Masking of speech by amplitude-modulated noise

Hans Å. Gustafsson and Stig D. Arlinger
Department of Technical Audiology, University Hospital, S-581 85 Linköping, Sweden

(Received 11 September 1991; revised 1 June 1993; accepted 21 August 1993)

The masking of speech by amplitude-modulated and unmodulated speech-spectrum noise has been evaluated by the measurement of monaural speech recognition in such noise on young and elderly subjects with normal-hearing and elderly hearing-impaired subjects with and without a hearing aid. Sinusoidal modulation with frequencies covering the range 2–100 Hz, as well as an irregular modulation generated by the sum of four sinusoids in random phase relation, was used. Modulation degrees were 100%, ±6 dB, and ±12 dB. Root mean-square sound pressure level was equal for modulated and unmodulated maskers. For the normal-hearing subjects, essentially all types of modulated noise provided some release of speech masking as compared to unmodulated noise. Sinusoidal modulation provided more release of masking than the irregular modulation. The release of masking increased with modulation depth. It is proposed that the number and duration of low-level intervals are essential factors for the degree of masking. The release of masking was found to reach a maximum at a modulation frequency between 10 and 20 Hz for sinusoidal modulation. For elderly hearing-impaired subjects, the release of masking obtained from amplitude modulation was consistently smaller than in the normal-hearing groups, presumably related to changes in auditory temporal resolution caused by the hearing loss. The average speech-to-noise ratio required for 30% correct speech recognition varied greatly between the groups: For young normal-hearing subjects it was −15 dB, for elderly normal-hearing it was −9 dB, for elderly hearing-impaired subjects in the unaided listening condition it was +2 dB and in the aided condition it was +3 dB. The results support the conclusion that within the methodological context of the study, age as well as sensorineural hearing loss, as such, influence speech recognition in noise more than what can be explained by the loss of audibility, according to the audiogram and the masking noise spectrum.

PACS numbers: 43.71.Lz, 43.71.Ky, 43.72.Dv

INTRODUCTION

One serious effect of noise is its negative impact on speech communication. The speech masking ability of a noise depends primarily on the relation between speech and noise intensities, i.e., on the speech-to-noise ratio (S/N). Usually, the noise intensity is expressed as a time-average level, e.g., the rms level measured with a time constant of 125 ms (i.e., the "fast" response of the sound level meter). However, the presence of amplitude modulation (AM) in the noise may also influence its speech interfering effect. Two noises of equal mean intensity, one containing AM and the other not, may well differ in speech masking. One might speculate that the presence of AM might increase the speech masking effect due to activation of neurons that are sensitive specifically to AM (Møller, 1971) and thus make them less capable of reacting to amplitude modulations in the speech signal. On the other hand, the intensity fluctuations might provide opportunities for the listener to pick up important fragments of the speech signal during the less loud intervals of the modulation cycle and thus provide a release of speech masking as compared to unmodulated noise.

A number of studies on speech recognition in AM noise have been published. The maskers used have generally been one or more of the following: Multitalker babble, speech-modulated white or speech-spectrum noise, single talker competition, or periodically modulated white or speech-spectrum noise, single talker competition, or periodically modulated white or speech-spectrum noise. Most often however, constant level (unmodulated) white or speech-spectrum noise has been used as a reference masker. The results from such studies can be summarized as follows.

When the masker is the speech from a single talker or noise modulated either periodically or by the speech of a single talker, speech intelligibility improves compared to when unmodulated noise is used (Carhart *et al.*, 1966, 1967, 1969; Dirks *et al.*, 1969; Festen and Plomp, 1990; Miller and Licklider, 1950; Wilson and Carhart, 1969), even if the modulated and unmodulated noises have equal average powers. In such cases, the instantaneous level of such maskers is either higher or lower than the average (rms) level during a relatively large portion of time. Presumably these results are explained by more speech information being detected during noise intervals of lower level than what is lost during noise intervals of higher level.

Multitalker noise or noise modulated by multiple talkers has been found to generally reduce speech intelligibility as compared to unmodulated noise (Carhart *et al.*, 1969; Danhauer and Leppler, 1979). Multitalker noise was more effective as a masker than random noise modulated by multiple talkers, and the effect was attributed to the perceptual similarity between this masker and the target signal (Carhart *et al.*, 1975; Young *et al.*, 1975).

 0001-4966/94/95(1)/518/12/$6.00 © 1994 Acoustical Society of America

Results contradicting those mentioned have also been reported. For example, Berry and Nerbonne (1972) and Horii *et al.* (1970) found that speech modulated by a single talker masks speech more than unmodulated noise does. Danhauer *et al.* (1985) found multitalker noise to mask speech less effectively than unmodulated noise. These discrepancies may be due to different ways of measuring masker power or failure to compensate for differences in average masker power.

Younger, normal-hearing subjects were used in all of the above studies. There have also been investigations concerning elderly and hearing-impaired subjects, which clearly show different results for these groups as compared to normal-hearing subjects. The advantage of a periodically modulated masker appears to be much less for hearing-impaired subjects (Wilson and Carhart, 1969). Also, for hearing-impaired subjects, a forward or backward speech masker results in equal or more masking than unmodulated noise does; contrary to the relation for normal-hearing (Festen and Plomp, 1990; Plomp, 1986). Elderly normal-hearing and elderly hearing-impaired listeners both have elevated sentence recognition thresholds in babble noise (Gelfand *et al.*, 1988). The use of hearing aids does not improve the performance (Tillman *et al.*, 1970; Welzl-Muller and Stephan, 1988).

In view of the conflicting results found in different studies with different masking noises and different subject populations, it was decided to investigate several listener groups with several different maskers in a single study. It includes three subject groups: Young, normal hearing, elderly, normal hearing and elderly, hearing impaired. Since hearing aids may affect the temporal characteristics of the sound transmitted, the third group was chosen to be hearing-aid users who were tested both with and without their hearing aids. Speech recognition scores in different kinds of amplitude-modulated noise were measured and compared to the score in unmodulated noise.

I. METHOD

A. Subjects

1. Young, normal-hearing subjects

Eleven young, normal-hearing individuals, 17–33 years, six males and five females, participated. For each subject, a test ear was chosen; five left ears and six right ears. The selection criteria were hearing threshold levels (HTLs) of no worse than 20 dB HL (*re*: ISO389, 1991) at all test frequencies 125–8000 Hz. The HTL difference between ears was not allowed to exceed 10 dB at any frequency. The average HTLs were all in the range from 0 to 5 dB HL.

2. Elderly, normal-hearing subjects

Twenty subjects, ten women and ten men, aged 54–69 years, participated. For each subject a test ear was chosen according to the audiogram. The maximum HTLs allowed were 20 dB HL in the frequency range 125–3000 Hz, 30 dB at 4 kHz, 35 dB at 6 kHz, and 40 dB at 8 kHz. These limits correspond approximately to the median HTLs for 60-year old men according to ISO 7029 (1984). Further, the limits correspond to the clinically normal range for frequencies up to and including 3 kHz, at the higher frequencies the limits correspond to a mild hearing loss. A maximum HTL difference of 10 dB between left and right ears at any frequency was allowed. Six left ears and 14 right ears were tested. The mean HTLs and standard deviations are shown in Fig. 1. Only at 8 kHz does the mean HTL exceed the normal range.

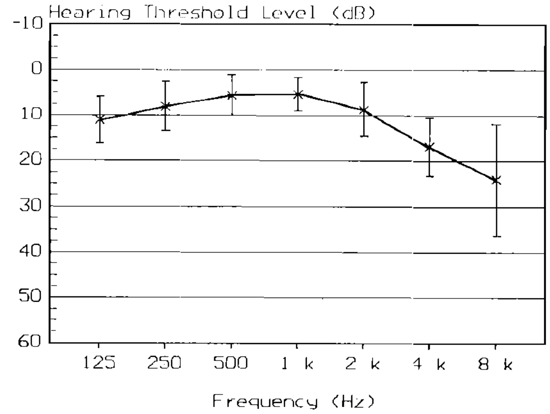

FIG. 1. Mean hearing threshold levels plus/minus one standard deviation for the group of elderly normal-hearing subjects.

3. Elderly, hearing-impaired subjects

Twenty subjects, nine men and eleven women, aged 55–70 years, participated. All subjects used hearing aids for at least one year (12 postauricular; 8 in-the-ear). Nine left ears and 11 right ears were tested. The mean HTLs and standard deviations are shown in Fig. 2. Figure 3 shows the average insertion gain for the aids when worn in the tested ears.

B. Noise

The masker was a speech-shaped random noise obtained by low-pass filtering a white noise with a slope of -6 dB/oct above 1000 Hz. This noise was presented either unmodulated or modulated in one of two ways: Sinusoidally by a single sinusoid or irregularly. For the latter, the

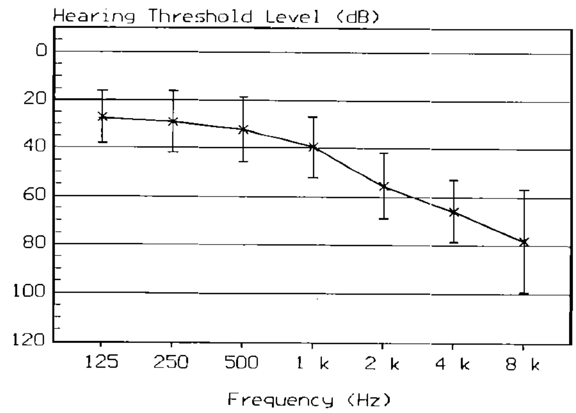

FIG. 2. Mean hearing threshold levels plus/minus one standard deviation for the group of elderly hearing-impaired subjects.

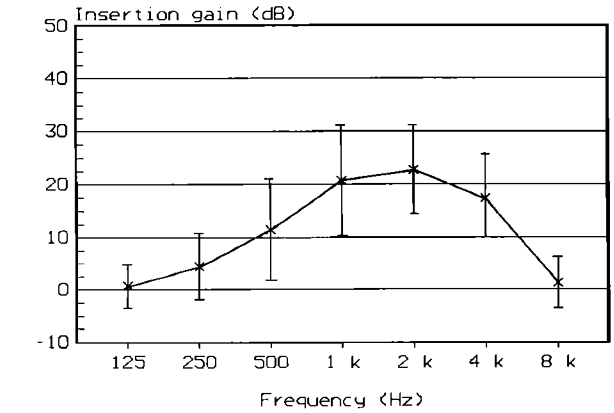

FIG. 3. Mean insertion gain plus/minus one standard deviation for the hearing aids used by the elderly, hearing-impaired test subjects.

sum of four sinusoids of equal amplitude in random phase relation was used as modulating signal. The same range of amplitude variation was used for single sinusoidal modulation as for the irregular modulation.

The degrees of modulation were either ±6 dB, ±12 dB, or ±100%, as illustrated in Fig. 4 on both linear (a) and logarithmic (b) ordinate scales. The logarithmic modulations ±6 dB and ±12 dB cause asymmetric sound pressure variations whereas the changes in sound pressure level are symmetric. The opposite applies to the linear ±100% modulation. The use of both linear and logarithmic modulation types in the study would therefore illustrate the relative importance for masking of the higher-level and the lower-level half intervals of the amplitude-modulated noise. Mathematically, the modulating signal, $u(t)$, was one of the following:

$$u_{100\%}(t)=1+\frac{1}{N}\sum_{i=1}^{N}\sin(2\Pi f_i t+\theta_i), \tag{1}$$

$$u_{6\ \mathrm{dB}}(t)=2^{(1/N)\sum_{i=1}^{N}\sin(2\Pi f_i t+\theta_i)}, \tag{2}$$

$$u_{12\ \mathrm{dB}}(t)=2^{(2/N)\sum_{i=1}^{N}\sin(2\Pi f_i t+\theta_i)}, \tag{3}$$

where $N=1$ for sinusoidal modulation and $N=4$ for irregular modulation. For unmodulated noise, $u(t)=1$.

In order to obtain the same rms sound pressure level (i.e., the same average power) for the modulated and unmodulated noises, the modulated noise had to be attenuated between 0.5 and 6 dB relative to the rms level of unmodulated noise, depending on modulation degree and whether sinusoidal or irregular modulation was used. The attenuation factors were calculated mathematically and were verified by measuring the rms level of the different modulated signals, after correction. The correction factors are listed in Table I. The rms sound pressure level was set individually for each subject. Figure 4(c) shows the resulting ranges of amplitude variation relative to the rms level of the unmodulated noise for the different types of modulation.

The frequencies of the modulating sinusoids were 2.1, 4.9, 10.2, and 19.9 Hz, covering the range of significant modulation frequencies in human speech (Fastl, 1987).

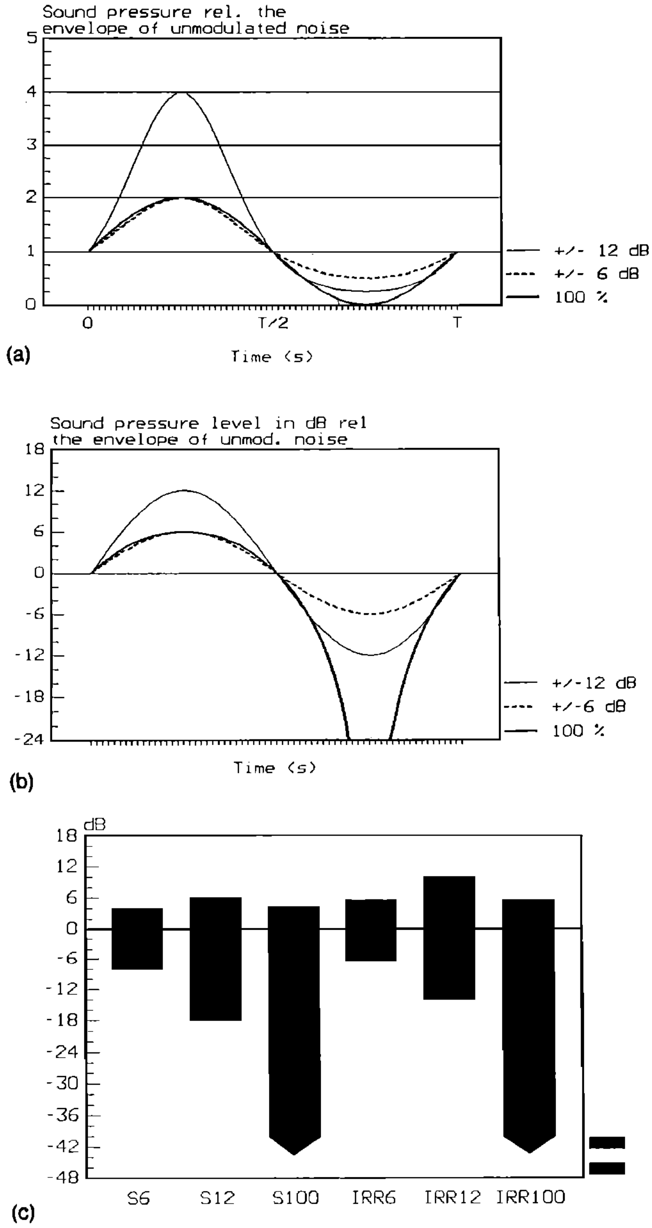

FIG. 4. The different modulation degrees illustrated for sinusoidal modulation in (a) linear scale, (b) logarithmic scale. The figures show the sound pressure envelope of the different modulated noises without rms compensation applied. On the y axis, 1 (in linear scale) and 0 dB (in log scale) correspond to the envelope of unmodulated noise. T is the period. In (c) is shown the envelope variation ranges for different modulations with rms correction applied. Zero dB is the level of the envelope of unmodulated noise. The actual modulation depth of 100% modulation is $-\infty$, indicated by downward pointing arrows.

They were chosen not to be even multiples of each other in order to assure maximum irregularity when added together to produce the irregular modulation condition. Five of the subjects in the experiment were also tested at the higher modulation frequencies of 50, 80, and 100 Hz with a modulation degree of ±12 dB. The different combinations of modulation parameters used are listed in Table II.

C. Speech

The speech material (Hagerman, 1982) consists of 11 lists with ten Swedish low-redundancy five-word sentences

TABLE I. Attenuation factors used to equate the rms sound pressure level of each modulated masker to that of unmodulated noise. N is the number of sinusoids in the modulating signal.

Modulation type	Attenuation (dB)
± 100%, $N=1$	1.8
$N=4$	0.5
± 6 dB, $N=1$	1.9
$N=4$	0.5
± 12 dB, $N=1$	6.1
$N=4$	2.0

in each. The sentences are spoken by a female voice. All sentences have the same structure: ⟨name⟩ ⟨verb⟩ ⟨number⟩ ⟨adjective⟩ ⟨noun⟩, e.g., "Peter held nine new boxes" (in Swedish "Peter höll nio nya lådor"). The same 50 words appear in all of the 11 lists but in different computer-edited combinations. There is a 7-s pause between consecutive sentences in a list during which time the subject is able to repeat the sentence. For each ten-sentence list, the number of words correctly recognized was noted and converted to a speech recognition score in percent.

Test tapes were cassette recordings copied from the master recording of the materials and included a 1000-Hz calibration tone. The level of the calibration tone was 3.4 dB below the speech level not exceeded 90% of the time, measured with the "fast" time constant of the sound level meter. All speech levels reported below refer to the level of this tone. The sound pressure level was kept at 65 dB for all tests except for testing the elderly, hearing-impaired subjects without their hearing aids. In this case, the speech was adjusted to a comfortable listening level (mean 87 dB SPL).

D. Test procedures

The main experiment used modulation frequencies in the 2- to 20-Hz range and comprised all subjects. A subgroup of five of the young, normal-hearing subjects participated in an additional experiment to gain information on the effect of higher modulation frequencies.

TABLE II. The maskers used in the main study. "Irregular" denotes the modulation where the modulating signal consists of the sum of four sinusoids in random phase relation. AM masker Nos. 2–13 were used in the main experiment and Nos. 14–16 in the additional experiment.

Masker no.	Modulation		Short form
1	none		UNMOD
2	irregular,	±6 dB	IRR6
3	irregular,	±12 dB	IRR12
4	irregular,	±100%	IRR100
5	sinusoidal,	±6 dB 4.9 Hz	S6/4.9
6		±12 dB 2.1 Hz	S12/2.1
7		±12 dB 4.9 Hz	S12/4.9
8		±12 dB 10.2 Hz	S12/10.2
9		±12 dB 19.9 Hz	S12/19.9
10		±100% 2.1 Hz	S100/2.1
11		±100% 4.9 Hz	S100/4.9
12		±100% 10.2 Hz	S100/10.2
13		±100% 19.9 Hz	S100/19.9
14		±12 dB 50 Hz	S12/50
15		±12 dB 80 Hz	S12/80
16		±12 dB 100 Hz	S12/100

Monaural speech recognition score was measured for the first 13 of the maskers listed in Table II. The S/N was individually chosen for each subject to achieve a 30% speech recognition score with the unmodulated noise masker. Pilot experiments revealed that AM offered a release of masking rather than an increase and thus the 30% criterion was adopted to obtain a test with the highest sensitivity possible. The order of masker presentation was randomly selected for each subject. The speech level was kept constant at 65 dB SPL and S/N adjustments were accomplished by varying the noise level in 3-dB steps. Each test session lasted for about 1 h.

The young and elderly, normal-hearing subjects were tested using headphones as follows: Three training lists were presented to aquaint the subject with the test situation and to reduce learning effects (Hagerman, 1982), using the unmodulated noise masker. The S/N was initially large enough for the subject to recognize all the words, but was gradually reduced until the speech recognition score was about 30%.

Two more lists (randomly selected for each subject) with the unmodulated noise masker were presented. The speech-to-noise ratios were chosen to yield speech recognition scores on both sides of 30%. The S/N corresponding to a speech recognition score of 30% was then calculated by linear interpolation. Keeping the S/N constant at the value just described, speech recognition scores for AM maskers 2–13 (Table II) were measured using one speech list for each masker. The same lists in the same order were used for each subject whereas the order of masker presentation was randomly selected for each subject. Finally, without altering the S/N, the score for unmodulated masking noise was determined once more using one randomly selected list for each subject. The mean of this score and the initially determined score was taken as the speech recognition score for unmodulated noise. This procedure was applied to compensate for learning or fatigue effects that might have occurred during the course of the experiment.

In addition, five of the young normal-hearing subjects were also tested using sinusoidal modulation frequencies of 50, 80, and 100 Hz. The S/N was determined in the same way as in the main experiment.

The elderly, hearing-impaired were tested both with and without their hearing aids using loudspeaker presentation. The same procedure as just described above was used with the following exceptions: The aided and unaided experiments were performed different test days. For half of the group the aided tests preceded the unaided, and for the other half the opposite order was used. The nontest ear was always plugged with an EAR foam plug. In the aided experiments, the speech level was 65 dB SPL as for the normal-hearing subjects. The subject was asked to adjust the volume control of his hearing aid to a setting that provided a comfortable loudness level of the speech signal presented from the loudspeaker. However, in the unaided experiment the speech level was individually adjusted to a

TABLE III. ANOVA table for the experiment on the young, normal-hearing subjects. The mean square for the interaction between maskers and subjects is used as the error estimate in all F ratios. SS=sum of squares, DF=degrees of freedom, MS=mean square. MOD is a collective denomination of the modulated maskers 2–13. Likewise, IRR designates maskers 2–4, S maskers 5–13, S12 maskers 6–9, and S100 maskers 10–13. m(...) denotes the mean score across maskers within parentheses. The comparison between modulation frequencies applies to the mean scores across ±12 dB and 100% modulation. In the interaction between modulation degree and frequency, comparisons are not made between scores but between score differences, indicated by a D. For example, D2.1 denotes the score difference between ±12 dB and 100% modulation at 2.1 Hz.

Comparison	SS	DF	MS	F	p
Between maskers	40379.52	12	3364.96	75.52	<0.001
Between subjects	4317.45	10	431.75	9.69	<0.001
Interaction	5347.64	120	44.56	···	···
Decomposition of SS between maskers:					
a/ UNMOD–m(MOD)	8763.92	1	8763.92	196.68	<0.001
b/ m(IRR)–m(S)	20281.71	1	20281.71	455.16	<0.001
Within IRR	1733.09	2	866.55	19.45	<0.001
c/ IRR6–m(IRR12&IRR100)	1590.95	1	1590.95	35.69	<0.001
d/ IRR12–IRR100	142.55	1	142.55	3.20	<0.01
Within S	9600.81	8	1200.10	26.93	<0.001
c/ S6/4.9–m(S12&S100)	6749.17	1	6749.17	151.46	<0.001
d/ m(S12)–m(S100)	35.64	1	35.64	0.80	N.S.
Between mod. freq.	1796.73	3	598.91	13.44	<0.001
e/ m(2.1&4.9)–m(10.2&19.9)	1313.64	1	1313.64	29.48	<0.001
f/ 2.1–4.9	384.09	1	384.09	8.62	<0.005
g/ 10.2–19.9	99	1	99	2.22	N.S.
h/ Interaction, degr&freq.	1019.27	3	339.76	7.62	<0.001
m(D2.1&D19.9)–m(D4.9&D10.2)	968.91	1	968.91	21.74	<0.001
D4.9–D10.2	48.09	1	48.09	1.08	N.S.
D2.1–D19.9	2.27	1	2.27	0.05	N.S.

level of comfortable loudness which was then used consistently for all unaided tests on that subject. The mean speech level chosen was 87 dB SPL (standard deviation 4.1 dB).

E. Equipment

The noise was generated by a signal processing computer based on the Texas Instruments TMS 32010 signal processor. The rms level of the AM noise was equated to that of the unmodulated noise by means of a built-in attenuator. Changes of modulation frequency and modulation degree were controlled by a personal computer. The noise signal and the speech signal from a Tandberg TCD 310 cassette recorder were routed to separate channels of a Madsen OB 822 audiometer, which was used to control the sound levels. From the audiometer, noise and speech were separately fed to a mixer, which allowed any combination of noise and speech levels to be set individually for either ear. The mixed signals were then fed to the Sennheiser HD 250 circumaural earphones. The earphone marked left was always placed over the test ear.

Since the elderly, hearing-impaired subjects were tested both with and without hearing aid the speech and noise had to be presented to them by means of a loudspeaker (anechoic frequency response flat within ±3 dB in the range 90 Hz–17 kHz). The test subject was seated in a chair in a soundproof room approximately 1 m in front of and facing the loudspeaker.

Sound level calibrations were accomplished for the headphone using a Bruel & Kjaer 4153 artificial ear with adapters DB 0843 and YJ 0304 and B&K 2209 precision sound level meter. For loudspeaker presentation the sound pressure levels were measured at the reference point (in the absence of the subject) by means of a B&K 4165 microphone, preamplifier B&K 2619, and microphone amplifier B&K 2607. For calibration of the equipment a calibrator type B&K 4230 was used.

II. RESULTS

A. Effects on speech masking by modulation parameters

Differences in speech recognition score between maskers were analyzed by analysis of variance (ANOVA). A two-way repeated mesures model was used, the two factors being subject and masker with subjects treated as a random factor. To further explore differences between maskers, the sum of squares for maskers was split in orthogonal parts as detailed in Table III and discussed below. An exception to this was the analysis of the additional experiment. Since this was only a supplementary experiment, orthogonal decomposition was abandoned in favor of the possibility of extracting as much information as possible. Each F test considering the comparison of two mean values corresponds to a two-tailed t test.

1. Young, normal-hearing subjects

Speech recognition scores obtained with the first 13 maskers of Table II on the young, normal-hearing subjects are shown in Fig. 5. The score with unmodulated masker is the mean of the scores measured before and after presentation of the modulated maskers. It is approximately 30% as expected from the pilot work used to set the noise level for this condition.

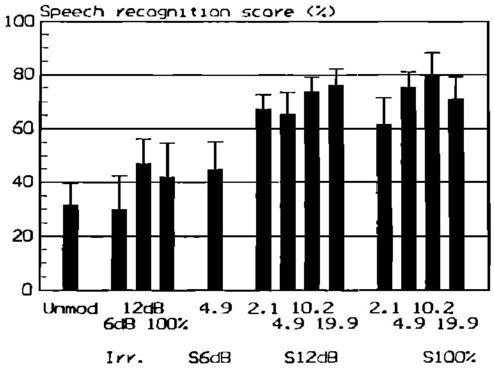

FIG. 5. Results of the main experiment on the group of young, normal-hearing subjects. The bars show the mean speech recognition score obtained with each masker. The horizontal line above each bar shows mean plus one standard deviation.

TABLE IV. ANOVA table for the additional experiment on five young, normal-hearing subjects. The mean square for the interaction between maskers and subjects is used as the error estimate in all F ratios. SS=sum of squares, DF=degrees of freedom, MS=mean square. LOW=the low-frequency maskers 2.1–19.9 Hz and HIGH=the high-frequency maskers 50, 80, and 100 Hz. m(...) denotes the mean score across maskers within parenthesis.

Comparison	SS	DF	MS	F	p
Between maskers	9052.80	7	1293.26	43.15	<0.001
Between subjects	77.60	4	19.40	0.65	N.S.
Interaction	839.20	28	29.97	···	···
UNMOD-m(HIGH)	627.27	1	627.27	20.93	<0.001
m(LOW)-m(HIGH)	4680.01	1	4680.01	156.15	<0.001
Within HIGH	746.13	2	373.07	12.45	<0.001
80–100	129.6	1	129.6	4.32	<0.05
50-m(80&100)	616.53	1	616.53	20.57	<0.001
m(80&100)-UNMOD	235.2	1	235.2	7.84	<0.01

Table III shows that the average speech recognition score with the amplitude-modulated maskers is significantly different from speech recognition in unmodulated noise (*a* in Table III), and furthermore that the speech masking effect of the irregularly modulated noise is significantly different from that of the sinusoidally modulated noise (*b*). Modulation depth of ±6 dB resulted in significantly less release of speech masking than ±12 dB or ±100% (*c*). However, there was no significant difference between the latter two degrees of modulation (*d*).

Significant differences between modulation frequencies were found. The average score with the two lower frequencies differs significantly from the average score with the two higher (*e*). The internal differences between the two lower and between the two higher frequencies are not significant (*f*,*g*). Further, there is a significant interaction between modulation degree and frequency for the ±12 dB and ±100% maskers (*h*). The interaction appears in such a way that ±12-dB modulation yields speech recognition scores about 5% greater than ±100% does, when modulation frequency is 2.1 or 19.9 Hz. When modulation frequency is 4.9 or 10.2 Hz, on the other hand, speech recognition with ±12-dB modulation is on the average 8% less than with 100%.

Figure 6 shows the results for the five subjects tested at the higher modulation frequencies 50, 80, and 100 Hz, with modulation degree ±12 dB. The results for 2–20 Hz (from the original experiment) is also shown for comparison. The ANOVA (Table IV) revealed significant differences ($p < 0.001$) between modulation frequencies and that the average score at high frequencies is significantly different both from the unmodulated noise score and from the average score at low frequencies. Further, the score at 50 Hz differs significantly from the average of the scores at 80 and 100 Hz, whereas the scores at 80 and 100 Hz do not differ. Finally, the average of the 80- and 100-Hz scores differs from the unmodulated score.

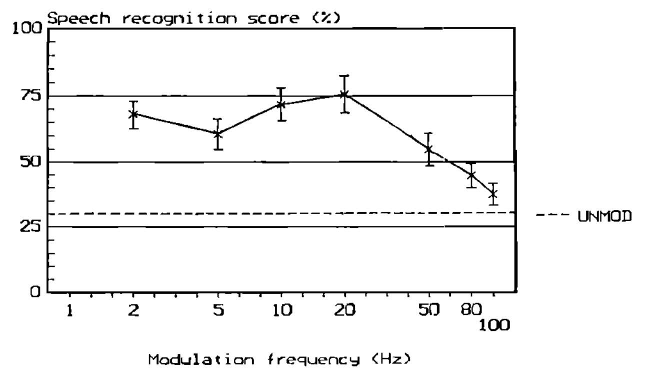

FIG. 6. Results of the additional experiment on five of the young, normal-hearing subjects. The mean ± one standard deviation is indicated at each modulation frequency tested.

2. Elderly, normal-hearing subjects

Figure 7 shows the mean speech recognition scores in the various AM noise maskers for the group of elderly subjects with normal audiograms. Table V presents the results of the ANOVA. When comparing the outcome with the results from the young, normal-hearing group the situation is very much the same both qualitatively and quantitatively. The differences concern the interaction between modulation degree and modulation frequency, which was

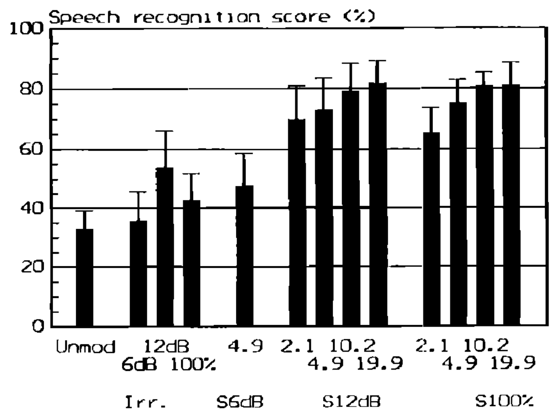

FIG. 7. Mean speech recognition scores for the group of elderly, normal-hearing subjects obtained with each masker. The horizontal line above each bar shows mean plus one standard deviation.

TABLE V. ANOVA table for the results obtained on the group of elderly normal-hearing subjects. The mean square for the interaction between maskers and subjects is used as the error estimate in all F ratios. SS=sum of squares, DF=degrees of freedom, MS=mean square. MOD is a collective denomination of the modulated maskers 2–13. Likewise, IRR designates maskers 2–4, S maskers 5–13, S12 maskers 6–9 and S100 maskers 10–13. m(...) denotes the mean score across maskers within parenthesis. The comparison between modulation frequencies applies to the mean scores across ±12 dB and 100% modulation. In the interaction between modulation degree and frequency, comparisons are not made between scores but between score differences, indicated by a D. For example, D2.1 denotes the score difference between ±12 dB and 100% modulation at 2.1 Hz.

Comparison	SS	DF	MS	F	p
Between maskers	78840.54	12	6570.04	155.83	<0.001
Between subjects	9984.66	19	525.51	12.46	<0.001
Interaction	9612.69	228	42.16	···	···
Decomposition of SS between maskers:					
UNMOD–m(MOD)	19480.01	1	19480.01	462.04	<0.001
m(IRR)–m(S)	36751.02	1	36751.02	871.68	<0.001
Within IRR	3283.2	2	1641.6	38.94	<0.001
IRR6–m(IRR12&IRR100)	2116.8	1	2116.8	50.21	<0.001
IRR12–IRR100	1166.4	1	1166.4	27.67	<0.001
Within S	19326.31	8	2415.79	57.30	<0.001
S6/4.9–m(S12&S100)	14263.21	1	14263.21	338.30	<0.001
m(S12)–m(S100)	4.9	1	4.9	0.12	N.S.
Between mod. freq.	4768.1	3	1589.37	37.70	<0.001
2.1–4.9	884.45	1	884.45	20.98	<0.001
10.2–19.9	42.05	1	42.05	1.00	N.S.
m(2.1&4.9)–m(10.2&19.9)	3841.6	1	3841.6	91.12	<0.001
Interaction, degr&freq.	290.1	3	96.7	2.29	<0.1
m(D2.1&D19.9)–m(D4.9&D10.2)	211.6	1	211.6	5.02	<0.05
D4.9-D10.2	2.45	1	2.45	0.06	N.S.
D2.1-D19.9	76.05	1	76.05	1.80	N.S.

not statistically significant for the group of elderly listeners with normal hearing but was significant for the young subjects (Table III).

3. *Elderly, hearing-impaired subjects*

The test results obtained without hearing aids (unaided) are shown in Fig. 8 and those obtained with hearing aids (aided) in Fig. 9. The maximum release of speech masking found, using sinusoidal modulation of ±12 dB at 10–20-Hz modulation frequency, increased the speech recognition score from 30% to close to 55%. This is significantly less than in the normal-hearing groups where the corresponding increase was from 30% to around 80% ($p < 0.001$, Student's t test). As shown in Table VI, there were no significant differences between unaided and aided conditions. However, there was again significant differences between the various maskers used. The differences generally correspond with those found in the elderly, normal-hearing group although they are much less pronounced here.

B. Speech-to-noise ratio for 30% correct speech recognition

The speech-to-noise ratio that yielded the criterion score of 30% correctly repeated words in nonmodulated noise varied considerably between the groups. The average S/N differed significantly between the three subject groups ($p < 0.001$, Student's t test) but not between aided and unaided listening for the elderly hearing impaired subjects (Fig. 10).

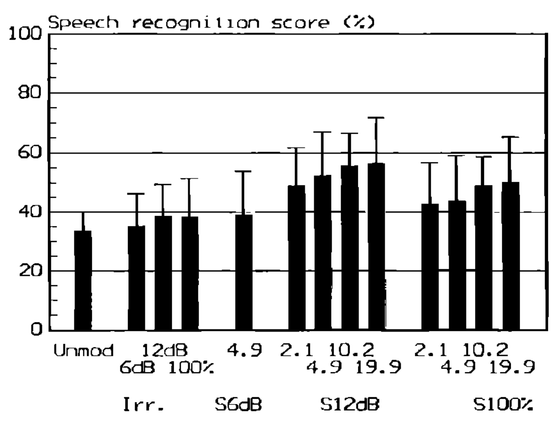

FIG. 8. Mean speech recognition scores (plus one standard deviation) for the group of elderly, hearing-impaired subjects for each masker in the unaided listening condition.

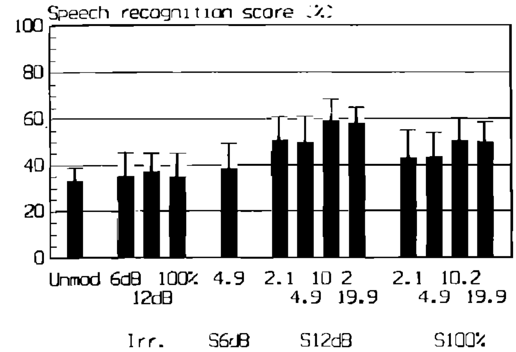

FIG. 9. As Fig. 8 but in the aided listening condition.

TABLE VI. ANOVA table for the results from the elderly hearing-impaired subjects. Individual error mean squares are used in the *F* ratios. SS=sum of squares, DF=degrees of freedom, MS=mean square. ERRMS=error mean square, DFERR=degrees of freedom for ERRMS. The *F* ratio is MS/ERRMS. AIDED denotes measurements with hearing aid and UNAIDED without. MOD is a collective denomination of the modulated maskers 2–13. Likewise, IRR designates maskers 2–4, S maskers 5–13, S12 maskers 6–9, and S100 maskers 10–13. m(...) denotes the mean score across maskers within parenthesis. The comparison between modulation frequencies applies to the mean scores across ±12 dB and 100% modulation. In the interaction between modulation degree and frequency, comparisons are not made between scores but between score differences, indicated by a D. For example, D2.1 denotes the score difference between ±12 dB and 100% modulation at 2.1 Hz.

Comparison	SS	DF	MS	DFERR	ERRMS	F	p
Between maskers	32758.89	25	1310.36	475	78.45	16.70	<0.001
Between subjects	24344.7	19	1281.3	475	78.45	16.33	<0.001
Interaction	37262.80	475	78.45	···	···	···	···
Decomposition of SS between maskers:							
m(AIDED)–m(UNAIDED)	0.28	1	0.28	19	664.43	0.00	N.S.
Between maskers, *unaided*	14201.15	12	1183.43	228	47.32	25.01	<0.001
UNMOD–m(MOD)	2800.62	1	2800.62	19	47.87	58.50	<0.001
m(IRR)–m(S)	5780	1	5780	19	75.11	76.95	<0.001
Within IRR	154.53	2	77.27	38	33.41	2.31	N.S.
Within S	5466	8	683.25	152	47.26	14.47	<0.001
S6/4.9–m(S12&S100)	2117.02	1	2117.02	19	82.60	25.63	<0.001
m(S12)–m(S100)	1890.62	1	1890.62	19	20.47	92.37	<0.001
Between mod. freq.	1413.08	3	471.02	57	59.78	7.38	<0.001
2.1–4.9	80	1	80	19	57.37	1.39	N.S.
10.2–19.9	22.05	1	22.05	19	53.52	0.41	N.S.
m(2.1&4.9)–m(10.2&19.9)	1311.02	1	1311.02	19	68.45	19.15	<0.001
Interaction, degr&freq.	45.27	3	15.09	57	31.88	0.47	N.S.
D2.1–D4.9	33.8	1	33.8	19	35.38	0.96	N.S.
D10.2–D19.9	0.45	1	0.45	19	31.92	0.01	N.S.
m(D2.1&D4.9)–m(D10.2&D19.9)	11.02	1	11.02	19	28.34	0.39	N.S.
Between maskers, *aided*	18557.46	12	1546.46	228	60.74	25.46	<0.001
UNMOD–m(MOD)	2977.66	1	2977.66	19	62.28	47.81	<0.001
m(IRR)–m(S)	8187.76	1	8187.76	19	85.26	96.03	<0.001
Within IRR	74.53	2	37.27	38	79.09	0.47	N.S.
Within S	7317.51	8	914.69	152	52.90	17.29	<0.001
S6/4.9–m(S12&S100)	2586.74	1	2586.74	19	59.53	43.46	<0.001
m(S12)–m(S100)	2387.02	1	2387.02	19	37.50	63.66	<0.001
Between mod. freq.	2317.48	3	772.49	57	62.97	12.27	<0.001
2.1–4.9	6.05	1	6.05	19	83.63	0.07	N.S.
10.2–19.9	16.2	1	16.2	19	45.67	0.35	N.S.
m(2.1&4.9)–m(10.2&19.9)	2295.22	1	2295.22	19	59.59	38.51	<0.001
Interaction, degr&freq.	26.28	3	8.76	57	45.76	0.19	N.S.
D2.1–D4.9	11.25	1	11.25	19	53.67	0.21	N.S.
D10.2–D19.9	1.8	1	1.8	19	25.8	0.07	N.S.
m(D2.1&D4.9)–m(D10.2&D19.9)	13.22	1	13.22	19	57.80	0.23	N.S.

III. DISCUSSION

A. Effects of modulation characteristics

For normal-hearing listeners, an amplitude-modulated noise masks speech no more (and usually less) than unmodulated noise of equal spectrum and rms level. Thus no evidence has been found for amplitude modulations of the masker to *increase* its speech-masking effect. The release of masking caused by the regular sinusoidal modulation is considerable, increasing the speech recognition scores from about 30% to about 80% with constant S/N ratio under optimal conditions. Assuming that Hagerman's (1982) psychometric function for the speech test material for normal hearing subjects is also valid for our subjects, the observed maximum release of masking corresponds to an improvement in S/N ratio of approximately 3 dB. For the irregular modulation, intended to bear a rough resemblance to the modulation spectrum of speech, the release of speech masking was relatively small. Generally, the results

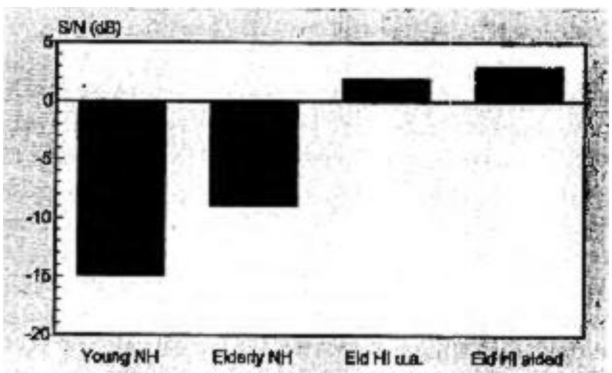

FIG. 10. Mean speech-to-noise ratio for 30% correct in unmodulated masking noise for the three groups of test subjects. NH=normal hearing, Eld HI u.a.=elderly, hearing impaired in unaided condition, Eld HI aided=elderly, hearing impaired in aided condition.

suggest that predictive methods for speech recognition such as the Articulation Index may overestimate the speech-masking effect of a noise containing significant amplitude modulations since those methods are based essentially on the average masker power in each frequency band without considering any modulations.

Festen and Plomp (1990) used broadband noise maskers with long-term average spectrum of either the male or the female voice used for the speech signal and with the intensity controlled by the speech signal envelope. The improvement in masked speech recognition threshold they found on young normal-hearing subjects was in the range 3 to 4 dB, which is of the same order of magnitude as our results for sinusoidal AM. In an earlier study (Festen and Plomp, 1986), where they used sinusoidal *intensity* modulation (100%) of the masker, normal-hearing test subjects on average improved their speech recognition threshold in noise by 5.5 dB in S/N ratio for the optimum modulation as compared to the unmodulated noise. Considering the difference between sinusoidal intensity and amplitude modulation, these results seem to agree well with ours.

Festen and Plomp (1990) suggested that comodulation masking release (CMR) might at least partly explain the release of masking found in AM masking noise. However, the results of a recent study (Grose and Hall, 1992) showed that whereas CMR may occur for speech detection, no suprathreshold effects were found (i.e., there was no influence of CMR on speech recognition scores above zero).

It appears that the release of masking is related to the presence of silent intervals of suffiently long duration to allow the normal-hearing listener to pick up important speech elements. This positive effect must be considerably greater than any negative effect caused by the increase in masking during the louder than average intervals of the AM noise.

1. The influence of modulation type

The two modulation types, irregular and sinusoidal modulation, produced very different results. This is most likely explained by the lower probability of noise intervals of sufficiently low-level and long duration to allow the listener to pick up meaningful fragments of the speech signal, the irregularity itself having no likely effect. Recall that the amplitude histograms differ between the two types of modulation, with the sinusoidal modulation showing higher relative occurrence of very low and very high amplitude than does the irregular modulation (Fig. 11). In addition, each period of the sinusoidal modulation guarantees an interval of relatively long duration with reduced noise amplitude. In contrast, when using irregular modulation the intervals with reduced amplitude noise are on average of much shorter duration. The range of variation of the instantaneous noise level was not very different between sinusoidal and irregular modulation for a given degree of modulation [Fig. 4(c)], while the speech-masking effect differed significantly, and therefore this characteristic is not very likely to contribute to the influence seen.

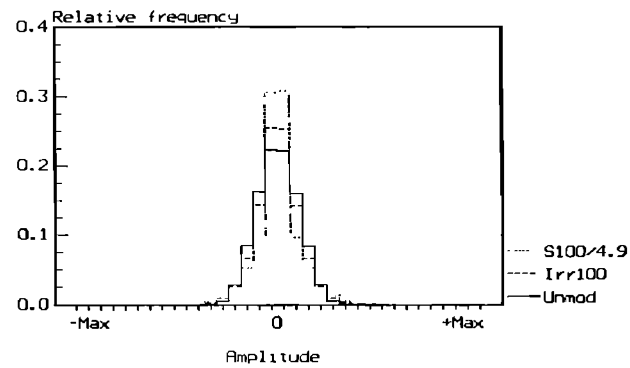

FIG. 11. Amplitude histograms for unmodulated noise (solid line), irregularly modulated noise 100% (dashed line) and sinusoidally modulated noise 100%, 4.9 Hz (dotted line). The histograms were obtained by sampling the signals at 20 kHz for about 10 s thus yielding a total of 204 800 sample per signal. The curves show the relative portion of samples in each amplitude class. ± Max is the input range of the A/D converter used during the sampling process.

2. The influence of modulation degree

The release of speech masking found at ±6-dB modulation was much smaller than at ±12 dB and ±100%. However, between ±12 dB and ±100% no significant differences existed for the normal-hearing groups. (In the group of elderly, hearing-impaired subjects these two modulation degrees did produce different speech masking effects. This is discussed below in section on the effect of hearing impairment.) These relations held for both irregular and sinusoidal modulation and are consequences of the modulation depths of the signals. Figure 4(c) shows the envelope variation range with rms correction applied, for each modulation type and degree. The more deeply modulated maskers produced the highest scores for the normal hearing subjects. This finding is in agreement with earlier studies using periodic modulation (Dirks *et al.*, 1969; Carhart *et al.*, 1966). The absence of significant difference between ±12 dB and ±100% modulation despite the large difference in modulation depth indicates that when a certain modulation depth is reached, very little is gained by increasing it further.

The "depth of the valleys" in the modulated noise evidently has a greater influence on the speech masking properties than the "height of the peaks." If not, irregular modulation ±12 dB would be a more effective masker than irregular modulation ±100%, which is not the case. It seems that the listener can take advantage of the less loud intervals of the noise to increase speech recognition, and that this advantage is not counterbalanced during the louder noise intervals.

3. The influence of modulation frequency

The effect of varying the modulation frequency is seen in Fig. 6 which applies to ±12-dB modulation. The scores are highest at 10.2 or 19.9 Hz and somewhat lower at 2.1 and 4.9 Hz. At frequencies above 20 Hz the scores seem to drop off and asymptotically reach the same score as for unmodulated noise. This general pattern seems to agree well with the results reported by Festen and Plomp (1986) using sinusoidal intensity modulation of the masker. The trend is supported by the significant differences between the average score with 2.1 and 4.9 Hz and the average score of 10.2 and 19.9 Hz as seen in all three groups of subjects (Figs. 5, 7, and 8).

The frequency dependence cannot be explained by the amplitude histograms, since these are mean values over a relatively long time (compared to the modulation period) and thus are equal for all sinusoidal modulations. Instead, the speech recognition score obtained with a certain AM masker modulation frequency is more likely to be related to the temporal characteristics of the masking noise, i.e., the number and duration of the low-level intervals that provide release of masking and high-level intervals that provide complete masking. At low modulation frequencies (below about 5 Hz) the high-level intervals are relatively long but sparse, which means that phonemes and syllables or even whole words may be masked, resulting in lower speech recognition. For frequencies between 10 and 20 Hz, the low-level intervals are shorter but more frequent, which partly counterbalances the loss of information during the higher level intervals and thus speech recognition increases. At higher frequencies (above about 30 Hz), the low-level intervals are very frequent but too short to aid speech recognition, which again decreases.

An alternative or additional interpretation of the increased speech masking effect of the AM noise at higher modulation frequency is in terms of auditory temporal resolution. The limited temporal resolution of the auditory system makes the masker envelope variations less audible and less useful with regard to the collection of informative fragments of the speech signal as the modulation frequency is increased. Considering the curve shown in Fig. 6 as representing a low-pass function, the cutoff frequency of this low-pass function can be estimated to be about 40 Hz. This agrees well with the approximately 50 Hz found by Viemeister (1979) and Formby (1985) in experiments concerning the detection of sinusoidal amplitude modulation of a wideband noise.

B. The effect of hearing impairment

The most striking effect of the hearing impairment found was the need for speech-to-noise ratios on average 11 to 12 dB better than the normal-hearing group of the same age in order to reach the criterion performance used in the study, i.e., 30% of the speech test material correct in unmodulated masking noise. A partial explanation for this finding is that the higher frequency range of the speech signal was inaudible due to the subjects' high-frequency loss and insufficient hearing-aid gain at frequencies above 2 kHz. However, Articulation Index calculations can account for no more than 4 dB of this difference in S/N. Therefore, a likely contributing factor is impaired spectral resolution, another type of distortion strongly correlated with sensorineural hearing loss (Ludvigsen, 1985).

The general effect of release of speech masking by amplitude modulation of the masking noise found among the elderly with hearing impairment was qualitatively very similar to the findings in the normal-hearing groups with the main difference being quantitative: The release of masking was much less for the hearing impaired. A likely explanation for this finding is the impaired temporal resolution commonly found in subjects with sensorineural hearing loss (Fitzgibbons and Wightman, 1982; Nelson and Freyman, 1987). Bacon and Viemeister (1985) noted a reduced sensitivity to detect amplitude modulation in broadband noise (i.e., the degree of modulation had to be larger than normal for detection), but the main influence of varying the modulation frequency was very similar in normal-hearing and hearing-impaired listeners. This finding also agrees with the present study. Impaired temporal resolution in the subjects with hearing loss could hypothetically consist of several components, one being increased post-masking from the louder into the less loud intervals of the masker and another related to a type of temporal integration, resulting in a reduced ability to retrieve important speech information during the short intervals of less loud masker. However, the present data provide no further insight into this.

In the group of elderly hearing-impaired subjects the pattern of unmasking did differ somewhat also qualitatively from that seen in the normal-hearing groups with regard to modulation type: Sinusoidal modulation of ±12 dB gave rise to more release of speech masking than ±100% for all four modulation frequencies tested (Figs. 8 and 9 and Table V). Figure 4 offers no evident explanation for this finding in terms of range of variation of instantaneous sound pressure level. We have no reasonable explanation to offer for this particular result.

C. The effect of age

In addition to the effect of hearing impairment on the degree of unmasking provided by amplitude modulation as discussed above, our study also showed large differences in S/N required by the subject groups to reach the criterion score of 30% correct in *un*modulated noise (Fig. 10). One unexpectedly large difference was that between the groups of young and of elderly normal-hearing subjects. As shown in Fig. 1 the average audiogram of the elderly group was essentially normal with mean hearing threshold levels of only 17 dB at 4 kHz and 24 dB at 8 kHz. Applying the Articulation Index to these average hearing threshold levels and the spectrum of the speech and the noise used in the study, the average hearing loss of the group accounted for less than 2 of the 6-dB group difference. The most reasonable explanation for the remaining difference of 4 dB is age-related differences in auditory signal processing between the two subject groups. Whether these depend on peripheral or central auditory pathways is not possible to infer from the present study.

The question of whether age as such has any effect on speech recognition through aging in central pathways or whether all differences found can be attributed to differences in peripheral auditory function has been discussed in several studies. Poulsen and Keidser (1991) found no age effect in speech recognition in noise when using monosyllabic test words with carrier phrase and a four alternative forced choice paradigm. Using sentence material in noise and a variety of cognitive tests, van Rooij and Plomp (1992) concluded that age-related changes in cognitive ability, related to speech recognition in noise are very unlikely; "age differences with respect to speech perception are most likely due to differences in auditive factors, notably differences in auditory sensitivity." However, Dubno *et al.* (1984) showed that the choice of speech material has a clear influence on age differences in speech perception. For bisyllabic test words and sentences with high predictability they found a difference in speech-to-noise ratio of 2 dB between a younger and an elderly group, whereas with low-predictability sentences the difference increased to 4 dB. Glasberg and Moore (1990) found a statistically significant correlation between speech recognition in noise and age when partialling out the average pure tone hearing loss, and concluded that "age emerges as a significant predictor of the ability to understand speech in noise."

Distorted speech of various types has been used in clinical tests of central auditory dysfunction. Korsan-Bengtsen (1973) found statistically significant differences between young and elderly normal-hearing listeners using interrupted speech (4, 7, and 10 interruptions per second) and time-compressed speech. Her elderly group was very similar to ours: Twenty subjects aged 50–60, otologically and neurologically normal with an average hearing threshold level at 4 kHz of 20 dB HL (our group 17 dB). Rodriguez *et al.* (1990) also evaluated a group of elderly subjects with close to normal pure tone audiograms. Their results indicate central auditory involvement related to age that can occur without concomitant decline in peripheral hearing sensitivity, cognitive function or linguistic competence.

The speech test material used in the present study is of the low predictability category. The task of repeating the five word sentences is also sensitive to changes in short-term memory functions. The conclusion from our study remains: Differences in auditory sensitivity explains only a part of the difference in speech-to-noise ratio found between young and elderly normal-hearing listeners. The remaining part of the difference is most likely due to aging effects which are not reflected by the pure tone audiogram.

D. Effect of wearing hearing aids

For the group mean results, no significant differences were found between aided and unaided conditions (Figs. 8 and 9). In a previous study, Arlinger *et al.* (1988) found that aided speech recognition thresholds in noise were superior to unaided for subjects using In-The-Canal hearing aids. The finding in the present study of a tendency toward the opposite result was not expected, and may be at least partially explained by the relatively poor average hearing aid gain at frequencies above 2 kHz (Fig. 3). The articulation index predicted a slightly better unaided recognition of the speech at the elevated level than the aided recognition at normal speech level, just as was found in the experiment reported here.

IV. CONCLUSIONS

Speech recognition is generally better in amplitude-modulated masking noise than in unmodulated noise of equal rms sound pressure level. The unmasking due to amplitude modulation increases with increasing modulation depth. For sinusoidal modulation, modulation frequencies in the 10- to 20-Hz range provide the largest release of masking.

The influence of the modulation parameters on the release of speech masking was similar for young and elderly subjects with normal hearing. However, elderly subjects with impaired hearing showed significantly smaller release of speech masking as a consequence of amplitude modulating the masking noise. With irregular AM no unmasking at all was recorded. No significant differences were found when this group of listeners was tested with and without hearing aid.

For all types of maskers used (unmodulated as well as amplitude-modulated noise), there were effects of both age and auditory function on the speech-to-noise ratio required to reach a criterion speech recognition score.

ACKNOWLEDGMENTS

The financial support by the National Swedish Environmental Protection Board is gratefully acknowledged. Valuable statistical advice was received by E. Leander, Ph.D., and E. Enqvist, Ph.D. Carl Ludvigsen, M.Sc., provided valuable help with measurements and a computer program for articulation index calculations.

Arlinger, S., Kinnefors, C., and Jerlvall, L. (1988). "A comparison between POGO and conventional hearing aid fitting," in *Hearing Aid Fitting*, edited by J. Hartvig Jensen (Proc. 13th Danavox Symposium, Copenhagen), pp. 97–113.

Bacon, S. P., and Viemeister, N. F. (1985). "Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners," Audiology **24**, 117–134.

Berry, R. C., and Nerbonne, G. P. (1972). "Comparison of the masking functions of speech-modulated and white noise," J. Acoust. Soc. Am. **51**, 121 (A).

Carhart, R., Tillman, T. W., and Johnson, K. R. (1966). "Binaural masking of speech by periodically modulated noise," J. Acoust. Soc. Am. **39**, 1037–1050.

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

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

Carhart, R., Johnson C., and Goodman, J. (1975). "Perceptual masking of spondees by combination of talkers," J. Acoust. Soc. Am. Suppl. 1, **58**, S35.

Danhauer, J. L., and Leppler, J. G. (1979). "Effects of four noise competitors on the California consonant test," J. Speech Hear. Disord. **44**, 354–362.

Danhauer, J. L., Doyle, P. C., and Lucks, L. (1985). "Effects of noise on NST and NU 6 stimuli," Ear Hear. **6**, 266–269.

Dirks, D. D., Wilson, R. H., and Bower, D. R. (1969). "Effect of pulsed masking on selected speech materials," J. Acoust. Soc. Am. **46**, 898–906.

Dubno, J. R., Dirks, D. D., and Morgan, D. E. (1984). "Effects of age and mild hearing loss on speech recognition in noise," J. Acoust. Soc. Am. **76**, 87–96.

Fastl, H. (1987). "A background noise for speech audiometry," Audiol. Acoust. **26**, 2–13.

Festen, J. M., and Plomp, R. (1986). "The extra effect of masker fluctuations on the SRT for hearing-impaired listeners," Proc. 12th Int. Congr. Acoust., Toronto, Vol. 1, Paper B11-4.

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

Fitzgibbons, P. J., and Wightman, F. L. (1982). "Gap detection in normal and hearing-impaired listeners," J. Acoust. Soc. Am. 72, 761–765.

Formby, C. (1985). "Differential sensitivity to tonal frequency and to rate of amplitude modulation of broadband noise by normally hearing listeners," J. Acoust. Soc. Am. 78, 70–77.

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

Glasberg, B. R., and Moore, B. C. J. (1990). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech," Scand. Audiol. Suppl. 32.

Grose, J. H., and Hall, J. W. (1992). "Comodulation masking release for speech stimuli," J. Acoust. Soc. Am. 91, 1042–1050.

Hagerman, B. (1982). "Sentences for testing speech intelligibility in noise," Scand. Audiol. 11, 79–87.

Horii, Y., House, A. S., and Hughes, G. W. (1970). "A masking noise with speech-envelope characteristics for studying intelligibility," J. Acoust. Soc. Am. 49, 1849–1856.

ISO 389 (1991). "Acoustics—Standard reference zero for the calibration of pure-tone air conduction audiometers," International Organization for Standardization, Geneva.

ISO 7029 (1984). "Acoustics—Threshold of hearing by air conduction as a function of age and sex for otologically normal persons," International Organization for Standardization, Geneva.

Korsan-Bengtsen, M. (1973). "Distorted speech audiometry," Acta Otolaryngol. (Stockh.), Suppl. 310.

Ludvigsen, C. (1985). "Relations among some psychoacoustic parameters in normal and cochlearly impaired listeners," J. Acoust. Soc. Am. 78, 1271–1280.

Miller, G. A., and Licklider J. C. R. (1950). "The intelligibility of interrupted speech," J. Acoust. Soc. Am. 22, 167–173.

Møller, A. R. (1971). "Unit responses in the rat cochlear nucleus to tones of rapidly varying frequency and amplitude," Acta Physiol. Scand. 81, 540–556.

Nelson, D. A., and Freyman, R. L. (1987). "Temporal resolution in sensorineural hearing-impaired listeners," J. Acoust. Soc. Am. 81, 709–720.

Plomp, R. (1986). "The hearing aid and the noise problem," Hear. Instrum. 37(6), 28–32.

Poulsen, T., and Keidser, G. (1991), "Speech recognition in young and elderly normal-hearing listeners in a closed response test," Scand. Audiol. 20, 245–249.

Rodriguez, G. P., DiSarno, N. J., and Hardiman, C. J. (1990). "Central auditory processing in normal-hearing elderly adults," Audiology 29, 85–92.

Tillman, T. W., Carhart, R., and Olsen, W. O. (1970). "Hearing aid efficiency in a competing speech situation," J. Speech Hear. Res. 13, 789–811.

van Rooij, J. C. G. M., and Plomp, R. (1992), "Auditive and cognitive factors in speech perception by elderly listeners. III. Additional data and final discussion," J. Acoust. Soc. Am. 91, 1028–1033.

Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. 66, 1364–1380.

Welzl-Muller, K., and Stephan, K. (1988). "Speech recognition performance with hearing-aid in noiseless and noisy conditions," Audiol. Akustik 27, 108–119.

Wilson, R. H., and Carhart, R. (1969). "Influence of pulsed masking on the threshold for spondees," J. Acoust. Soc. Am. 46, 998–1010.

Young, L. L., Jr., Parker, C., and Carhart, R. (1975). "Effectiveness of speech and noise maskers on numbers embedded in continuous discourse," J. Acoust. Soc. Am. Suppl. 1 58, S35.