrum of the signal to the reduced dynamic range of the impaired ear. In such a hearing aid, the speech signal in the different frequency bands is favorably presented within the dynamic range of hearing so that maximal audibility of the speech can be obtained. At the same time those frequency bands that do not contribute to intelligibility, because the speech in these bands is masked by noise, should be attenuated. This may improve listening comfort and reduce the potential effect of spread of masking to other frequency regions (cf. Festen et al., 1990; Kates, 1990), without eliminating any useful contribution to speech intelligibility. Because in practice noise spectra vary in time, the amplitude-frequency response of such a hearing aid will vary along with the current noise spectrum, preferably taking about 0.25-0.5 s to adapt itself to (gross) changes in the acoustical situation (Plomp, 1988). For shorter durations, as used in syllabic compression, no conclusive results have been reported (cf. Braida et al., 1979). However, short attack and release times of compression will reduce the intensity contrasts of speech within each frequency channel, and this may affect the speech-reception threshold negatively (Plomp, 1988).

With wideband rather than frequency-dependent gain control, a strong low-frequency noise would automatically reduce the gain for all frequencies equally, easily resulting in reduced audibility of higher-frequency speech components. Particularly when the listener's hearing levels differ over the frequency range, as in a high-frequency hearing loss, adjusting a wideband amplification factor does not provide an adequate adaptation of the incoming signal to the reduced dynamic range of the hearing-impaired (Plomp, 1988).

Because the variations of the amplitude-frequency response affect both speech and interfering noise, speech-to-noise ratios in the different frequency bands remain unchanged. However, in spite of a constant speech-to-noise ratio, the speech-reception threshold (SRT) in noise may still be negatively affected by the shape of the amplitude-frequency response or by its rate of change. This problem was investigated in earlier studies with normal-hearing and hearing-impaired listeners (van Dijkhuizen et al., 1987 and 1989), where the noise had a spectrum identical to the long-term average spectrum of the sentences. This corresponds to the frequent situation in listening where the human voice itself is the interfering sound. For normal-hearing listeners van Dijkhuizen et al. (1987) found that the SRT for sentences in noise is practically unaffected when the slope of the amplitude-frequency response is steady-state within a range between -7 and +10 dB/oct, or when this slope is slowly varying in time between -5 and +5 dB/oct. Further steepening the falling slope of the amplitude-frequency response appears to increase the risk of spread of masking. For hearing-impaired listeners van Dijkhuizen et al. (1989) found that, in conditions with steady-state and slowly varying amplitude-frequency responses, the increase of masked SRT is less than 2 dB, provided the negative slope of the speech and noise spectrum is not steeper than -3 dB/oct relative to the line bisecting the ear's dynamic range. Steeply rising amplitude-frequency responses do not have a critical effect on the SRT. Because most disturbing noises

in the environment have their strongest components in the lower frequencies (cf. Kryter, 1970), adjusting a rising hearing aid response will also be most appropriate for the purpose of improving listening comfort.

The robustness of speech intelligibility for normal and impaired listeners under variations in the amplitude-frequency response, referred to above, is promising for the success of frequency-dependent automatic gain control in hearing aids. The question that follows is: how large is the beneficial effect on intelligibility, by reduction of spread of masking, of adjusting the amplitude-frequency response in situations of seriously disturbing noise with a narrow-band maximum? Spread of masking caused by such a masker spectrum particularly manifests itself from lower to higher frequencies (upward spread of masking), and tends to grow progressively with signal level (cf. Bilger and Hirsh, 1956). Several investigators have reported greater than normal upward spread of masking in listeners with a sensorineural hearing impairment (cf. Jerger et al., 1960; Rittmanic, 1962), although others have stressed that results may vary considerably among listeners, and depend strongly on the measure used to describe excessive spread of masking (Martin and Pickett, 1970; Tyler, 1986).

This paper presents the results of a study on the above-mentioned question. This study was carried out for groups of normal-hearing and sensorineurally hearing-impaired listeners by investigating the effect of adapting frequency-selective versus wideband compression on the masked SRT of short meaningful sentences. The noise had a spectrum identical to the long-term average spectrum of the speech, except that in one low-frequency octave band its level was substantially raised. Speech components within that noise band could not contribute to intelligibility. Compression was applied to the combined signal of speech and noise.

For optimal speech understanding both in quiet and in noise, speech should be presented above threshold level and below the level of uncomfortable loudness over a frequency range that is as wide as possible (cf. Skinner et al., 1982). Therefore, as the basic "input" to all frequency-selective and wideband conditions, we used speech and noise spectrally shaped for each individual listener according to the line that bisects his/her dynamic range of hearing, limited by the threshold of hearing and the level of uncomfortable loudness (cf. Levitt, 1978; van Dijkhuizen et al., 1989). In a single low-frequency octave band, however, the noise level was increased by 20 dB relative to this bisector line. This increase was either steady-state (Experiment 1) or time-varying (Experiment 2). Compression was applied by adjusting the gain in relation to the change in SPL of the total input signal in the specified frequency band (narrow or wide), and, as a result, will be steady-state in Experiment 1 and time-varying in Experiment 2. The compression factor was varied as an experimental parameter.

According to the concept of the Articulation Index (French and Steinberg, 1947), speech intelligibility is proportional to the weighted sum of speech-to-noise ratios in a series of filter bands covering the speech spectrum. For steep masker spectra, the calculation of the AI includes corrections for effects of spread of masking, validated for a set of experimental data from normal-hearing listeners (Kryter, 1962b). For conditions of speech presented in low-pass noise with a level 12 dB above the speech level, Festen et al. (1990) used the AI to illustrate the beneficial effect of selective attenuation of speech and noise in the pass-band of the noise to a level below the onset of upward spread of masking. According to articulation

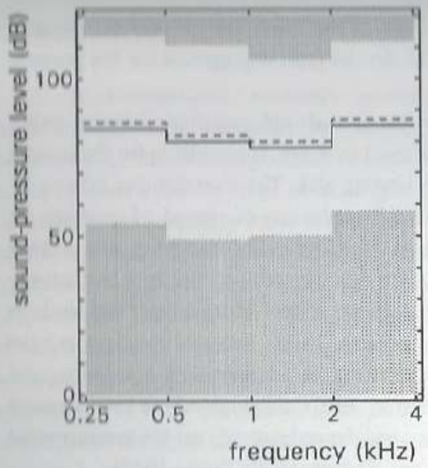


Fig. 1a. Example of how spectra for speech (dashed curve) and noise (solid curve) were each shaped according to the bisector line (solid curve) of the listener's dynamic range of hearing (non-shaded area). Spectra are expressed in octave-band levels. Speech-to-noise ratio (as determined by the adaptive threshold-estimation procedure) in this example is 2 dB. SPLs in the four octave bands from low to high are for speech: 86, 82, 80, and 87 dB, and for noise: 84, 80, 78, and 85 dB, respectively. SPLs for the total signal, obtained by summing the intensities of speech and noise, are 88.1, 84.1, 82.1, and 89.1 dB, for the four octave bands from low to high, respectively, and 92.8 dB for the wideband signal.

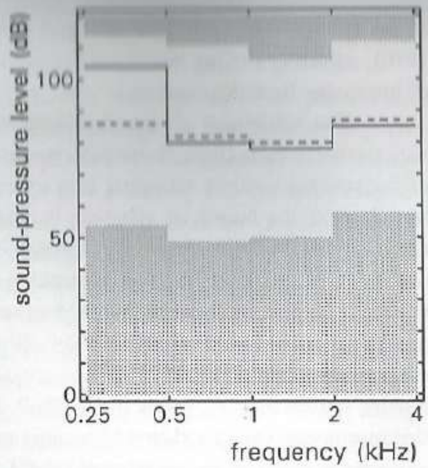


Fig. 1b. As Fig. 1a, however, with the noise in the 0.25-0.5kHz band raised by 20 dB relative to the bisector. SPLs in the four octave bands from low to high are for speech: 86, 82, 80, and 87 dB, and for noise: 104, 80, 78, and 85 dB, respectively. SPLs for the total signal are now: 104.1, 84.1, 82.1, and 89.1 dB, for the four octave bands from low to high, respectively, and 104.3 dB for the wideband signal. The increase in signal level can be determined by subtracting the corresponding levels in Fig. 1a, and is 104.1-88.1 = 16 dB for the noisy octave band 0.25-0.5kHz, and 104.3-92.8 = 11.5 dB for the wideband signal. For a compression factor of 0%, the increase in signal level is not compressed, and thus, the spectra presented in this figure remain unchanged.

theory, speech components in this noise band have no useful contribution to speech perception. For an overall SPL of the noise (cut-off frequency 1 kHz) between 80 and 100 dB, they calculated a predicted improvement of intelligibility that is comparable to a wideband increase of speech-to-noise ratio between about 1.5 and 7.5 dB. For hearing-impaired listeners it was shown (van Dijkhuizen et al., 1989) that, particularly in conditions with steep negatively sloping spectra of speech and noise, the spread-of-masking corrections in the calculation of the AI are far too small to account for the observed variability in masked SRT. It was concluded that, because articulation theory was designed to describe normal hearing, at least part of the variability in the data should be assigned to greater than normal upward spread of masking in impaired ears. For hearing-impaired listeners, therefore, the expected

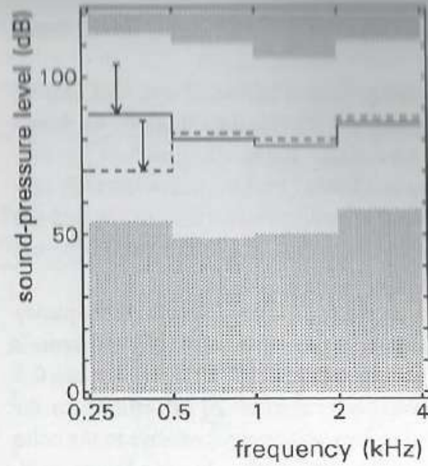


Fig. 1c. As Fig. 1b, with 100% frequency-selective compression applied to speech and noise. Because the increase of signal level in the noisy octave band 0.25-0.5kHz is 16 dB (see legends of Fig. 1b), the compression applied to the signal in this band is also 16 dB (indicated by the length of the arrows).

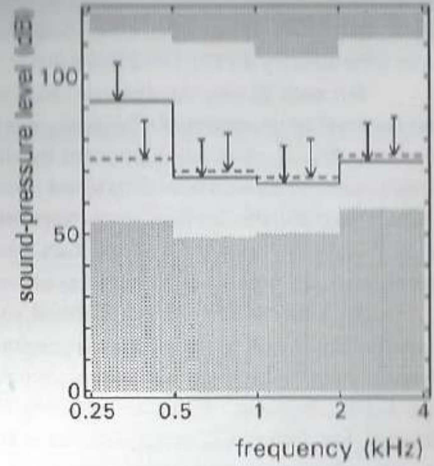


Fig. 1d. As Fig. 1b, with 100% wideband compression applied to speech and noise. Because the increase of the wideband signal level is 11.5 dB (see legends of Fig. 1b), the compression applied to the wideband signal is also 11.5 dB (indicated by the length of the arrows).

benefit of selective compression of the frequency region of the noise may be larger than for normal-hearing listeners, but will depend strongly on the individual listener's upward spread of masking.

Several studies, aimed at evaluating a single sophisticated technique of adaptive filtering, have been carried out in the past, reporting modest improvements of speech intelligibility in low-frequency noise (van Tasell et al., 1988; Stein and Dempesy-Hart, 1984; Ono et al., 1983). However, experiments on systematically and exclusively varying the gain to be applied, with speech and noise at all frequencies and in all conditions presented well within the listener's auditory range, have, to our knowledge, not been reported in the literature.

I. EXPERIMENT 1: STEADY-STATE CONDITIONS

A. Method

Various steady-state spectra for speech and noise were each realised by a set of four parallel octave filters of 0.25-0.5, 0.5-1, 1-2, and 2-4 kHz, programmed digitally in a TMS 320-10 signal processor (sampling rate 10kHz). Six-pole elliptic filters with slopes of approximately 40 dB/oct were used. Amplification factors in the different frequency bands were controlled by a PDP 11/10 computer.

For each listener the dynamic range of hearing, limited by the threshold of hearing and the level of uncomfortable loudness, was measured prior to testing using 300-ms bursts of octave-filtered white noise separated by silent intervals of 100 ms. For each of the four octave bands, thresholds were determined according to a Békésy tracking procedure (step size 1 dB). Uncomfortable loudness levels were determined according to a procedure in which the noise level is increased by 1 dB for each new 300-ms burst, and the listener has to push a button when the noise is experienced as uncomfortably loud.

As a baseline for all experimental conditions we used a fixed amplitude-frequency response that shapes speech and noise spectra according to the bisector line of the listener's dynamic range of hearing. In a single octave band, between 0.25 and 0.5 kHz or between 0.5 and 1 kHz, however, the noise level was increased by a constant 20 dB relative to the bisector. For each new sentence, the overall level of speech was changed relative to the noise in an adaptive procedure in order to determine the SRT. In conditions with frequency-selective compression, amplification in each of the four octave bands was independently adjusted to changes in signal level in the respective band. In conditions with wideband compression, a single amplification factor, that affected all frequencies between 0.25 and 4 kHz equally, was adjusted to changes in the wideband signal level. For each speech-to-noise ratio during the adaptive threshold-estimation procedure, amplification factors were computed once on the basis of the level of the total signal (consisting of speech plus noise), measured in octave bands or wideband (depending on the experimental condition), and held constant for the duration of the sentence. Amplification factors were such that the level increase due to the extra 20 dB of noise was not compressed (0%), compressed to its half (50% in dB), or fully compressed (100%), over the specified band (octave or wide). Examples of speech and noise spectra are given in Figs. 1a to d.

Two different octave bands containing the extra 20 dB of noise (0.25-0.5kHz and 0.5-1kHz), two types of gain control (frequency-selective and wideband), and three compression factors (0, 50, and 100%), gave ten different experimental conditions (0% compression conditions were equal for both types of gain control). In the reference

Table I. Mean pure-tone air-conduction thresholds with their standard deviations, expressed in dB HL (12 hearing-impaired listeners).

	Frequency in kHz				
	0.25	0.5	1	2	4
Mean	30	40	51	53	68
sd	15	13	9	10	13

Table II. Octave-band levels (in dB SPL) representing uncomfortable loudness (first row), and hearing threshold (second row), for the individual normal-hearing and hearing-impaired subjects, determined with bursts of octave-filtered noise.

NORMAL-HEARING					HEARING-IMPAIRED				
Subj. nr.	Frequency in kHz 0.25-0.5	0.5-1	1-2	2-4	Subj. nr.	Frequency in kHz 0.25-0.5	0.5-1	1-2	2-4
1	107	105	100	97	1	98	98	92	93
	16	8	7	10		38	37	42	59
2	108	103	100	102	2	122	109	98	106
	16	6	2	4		63	59	57	74
3	111	99	96	95	3	117	117	110	110
	17	10	8	13		38	36	41	52
4	116	115	98	96	4	120	118	117	110
	13	6	2	6		58	49	38	46
5	122	121	110	108	5	117	113	115	116
	13	4	6	15		58	60	69	63
6	119	114	109	106	6	118	115	113	118
	16	8	4	10		60	56	58	68
7	121	120	120	119	7	96	94	95	95
	13	6	4	11		32	35	52	57
8	122	122	113	110	8	119	112	111	118
	15	6	2	3		61	56	42	49
9	118	115	113	110	9	120	120	118	117
	12	4	5	12		36	36	39	46
10	120	119	119	118	10	121	119	113	116
	8	4	4	10		48	49	50	55
11	111	113	115	107	11	101	92	92	98
	11	6	4	5		26	23	40	40
12	107	98	93	94	12	107	107	107	107
	14	11	3	4		38	30	41	60
M	115	112	107	105	M	113	110	107	109
	14	7	4	9		46	44	47	56
sd	6	9	9	9	sd	10	10	10	9
	3	2	2	4		13	12	10	10

("bisector") condition, speech and noise were shaped according to the dynamic-range bisector. In a final ("normal") condition, we used speech without spectral shaping according to the dynamic-range bisector, masked by a noise with the same spectrum as the speech. In this condition, overall noise level was such that in all octave bands levels were at least 20 dB above the listener's threshold.

The speech material consisted of 12 lists of 11 short (8 or 9 syllables) everyday Dutch sentences spoken by a female speaker (Plomp and Mimpen, 1979a). Because only 130 sentences were available, two sentences were presented twice in different lists. Sentences and noise were recorded on different tracks of a magnetic tape, so that their overall levels could be independently varied and they could be independently fed into a TMS 320-10 signal processor, before they were mixed. Amplification factors in the different filter bands were updated for each new sentence in order to realise the appropriate shaping of speech and noise spectra together with the desired compression.

Twelve normal-hearing listeners (age between 20 and 29 years), and twelve listeners with a sensorineural hearing impairment (age between 36 and 70 years), participated in the experiment as paid volunteers. For the normal-hearing group, pure-tone thresholds between 0.25 and 4 kHz were below 15 dB HL. For the hearing-impaired group, the hearing loss for pure tones averaged over 0.5, 1, and 2 kHz was between 39 and 57 dB for their best ear, and air-bone gap was less than 10 dB between 0.25 and 4 kHz. Average pure-tone thresholds with standard deviations for this group are given in Table I. Performance scores for monosyllables in quiet reached at least 90% intelligibility. Hearing thresholds and levels of uncomfortable loudness for octave bands of noise are given in Table II for individual listeners of both groups. The auditory range, averaged over the normal-hearing listeners, was 101, 105, 103, and 96 dB, for the octave bands from low to high, respectively. Averaged over the hearing-impaired listeners, this range was 67, 66, 60, and 53 dB, respectively.

Sentences in noise were presented monaurally to the listener's best ear over a headphone in a soundproof room. The twelve lists of sentences were presented in a fixed order. The twelve experimental conditions were distributed over the lists according to a digram-balanced design per twelve listeners (Wagenaar, 1969) in order to avoid the effects of learning and fatigue in the average results. The speech-reception threshold in noise represents the speech-to-noise ratio (in the frequency bands not containing the extra 20 dB of noise) at which the listener reproduced 50% of the sentences without a single error. Speech-to-noise ratio was varied by adjusting the level of the sentences relative to the level of the noise in an adaptive up-down procedure with a step size of 2 dB (Plomp and Mimpen, 1979a).

Octave-band levels in the "normal" and the "bisector" conditions are given in Table III for the individual normal-hearing and hearing-impaired listeners. Because levels for speech vary with speech-to-noise ratio, only levels for noise are presented. Octave-band levels for noise in the conditions with the extra noise in the 0.25-0.5kHz or 0.5-1kHz band, without the effect of compression, were equivalent to those in the "bisector" condition after adding 20 dB to the appropriate band level. The actual compression, expressed in dB, in the different frequency-selective and wideband conditions, can be read from Table IV. The compression as given in this table is computed for the speech-to-noise ratio representing the measured SRT in noise. Note that in wideband conditions, the compression, when expressed in dB, is often considerably smaller than in the corresponding frequency-selective conditions. This is because the increase of the wideband signal level caused by the extra local 20 dB of noise, may be low due to the presence of intense sounds in frequency bands other than the noise band. For each listener, speech and noise were in all experimental conditions well within the range between threshold level and the level of uncomfortable loudness in all frequency bands. As long as this condition is fulfilled, we may assume that in all conditions it is the masking effectiveness of the noise that determines the SRT.

Table III. Octave-band levels for noise (in dB SPL) in the "normal" condition (first row: Experiment 1; second row: Experiment 2) and the "bisector" condition (third row), for the individual normal-hearing and hearing-impaired listeners. Because in the "normal" condition the amplitude-frequency response is flat, subtraction of "normal" levels from the levels in the "bisector" condition gives the actual amplitude-frequency responses in the latter condition.

	NORMAL-HEARING					HEARING-IMPAIRED			
Subj.	Frequency in kHz				Subj.	Frequency in kHz			
nr.	0.25-0.5	0.5-1	1-2	2-4	nr.	0.25-0.5	0.5-1	1-2	2-4
1	46	42	33	30	1	95	91	82	79
	43	41	35	30		92	90	84	79
	62	57	54	54		68	68	67	76
2	40	36	27	24	2	110	106	97	94
	37	35	29	24		107	105	99	94
	62	55	51	53		93	84	78	90
3	49	45	36	33	3	88	84	75	72
	46	44	38	33		85	83	77	72
	64	55	52	54		78	77	76	81
4	42	38	29	26	4	82	78	69	66
	39	37	31	26		79	77	71	66
	65	61	50	51		89	84	78	78
5	51	47	38	35	5	102	98	89	86
	48	46	40	35		97	95	89	84
	68	63	58	62		88	87	92	90
6	46	42	33	30	6	104	100	91	88
	43	41	35	30		101	99	93	88
	68	61	57	58		89	86	86	93
7	47	43	34	31	7	93	89	80	77
	44	42	36	31		90	88	82	77
	67	63	62	65		64	65	74	76
8	39	35	26	23	8	85	81	72	69
	36	34	28	23		82	80	74	69
	69	64	58	57		90	84	77	84
9	48	44	35	32	9	82	78	69	66
	45	43	37	32		79	77	71	66
	65	60	59	61		78	78	79	82
10	46	42	33	30	10	91	87	78	75
	43	41	35	30		88	86	80	75
	64	62	62	64		85	84	82	86
11	41	37	28	25	11	76	72	63	60
	38	36	30	25		73	71	65	60
	61	60	60	56		64	58	66	69
12	40	36	27	24	12	96	92	83	80
	37	35	29	24		93	91	85	80
	61	55	48	49		73	69	74	84
M	45	41	32	29	M	92	88	79	76
	42	40	34	29		89	87	81	76
	65	60	56	57		80	77	77	82
sd	4	4	4	4	sd	10	10	10	10
	4	4	4	4		10	10	10	10
	3	3	5	5		11	10	7	7

Table IV. Compression (in dB) in Experiment 1 (steady-state conditions), as a function of noise band, type of gain control, and compression factor. The compression as given in this table is computed for the speech-to-noise ratio representing the (measured) SRT in noise. Values are presented for individual normal-hearing and hearing-impaired listeners.

NORMAL-HEARING	NOISE BAND 0.25-0.5kHz				NOISE BAND 0.5-1kHz			
	TYPE OF GAIN CONTROL: wideband		selective		wideband		selective	
Subj. nr.	COMPRESSION FACTOR 50%	100%	50%	100%	50%	100%	50%	100%
1	7	15	9	16	5	12	8	18
2	7	15	8	17	5	8	9	18
3	7	17	9	18	4	8	9	18
4	7	15	8	17	5	11	9	18
5	6	12	7	16	3	5	8	18
6	7	16	8	15	4	8	9	18
7	6	13	8	15	5	10	9	18
8	7	16	9	18	6	11	9	17
9	5	14	7	15	4	8	8	18
10	5	12	8	16	5	8	9	17
11	5	14	8	17	5	11	8	18
12	7	16	8	17	5	10	9	17
M	6.3	14.6	8.1	16.4	4.7	9.2	8.7	17.8
sd	0.9	1.6	0.7	1.1	0.8	2.0	0.5	0.5

HEARING-IMPAIRED	NOISE BAND 0.25-0.5kHz				NOISE BAND 0.5-1kHz			
	TYPE OF GAIN CONTROL: wideband		selective		wideband		selective	
Subj. nr.	COMPRESSION FACTOR 50%	100%	50%	100%	50%	100%	50%	100%
1	3	4	7	14	2	5	7	16
2	4	9	7	12	1	2	6	17
3	3	7	6	16	3	5	7	16
4	8	15	9	17	5	9	9	17
5	5	8	8	15	5	6	8	18
6	2	6	6	11	3	5	8	16
7	2	4	7	16	2	4	8	18
8	7	10	6	15	3	6	6	15
9	2	9	7	12	4	6	8	17
10	3	6	6	10	3	4	6	16
11	4	6	8	16	2	3	8	17
12	2	4	6	15	1	3	7	15
M	3.8	7.3	6.9	14.1	2.8	4.8	7.3	16.5
sd	2.0	3.2	1.0	2.3	1.3	1.9	1.0	1.0

B. Results

SRT-in-noise values for individual normal-hearing and hearing-impaired listeners are given in Table V. Mean thresholds for the two groups of subjects are plotted in Fig. 2 (upper panel: normal-hearing group; lower panel: hearing-impaired group).

It appears that the lower SRTs in noise were found in the "bisector" and "normal" conditions, which were the two conditions without the extra 20 dB of noise in the 0.25-0.5 or 0.5-1 kHz band. For the hearing-impaired group, however, SRT in the "normal" condition was clearly worse than in the "bisector" condition. Among the conditions with the extra local 20 dB of noise, the highest SRTs are generally found for conditions without (0%) compression or with wideband compression, for both groups of listeners.

An analysis of variance was carried out on the SRT data for frequency-selective and wideband compression conditions relative to the SRT for 0% compression. Results, presented in Table VI for the normal-hearing and hearing-impaired groups, showed that for both groups, the effect of type of gain control is significant. This effect is even highly significant for the hearing-impaired group. For this group, mean SRT in conditions with wideband compression remains practically unchanged relative to the SRT for 0% compression.

Figure 2 shows that, for both listener groups, the beneficial effect of frequency-selective over wideband compression was on the average small with the 20-dB increase of noise in octave band 0.25-0.5kHz. With the extra noise in octave band 0.5-1kHz, however, the curves for the two types of gain control diverged for increasing compression. In conditions with frequency-selective compression, the largest decrease in mean SRT relative to the SRT for 0% compression occurred for a compression factor of 100%, with, however, a substantial spread among subjects, particularly for the hearing-impaired group. Figure 2 also shows that masked SRT in conditions with noise band 0.25-0.5kHz was for both groups generally somewhat higher than in conditions with noise band 0.5-1kHz.

Comparison of the SRT-data for the two groups of listeners in Table V indicates that mean SRT in noise was in all experimental conditions higher for the hearing-impaired group, with larger standard deviations.

II. EXPERIMENT 2: TIME-VARYING CONDITIONS

A. Method

The same baseline condition of speech and noise was used as in Experiment 1. This time the noise level in the octave band 0.25-0.5kHz or 0.5-1kHz slowly increased by 20 dB relative to the bisector during presentation of the sentence. The noise increment, realised according to a linear function along a dB-scale, started at the beginning of each sentence and was completed after 1 s. During the remainder of the sentence, noise level was kept constant. Average duration of the sentences was 1.6 s. Compression was set in the same manner as in Experiment 1. However, because the total signal level was rising in time along with the 20-dB increase of noise, compression, expressed in dB, also increased in time, and did so simultaneously with the signal level. For simplicity of realisation, the increase in SPL of the

Table V. SRT in noise, expressed in dB speech-to-noise ratio (in those frequency bands not containing the extra 20 dB of noise) for Experiment 1 (steady-state conditions), as a function of noise band, type of gain control, and compression factor. The last two columns give the SRT for the two conditions without the extra 20 dB of noise. Results are presented for individual normal-hearing and hearing-impaired listeners.

NORMAL-HEARING	NOISE BAND 0.25-0.5kHz					NOISE BAND 0.5-1kHz						
	TYPE OF GAIN CONTROL											
	--	wideband		selective		--	wideband		selective			
Subj.	COMPRESSION FACTOR										bi-	normal
nr.	0%	50%	100%	50%	100%	0%	50%	100%	50%	100%	sector	
1	4.0	2.0	0.5	-0.5	1.5	1.0	1.0	-3.5	1.0	-1.5	-2.5	-4.0
2	1.5	2.5	1.5	0.5	1.0	-0.5	-1.5	0.5	-1.5	-3.0	-3.5	-6.0
3	6.5	3.0	-3.5	-1.0	-2.0	-0.5	0.0	-2.5	-3.0	-3.0	-5.0	-4.0
4	5.5	1.0	0.0	2.5	1.0	3.0	2.0	0.5	0.0	-1.0	-4.5	-2.5
5	9.0	4.5	4.5	5.5	1.5	5.5	6.0	8.5	1.0	-2.0	-0.5	-3.0
6	7.5	3.5	0.0	3.5	3.0	6.0	3.5	3.0	-1.5	-1.0	-2.5	-2.5
7	1.0	1.5	0.0	2.0	3.0	0.5	-0.5	-1.5	0.0	-2.5	-3.0	-5.0
8	0.0	3.0	-1.5	-1.0	-3.0	1.5	-1.0	0.0	-1.0	0.0	-3.5	-3.5
9	6.0	6.0	1.0	5.0	4.0	3.5	2.0	2.5	1.5	-1.5	-3.5	-4.5
10	7.0	4.5	0.0	1.5	2.0	0.0	0.0	3.0	-1.5	0.0	-4.0	-1.5
11	3.5	4.5	-1.5	1.0	0.0	3.0	1.5	1.0	2.5	-1.5	-2.5	-2.0
12	5.5	4.0	0.0	1.5	0.0	1.5	2.0	0.0	-0.5	0.5	-4.5	-5.0
M	4.8	3.3	0.1	1.7	1.0	2.0	1.3	1.0	-0.3	-1.4	-3.3	-3.6
sd	2.8	1.5	1.9	2.1	2.0	2.2	2.1	3.1	1.6	1.2	1.2	1.4

HEARING-IMPAIRED	NOISE BAND 0.25-0.5kHz					NOISE BAND 0.5-1kHz						
	TYPE OF GAIN CONTROL											
	--	wideband		selective		--	wideband		selective			
Subj.	COMPRESSION FACTOR										bi-	normal
nr.	0%	50%	100%	50%	100%	0%	50%	100%	50%	100%	sector	
1	7.5	5.5	7.5	6.0	5.0	4.5	7.5	6.0	6.0	2.5	-2.0	0.5
2	12.0	9.5	8.5	6.0	7.0	12.5	10.0	10.0	6.5	1.0	0.5	5.0
3	12.0	9.5	7.0	8.0	1.5	7.5	5.5	9.0	4.0	2.5	1.0	4.0
4	0.0	1.0	1.5	-1.0	1.0	1.0	1.0	2.5	-1.0	0.5	-4.0	-1.5
5	0.5	0.0	3.5	1.0	3.0	2.5	-2.0	4.5	0.5	-1.5	-4.5	-4.0
6	7.5	10.5	8.5	6.5	8.5	5.0	3.0	6.0	3.5	2.5	0.0	-3.0
7	6.5	1.5	1.5	6.0	2.5	3.0	5.0	3.0	2.5	-1.0	-1.5	2.0
8	11.5	3.0	8.0	8.0	3.5	6.0	7.5	6.5	7.0	3.5	-1.0	3.0
9	4.0	10.0	3.0	6.0	7.0	1.0	5.5	7.0	3.0	-0.5	-1.5	3.5
10	9.0	9.5	8.5	7.0	10.0	14.5	10.0	11.0	7.0	2.5	3.5	11.0
11	3.5	4.5	6.0	0.5	2.0	1.0	-0.5	4.0	0.5	-0.5	-1.0	-1.5
12	8.0	5.0	6.0	7.5	3.5	3.0	6.0	1.5	5.0	3.0	-2.5	2.0
M	6.8	5.8	5.8	5.1	4.5	5.1	4.9	5.9	3.7	1.2	-1.1	1.8
sd	4.1	3.9	2.7	3.1	2.9	4.4	3.9	3.0	2.7	1.8	2.2	4.1

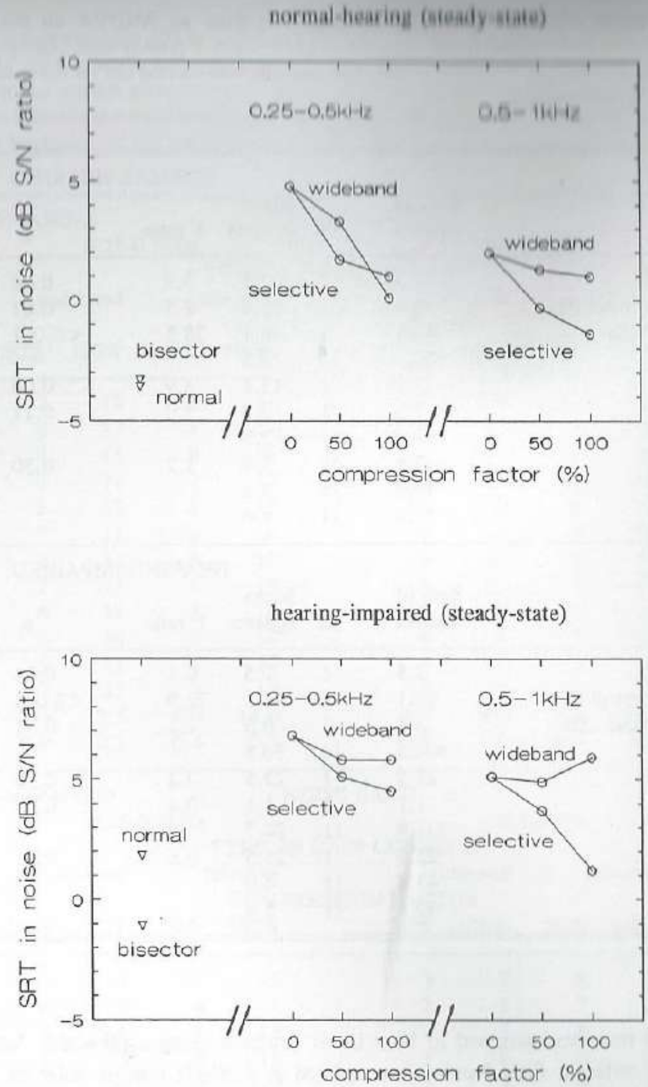


Fig. 2. Mean SRT in noise for the normal-hearing group (12 listeners, upper panel) and the hearing-impaired group (12 listeners, lower panel) in Experiment 1 (steady-state conditions) as a function of compression factor. Parameters are the octave band containing the extra noise and type of gain control. Mean SRTs for the two conditions without the extra noise are represented by triangles.

TABLE VI. Mean effects and first-order interactions from an ANOVA on the results for frequency-selective and wideband compression in Experiment 1 (steady-state conditions), for the normal-hearing and hearing-impaired group. The analysis was carried out on the masked SRT re the SRT for 0% compression.

NORMAL-HEARING					
Source	Sum of squares	df	Mean squares	F ratio	p
Noise band (N)	47.5	1	47.5	3.8	0.08
Type of gain control (G)	26.6	1	26.6	8.5	0.01
Compression factor (C)	46.1	1	46.1	28.3	<0.001
Subjects (S)	159.5	11	14.5		
N x G	13.1	1	13.1	3.9	0.08
N x C	7.3	1	7.3	3.0	0.11
N x S	137.9	11	12.5		
G x C	2.8	1	2.8	1.2	0.30
G x S	34.3	11	3.1		
C x S	17.9	11	1.6		

HEARING-IMPAIRED					
Source	Sum of squares	df	Mean squares	F ratio	p
Noise band (N)	2.5	1	2.5	0.1	0.76
Type of gain control (G)	91.1	1	91.1	22.9	<0.001
Compression factor (C)	6.3	1	6.3	1.2	0.31
Subjects (S)	404.3	11	36.8		
N x G	23.5	1	23.5	7.2	0.02
N x C	1.1	1	1.1	0.4	0.55
N x S	271.8	11	24.7		
G x C	25.5	1	25.5	9.6	0.01
G x S	43.8	11	4.0		
C x S	59.4	11	5.4		

total signal in time was assumed to be a linear function along a dB-scale. Amplification factors in the different filter bands were updated at a 30-Hz rate in order to realise the spectral shaping of speech and noise together with the desired compression.

Further details of the experimental conditions, apparatus, and stimulus presentation are as in Experiment 1. Again, 12 lists of 11 short everyday Dutch sentences were used. However, the sentences were different from those in Experiment 1, and were spoken by a male talker (Plomp and Mimpen, 1979a). Two of the 130 sentences were presented twice in different lists. In this experiment, the same subjects participated as in Experiment 1. It follows that octave-band levels for noise in the "bisector" condition were also the same (see Table III). Octave-band levels for noise in the "normal" condition differ from those in

Table VII. Compression (in dB) in Experiment 2 (time-varying conditions), as a function of noise band, type of gain control, and compression factor. The compression as given in this table is computed for the speech-to-noise ratio representing the (measured) SRT in noise. Values are presented for individual normal-hearing and hearing-impaired listeners. Asymptotic values reached 1 s after the beginning of the sentence are given.

NORMAL-HEARING	NOISE BAND 0.25-0.5kHz				0.5-1kHz			
	TYPE OF GAIN CONTROL							
	wideband		selective		wideband		selective	
Subj.	COMPRESSION FACTOR							
nr.	50%	100%	50%	100%	50%	100%	50%	100%
1	7	13	8	17	5	9	9	17
2	8	17	8	17	5	8	9	17
3	7	15	8	16	4	8	9	18
4	7	11	7	17	5	10	8	18
5	5	13	8	17	4	8	9	17
6	7	12	8	16	4	8	9	16
7	7	14	8	17	5	10	9	17
8	7	15	9	17	6	10	9	17
9	6	14	8	16	4	8	8	17
10	6	10	8	17	4	8	8	17
11	5	11	8	17	6	12	9	18
12	7	15	8	15	5	9	8	17
M	6.6	13.3	8.0	16.6	4.8	9.0	8.7	17.2
sd	0.9	2.1	0.4	0.7	0.8	1.3	0.5	0.6

HEARING-IMPAIRED	NOISE BAND 0.25-0.5kHz				0.5-1kHz			
	TYPE OF GAIN CONTROL							
	wideband		selective		wideband		selective	
Subj.	COMPRESSION FACTOR							
nr.	50%	100%	50%	100%	50%	100%	50%	100%
1	3	4	7	15	3	7	8	17
2	6	9	6	11	2	2	7	16
3	4	6	7	15	4	7	8	16
4	8	15	8	18	5	9	8	17
5	5	10	9	16	5	9	9	18
6	4	7	6	14	2	3	7	15
7	2	3	8	14	3	4	8	18
8	7	11	8	16	3	6	7	16
9	3	9	8	15	4	7	8	16
10	5	6	7	8	2	7	7	14
11	4	8	8	18	3	5	9	19
12	2	5	8	17	1	3	7	17
M	4.4	7.8	7.5	14.8	3.1	5.8	7.8	16.6
sd	1.9	3.3	0.9	2.9	1.2	2.3	0.8	1.4

Table VIII. SRT in noise, expressed in dB speech-to-noise ratio (in those frequency bands not containing the extra 20 dB of noise) for Experiment 2 (time-varying conditions), as a function of noise band, type of gain control, and compression factor. The last two columns give the SRT for the two conditions without the extra noise. Results are presented for individual normal-hearing and hearing-impaired listeners.

NORMAL-HEARING	NOISE BAND 0.25-0.5kHz					0.5-1kHz						
	TYPE OF GAIN CONTROL --	wideband		selective		--	wideband		selective			
Subj. nr.	COMPRESSION FACTOR 0%	50%	100%	50%	100%	0%	50%	100%	50%	100%	bi-sector	normal
1	3.5	2.5	4.0	3.5	0.5	1.5	0.0	2.0	-2.0	0.0	-3.5	1.0
2	5.0	0.5	-2.5	2.5	-0.5	0.0	-2.5	1.0	0.0	0.0	-3.0	-2.0
3	2.0	4.5	1.5	2.5	2.0	1.0	-3.0	-0.5	-0.5	-3.0	2.5	-4.5
4	5.0	3.0	7.0	4.0	0.5	6.0	1.5	3.5	1.5	-1.0	-0.5	-3.0
5	6.0	6.5	4.0	3.5	0.0	6.0	2.5	4.5	0.0	0.5	-2.0	-1.5
6	8.0	4.0	6.0	3.5	2.0	3.0	4.0	2.5	0.0	1.5	-1.0	-5.0
7	2.5	0.0	-1.0	0.5	-0.5	0.5	-1.5	-0.5	-1.5	-0.5	-2.0	-4.0
8	3.5	1.5	0.0	0.0	0.0	0.5	0.0	2.5	0.0	-0.5	-3.0	-2.0
9	7.5	3.5	1.0	2.5	2.0	3.0	1.5	3.0	2.0	-0.5	-2.5	-1.5
10	7.5	2.5	4.5	3.0	1.0	0.0	5.0	3.0	1.0	0.5	0.0	-4.0
11	4.0	3.5	3.0	1.0	0.5	1.0	-1.0	-0.5	-0.5	-1.5	-1.5	-3.0
12	4.5	2.5	1.0	1.0	3.0	3.0	-0.5	2.0	2.5	-0.5	-2.5	-0.5
M	4.9	2.9	2.4	2.3	0.9	2.1	0.5	1.9	0.2	-0.4	-1.6	-2.5
sd	2.0	1.8	2.9	1.3	1.1	2.1	2.5	1.7	1.3	1.1	1.7	1.8

HEARING-IMPAIRED	NOISE BAND 0.25-0.5kHz					0.5-1kHz						
	TYPE OF GAIN CONTROL --	wideband		selective		--	wideband		selective			
Subj. nr.	COMPRESSION FACTOR 0%	50%	100%	50%	100%	0%	50%	100%	50%	100%	bi-sector	normal
1	8.0	5.5	8.0	4.5	3.0	5.0	4.0	2.0	1.0	-0.5	-4.5	-0.5
2	11.0	5.5	9.0	8.0	8.5	11.5	6.0	10.5	5.0	2.5	3.0	9.5
3	8.0	5.5	7.5	4.5	4.0	6.5	2.5	5.5	2.5	2.5	-1.5	1.0
4	0.0	1.0	1.5	0.5	-2.0	1.0	1.0	2.0	1.0	-0.5	-2.5	-4.0
5	-0.5	1.0	0.0	-2.5	1.5	0.5	-1.0	-1.0	-2.5	-1.0	-4.5	-4.0
6	5.0	6.0	6.5	8.0	4.5	10.0	7.5	9.5	4.0	3.5	0.5	4.5
7	1.5	3.0	4.5	1.5	4.5	-1.5	-0.5	2.0	0.5	-1.5	-2.0	-0.5
8	3.0	3.5	6.5	2.0	1.5	4.0	6.5	6.0	4.5	1.5	2.5	3.0
9	3.0	6.5	1.5	2.5	3.0	3.5	2.5	5.0	1.5	1.5	-1.0	1.0
10	6.5	4.0	9.5	4.0	12.0	10.0	10.5	6.5	4.5	5.0	5.0	6.0
11	3.0	4.5	2.0	0.5	-2.0	-1.5	-2.0	-1.0	-2.0	-5.5	-2.0	-2.5
12	2.0	6.0	2.0	2.0	0.5	-1.0	2.5	2.5	4.5	1.0	-1.0	7.0
M	4.2	4.3	4.9	3.0	3.3	4.0	3.3	4.1	2.0	0.7	-0.7	1.7
sd	3.5	1.9	3.4	3.1	4.0	4.7	3.8	3.7	2.6	2.8	2.9	4.4

Experiment 1 because a voice with a different spectrum was used (see also Table III). Asymptotic band levels for the conditions with the extra noise in the 0.25-0.5kHz or 0.5-1kHz band, reached 1 s after the beginning of the sentence, and without the effect of compression, can be inferred from Table III in the same manner as in Experiment 1. The actual compression, expressed in dB, holding for the speech-to-noise ratio at threshold, in the different frequency-selective and wideband conditions can be read from Table VII.

B. Results

Masked SRTs for individual normal-hearing and hearing-impaired listeners are given in Table VIII. Mean thresholds for the two groups of subjects are plotted in Fig. 3 (upper panel: normal-hearing group; lower panel: hearing-impaired group).

An analysis of variance, carried out on the SRT data for frequency-selective and wideband compression conditions relative to the SRT for 0% compression (see Table IX), showed, for both groups of listeners, a significant effect of the type of gain control. As was found in Experiment 1 for steady-state conditions, this effect is most significant for the hearing-impaired group.

Comparison of Figs. 2 and 3, and of Tables V and VIII, shows that the same trends as observed in the SRT data for steady-state conditions apply to the data for time-varying conditions. However, in time-varying conditions, the increase of masked SRT relative to the SRT for the "bisector" condition was for both groups of listeners, somewhat smaller than in steady-state conditions.

III. GENERAL DISCUSSION AND CONCLUSIONS

The two experiments reported in the previous section, show that, in conditions with a 20-dB increase of noise in a single octave band, frequency-selective compression gives lower SRTs than wideband compression, for both the normal-hearing and hearing-impaired listener groups. The effect of frequency-selective versus wideband compression is comparable for steady-state and time-varying conditions. In this discussion, reference will be made to both experiments, unless stated otherwise.

The beneficial effect of frequency-selective over wideband compression is given by the difference in SRT between the two types of gain control, and is most apparent with the 20-dB increase of noise in octave band 0.5-1kHz. For this noise band, frequency-selective compression is on the average 1.7 and 2.7 dB more beneficial than wideband compression, for normal-hearing and hearing-impaired listeners, respectively.

There are three factors that, together, explain why masked SRT in conditions with the extra local 20 dB of noise is high compared to the SRT in the "bisector" (reference) condition, particularly in conditions without compression and with wideband compression:

[1] *Masking related to the speech-to-noise ratio* in the octave band with the extra 20 dB of noise. Only in the "bisector" and "normal" conditions, the noise has a spectrum identical to the long-term average spectrum of the sentences. Because in these conditions the speech-to-noise ratio is the same in all frequency bands, speech components in all bands

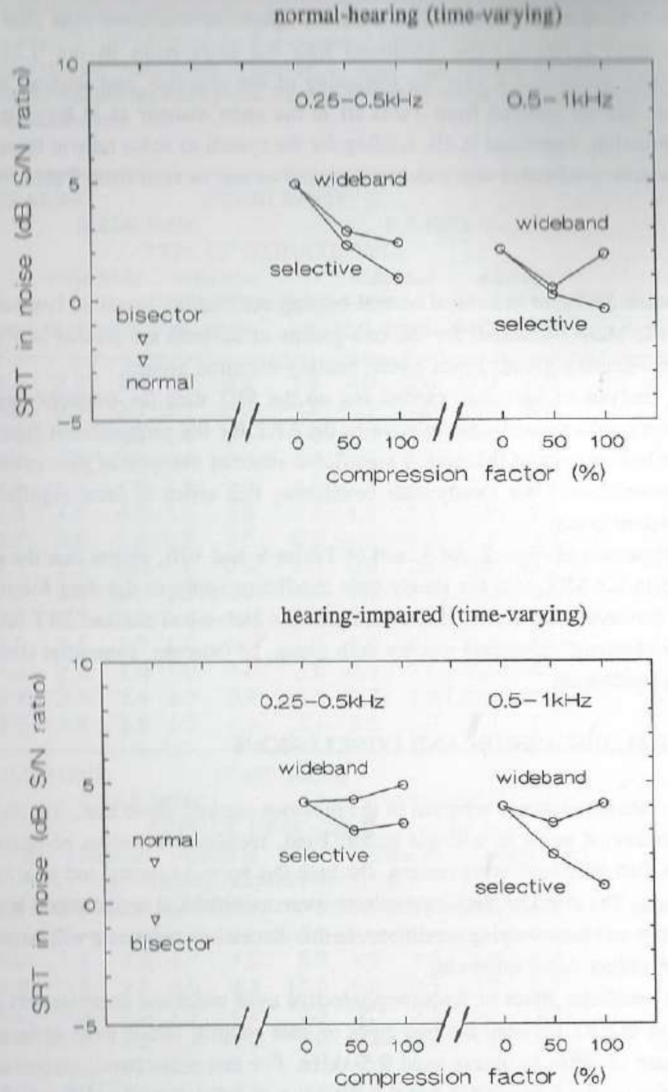


Fig. 3. Mean SRT in noise for the normal-hearing group (12 listeners, upper panel) and the hearing-impaired group (12 listeners, lower panel) in Experiment 2 (time-varying conditions) as a function of compression factor. Parameters are the octave band containing the extra noise and type of gain control. Mean SRTs for the two conditions without the extra noise are represented by triangles.

Table IX. Mean effects and first-order interactions from an ANOVA on the results for frequency-selective and wideband compression in Experiment 2 (time-varying conditions), for the normal-hearing and hearing-impaired group. The analysis was carried out on the masked SRT re the SRT for 0% compression.

Source	Sum of squares	df	Mean squares	F ratio	p
NORMAL-HEARING					
Noise band (N)	36.3	1	36.3	1.7	0.21
Type of gain control (G)	32.7	1	32.7	8.5	0.01
Compression factor (C)	2.0	1	2.0	3.9	0.07
Subjects (S)	59.9	11	5.4		
N x G	0.4	1	0.4	0.6	0.45
N x C	10.7	1	10.7	5.8	0.03
N x S	228.9	11	20.8		
G x C	12.8	1	12.8	10.6	0.01
G x S	42.2	11	3.8		
C x S	5.7	11	0.5		
HEARING-IMPAIRED					
Noise band (N)	34.4	1	34.4	3.5	0.09
Type of gain control (G)	93.0	1	93.0	74.4	<0.001
Compression factor (C)	0.4	1	0.4	0.1	0.75
Subjects (S)	307.0	11	27.9		
N x G	6.3	1	6.3	1.6	0.24
N x C	1.4	1	1.4	0.4	0.57
N x S	109.3	11	9.9		
G x C	7.3	1	7.3	2.1	0.17
G x S	13.8	11	1.3		
C x S	46.0	11	4.2		

contribute to intelligibility. However, in conditions with the extra 20 dB of noise in a single octave band, the decrease of speech-to-noise ratio in the noisy octave band causes a loss of useful speech information.

[2] Masking of the higher frequency components of the speech signal by intense lower frequency sounds in the octave band with the 20-dB increase of noise, i.e. *upward spread of masking*. For a low-frequency noise with an SPL of 80 dB, the slope of the masking pattern for normal-hearing listeners can be estimated to be between -20 and -30 dB/oct (cf. Ehmer, 1959). For sensorineurally hearing-impaired listeners in similar acoustical conditions, Jerger et al. (1960) found that this slope is only about -6 dB/oct. For our masker conditions, this

suggests that for normal-hearing listeners upward spread of masking may be limited to one octave band adjacent to the noisy octave band, whereas for hearing-impaired listeners, it may affect all bands higher in frequency than the noisy octave band. The masker slope for hearing-impaired listeners suggests that spread of masking may even occur in the "normal" condition without the extra 20 dB of noise, where the spectral slope of speech and noise is between -3 and -9 dB/oct. This may explain why, for these listeners, masked SRT in the "normal" condition is high compared to SRT in the "bisector" condition, where the average spectral slope of speech and noise is practically flat (see Table III).

It may be inferred from the above, that particularly in conditions with 0% compression, where the most intense low-frequency maskers were used, the negative effect of upward spread of masking would be expected to be much larger for the hearing-impaired group than for the normal-hearing group. Nevertheless, the corresponding increase in SRT relative to the SRT in the respective "bisector" condition is on the average equally large for the normal-hearing and the hearing-impaired group (see Tables VI and VII). The generally larger interindividual spread for the latter group, on the other hand, agrees with the notion that there exists considerable variability among hearing-impaired listeners in their individual upward spread of masking (cf. Martin and Pickett, 1970).

[3] *Variation of speech and noise level in time*. Whereas the first two factors refer to loss of information, a third factor may be added that refers to the possible distraction of the listener, in time-varying conditions, by the slow changes in signal levels during sentence presentation. These changes in signal levels result from the 20-dB increase of noise alone (in conditions without compression), or from the combined changes in the level of speech and noise (in conditions with frequency-selective or wideband compression).

However, a comparison of Tables VI (steady-state conditions) and VII (time-varying conditions) shows that the SRT, after subtracting SRT in the respective "bisector" condition in order to correct for speaker-related differences, is generally lower in time-varying than in steady-state conditions. The average difference is 1.5 to 2 dB for both groups of listeners. This suggests that, in time-varying conditions, speech understanding is facilitated by the less-masked first part of the sentence, without being critically affected by the variation itself. This result is a prerequisite for the succes of long time constants for the gain to adapt itself to changing acoustical situations (see Introduction).

Because compression leaves the speech-to-noise ratio in the different frequency bands intact, the effect of masking caused by the low speech-to-noise ratio in the band with the extra 20 dB of noise, referred to in [1], is the same for all compression conditions. Given this constant role of masking related to the speech-to-noise ratio, the large threshold-differences among the various compression conditions must, at least in the steady-state experiment, be attributed to the greater or smaller effect in these conditions of upward spread of masking (see [2]).

Since in conditions with frequency-selective compression, exclusively the frequency band containing the excessive 20 dB of noise is attenuated, the frequency range of speech contaminated by the spread of masking will shrink with each dB of attenuation. It can be seen in Table IV that the attenuation corresponding to 100% compression was maximally about 17

dB, in both groups of listeners. There are indications in the literature (Ehmer, 1959; Gagné, 1988), referring to normal-hearing listeners, that for a reduction in the SPL by this order of magnitude, the slope of the internal masker curve for a low-frequency masker may be steepened by 10 to 20 dB/oct. This gives an additional reduction of the effect of upward spread of masking in conditions with frequency-selective compression.

Wideband compression, unlike frequency-selective compression, leaves the relation among the different band levels unchanged. As a result, a reduction of upward spread of masking may only be possible through the above-mentioned steepening of the masker slope, provided the attenuation of the signal is substantially large. However, as was explained earlier, the compression in decibels in wideband conditions is generally small (see also Tables IV and VII). Therefore, the limited benefit observed for wideband compression, with even a lack of any systematic effect for the hearing-impaired group, is not surprising. It can be seen in Tables IV and VII that for the hearing-impaired group, the compression in dB in wideband conditions may be particularly small. This is because the increase of the wideband signal level, caused by the extra local 20 dB of noise, was extra limited due to the higher speech-to-noise ratios at threshold and the more strongly amplified high-frequency bands (see Table III) for this subject group.

The differences between the two types of gain control, referred to above, suggest that the more effective suppression of upward spread of masking in conditions with frequency-selective compression is responsible for the superiority of frequency-selective over wideband compression. In frequency-selective conditions, the decrease in mean SRT relative to the SRT for 0% compression is, for both listener groups, maximal for a compression factor of 100%, and may be as large as 3.5 to 4 dB. It appears that the reduction of upward spread of masking is most effective when the increase of signal level in the noisy octave band is fully compressed.

Because in conditions with 100% frequency-selective compression, octave band levels of the total signal (consisting of speech plus noise) are identical to those in the "bisector" condition (provided the speech-to-noise ratio during the adaptive threshold-estimation procedure is the same), it may be assumed that upward spread of masking is suppressed to the same level as in the "bisector" condition. This suggests that the difference in SRT between these conditions exclusively reflects the loss of speech information due to the decreased speech-to-noise ratio in the octave band with the extra 20 dB of noise (see [1]).

Figs. 2 and 3 show that the difference in SRT between 100% frequency-selective compression conditions and the respective "bisector" condition is larger for noise band 0.25-0.5kHz than for noise band 0.5-1kHz, for both groups of listeners. It can be seen that, for the latter noise band, 100% frequency-selective compression even restores mean SRT to within 2 dB above the SRT in the "bisector" condition. It appears that information loss in octave band 0.25-0.5kHz is more critical than a loss in octave band 0.5-1kHz. This agrees with a finding by Studebaker et al. (1987), that for continuous discourse, the frequency region of maximum perceptual importance is between 0.4 and 0.5 kHz. The more critical loss in the 0.25-0.5kHz band is seen for all of the corresponding five compression conditions in the finding that SRT is on the average 1 to 2 dB higher than for the 0.5-1kHz band, for both groups of listeners. It should be remembered that loss of speech information in the noisy octave band itself affects SRT in all compression conditions in the same way.

In summary, we may conclude that, in conditions with a strong increase of the noise level in a single octave band, adapting a frequency-selective amplification factor is superior to adapting a single wideband amplification factor, for normal-hearing and, more strongly, for hearing-impaired listeners. This applies both when the increased noise level is steady-state and when the increase develops in time. For both groups of listeners, frequency-selective compression of the signal in the noisy octave band gives a maximum decrease in masked SRT of 3.5 to 4 dB (relative to the SRT for 0% compression) for a compression factor of 100%.

CHAPTER 5

CONTROLLING THE GAIN IN A FOUR-CHANNEL HEARING AID BY THE MINIMA IN THE TEMPORAL ENVELOPE OF THE SOUND

Janette N. van Dijkhuizen, Joost M. Festen, and Reinier Plomp
submitted for publication in: Audiology

ABSTRACT

Conventional automatic gain-control (AGC) acts upon the average sound level, irrespective of whether there is a speech signal or not. This has the disadvantage that background noise during periods without a speech signal is amplified to levels experienced as "noisy" by the listener. This annoyance can be reduced by using the level of the minima in the sound envelope rather than the average level to control the gain. Such a gain control will not affect a speech signal because of its intensity fluctuations. However, masking noise reduces these fluctuations and activates the control. A multichannel version of this control can selectively attenuate only the frequency bands in which noise is present. This paper studies the effectiveness of a four-channel AGC system, in which the frequency-dependent amplification factor is automatically controlled by the minima of the temporal sound envelope in the respective frequency channel. The effect of a condition with the gain control in all channels switched off is the reference. Results for 10 listeners with a sensorineural hearing impairment show that, for various sounds frequently interfering in practice and with spectra that are roughly comparable to that of the speech signal, the condition with gain control does not affect the speech-reception threshold (SRT) in noise, but substantially reduces the subjective impression of noisiness when no speech communication takes place. This holds particularly for interfering sounds with a more or less continuous character, like stationary noise or music, where the AGC is most active. For these sounds it was found that for a given increase in input level, the corresponding growth in perceived noisiness is equivalent to the growth in perceived noisiness produced by only about one-fifth of that increase (in decibels) in a condition without gain control.

INTRODUCTION

An essential condition for optimal speech intelligibility by the hearing impaired is that the speech signal is presented to the ear at such a level that its information-bearing fluctuations are audible over a frequency span which is as wide as possible. This condition can be met by means of a multichannel automatic gain-control (AGC) hearing aid, in which, under various acoustical situations, the frequency-dependent amplification adapts itself to the level and spectrum of the incoming sound.

Usually, AGC circuits are designed in such a way that they act on the average level of the incoming sound. This means that they do not discriminate between speech and interfering noise. As a result, during periods without speech, the hearing aid amplifies background noise up to a level not wanted by the hearing-impaired person. The complaint that the hearing aid is so "noisy" is a major disadvantage of the application of AGC.

In the opinion of the authors, this disadvantage can be reduced substantially by using the level of the minima in the temporal sound envelope rather than the average level to control the gain. Since the information in speech is carried by its (rapid) intensity fluctuations in the various frequency bands, there is no need to amplify sounds as long as these fluctuations are absent; in that case no speech signal is present or, if present, masked by interfering noise. By using the minima in the sound envelope as an AGC criterion, the speech signal can be presented at an appropriate level whereas noise alone or masked speech will be attenuated. As most interfering sounds, including voice babble and music, vary over a much narrower range of amplitudes than the signal of a single voice, the discomfort of AGC will be substantially reduced. Of course, such an AGC will be most effective in case of steady-state noises varying only slowly in time.

In this article a study into the effectiveness of an amplification controlled by the minima in the temporal sound envelope, rather than by the average level, is described. Because the shape of the noise spectrum may vary, the maximal effect can be expected for a frequency-dependent AGC. In our case, this is realised by splitting up the speech-frequency range in four channels. An additional advantage of such a system is that it will reduce the effect of spread of masking to other frequency regions if intense frequency-limited interfering sounds are present (cf. Festen et al., 1990; Zurek and Rankovitz, 1990; van Dijkhuizen et al., 1991a).

The amplitude-frequency response of such a hearing aid will change in time along with variations of the background noise, preferably taking about 0.5 s to adjust itself. No conclusive results have been reported in favor of much shorter time constants as used in syllabic compression (cf. Braida et al., 1979). An argument against short attack and release times is that they will reduce the information-bearing fluctuations of speech within each frequency channel, and this will negatively affect the speech-reception threshold (Plomp, 1988). A time constant of about 0.5 s, as suggested above, seems to be a good compromise to preserve the intensity fluctuations within syllables whereas background noise, usually varying slower in time, is suppressed.

Because the variations of the amplitude-frequency response affect both speech and interfering sounds, speech-to-noise ratios in the different frequency bands remain unchanged.

Nevertheless, the speech-reception threshold (SRT) in noise may still be affected by the shape of the amplitude-frequency response or by its rate of change. This problem was investigated in earlier studies with normal-hearing and hearing-impaired listeners (van Dijkhuizen et al., 1987 and 1989). In these studies, the noise had a spectrum identical to the long-term average spectrum of the sentences, as for frequent listening situations where the human voice itself is the interfering sound (e.g. a second speaker, speech babble). Because speech-to-noise ratio was the same in all frequency bands, speech components in all bands could contribute to intelligibility. For normal-hearing listeners van Dijkhuizen et al. (1987) found that the SRT for sentences in noise is almost unaffected when the slope of the amplitude-frequency response is steady-state within a range between -7 and +10 dB/oct, or when this slope is slowly changing in time between -5 and +5 dB/oct, or vice versa (transition time between 0.125 and 2 s). Further steepening the falling slope of the amplitude-frequency response appears to increase the risk of upward spread of masking. For hearing-impaired listeners van Dijkhuizen et al. (1989) found that, in conditions with steady-state amplitude-frequency responses or with responses that are slowly changing in time (transition time 0.25 or 1 s), the increase of masked SRT is at most 2 dB as long as the negative slope of the speech and noise spectrum is not steeper than -3 dB/oct relative to the line bisecting the ear's dynamic range. Steeply rising amplitude-frequency responses do not have a critical effect on the SRT. Because in practice most disturbing sounds have their strongest components in the lower frequencies (cf. Kryter, 1970), a rising rather than a falling hearing aid response will also be most appropriate for the purpose of reducing the discomfort caused by such sounds.

The relative insensitivity of masked SRT to variations of the amplitude-frequency response, reported in the above studies for both normal-hearing and hearing-impaired listeners, is an important requisite for the effectiveness of frequency-dependent automatic gain-control in hearing aids. A follow-up study (van Dijkhuizen et al., 1991a) investigated the beneficial effect on intelligibility, by a reduction of spread of masking, of adjusting the amplitude-frequency response in situations of intense interfering noise with a narrow-band maximum. The effect of adjusting a single wideband amplification factor was the reference. The noise had a spectrum identical to the long-term average spectrum of the speech, except that in one low-frequency octave band its level was increased by 20 dB. Hence, speech components within that noise band could not contribute to intelligibility. For both normal-hearing and hearing-impaired listeners it was found that frequency-dependent compression of the signal in the octave band with the 20-dB increase of noise is more beneficial than wideband compression, and gives a decrease in SRT of maximally about 4 dB relative to the SRT obtained without compression. This applies both when the increased noise level is steady-state and when the increase of noise level, together with the gain, develops slowly in time (transition time 1 s).

The results reported in the above studies are promising for the success of slow-acting frequency-dependent AGC in hearing aids. In these studies, masker spectra as well as variations in the amplitude-frequency response were still experimentally controlled. The aim of the present paper is to study the practical merits of multichannel AGC in which the fast fluctuations typical of the speech signal are preserved whereas the level of the slower varying background noise is selectively reduced. The following question has been studied: what is, for various interfering sounds commonly found in practice, the effect of automatic control of the frequency-dependent gain by the minima in the temporal sound envelope on (1) speech intelligibility, and (2) comfort of listening, as determined by the subjective impression of noisiness?

The present experiments were carried out for listeners with a sensorineural hearing impairment. In *Experiment 1*, we compared, for five conditions of sounds frequently interfering in normal listening situations, the masked SRT of short meaningful sentences with and without automatic control of the gain by the minima. As was referred to earlier, speech level contrasts should, in order to contribute to intelligibility, be presented above threshold level and below the level of uncomfortable loudness in all frequency bands (cf. Skinner et al., 1982). Therefore, in all conditions with and without gain control, amplification factors in the different frequency bands were adjusted in such a way that the speech signal was presented at a level warranting 100% intelligibility in quiet, with speech components well above threshold level and below the level of uncomfortable loudness at all frequencies. In conditions with gain control, additionally, the amplification per frequency band was automatically reduced if, by an increase of the noise level, the minima in the signal envelope exceeded a critical level. Adaptation of the amplification was such that the increase of minima relative to this onset level was fully reduced.

In *Experiment 2*, we determined, for the same listeners that participated in Experiment 1, the (wideband) attenuation of the interfering sound required in conditions without gain control in order to equate the perceived noisiness with and without gain control. This was done for the five different types of interfering sound. According to Kryter (1970), perceived noisiness is "the subjective impression of the unwantedness of a not unexpected, nonpain or fear-provoking sound as part of one's environment". Noisiness tends to grow with sound level, although other physical properties of the sound, such as spectral contents and duration, as well as its psychological meaning to the listener, may also play a role (cf. Kryter, 1970). In order to reduce the sensation of noisiness in practice, however, the hearing-aid user only has access to a manual gain-control for the wideband signal level. The reason why we did not use a perceptual scale to measure the difference in noisiness, but instead determined the level adjustment for equal noisiness, was that this kept the task for the subject as simple as possible. For each listener, the amplification factors with the gain control switched off, as well as the onset level of gain control, were, for each of the different frequency bands, the same as in Experiment 1.

Many experiments on the effect of adaptive filtering techniques, aimed at reducing the disturbing effects of noise in normal speech communication situations, have been carried out in the past (e.g. Ono et al., 1983; Stein and Dempesy-Hart, 1984; Graupe et al., 1987; van Tasell et al., 1988; Klein, 1989; Stein et al., 1989; Tyler and Kuk, 1990). However, filtering easily results in the loss of speech components that may still contribute to intelligibility. Because in all our experimental conditions, gain is adjusted to present, under all acoustical conditions, contributing speech fluctuations well within the listener's auditory range at all frequencies, the present experiments cannot easily be related to experiments on the effect of filtering. Neuman and Schwander (1987) evaluated an adaptive filtering system in which the importance of preserving speech information was recognized. In the evaluation,

however, no conditions of critical speech-to-noise ratios were included.

I. EXPERIMENT 1 : EFFECT ON SPEECH INTELLIGIBILITY

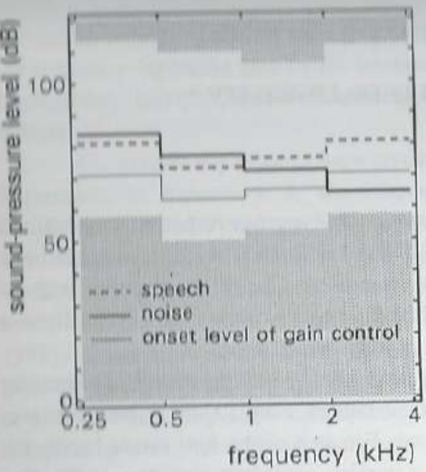
A. Method

Various fixed and dynamically varying amplitude-frequency responses were realised by means of four parallel octave filters of 0.25-0.5, 0.5-1, 1-2, and 2-4 kHz. Six-pole elliptic filters with slopes of approximately 40 dB/oct were used. The filters, together with an algorithm that controlled the gain in the individual filter bands, were realised digitally in a programmable TMS 320E15 signal processor (sampling rate 10 kHz).

For each listener the dynamic range of hearing, limited by the threshold of hearing and the level of uncomfortable loudness, was first measured using 300-ms bursts of octave-band noise separated by silent intervals of 100 ms. For each of the four octave bands, the threshold was determined according to a Békésy tracking procedure (step size 1 dB). The level of uncomfortable loudness per octave band was determined according to a procedure in which the noise level is increased by 1 dB for each new 300-ms burst, and the listener has to push a button when the noise was experienced as uncomfortably loud.

Prior to the masked SRT, the SRT in quiet was determined, representing the level at which 50% intelligibility is obtained, for speech that was spectrally shaped according to the listener's threshold-of-hearing for octave-band-filtered noise. In all experimental conditions testing masked SRT, i.e. both with and without gain control, the amplification in the different filter bands was adjusted to shape the long-term average spectrum of the speech signal according to the listener's threshold-of-hearing, and raise its overall level 20 dB above the (measured) SRT in quiet. For the same sentence material as used in the present experiment (but not spectrally shaped), Plomp and Mimpen (1979a) found that normal-hearing listeners obtain 100% intelligibility in quiet at presentation levels of 3 to 4 dB above the SRT and higher. This suggests that, even for the hearing-impaired listeners in this study, the above adjustment of amplification should easily provide 100% intelligibility in quiet. For each new sentence, the overall level of the masker was changed relative to the level of the speech in an adaptive up-down procedure in order to determine masked SRT, as will be explained below.

In conditions with gain control, additionally, the amplification per octave band was automatically reduced if the level of the noise in the respective filter-band, as estimated by the level of the envelope minima, exceeded the onset level of gain control in that band. The onset level was chosen 10 dB below the long-term average band level of the speech. The amplification factor was reduced to such an extent that the increase of minima relative to the onset level of gain control was fully compressed, and was computed as follows. Because the amplification has to be inversely related to the level of the minima, the inverse of the amplitude envelope was calculated for each filter-band output, and from this signal the peak values were determined in order to find the amplification required in that band to attach the minima to the desired level. The amplification per band was smoothed by a low-pass filter in order to avoid abrupt changes. The resulting attack time of compression with an increase



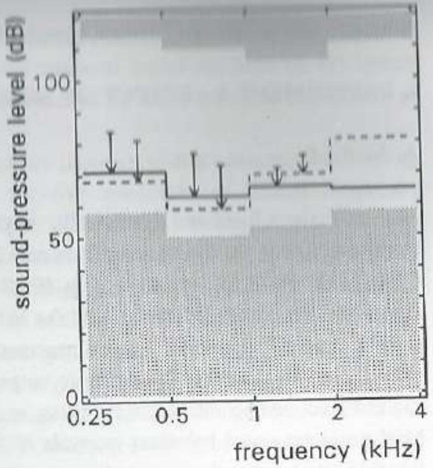


Fig. 1a. Example of how the spectra for speech and noise were shaped in Experiment 1 for a condition without gain control. Spectra are expressed in octave-band levels. The listener's dynamic range of hearing, limited by the threshold of hearing and the level of uncomfortable loudness, is represented by the non-shaded area. The amplification in the different frequency bands was adjusted in two steps: first, the long-term average spectrum of the speech was shaped according to the listener's threshold-of-hearing, and second, its overall level was raised 20 dB above the SRT in quiet for this speech signal. The spectrum of the noise was shaped by the same transmission as the speech, and the level of the noise was varied in order to determine masked SRT. In this example, SPLs in the four octave bands from low to high are for speech: 81, 73, 76, and 81 dB, and for noise: 84, 77, 72, and 65 dB, respectively.

Fig. 1b. As Fig 1a, however, for a condition with gain control. Per octave band, the amplification was automatically reduced if, with the gain control switched off, the level of the envelope minima exceeded the onset level of gain control, 10 dB below the level of the speech. With the gain control switched off (see Fig. 1a), the level of the envelope minima, for simplicity assumed to be equal to the level of the noise, exceeded the onset level of gain control by 13, 14, and 6 dB, in the 0.25-0.5, 0.5-1, and 1-2 kHz band, and was below this onset level in the 2-4kHz band. In the condition with gain control, the reduction of amplification was such that the increase of minima relative to the onset level of gain control was fully compressed, as shown by the length of the arrows.

of noise level was about 575 ms in all four octave bands. In all experimental conditions with and without gain control, amplification was applied to the total signal, consisting of speech and masker. Examples of how the spectra of speech and noise were shaped are given in Figs. 1a and b.

The following sounds were used as maskers: (1) running speech by a single male talker, (2) speech babble, consisting of running speech by two male and two female talkers, (3) noise with a spectrum identical to the long-term average spectrum of the target sentences, (4) noise recorded inside a car driving at a constant 80 km/h on a smooth road, and (5)

music, consisting of piano and background orchestra. Figure 2 shows long-term average spectra, in one-third-octave bands, for the five different maskers and the target speaker. It appears that all maskers have falling spectral slopes which are grossly similar to that of the target speech.

Five maskers (single speaker, speech babble, speech noise, car noise, and music) and two gain modes (with and without gain control) give ten experimental conditions. Together with two conditions testing the SRT in quiet (also with and without gain control), this gives a total of twelve different experimental conditions. In the quiet conditions, speech was spectrally shaped according to the listener's threshold-of-hearing for octave-band-filtered noise, and its overall level was changed for each new sentence in an adaptive threshold-estimation procedure. This time, the onset level of gain control in each frequency band was set equal to the listener's threshold-of-hearing level in that band. The SRT in quiet measured without gain control was adopted as the reference for conditions testing masked SRT, as referred to above.

The speech material consisted of twelve lists of 11 short (8 or 9 syllables) everyday Dutch sentences spoken by a female talker (Plomp and Mimpen, 1979a). Because only 130 sentences were available, two sentences were presented twice in different lists. Sentences and maskers were separately stored on computer disk, so that their overall levels could be varied independently before they were mixed. The mixed signal of speech and masker, or (in quiet conditions) the speech signal alone, was presented through a loudspeaker in a room with

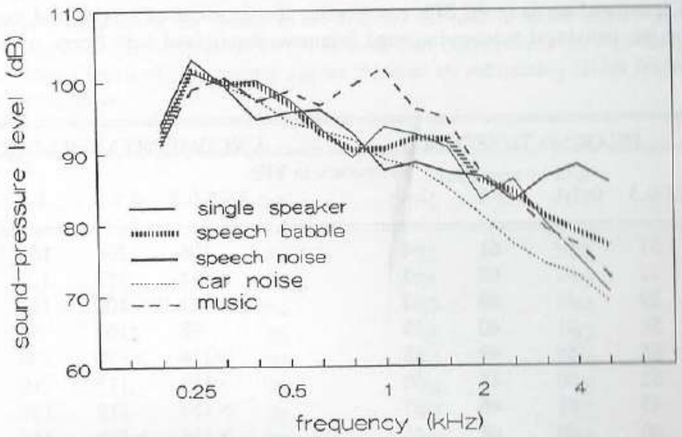


Fig. 2. Long-term average spectra, in one-third-octave bands, for the five different maskers used in Experiments 1 and 2. The long-term average spectrum of the target sentences as used in Experiment 1 coincides with that of speech noise. Spectra are shown for equal overall SPL.

Table I. Mean pure-tone air-conduction thresholds in dB HL for the ten hearing-impaired listeners participating in the Experiments 1 and 2, with standard deviations.

	Frequency in kHz 0.25	0.5	1	2	4
M	25	37	48	51	58
sd	10	12	12	9	15

dimensions of 6m x 7m x 7m, and reverberation time of about 0.4 s for frequencies between 0.25 and 4 kHz. Background noise level in this room was very low (about 35 dBA). A microphone was located at a distance of 1.5 m from the loudspeaker, corresponding to the critical distance where direct and reverberant energy densities are equal (Kuttruff, 1973). Loudspeaker and microphone were both 1.3 m above floor level. In masked conditions, speech level at the position of the microphone was a constant 65 dBA. The sound signal picked up by the microphone was digitized, and fed into a TMS 320E15 signal processor for filtering and adjustment of the frequency-dependent gain per individual listener. Sentences in quiet or in noise, depending on the experimental condition, were presented monaurally to the listener's best ear via a hearing aid telephone. Listeners were seated in an adjacent soundproof

Table II. Octave-band levels in dB SPL representing the threshold of hearing and uncomfortable loudness for the individual hearing-impaired listeners, determined with bursts of octave-band-filtered noise.

Subj. nr.	HEARING THRESHOLD 0.25-0.5	0.5-1	1-2	2-4	UNCOMFORTABLE LOUDNESS 0.25-0.5	0.5-1	1-2	2-4
	Frequency in kHz							
1	57	59	61	64	106	104	101	99
2	52	61	69	65	>114	113	114	103
3	29	42	69	62	111	109	111	117
4	58	41	40	59	98	103	98	101
5	66	56	49	58	>114	>120	120	117
6	52	50	45	60	>114	117	110	103
7	45	41	46	57	>114	119	126	>123
8	60	69	68	45	>114	>120	116	106
9	61	57	49	43	102	97	103	96
10	57	44	54	48	>114	>120	126	105
M	54	52	55	56				
sd	10	10	11	8				

room. This separate room for the listening experiment was used in order to avoid direct radiation from the loudspeaker to the listener's ear. Maximum output levels of the system, measured in a 2-cc coupler (Brüel & Kjaer DB 0138), were, with the gain control switched off, 114, 120, 127, and 123 dB SPL for the octave bands from low to high, respectively.

Ten listeners with a sensorineural hearing impairment (age between 27 and 65 years) participated in the experiment as paid volunteers. The hearing loss for pure tones averaged over 0.5, 1, and 2 kHz was between 35 and 55 dB for their best ear, and the air-bone gap was always less than 10 dB between 0.25 and 4 kHz. Mean pure-tone air-conduction thresholds with their standard deviations are given in Table I. Performance scores for monosyllables in quiet reached at least 90%. Hearing thresholds and levels of uncomfortable loudness for octave bands of noise are given in Table II for the individual listeners. In the 0.25-0.5, 0.5-1, and 2-4 kHz band, uncomfortable loudness levels exceeded the maximum output of the system for 6, 3, and 1 of the 10 listeners, respectively.

Twelve lists of sentences were presented in a fixed order. Whereas the two experimental conditions testing the SRT in quiet were alternated over the first two lists, the ten conditions testing masked SRT were distributed over the remaining ten lists according to a digram-balanced design per ten listeners (Wagenaar, 1969) in order to avoid effects of learning and fatigue in the average results. The SRT represents the speech level in quiet conditions or the speech-to-noise ratio in masked conditions, at which the listener reproduced 50% of the sentences correctly. Speech level or speech-to-noise ratio, depending on the experimental condition, was varied in an adaptive up-down procedure with a step-size of 2

Table III. Octave-band levels for speech in dB SPL for the individual hearing-impaired listeners in the conditions in which the masked SRT was measured (Experiment 1), with the gain control switched off. Onset levels of gain control can be obtained by subtracting 10 dB from the octave-band levels in this table.

Subj. nr.	Frequency in kHz 0.25-0.5	0.5-1	1-2	2-4
1	83	85	87	90
2	82	91	99	95
3	54	67	94	87
4	80	63	62	81
5	93	83	76	85
6	76	74	69	84
7	71	67	72	83
8	80	89	88	65
9	87	83	75	69
10	77	64	74	68
M	78	77	80	81
sd	10	11	12	10

dB (Plomp and Mimpen, 1979a). Variations in the speech-to-noise ratio were applied by adjusting the level of the masker relative to the level of the speech.

Octave-band levels for speech (with the gain control switched off) in the conditions in which the masked SRT was measured, are given in Table III for the individual listeners. Because levels for noise vary with speech-to-noise ratio and with the specific masker used, only levels for speech are presented. Onset levels of gain control can be obtained by subtracting 10 dB from the band levels in this table.

B. Results

Figure 3 shows the long-term average sound level in a condition with gain control as a function of the level without gain control, for the target speech and the five different maskers, represented for each of the four octave bands separately. It can be seen that below the onset level of gain control, sound level in the condition with gain control is equal to that in the condition without gain control. The deviations from the linear relation (diagonals) are the result of the reduction of amplification when the level of the envelope minima exceeds the onset level of gain control. As had to be expected, this deviation is largest for the more

Table IV. Speech-reception thresholds for individual hearing-impaired listeners, in Experiment 1. Column 2: SRT in quiet for a condition with gain control relative to a condition without gain control, expressed in dB (measured at the position of the hearing-aid microphone). Columns 3 to 12: SRT in noise, expressed in dB speech-to-noise ratio (at the position of the microphone) for the five different maskers. Gain mode is indicated by - or + for conditions without or with gain control, respectively.

		MASKER									
	Quiet	Single speaker		Speech babble		Speech noise		Car noise		Music	
		GAIN MODE									
Subj. nr.		-	+	-	+	-	+	-	+	-	+
1	2.0	-3.0	-4.0	4.0	4.5	-1.0	1.5	-1.5	-2.5	-2.0	-2.0
2	-2.0	-1.5	-1.0	5.0	5.0	-1.0	1.0	-2.0	-2.0	0.0	2.0
3	2.0	3.5	4.5	6.0	5.5	5.5	3.5	3.0	-1.0	3.0	-3.0
4	0.5	-2.5	-2.5	-0.5	5.5	1.5	2.5	-4.5	0.0	-4.0	-1.0
5	0.0	-1.0	-3.5	4.0	1.5	1.5	2.0	1.5	1.5	-1.5	0.5
6	1.0	-1.0	-3.0	1.0	0.0	0.0	-0.5	-4.0	-2.0	-2.5	-3.0
7	-6.0	-2.0	-0.5	3.5	4.0	0.0	1.5	-2.0	-0.5	-1.5	-1.5
8	2.0	0.5	0.5	5.0	4.0	2.5	3.5	1.0	1.0	1.5	2.0
9	-6.0	7.5	6.0	8.0	7.5	6.0	4.5	3.0	3.0	5.5	5.5
10	-1.5	-6.0	-1.0	3.0	2.0	-1.0	-2.0	-3.5	-3.0	-4.0	-1.5
M	-0.8	-0.6	-0.5	3.9	4.0	1.4	1.8	-0.9	-0.6	-0.6	-0.2
sd	3.1	3.7	3.3	2.4	2.2	2.6	1.9	2.8	1.9	3.1	2.7

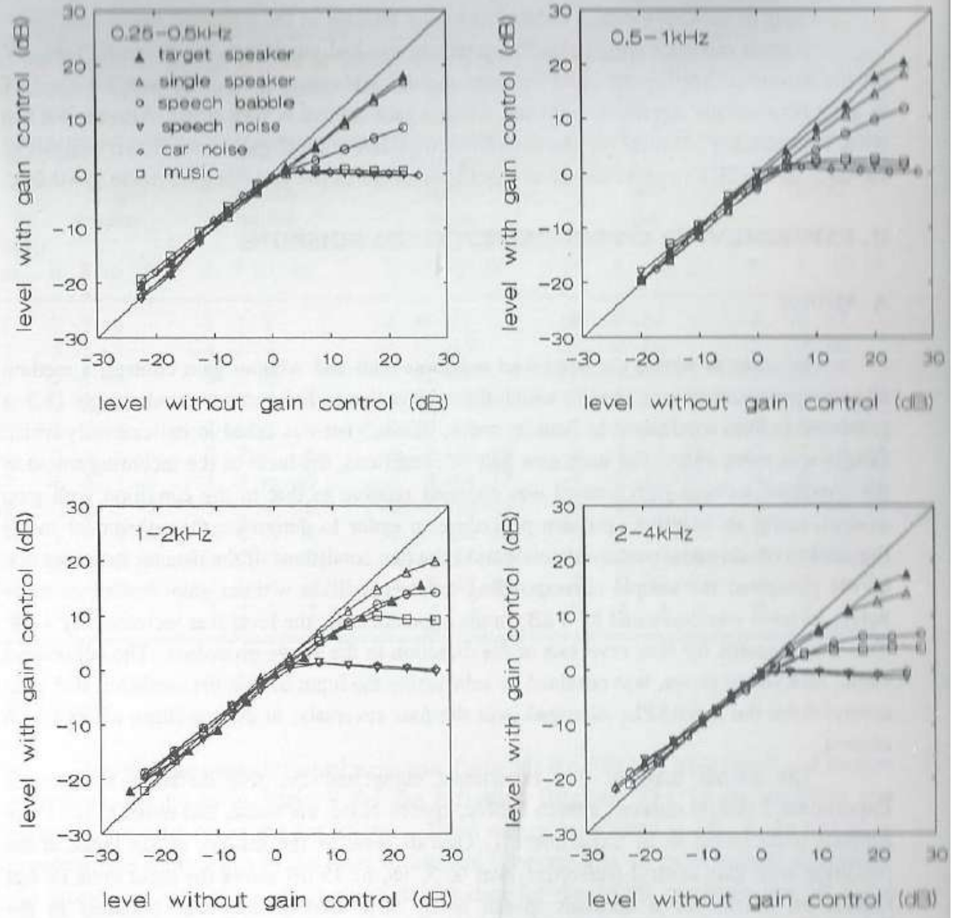


Fig. 3. Long-term average sound level in a condition with gain control as a function of the level witho gain control, for the target speaker and five masking signals, represented for each of the four octave band Levels are expressed in dB relative to the onset level of gain control. The diagonals represent the line relation between the levels in the conditions with and without gain control, for signals in which the lev of the envelope minima remains below the onset level of gain control.

continuous sounds: speech noise, car noise, and music. As can be seen in Fig. 3, the effect of gain control is very small for speech by one speaker. These diagrams demonstrate the effectiveness of an amplification controlled by the minima in the temporal sound envelope.

Speech-reception thresholds for quiet and masked conditions are given in Table IV for the individual hearing-impaired listeners. As this table shows, the difference for the SRT in quiet between the conditions with and without gain control is very small. Also the SRT in noise is practically identical for the conditions with and without gain control. An analysis of variance on the SRT in noise shows no significant effect of the variable gain mode ($p=0.60$).

II. EXPERIMENT 2 : EFFECT ON PERCEIVED NOISINESS

A. Method

In order to equate the perceived noisiness with and without gain control, a method of paired comparison was used in which the subject listened to a short sound sample (3.5 s) presented in both conditions, in random order. The subject was asked to indicate only which sample was more noisy. For each new pair of conditions, the level of the incoming sound in the condition without gain control was changed relative to that in the condition with gain control, using an adaptive up-down procedure in order to determine the adjustment in dB required to obtain equal perceived noisiness in the two conditions. If the listener indicated that he/she perceived the sample corresponding to the condition without gain control as more noisy, its level was decreased by 4 dB; in the opposite case, the level was increased by 4 dB. This was repeated for four reversals of the direction in the above procedure. The adjustment value, referred to above, was obtained by subtracting the input SPL in the condition with gain control from the input SPL, averaged over the four reversals, in the condition without gain control.

The sounds used in this experiment comprised the five different maskers of Experiment 1 (single speaker, speech babble, speech noise, car noise, and music). The same subjects participated as in Experiment 1. Overall level of the masker at the input in the condition with gain control (reference) was 0, 5, 10, or 15 dB above the input level of that masker at which, for a constant speech level, 50% intelligibility was obtained in the corresponding condition in Experiment 1. For each listener, the amplification with the gain control switched off, and the onset level of gain control, were the same as in Experiment 1, in each frequency band. Five different masker sounds and four levels (0, 5, 10, and 15 dB) give a total of twenty different experimental conditions. Levels and maskers within each level were, for each listener, presented in random order. Further details of the apparatus, processing algorithm, and stimulus presentation were as in Experiment 1.

Table V. Level adjustment in dB required in conditions without gain control in order to obtain the same perceived noisiness as in conditions with gain control, for the individual hearing-impaired listeners, in Experiment 2. Values are presented for five maskers and four levels of the masker at the input in the condition with gain control (reference). Masker levels are expressed in dB relative to the input level of that masker at which, for a constant speech level, 50% intelligibility was obtained in the condition with gain control in Experiment 1.

	MASKER																			
	Single speaker				Speech babble				Speech noise				Car noise				Music			
Subj.	RELATIVE INPUT LEVEL (in dB)																			
nr.	0	5	10	15	0	5	10	15	0	5	10	15	0	5	10	15	0	5	10	15
1	-2	-3	-6		-1	-2	-5		-6	-10	-16		-6	-9	-14		-3	-5	-7	
2	-3	-2	-3		-4	-4	-4		-6	-13	-15		-5	-9	-14		-2	-7	-10	
3	-2	-1	-4		-3	-3	-3		-4	-11	-15		-11	-15	-19		-8	-14	-17	
4	0	-2	-3	-5	-1	0	-6	-6	-1	-2	-10	-12	-7	-6	-12	-14	-2	-5	-4	-15
5	-2	-2	-4		-2	-2	-8		-4	-6	-11		-2	-6	-10		-2	-2	-6	
6	-3	-2	-3	-6	-2	-7	-7	-8	-6	-10	-17	-19	-2	-3	-13	-18	-3	-1	-4	-14
7	-6	-2	-2	-9	1	-3	-6	-10	-3	-4	-14	-17	-1	-2	-13	-14	-4	-5	-7	-13
8	-2	-2	-2	-4	-1	-2	-6	-6	-2	-6	-10	-14	-5	-8	-14	-16	-2	-7	-10	-15
9	1	1	-6	-8	-2	-6	-6	-6	-2	-6	-10	-15	-2	-2	-9	-14	-3	-3	-6	-11
10	1	1	-2	-7	-1	-2	-7	-8	-6	-13	-15	-18	-6	-10	-13	-17	-5	-8	-10	-17
M	-2	-1	-3	-6	-2	-3	-6	-7	-4	-8	-13	-16	-5	-7	-13	-16	-3	-6	-8	-14
sd	2	1	2	2	1	2	2	2	2	4	3	3	3	4	3	2	2	4	4	2

B. Results

Level adjustments for equal noisiness, found for the different input levels and masker conditions, are given in Table V for the individual listeners. The 15-dB condition was measured for only six of the ten listeners. Mean values are plotted in Fig. 4. As had to be expected, the mean value of the level adjustment is negative in all experimental conditions. An analysis of variance without and with the data for the 15-dB condition (for 10 and 6 listeners, respectively), showed that the effect of the variables level and masker, as well as the interaction of these variables, are highly significant ($p<0.001$).

Whereas in the 0-dB condition, mean level adjustments between -2 and -5 dB were found, the curves for the different maskers diverge with increasing level. For every 5-dB increase in the level of the incoming sound, the attenuation required in a condition without gain control in order to obtain the same perceived noisiness as in a condition with gain control, increases by only about 1.5 dB for a single speaker, and by as much as about 4 dB for speech noise, car noise, and music. Stated differently, for the latter sounds, an increase in input level of 5 dB in a condition with gain control is, in terms of perceived noisiness, equivalent to an increase in input level of only about 1 dB in a condition without gain control.

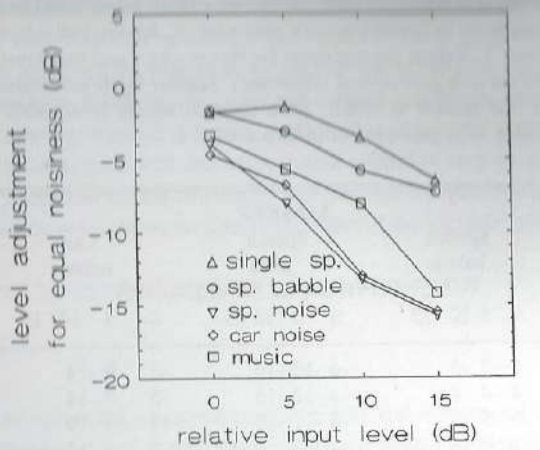


Fig. 4. Mean level adjustment in dB required in conditions without gain control in order to obtain the same perceived noisiness as in conditions with gain control, in Experiment 2 (10 listeners; 15-dB condition: 6 listeners). Values are plotted as a function of the level of the masker at the input in the condition with gain control (reference). Masker levels are expressed in dB relative to the input level of that masker at which, for a constant speech level, 50% intelligibility was obtained in the condition with gain control in Experiment 1. Masker type is the parameter.

III. GENERAL DISCUSSION AND CONCLUSIONS

The results of *Experiment 1* show that, in various conditions of sounds frequently interfering in practical speech communication situations, equal SRTs are found with and without the gain controlled automatically by the minima in the sound envelope.

As the amplification applied in conditions with and without gain control affects both speech and masker, this result had to be expected. An automatic reduction of the amplification in masked frequency regions, in conditions with gain control, leaves speech-to-noise ratio within the different frequency bands unchanged. It follows that such a reduction of amplification will only weaken the masking effectiveness of the noise by reducing the potential effect of spread of masking from intense narrowband sounds to neighboring speech bands, which may particularly manifest itself from lower to higher frequencies (upward spread of masking). However, because in our experimental conditions, speech and masker had roughly comparable spectra, which, after shaping, were on the average practically flat, the chance of spread of masking was virtually absent.

More importantly, however, the results show that the signal attenuation in conditions with the frequency-dependent gain controlled by the minima in the sound envelope does not make useful speech components inaudible. Also, speech intelligibility is not harmed by the rate and range of variations in the amplitude-frequency response.

The results of *Experiment 2* show that, for the different sounds that were used as maskers in Experiment 1, conditions without gain control require lower levels of the incoming signal compared to conditions with gain control in order to give the same perception of noisiness. For the *same* level at the input, this means that, by using the minima in the sound envelope to control the amplification per frequency band, interfering sounds are perceived as less noisy, and thus, as less annoying. In practice, this has the advantage that the hearing-impaired listener provided with such a gain control will be much less in need of manually readjusting the hearing-aid's characteristics along with the acoustical condition. This need is illustrated by the finding of Ringdahl et al. (1990) that, in various everyday listening environments, listeners with a sensorineural hearing impairment prefer the use of a hearing aid in which they can change between different frequency responses, depending on the acoustical background, to the use of a hearing aid with a single, fixed, frequency response.

The wideband attenuation in conditions without gain control relative to conditions with gain control in order to correct for the greater perceived noisiness in the first, increases with the level of the incoming sound. From this result, it may be inferred that the difference in noisiness between conditions with and without gain control becomes larger with input level, due to the increasing effect of gain control. The rate of increase in the adjusted attenuation in conditions without gain control depends strongly on the masker used, and was maximally about 4 dB for every 5-dB increase in level for poorly fluctuating sounds like speech-spectrum noise, car noise, and music. This reflects the effective compression of these signals in conditions with gain control (see Fig. 3). For these sounds, the change in perceived noisiness produced by a 5-dB variation in input level in conditions with gain control, is equivalent to the change in perceived noisiness produced by only about 1-dB variation in input level in conditions without gain control.

For normal-hearing listeners Wilson (1963) showed that an increase of 25 dB in the overall level of automobile noise is sufficient to cause a change in the subjective rating of noisiness from "quiet" to "extremely noisy". By an extrapolation of our results on noisiness, it may be inferred that, for the hearing-impaired group tested, the growth in perceived noisiness produced by a 25-dB increase in the input level of stationary noise in a condition with gain control, would be equivalent to the growth in noisiness produced by an increase in input level of only 5 dB in a condition without gain control. The qualitative ratings of noisiness by normal-hearing listeners, referred to above, suggest that, in terms of actual listening comfort, this represents a substantial improvement for the condition with automatic frequency-dependent control of the amplification compared to the condition without such a gain control.

In summary, we may conclude that, in various conditions of everyday interfering sounds with spectra that do not deviate much from the spectrum of speech, automatic frequency-dependent control of the amplification by the minima in the temporal sound envelope does not affect the SRT in noise, but considerably reduces the sensation of noisiness when no speech communication takes place. This holds most strongly for sounds with a non-fluctuating temporal envelope in the individual frequency bands of the gain control. For these sounds, the growth in perceived noisiness produced by a given increase in input level is equivalent to the growth in noisiness produced by only about one-fifth of that increase (in decibels) in conditions with a fixed amplitude-frequency response.

CHAPTER 6

SUMMARY

People with a sensorineural hearing impairment often complain about having difficulties with understanding speech in the presence of interfering sounds. These difficulties bring about the need for a higher speech-to-noise ratio compared to normal-hearing people. Attempts in the field of hearing-aid research to improve the speech-to-noise ratio have as yet not been successful in critical listening conditions. For the time being, therefore, the hearing-aid user will be most helped when we concentrate on a straightforward signal treatment aimed at presenting the total signal of speech in noise as favorably as possible to the listener's ear, independently of signal level.

A prerequisite for optimal speech reception is that speech is presented at such a level that its information-bearing fluctuations are audible for the hearing-impaired listener over a frequency range that is as wide as possible. This can be achieved by a multichannel automatic gain-control (AGC) hearing aid in which, under the variable acoustical conditions, the frequency-dependent gain adapts itself to the current level and spectrum of the incoming sound. Such a hearing aid is particularly beneficial for listeners with a sloping audiogram. The effectiveness of such a hearing aid can be improved further by using the level of the minima in the temporal envelope of the signal as an AGC-criterion. In this way, the rapid intensity fluctuations of speech can be preserved, whereas sounds lacking such fluctuations, like (speech masked by) noise, are selectively suppressed. This will improve listening comfort and may reduce the effect of spread of masking in conditions with intense narrowband noise. The adjustment of amplification in the various frequency bands, as a response to changes in the acoustical situation, should preferably take about 0.5 s. With such a time constant, intensity fluctuations within syllables are left unaffected, whereas the gain effectively adapts itself to the usually much slower variations in background noise. The main question of this study was: how effective is this concept of signal treatment, aiming at presentation of the signal, under the variable acoustical conditions, at a favorable level within the listener's dynamic-range of hearing?

A hearing aid designed as outlined above, has an amplitude-frequency response that varies in time. These variations leave the speech-to-noise ratios in the various frequency bands unaffected, but may affect the speech-reception threshold (SRT) in noise in a negative way by their range and rate. Because many hearing-impaired people already have difficulties with understanding speech in the presence of interfering sounds, it is clear that a hearing aid can only be effective if such a deterioration of the SRT in noise is avoided. Therefore, in the *first part* of this study, presented in the Chapters 2 and 3, we investigated the effect of varying the amplitude-frequency response on the masked SRT of short everyday sentences. As a masking sound we adopted noise with a spectrum identical to the long-term average spectrum of the sentences. This spectrum is common to all those conditions in which human voices produce the interfering sound (e.g. second speaker or speech babble).

In Chapter 2, experiments are described for groups of normal-hearing subjects. This

allows interpretation without the complicating factor of hearing impairment. In the first experiment (20 subjects), various steady-state slopes of the amplitude-frequency response were applied during presentation of the sentences. In the second experiment (20 subjects), a single change in this slope was given halfway through the sentence, with five different rates of change. In the third experiment (10 subjects), the slope of the amplitude-frequency response was varied continuously. Results show that the SRT in noise is constant for steady-state slopes of the amplitude-frequency response within a range between about -7 and +10 dB/oct, or when a single transition in slope is given between +5 and -5 dB/oct, and vice versa. Varying the transition time over a range from 2 s down to 1/8 s does not appear to have a substantial effect. Continuous variation of the response slope between +5 and -5 dB/oct, however, gives a gradual increase of the SRT with increase of the frequency of slope variation. For variation frequencies of up to 1 Hz, the increase in threshold is less than 1.8 dB.

In Chapter 3, a study similar to that in Chapter 2 is described, except that this time the subjects were 20 listeners with a sensorineural hearing impairment. Pure-tone hearing loss, averaged over 0.5, 1, and 2 kHz, was between 30 and 55 dB. As a baseline in all experimental conditions, we used speech and noise spectrally shaped according to the line bisecting the dynamic-range of hearing for each individual listener. All variations in the frequency response were applied relative to this baseline. Results show that in conditions with a steady-state amplitude-frequency response or with a response that changes once halfway through the sentence with a transition time of 0.25 or 1 s, the increase in masked SRT is less than 2 dB, provided that the falling amplitude-frequency response does not exceed roughly -3 dB/oct relative to the dynamic-range bisector. Further steepening the slope increases the risk of upward spread of masking. A rising slope of the amplitude-frequency response, on the other hand, can be steepened without an effect on the SRT. Because interfering sounds in everyday conditions predominantly have a low-frequency spectrum, hearing aids with a rising rather than a falling frequency response will also be most appropriate for the purpose of reducing the discomfort caused by such sounds.

The stability of the masked SRT with variations of the amplitude-frequency response, as reported in the Chapters 2 and 3 for normal-hearing and hearing-impaired listeners, respectively, constitutes a first requirement for an effective frequency-dependent AGC in hearing aids. The greatest benefit in terms of the SRT from frequency-dependent control of the amplification has to be expected in conditions where the spectrum of the noise deviates strongly from that of the speech. In these conditions, frequency-dependent amplification may reduce upward spread of masking. As a next step, we investigated the upper limit of this benefit by adjusting the frequency-dependent amplification in conditions of intense frequency-limited interfering noise. This was done in the *second part* of this study, presented in Chapter 4, for 12 normal-hearing and 12 hearing-impaired listeners. For the latter group, pure-tone hearing loss, averaged over 0.5, 1 and 2 kHz, was between 39 and 57 dB. Speech and noise were both spectrally shaped according to the listener's dynamic-range bisector; however, the level of the noise in one octave band (0.25-0.5 or 0.5-1 kHz) was increased by 20 dB. In the first experiment, the increased noise level was fixed. In the second experiment, the increase of noise level, together with the applied amplification, developed slowly in time (transition time 1 s). The results for both normal and impaired listeners show that frequency-selective

compression of the signal in the octave band with the 20-dB increase of noise is more beneficial than wideband compression, and gives a decrease in SRT of 4 dB, on the average, relative to a condition without compression. This applies both to steady-state and time-varying conditions.

The results reported in the Chapters 2-4 are promising for the success of slow-acting frequency-dependent AGC in hearing aids. As a *third part* of this study, Chapter 5 presents the ultimate validation of the above-outlined concept for signal treatment by an investigation into the effectiveness of automatic frequency-dependent control of amplification by the level of the minima in the temporal sound envelope. Subjects were 10 hearing-impaired listeners with a pure-tone hearing loss, averaged over 0.5, 1, and 2 kHz, between 35 and 55 dB. In the first experiment, we compared the SRT in noise with and without gain control. In the second experiment, we investigated the reduction in subjective noisiness by measuring the wideband attenuation required for conditions without gain control in order to obtain a noisiness equal to that for conditions with gain control. Results of Chapter 5 show that the condition with gain control does not affect the SRT for sentences in the presence of everyday interfering sounds, having spectra that are roughly comparable to that of the speech signal; however, it substantially reduces the perceived noisiness when no speech communication takes place. In line with our expectations, the effect of the gain control on the signal was very small when a single speaker was present, and was greatest in case of sounds with a more or less continuous character (e.g. stationary noise, music). For the latter sounds it was found that for a given increase in input level, the corresponding growth in perceived noisiness is equivalent to the growth in perceived noisiness produced by only about one-fifth of that increase in input level (in decibels) in a condition without gain control.

In conclusion, the data reported in this study have shown that there are good perspectives for a hearing aid with multichannel automatic gain-control in which the frequency-dependent amplification preserves the intensity fluctuations of speech and reduces those parts of the sound lacking such fluctuations. It appears that such a gain control is able to give an improvement of listening comfort under various conditions of the background noise, without negatively affecting speech intelligibility by its variation of the amplitude-frequency response. Additionally, in conditions with intense narrowband noise, speech intelligibility is even improved by the reduction of the effects of upward spread of masking. In short, the above multichannel AGC hearing aid can, under the variable acoustical conditions, optimize speech intelligibility while minimizing the disturbing effects of noise, without the need of a manual readjustment of the hearing-aid characteristics. This is particularly advantageous for elderly hearing-aid users with reduced manual dexterity. It appears that, along the lines suggested in this study, next-generation hearing aids can be designed which are more effective than the present ones.

SAMENVATTING

Veel slechthorenden hebben moeite met het verstaan van spraak in een rumoerige omgeving. Hierdoor hebben zij veelal behoefte aan een hogere spraak-ruisverhouding dan normaalhorenden. Pogingen om door middel van signaalbehandeling de spraak-ruisverhouding te verbeteren zijn tot nog toe, met name in kritische lawaaicondities, niet erg succesvol gebleken. Voorlopig lijkt de slechthorende het meest gebaat te zijn bij een hoortoestel dat ervoor zorg draagt dat het totale signaal van spraak in ruis zo gunstig mogelijk aan het oor wordt aangeboden.

Om spraak te kunnen verstaan is het een vereiste dat de informatie-dragende fluctuaties van het spraaksignaal voldoende kunnen worden waargenomen. Voor slechthorenden houdt dit in dat de spraak versterkt aangeboden dient te worden zodat deze fluctuaties over een zo breed mogelijk frekwentiegebied hoorbaar zijn. Met name wanneer het gehoor in diverse frekwentiegebieden een verschillende mate van ongevoeligheid vertoont, zal een hoortoestel gewenst zijn waarvan de versterking in meerdere frekwentiekanalen automatisch wordt bijgeregeld. In een dergelijk toestel kan de frekwentie-afhankelijke versterkingsfactor zich, onder de steeds wisselende akoestische omstandigheden, aanpassen aan het actuele niveau en spectrum van het binnenkomende signaal.

In hoortoestellen met automatische sterkteregeling is het gebruikelijk dat de versterking wordt geregeld op grond van het gemiddelde signaalniveau, zonder dat hierbij onderscheid wordt gemaakt tussen spraak en stoorsignaal. Dit betekent dat gedurende passages zonder spraaksignaal, achtergrondruis wordt versterkt tot een niveau dat door de slechthorende als hinderlijk wordt ervaren. Dit bezwaar van een automatische sterkteregeling in hoortoestellen kan aanzienlijk worden gereduceerd wanneer, in plaats van het gemiddelde signaalniveau, het niveau van de "minima" (dalen) in de signaalomhullende als regelparameter voor de versterking wordt gebruikt. Hierdoor kunnen de snelle intensiteitsfluctuaties van spraak worden behouden, terwijl geluidspassages waar dergelijke fluctuaties ontbreken, zoals in het geval van (spraak gemaskeerd door) ruis, kunnen worden onderdrukt. Dit zal het luistercomfort bevorderen.

De instelling van de versterking in de verschillende frekwentiebanden vindt bij voorkeur plaats in een tijd van ongeveer 0.5 sec. Met een dergelijke tijdconstante worden namelijk de intensiteitsfluctuaties binnen de duur van een lettergreep onaangetast gelaten, terwijl de versterking zich toch effectief kan aanpassen aan de veelal tragere variaties in achtergrondruis.

De hoofdvraag van dit proefschrift is: hoe effectief is bovengenoemde benadering van signaalbehandeling, die er op gericht is het binnenkomende signaal, onder de wisselende akoestische omstandigheden, zo gunstig mogelijk aan te bieden binnen het gehoorbereik van de luisteraar?

De frekwentie-karakteristiek van zo'n hoortoestel zal in de tijd variëren. Zulke variaties zullen de spraak-ruisverhouding in de verschillende frekwentiekanalen ongewijzigd laten. De vraag is echter of het gehoor wel bestand is tegen een dergelijke "levende" frekwentie-karakteristiek. Het *eerste gedeelte* van dit proefschrift, bestaande uit de

Hoofdstukken 2 en 3, bestudeerde daarom het effect van schommelingen in de helling van de frekwentie-karakteristiek op de spraakverstaanvaardigheidsdrempel (SRT) voor korte alledaagse zinnen in rumoer. Als maskeerder gebruikten we ruis met hetzelfde gemiddelde frekwentie-spectrum als de zinnen. Dit maskeerspectrum werd gekozen omdat in alledaagse luistersituaties de menselijke stem zelf een van de meest voorkomende stoorsignalen is.

In Hoofdstuk 2 worden experimenten met groepen normaalhorende proefpersonen beschreven. Dit maakt een interpretatie mogelijk zonder dat daarbij complicerende invloeden van slechthorendheid worden betrokken. In een eerste experiment (20 proefpersonen) werden verschillende constante hellingen van de frekwentie-karakteristiek getest. In een tweede experiment (20 proefpersonen) maakte de helling van deze karakteristiek een eenmalige omslag halverwege de zin, waarbij vijf verschillende omslagtijden werden toegepast. In een derde experiment (10 proefpersonen) werd de helling van de frekwentie-karakteristiek continu gevarieerd. Resultaten tonen aan dat de SRT in ruis vrijwel ongewijzigd blijft wanneer de frekwentie-karakteristiek constant is met een helling binnen een bereik van circa -7 tot +10 dB/oct, of wanneer deze helling een eenmalige omslag maakt van +5 naar -5 dB/oct, of omgekeerd. De keuze van de omslagtijd tussen 2 en 1/8 sec lijkt de SRT niet noemenswaardig te beïnvloeden. Een continue schommeling van de helling in de frekwentie-karakteristiek tussen +5 en -5 dB/oct geeft daarentegen een geringe drempelverhoging te zien naarmate de tijd waarin de variatie zich voltrekt, korter wordt. Voor variatiefrekwenties tot 1 Hz is de drempelverhoging kleiner dan 1.8 dB.

Hoofdstuk 3 beschrijft een vergelijkbaar onderzoek als Hoofdstuk 2, maar nu uitgevoerd voor 20 proefpersonen met een perceptief gehoorverlies. Gehoorverlies voor tonen, gemiddeld over 0.5, 1, en 2 kHz, lag tussen 30 en 55 dB. Als basisinstelling in alle experimentele condities gebruikten we spraak en ruis waarvan de spectra waren gevormd volgens de lijn die het gebied tussen gehoordrempel en niveau voor onaangename luidheid, doormidden deelt ("bisector"). Alle variaties in de helling van de frekwentie-karakteristiek werden uitgevoerd ten opzichte van deze basisinstelling. Resultaten laten zien dat wanneer de frekwentie-karakteristiek constant is, of wanneer deze tijdens de zin een eenmalige omslag maakt in een tijd van 0.25 of 1 sec, de verhoging in SRT minder is dan 2 dB. Voorwaarde is echter dat de helling van de frekwentie-karakteristiek niet steiler is dan ongeveer -3 dB/oct ten opzichte van de bisector. Voor hellingen steiler dan -3 dB/oct kan sterke maskering naar hogere frekwenties optreden. De steilheid van een positieve helling van de karakteristiek mag daarentegen vergroot worden zonder dat dit de SRT beïnvloedt. In alledaagse luistercondities vertonen stoorsignalen overwegend laagfrekwente spectra. Daarom zal een hoortoestel met een positief hellende karakteristiek, behalve uit het oogpunt van spraakverstaan, ook uit het oogpunt van luistercomfort de voorkeur genieten boven een hoortoestel met een negatief hellende karakteristiek.

Zoals is aangetoond voor normaal- en slechthorende proefpersonen in de Hoofdstukken 2 en 3, is de SRT in ruis betrekkelijk ongevoelig voor variaties in de frekwentie-karakteristiek. Dit betekent dat is voldaan aan een belangrijke randvoorwaarde voor succes van een meerkanalige automatische versterkingsregeling in hoortoestellen. Verwacht mag worden dat het voordeel van een frekwentie-afhankelijke versterkingsregeling, in termen van spraakverstaan, het grootst is in die situaties waar het spectrum van de

maskeerder duidelijk afwijkt van dat van de spraak. In zulke condities kan een selectieve verzwakking van een gemaskeerde frekwentieband een vermindering geven van sterke maskeringseffecten naar aangrenzende (hoger gelegen) frekwentiegebieden (upward spread of masking). In het *tweede gedeelte* van dit proefschrift, beschreven in Hoofdstuk 4, onderzochten we de mate waarin een frekwentie-selectieve aanpassing van de versterking de SRT voordelig beïnvloedt in condities met spectraal-locale stoorruis. Dit onderzoek werd uitgevoerd voor 12 normaalhorende en 12 slechthorende luisteraars. Voor de laatste groep lag het gehoorverlies voor tonen, gemiddeld over 0.5, 1, en 2 kHz, tussen 39 en 57 dB. Opnieuw werden de spectra van spraak en ruis voor iedere proefpersoon gevormd naar de bisector van het gehoorbereik. Daarbij werd echter in één octaafband (0.25-0.5 of 0.5-1 kHz) het niveau van de ruis met 20 dB verhoogd. In een eerste experiment was de verhoging in ruisniveau constant. In een tweede experiment nam het ruisniveau toe gedurende 1 sec, gelijktijdig met het aanbieden van de zin. Voor zowel normaal- als slechthorende luisteraars laten de resultaten zien dat een selectieve verzwakking van het signaal in de octaafband met de 20-dB toename van de ruis een grotere winst oplevert dan een breedbandige verzwakking. Afname in de gemiddelde drempel is 4 dB, zowel in constante als in de tijd variërende condities.

De resultaten van de Hoofdstukken 2-4 zijn bemoedigend ten aanzien van de toepassing van een trage frekwentie-afhankelijke versterkingsregeling in hoortoestellen. In het *derde gedeelte* van dit proefschrift, beschreven in Hoofdstuk 5, onderzochten we tenslotte de effectiviteit van een frekwentie-afhankelijke versterking die automatisch wordt geregeld op grond van het niveau van de minima in de signaalomhullende. Proefpersonen waren 10 slechthorende luisteraars met een gehoorverlies voor tonen, gemiddeld over 0.5, 1, en 2 kHz, tussen 35 en 55 dB. In een eerste experiment werd de SRT met en zonder versterkingsregeling gemeten. In een tweede experiment onderzochten we in welke mate het geluidvolume in condities zonder versterkingsregeling moet worden bijgesteld teneinde eenzelfde hinderlijkheid van het geluid te verkrijgen als in condities met versterkingsregeling. Resultaten tonen aan dat de conditie met versterkingsregeling de SRT onaangetast laat in de aanwezigheid van diverse alledaagse stoorgeluiden waarvan de spectra lijken op het spectrum van de spraak. Echter, in situaties waar geen spraakcommunicatie plaatsvindt, wordt onder een dergelijke regeling de hinderlijkheid van het stoorgeluid aanzienlijk gereduceerd. Zoals verwacht mocht worden, is het effect van de versterkingsregeling op het signaal zeer klein in het geval van één enkele spreker, en is dit effect het grootst voor geluiden met een min of meer continu karakter (bijv. stationaire ruis en muziek). Voor het laatste type stoorgeluid werd gevonden dat een toename in ingangsniveau wel vijf maal zo groot moet zijn (in decibels) als in een conditie zonder versterkingsregeling teneinde eenzelfde toename in geluidshinder te veroorzaken.

Op grond van de resultaten van dit proefschrift mogen we concluderen dat er goede perspectieven bestaan voor een hoortoestel met een meerkanalige automatische volumeregeling waarin de frekwentie-afhankelijke versterking de intensiteitsfluctuaties van spraak ongewijzigd laat, en de meer stationaire signaalpassages onderdrukt. Een dergelijke versterkingsregeling geeft onder diverse stoorlawaaicondities een verbetering in luistercomfort zonder dat daarbij het spraakverstaan negatief wordt beïnvloed door de schommelingen in de frekwentie-karakteristiek. In condities waar het spectrum van de ruis een sterk locaal maximum vertoont, wordt bovendien, door vermindering van maskering naar aangrenzende frekwentiegebieden, een winst in spraakverstaan behaald. Samenvattend geeft genoemde meerkanalige automatische versterkingsregeling, in verschillende akoestische omstandigheden, maximaal spraakverstaan bij een minimaal storend achtergrondlawaai, zonder dat de luisteraar daarbij zelf de hoortoesteleigenschappen hoeft bij te stellen. Hiermee is de weg geopend naar een nieuw, verbeterd type hoortoestel.

REFERENCES

ANSI (1969). ANSI S3.5-1969, "American National Standard methods for the calculation of the articulation index" (American National Standards Institute, New York).

Bilger, R.C., and Hirsh, I.J. (1956). "Masking of tones by bands of noise," J.Acoust.Soc.Am. 28, 623-630.

Braida, L.D., Durlach, N.I., Lippmann, R.P., Hicks, B.I., Rabinowitz, W.M., and Reed, C.M. (1979). "Hearing aids - A review of past research on linear amplification, amplitude compression, and frequency lowering," ASHA Monogr. No. 19.

Ehmer, R.H. (1959). "Masking by tones versus noise bands," J.Acoust.Soc.Am. 31, 1253-1256.

Festen, J.M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for hearing-impaired and normal hearing," J.Acoust.Soc.Am. 88, 1725-1736.

Festen, J.M., van Dijkhuizen, J.N., and Plomp, R. (1990). "Considerations on adaptive gain and frequency response in hearing aids," Acta Otolaryngol. Suppl. 469, 196-201.

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

Gagné, J.P. (1988). "Excess masking among listeners with a sensorineural hearing loss," J.Acoust.Soc.Am. 83, 2311-2321.

Graupe, D., Grosspietsch, J.K., and Basseas, S.P. (1987). "A single-microphone-based self-adaptive filter of noise from speech and its performance evaluation," J. Rehabil. Res. Dev. 24 (4), 119-126.

Haggard, M.P., Lindblad, A.C., and Foster, J.H. (1986). "Psychoacoustical and audiometric prediction of auditory disability for different frequency responses at listener-adjusted presentation levels," Audiology 25, 277-298.

Jerger, J.F., Tillman, T.W., Peterson, J.L. (1960). "Masking by octave bands of noise in normal and impaired ears," J.Acoust.Soc.Am. 32, 385-390.

Kates, J.M. (1990). "Hearing aid signal-processing system," J.Acoust.Soc.Am. 87, 2272 (patent).

Klein, A.J. (1989). "Assessing speech recognition in noise for listeners with a signal processor hearing aid," Ear and Hear. 10, 50-57.

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

Kryter, K.D. (1962b). "Validation of the Articulation Index," J.Acoust.Soc.Am. 34, 1698-1702.

Kryter, K.D. (1970). The effects of noise on man (Academic Press, New York), Ch.9.

Kuttruff, H. (1973). Room acoustics (Applied Science Publishers Ltd., London).

Levitt, H. (1978). "Methods for the evaluation of hearing aids," Scand. Audiol. Suppl. 6, 199-240.

Lutman, M.E., and Clark, J. (1986). "Speech identification under simulated hearing-aid frequency response characteristics in relation to sensitivity, frequency resolution, and temporal resolution," J.Acoust.Soc.Am. 80, 1030-1040.

Martin, E.S. and Pickett, J.M. (1970). "Sensorineural hearing loss and upward spread of masking," J. Speech Hear. Res. 13, 426-437.

Neuman, A.C., and Schwander, T.J. (1987). "The effect of filtering on the intelligibility and quality of speech in noise," J. Rehabil. Res. Dev. 24(4), 127-134.

Ono, H., Kanzaki, J., and Mizoi, K. (1983). "Clinical results of hearing aid with noise-level-controlled selective amplification," Audiology 22, 494-515.

Pavlovic, C.V. (1987). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," J.Acoust.Soc.Am. 82, 413-422.

Plomp, R., and Mimpen, A.M. (1979a). "Improving the reliability of testing the speech-reception threshold for sentences," Audiology 10, 320-334.

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

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

Plomp, R., Anema, P.C., and van Dijkhuizen, J.N. (1986). "Towards a hearing aid with multichannel automatic gain control," *Proceedings of 12th International Congress on Acoustics*, Toronto, B4-1.

Plomp, R. (1988). "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function," J.Acoust.Soc.Am. 83, 2322-2327.

Ringdahl, A., Eriksson-Mangold, M., Israelsson, B., Lindkvist, A., and Mangold, S. (1990). "Clinical trials with a programmable hearing aid set for various listening environments," Brit. J. of Audiol. 24, 235-242.

Rittmanic, P.A. (1962). "Pure tone masking by narrow noise bands in normal and impaired ears," J. Aud. Res. 2, 287-304.

Skinner, M.W., Pascoe, D.P, Miller, J.D., and Popelka, G.R. (1982). "Measurements to determine the optimal placement of speech energy within the listener's auditory area: A basis for selecting amplification characteristics," The Vanderbilt hearing-aid report, Monographs in Contemporary Audiology, p. 161-169.

Stein, L.K., and Dempesy-Hart, D. (1984). "Listener-assessed intelligibility of a hearing aid self-adaptive noise filter," Ear and Hear. 5, 199-204.

Stein, L.K., McGee, T. and Lewis, P. (1989). "Speech recognition measures with noise suppression hearing aids using a single-subject experimental design," Ear and Hear. 10, 375-381.

Studebaker, G.A., Pavlovic C.V., and Sherbecoe, R.L. (1987). "A frequency importance function for continuous discourse," J.Acoust.Soc.Am. 81, 1130-1138.

Tyler, R.S. (1986). "Frequency resolution in hearing-impaired listeners," in: Frequency selectivity in hearing, edited by B.C.J. Moore (Academic Press, London), pp. 309-371.

Tyler, R.S., and Kuk, F.K. (1989). "The effects of 'noise suppression' hearing aids on consonant recognition in speech-babble and low-frequency noise," Ear and Hear. 10, 243-249.

Van Dijkhuizen, J.N., Anema, P.C., and Plomp, R. (1987). "The effect of varying the slope of the amplitude-frequency response on the masked speech-reception threshold of sentences," J.Acoust.Soc.Am. 81, 465-469.

Van Dijkhuizen, J.N., Festen, J.M., and Plomp, R. (1989). "The effect of varying the amplitude-frequency response on the masked speech-reception threshold of sentences for hearing-impaired listeners," J.Acoust.Soc.Am. 86, 621-628.

Van Dijkhuizen, J.N., Festen, J.M., and Plomp, R. (1991a). "The effect of frequency-selective attenuation on the speech-reception threshold of sentences in conditions of low-frequency noise," accepted for publication in J.Acoust.Soc.Am.

Van Dijkhuizen, J.N., Festen, J.M., and Plomp, R. (1991b). "Controlling the gain in a four-channel hearing aid by the minima in the temporal envelope of the sound," submitted for publication in Audiology.

Van Tasell, D.J., Larsen, S.Y., and Fabry, D.A. (1988). "Effects of an adaptive filter hearing aid on speech recognition in noise by hearing-impaired subjects," Ear and Hear. 9, 15-21.

Wagenaar, W.A. (1969). "Note on the construction of digram-balanced latin squares," Psych. Bull. 72, 384-386.

Wilson, A.H. (1963). Noise (Her Majesty's Stationary Office, London).

Zurek, P.M., and Rankovic, C.M. (1990). "Potential benefits of varying the frequency-gain characteristic for speech reception in noise," J.Acoust.Soc.Am. 87, S1.