The role of envelope fluctuations in spectral masking

Citation for published version (APA):
Heijden, van der, M. L., & Kohlrausch, A. G. (1995). The role of envelope fluctuations in spectral masking. *Journal of the Acoustical Society of America*, *97*(3), 1800-1807. https://doi.org/10.1121/1.412056

DOI:
10.1121/1.412056

Document status and date:
Published: 01/01/1995

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 03. Dec. 2024

The role of envelope fluctuations in spectral masking

Marcel van der Heijden and Armin Kohlrausch
Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 8 August 1994; revised 20 November 1994; accepted 2 December 1994)

Two experiments are reported in this study. In the first experiment the masking effect of five different types of narrow-band maskers was compared. The masker was either a tone, a narrow-band Gaussian noise, or a multiplication noise obtained by multiplying a sinusoid with a low-pass Gaussian noise. The noise maskers had a bandwidth of either 20 or 100 Hz. In all cases the masker had a center frequency of 1.3 kHz and a duration of 500 ms. Five-point growth-of-masking functions were measured using a 2-kHz tonal target with a duration of 400 ms, temporally centered in the masker. Six subjects participated in the experiment. Although considerable intersubject differences were observed, the data of all subjects showed several common trends. First, the tonal maskers produced more masking than the noise maskers. Second, Gaussian noise maskers produced more masking than multiplication noise maskers of the same bandwidth. Finally, 100-Hz-wide noise maskers produced more masking than 20-Hz-wide maskers of the same type. Differences in masked thresholds between the various masker types generally increased with masker level, and exceeded 25 dB in some conditions. The results are discussed in terms of masker envelope fluctuations. In the second experiment the masking effect was investigated for a bandpass noise at 1.3 kHz, with regular zero crossings, but with the envelope characteristics of a 100-Hz-wide Gaussian noise. Five-point growth-of-masking functions were measured using a tonal target of 2 kHz. Masked thresholds produced by this hybrid masker were not significantly different from those produced by a 100-Hz Gaussian masker, but differed significantly from those produced by 100-Hz-wide multiplication noise. This result indicates that differences in masking between Gaussian and multiplication noise are not due to their different fine structure, but to their different envelope statistics.

PACS numbers: 43.66.Dc, 43.66.Mk, 43.66.Ba [WJ]

INTRODUCTION

The masking behavior of narrow-band signals is generally believed to reflect the spectral selectivity of the hearing system (e.g., Schafer *et al.*, 1950). When the masker is a pure tone, however, measurements can be plagued by "false cues" such as beats and combination tones (e.g., Wegel and Lane, 1924). As a remedy, one often uses a narrow-band noise masker instead of a sinusoid (Egan and Hake, 1950). It is tacitly assumed that, apart from the elimination of the forementioned artefacts, this substitution has no effect on the masked thresholds. Under the assumptions of the so called power-spectrum model of masking (Fletcher, 1940; Patterson, 1974), it is indeed true that equally intense maskers of subcritical bandwidth which are centered around the same frequency, will produce equal masking. This model states that the detectability of a signal in the presence of a masker is fully determined by the signal-to-masker ratio after a linear filter process. As long as the masker bandwidth does not exceed the width of the auditory filters, it is obvious that the predicted amount of masking only depends on the masker's level and center frequency.

On the other hand, the substitution of a narrow band of noise for a tone introduces fluctuations in both envelope and fine structure. These fluctuations may play a role in the masking produced by these stimuli. In fact, it has been reported by several investigators that, under suitable spectral conditions, masker envelope fluctuations can cause a *release* from masking (Buus, 1985; Mott and Feth, 1986).

The main goal of the present study is to determine what aspects of a narrow-band masker, apart from intensity and center frequency, determine the masker's effectiveness. These "secondary aspects" include bandwidth, regularity of fine structure, and fluctuations of the envelope. Their importance can be assessed from a comparison between the effect of maskers that differ systematically in these aspects.

As a tool for comparing the different masker types, we have chosen growth-of-masking functions (masked threshold as a function of masker level). There are two reasons for this choice. First, in most studies which directly compare different types of narrow-band noise maskers, masked thresholds are measured as a function of some spectral parameter of the masker or the target (Buus, 1985; Mott and Feth, 1986; Moore and Glasberg, 1987). Thus few data are available which directly illustrate the level dependence of the differences between the maskers under consideration. Second, a comparison between data from different studies on narrow-band masking suggested that there exists a considerable difference between tone-on-tone and noise-on-tone growth-of-masking functions (van der Heijden and Kohlrausch, 1994).

In experiment I, we compare the masking effect of a tonal masker and four different noise maskers that differ in bandwidth and noise statistics. The bandwidth was either 20 or 100 Hz. The statistics were determined by the method of noise generation; the noise masker was either a bandpassed Gaussian noise or a multiplication noise, obtained by multiplying a sinusoid with a low-pass Gaussian noise. Both types of bandpass noise have been extensively used in masking experiments. There are two aspects in which Gaussian and

0001-4966/95/97(3)/1800/8/$6.00

multiplication noise differ. First, multiplication noise has regular zero crossings, while Gaussian noise has irregular zero crossings. Second, the two types of noise have different envelope distributions: The envelope of Gaussian noise has a Rayleigh distribution, while the envelope distribution of multiplication noise is the positive half of a Gaussian distribution.

Experiment II is designed to discriminate between the two aspects in which Gaussian noise and multiplication noise differ. For this purpose we constructed a "hybrid" masker which has regular zero crossings but a Rayleigh-distributed envelope. A comparison of the masking effectiveness of this hybrid masker with the two other noise maskers will reveal which aspect of the maskers is responsible for masking differences. If masking functions of the hybrid masker coincide with those of multiplication noise, then clearly the difference in masking is due to the regularity of the fine structure. Conversely, if the hybrid masker and Gaussian noise produce equal masking, then it is the envelope distribution that makes the difference.

I. EXPERIMENT I

A. Method

1. Stimuli

Simultaneous growth-of-masking functions (signal level at threshold versus masker level) were obtained for five types of maskers. The maskers were either a sinusoid (T), a bandpassed Gaussian noise (G), or a 'multiplication noise' (M), obtained by multiplying a sinusoid with a low-pass Gaussian noise (the bandwidth of this stimulus is two times the bandwidth of the original low-pass noise). The noise maskers G and M had bandwidths of either 20 or 100 Hz and were centered at 1.3 kHz. For each masker type, masked thresholds were measured at masker levels of 60, 66, 72, 78, and 84 dB SPL. All maskers had a duration of 500 ms. The signal was a 2-kHz sinusoid with a duration of 400 ms, temporally centered in the masker. Both the target and the masker were provided with 20-ms Hanning ramps. A bandpassed Gaussian noise with a lower cutoff frequency of 500 Hz and a higher cutoff frequency of 800 Hz was added to each masker interval to prevent subjects from using distortion products at 600 Hz (cubic difference tone) and 700 Hz (quadratic difference tone) as a detection cue. Its total level was always 25 dB below the total level of the 1.3-kHz masker. Data of Zwicker (1979) indicate that this level is sufficient to mask the forementioned distortion products. This was confirmed by a series of pilot studies.

All stimuli were digitally generated at a 32-kHz sampling rate and played out using the built-in 16 bit D/A converters of a Silicon Graphics Iris computer. Before each run a 4-s circular noise buffer was calculated according to the masker specifications of the run. From this buffer, 500-ms samples were drawn randomly from that buffer for each stimulus. A new buffer was calculated for each run.

Bandlimited Gaussian noise was produced as follows. First, a 4-s buffer of wideband Gaussian noise was obtained by drawing independent samples from a Gaussian distribution. A discrete Fourier transform was applied to this buffer (leading to a spectrum with a spacing of 0.25 Hz between the components). The undesired spectral components were set to zero and an inverse Fourier transform yielded a 4-s buffer of bandlimited Gaussian noise. In this way, the long-term spectrum of the cyclic noise buffer had infinitely steep spectral edges. The multiplication of noise buffers with a sinusoid in order to obtain multiplication noise did not affect the shape of the spectral edges. The steepness of the spectral edges of the actual noise maskers, as presented in each interval, was only limited by their duration of 500 ms and the use of Hanning ramps of 20 ms.

Stimuli were presented diotically via a Telephonics TDH 49 headset mounted in fluid-filled circumaural cushions. The earphone had previously been calibrated by means of a probe microphone placed at the ear canal entrance of several subjects. The response was flat ±3 dB over the range 500 to 6000 Hz as measured at the ear canal entrance.

2. Procedure

Masked thresholds were determined with a three-interval forced-choice adaptive procedure (Levitt, 1971). Each trial contained three 500-ms observation intervals separated by 200 ms. The masker occurred in all three intervals. The target occurred randomly but with equal probability in one of the three intervals. After the subject's response was collected, a 300-ms pause preceeded the next trial. Correct-answer feedback was provided on a computer screen. Each trial block began with the target about 20 dB above masked threshold. After two consecutive correct responses at the same target level, it was decreased and for each incorrect response it was increased. This procedure tracks the 70.7% correct point of the psychometric function. The step size was 8 dB at the beginning of each block, was reduced to 4 dB after the second reversal, and to 2 dB after the fourth reversal. Using 2-dB steps, ten more reversals were obtained. The threshold for a block was estimated by taking the median of the signal levels of these ten last reversals. Thresholds reported in this paper are the averages of three single threshold estimates. If the standard deviation of the first three estimates exceeded 3 dB, more blocks were measured, until the standard deviation of the last three estimates did not exceed 3 dB. These last three estimates were averaged.

Absolute thresholds at 2 kHz were measured using a similar procedure. In this measurement, observation intervals were marked by a very weak 4-kHz tone (about 20 dB above absolute threshold) presented for 200 ms just prior to the intervals.

Subjects were tested in a single-walled sound-attenuating booth placed within a larger sound-attenuated room.

3. Subjects

Six subjects participated in the experiment. Four subjects A to D were paid an hourly wage for their services. Subjects E and F were the authors. All subjects reported normal hearing. Absolute thresholds at 2 kHz, measured as described above, were 8, 8, 0, 11, 9, and 5 dB SPL for subjects A, B, C, D, E and F, respectively.

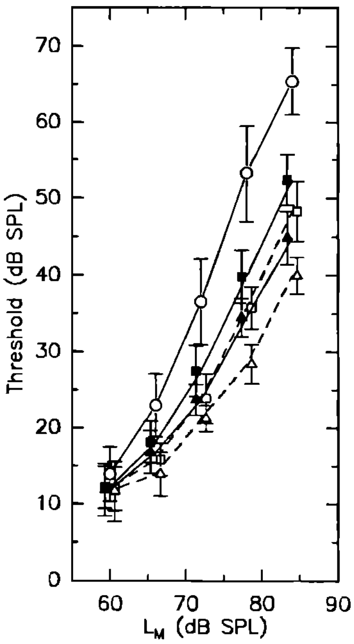

FIG. 1. Simultaneous growth-of-masking functions measured with various narrow-band maskers centered at 1.3 kHz. The target was a 2-kHz tone. The five curves indicate the results for five different maskers: a sinusoid (open circles), a 100-Hz-wide Gaussian noise (filled squares), a 20-Hz-wide Gaussian noise (open squares), a 100-Hz-wide multiplication noise (filled triangles), and a 20-Hz-wide multiplication noise (open triangles). Data points measured with 20-Hz-wide maskers are connected with dashed lines. Averaged data of six subjects are plotted; error bars indicate ± one standard deviation across the values of the six subjects. For visual clarity some datapoints have been shifted horizontally; the lines connecting the points, however, have been left unchanged.

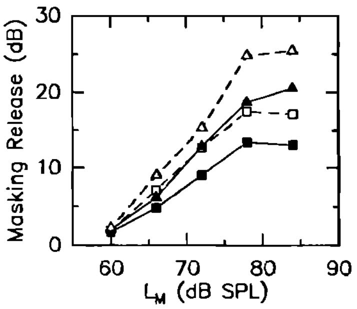

FIG. 2. Release of masking for noise maskers, plotted as a function of masker level. The ordinate indicates the difference in masked thresholds produced by a tone and a noise masker. The curves are based on the averaged data shown in Fig. 1.

B. Results

Figure 1 presents the averaged growth-of-masking functions (from here: "masking functions") for the different masker types. The error bars in this figure indicate plus and minus one across-subject standard deviation. The intersubject differences are largest for the conditions with a tonal masker (open circles); the standard deviation for the averaged tone-on-tone masking functions range from 3 to 6 dB (standard deviations of all individual data never exceed 3 dB). The data for noise maskers (squares and triangles) show less intersubject differences: The standard deviations of these data never exceed 4 dB. Despite the considerable intersubject differences, the data of all subjects show the same trends. These trends will now be described on the basis of the averaged data.

The masking function for a tonal masker (circles, solid lines) shows a considerably nonlinear, expansive behavior. The steepest part is the middle part of the function: when the masker level is increased from 66 to 78 dB SPL, the threshold increases by 30.3 dB. This corresponds to a slope of 2.5 dB/dB. At the highest and lowest masker levels the function is less steep. These observations agree with those of Schöne (1979), who measured tone-on-tone masking functions for various spectral masker-target separations.

At all levels the tonal masker produces more masking than each of the noise maskers. All four masking functions measured with noise maskers are less steep in the middle range compared to the tonal masking function, and do not show a flattening toward higher masker levels. When the level of a noise masker is increased from 66 to 84 dB, the thresholds increase by 26 to 34 dB, depending on the noise type. This corresponds to a slope of 1.4 to 1.9 dB/dB. At the lowest masker level (60 dB SPL) all types of noise maskers produce the same amount of masking (threshold at 12 dB SPL); the threshold for a tonal masker at 60 dB SPL is somewhat higher (14 dB SPL). The average absolute threshold for the 2-kHz target is 7 dB SPL.

In order to analyze the differences between the four noise maskers, a four-way analysis of variance was performed on the thresholds for noise maskers. The four factors of the analysis were masker level, noise statistics (Gaussian versus multiplication noise), masker bandwidth, and subject. It was found that all main effects and the interactions of level with each of the other factors were significant ($p<0.0001$). The root-mean-square error ("measurement error," as distinct from the "inter subject spread" shown in Fig. 1) was found to be 2.0 dB. The results of the analysis seem to justify a detailed description of the differences between the various noise maskers.

At higher masker levels the curves for different noise types diverge and the differences between the various noise types generally increase with masker level. Both bandwidth and statistics of the noise maskers affect thresholds. For both bandwidths, the Gaussian noise produces more masking than the corresponding multiplication noise. For both types of maskers, the 100-Hz-wide noise produces more masking than the 20-Hz-wide noise. The differences in masking efficiency between tone and noise maskers and between noise maskers among themselves are clearly illustrated by plotting the "release from masking" for the noise maskers as a function of masker level. This release is defined as the difference in masked thresholds for tonal and for noise maskers of equal level. Figure 2 shows this masking release for the average data. At the lowest masker level all noise maskers show a small release of 2 dB. The amount of release increases with masker level. At the highest masker level (84 dB SPL) the release from masking is 13 to 25 dB, depending on the noise type. The rate of increase is different for the different masker types. At the highest masker levels, the release from masking plateaus because the slope of the tonal masking function decreases, approximating the slope of the noise masking functions.

C. Discussion

Before discussing the results of experiment I, we want to emphasize that the effect of masker fluctuation depends markedly on the spectral relationship between masker and target. In the experiments of the present study the target frequency is well above the masker frequency. When this is not the case, the introduction of masker fluctuations can actually cause an *increase* of masking. This increase has been reported for target frequencies above, but close to the masker frequency (Johnson-Davies, 1981), for targets centered in narrow-band noises (Bos and de Boer, 1966; Hartmann and Pumplin, 1988), and for target frequencies below the masker frequency (Mott and Feth, 1986; Wright, 1992). When the target frequency is close to the masker frequency, interactions between masker and target can play a role in the observed effects.

In the present experiment, the spectral distance between masker and target is sufficiently large (about 700 Hz) to make the influence of direct masker–target interactions (beats) very unlikely. In addition, the 1.3-kHz masker is in the region of the "shallow tail" of tuning curves measured at 2 kHz (Stelmachowicz and Jesteadt, 1984). This means that the amount of masking is fairly insensitive to the masker frequency. Since all maskers used here have the same center frequency, and the spectra are only different in terms of bandwidth, spectral influences on the observed masked thresholds would in any event be second-order effects. The fact that both spectral effects and direct masker–target interaction play a minor role in the observed masking differences, points to the role of differences in the character of masker fluctuations. Before investigating this aspect in detail, we will now briefly discuss the various properties of the results of experiment I, and relate them to relevant studies reported in the literature.

The fact that masker fluctuations (both deterministic and stochastic) can give rise to a release from masking, has been reported by several authors (Buus, 1985; Mott and Feth, 1986; Moore and Glasberg, 1987). Buus (1985) observed the effect with both deterministic maskers (two-tone complexes) and stochastic maskers (narrow-band noise). For both masker types he used targets with a frequency at least 1.5 times the frequency of the highest masker component. He pointed out that the release from masking occurring with a stochastic masker is at odds with energy-detection models of masking (e.g., Green and Swets, 1966), which predict an *increase* of masking due to the introduction of masker uncertainty.

The differences in masking between the various noise maskers mainly appear as different slopes of the masking functions. This level dependency of masking differences is clearly demonstrated in Fig. 2, which shows the increase of masking release (taking tone-on-tone masking as a reference) with the level of the different noise maskers. A similar level dependency has been reported by Buus (1985), who compared masking by two-tone complexes and tonal maskers. The tonal masker and the upper component of the two-tone masker had the same frequency of 1075 Hz. For a masker level of 70 dB SPL, the maximum release from masking was 10 dB, whereas with a masker level of 90 dB SPL the maximum release was 25 dB. This across-level comparison, however, is somewhat obscured because he used different target frequencies for the different masker levels: The target frequency was 1600 to 2150 Hz for the 70-dB masker and 3400 to 4800 Hz for the 90-dB masker. The level dependency of the masking release is also consistent with results reported by Moore and Glasberg (1987). They used 1-kHz sinusoids, amplitude modulated at a rate of 8 Hz, with various modulation depths, to mask a tonal target with a frequency of 1.8 kHz. Three masker levels of 70, 75, and 80 dB SPL were applied. For a modulation depth of 0% (tonal masker) masked thresholds increased by as much as 24 dB when the masker level was increased by 10 dB. For the 100% modulated maskers, masked threshold only increased by 8 dB.

Due to a lack of data in the literature that focus on level dependence, we have not found studies reporting the "saturation" of masking release at high masker levels (above 80 dB SPL) as apparent from Fig. 2. This effect, however, might be inferred indirectly by comparing tone-on-tone masking data with narrow-band-noise-on-tone masking data from different studies, as has been done in van der Heijden and Kohlrausch (1994).

As the various growth-of-masking functions obtained in the present study differ in their slope, they cannot be transformed into each other by either a horizontal or a vertical shift. The former would correspond to a constant difference in "effective masker level," the latter to a constant amount of masking release (with a lower limit set by the absolute threshold of the signal). A more detailed discussion of the possible connections between the different growth-of-masking functions can be found in Sec. III of the present study.

The effect of bandwidth agrees with Buus' (1985) finding that the release from masking due to beats decreases with increasing beat rate. The highest rate at which Buus (1985) observed a release was 160 Hz. It is also consistent with many data on masking-period patterns, which show a deeper "modulation" of the masked threshold at low repetition rates of a pulsed or modulated masker (e.g., Zwicker, 1976). The bandwidth effect suggests a limited temporal resolution of the detection process. It is improbable that this limitation is due to filtering in the inner ear; an auditory filter centered at 2 kHz with an estimated bandwidth of 240 Hz (Glasberg and Moore, 1990) is unlikely to significantly affect the envelope modulations of a 100-Hz-wide masker. A thorough evaluation of temporal effects would include aspects of nonsimultaneous masking, duration effects, and modulation detection (cf. Fastl, 1975; Zwicker, 1976) and is beyond the scope of this paper which concentrates on the statistics of masker envelope fluctuations.

A difference in masking efficiency between multiplication noise (M) and Gaussian noise (G) has been reported previously by Mott and Feth (1986). They compared masking patterns of a tone, narrow-band noise with a flat temporal envelope (FM noise), Gaussian noise and multiplication noise. All maskers had a level of 75 dB SPL and were centered around 1500 Hz; they had bandwidths between 50 and 60 Hz. These four maskers have the following characteristics of fine structure and envelope. For the tone and the FM noise, the envelope is constant, while for the other two noise types, it is stochastically varying. The fine structure (i.e., the temporal distance between subsequent zero crossings) is regular for the tone and the multiplication noise, while it is fluctuating for the FM and the Gaussian noise. Mott and Feth (1986) found no significant difference in masking between the tone and the FM noise. In the region from 2 to 3 kHz the thresholds were highest for the tonal masker, about 10 dB lower for the Gaussian noise, and another 5 dB lower for the multiplication noise. These observations agree qualitatively with those of the present study.

Mott and Feth (1986) attributed the difference in masking between *M* and *G* noise to differences in fine structure; the multiplication noise has regular zero crossings, while the Gaussian noise does not. By concentrating on this aspect, they ignored another distinction between the two types of noise, namely, their envelope statistics. The envelope of Gaussian noise has a Rayleigh distribution, whereas the envelope distribution of multiplication noise is the positive half of a Gaussian distribution with zero mean (the envelope being the absolute value of the low-passed Gaussian noise used as the multiplicator). The latter has a maximum at zero (corresponding to the zero crossings of the low-passed noise used as the multiplicator), while in Gaussian noise the probability that the envelope is zero vanishes (cf. Fig. 3). Any strategy using minima in the envelope to improve detection will be sensitive to the differences of envelope statistics. This will be true regardless of the exact nature of the proposed strategy. In the second experiment reported in this paper, we investigate which aspect of the noise makes *M* a less effective masker than *G*: its different envelope distribution or its regular fine structure.

II. EXPERIMENT II

A. Method

1. *Stimuli*

Simultaneous growth-of-masking functions (signal level at threshold versus masker level) were obtained for a narrowband masker (*GM*) with envelope characteristics of Gaussian band-pass noise, but with regular zero crossings. Apart from the masker type, the present experiment replicated the measurements of experiment I (including the bandpassed Gaussian noise used to mask distortion products) as described in Sec. I A.

The masker (*GM*) was calculated as follows: First, a 100-Hz-wide Gaussian noise centered at 1.3 kHz was calculated. Next, the Hilbert envelope of this noise was extracted. Since the spectrum of this envelope is not bandlimited (Lawson and Uhlenbeck, 1950, section 3.8), all spectral components of the envelope above 100 Hz were set to zero. Finally, the low-passed envelope was multiplied by a 1.3-kHz sinusoid.

The envelope distributions of Gaussian noise, multiplication noise and "hybrid" noise are plotted together in Fig. 3. It can be observed from this figure that the envelope distribution of the *GM* masker is practically identical to that of Gaussian noise. In addition, the envelope *spectra* of *G* and *GM* noise are practically identical.

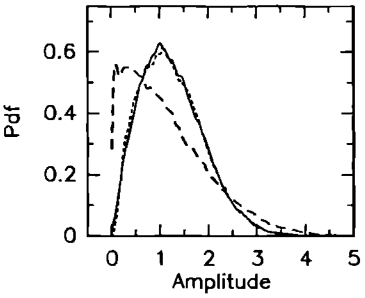

FIG. 3. Probability distributions of the envelope of three types of noise: Gaussian noise (solid line), multiplication noise (long dashes), and "hybrid noise" (see text, short dashes). The curves were calculated by averaging the envelope distributions of 100 independent 1-s samples of noise.

This identity of envelope characteristics can only be achieved at the cost of differences in the long-term power spectra. The long-term spectrum of the *GM* masker has the following properties. It is symmetric around the center frequency of 1300 Hz and has a total width of 200 Hz. The spectrum has a peak at the center frequency and decays approximately linearly on a power scale (cf. Lawson and Uhlenbeck, 1950, Fig. 3.14). It is unlikely that these spectral differences as such influence the masking behavior of the noise. First, the differences are confined to a spectral range between 1200 and 1400 Hz, whereas the signal frequency is 2 kHz. Second, as argued in Sec. I C, symmetrical spectral changes in the masker would in any case be second order effects in the current setup of the experiment.[1]

2. *Procedure and subjects*

Methods of data collection and processing are identical to those of experiment I. Three of the subjects that participated of experiment I (subjects D, E, and F) participated in this experiment.

B. Results and discussion

Figure 4 presents average masking functions for *GM* maskers, combined with data taken from experiment I. These reference data are masking functions of subjects D, E, and F, measured with 100-Hz-wide *G* and *M* maskers (cf. Fig. 1). For all three subjects the *GM* masking function coincides with the *G* masking functions more than it does with the *M* masking function. A three-way analysis of variance was performed on the results of experiment II together with the 100-Hz bandwidth data of experiment I. The three factors of the analysis were masker level, envelope, and fine structure (the hybrid masker *GM* fell in the same envelope class as the Gaussian noise but in the same fine structure class as the multiplication noise; Gaussian and multiplication noise differed both in envelope and fine structure). The analysis showed a significant effect of envelope [$F(1,120)=32.79$, $p<0.0001$] and its interaction with level [$F(4,120)=8.24$, $p<0.0001$]. Fine structure and its interaction with level were not significant ($p>0.5$). These observations clearly indicate that differences in envelope statistics rather than differences

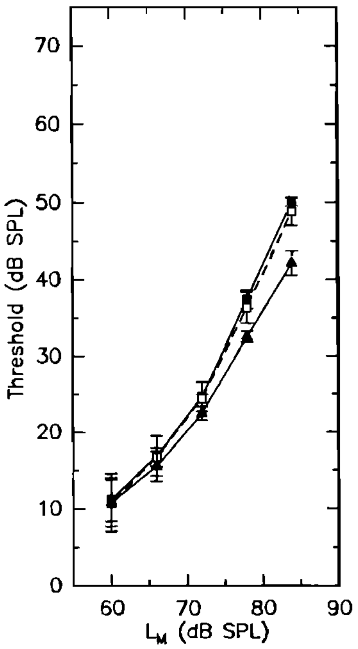

FIG. 4. Simultaneous growth-of-masking functions measured with three narrow-band maskers centered at 1.3 kHz. The target was a 2-kHz tone. The three curves indicate the results for three different maskers: 100-Hz-wide Gaussian noise (filled squares), "hybrid noise" (see text, open squares), and 100-Hz-wide multiplication noise (filled triangles). Averaged data of three subjects are plotted; error bars indicate ± one standard deviation.

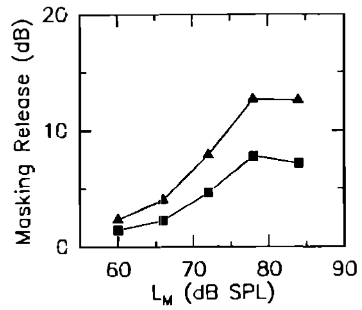

FIG. 5. Predicted release of masking for noise maskers, plotted as a function of masker level. The ordinate indicates the difference in masked thresholds produced by a tone and a noise masker (cf. Fig. 2). The curves show the prediction of the transform-and-average model described in the text. Triangles indicate the release for multiplication noise and squares the release for Gaussian noise.

in fine structure are responsible for the differences in masking produced by Gaussian and multiplication noise.

III. SUMMARY AND GENERAL DISCUSSION

In Experiment II it is found that differences in envelope statistics rather than fine structure are responsible for the different masking behaviour of Gaussian and multiplication noise. This being established, we can now summarize the factors that determine the masking differences between the various narrow-band maskers used in the present study:

—level: differences between the various masker types increase with masker level; at low masker levels the masked thresholds converge.

—bandwidth: masked thresholds increase with masker bandwidth.

—envelope statistics: the occurrence of marked envelope minima makes a masker less effective.

Although the fluctuating maskers of the present study are stochastic, it should be noted that each of the listed effects has its equivalent in conditions with deterministic maskers. As for the level and bandwidth effects, such equivalences have been discussed in Sec. I C. The influence of envelope statistics is analogous to the effect of modulation depth on the amount of masking by an amplitude-modulated masker (Moore and Glasberg, 1987). These equivalences indicate that an explanation of the effects should not be dependent on the *stochastic* nature of the fluctuations.

As an illustration of the influence of masker envelope statistics, let us consider the following oversimplified method for predicting masking by a fluctuating masker from masking by a tone. It is very similar to the procedure applied by Fastl (1975), but unlike his approach we ignore forward and backward masking effects. First, the time-varying envelope of the masker is extracted. At every instant the masker is assumed to produce an "instantaneous masked threshold." This threshold is simply the masked threshold for a tonal masker at a level equal to the instantaneous power of the fluctuating masker. These masked thresholds are derived from the averaged tone-on-tone masking data in Fig. 1 by a simple linear interpolation. Finally, the averaging of "instantaneous thresholds" leads to the predicted masked threshold. The rationale behind this procedure lies in the assumption that the masked thresholds are a measure of masker excitation and that the threshold is determined by the average excitation of the fluctuating masker (Zwicker, 1970). Mathematically, the procedure boils down to first transforming the instantaneous intensity of a fluctuating stimulus before averaging it. The scale in which the "instantaneous masked thresholds" are expressed and averaged is crucial for the predictions; like Fastl (1975) we express them in decibels. Compared with a power scale this attaches more weight to the envelope minima.

The procedure just described is an oversimplification in the sense that it focuses on the envelope statistics and ignores effects of temporal resolution. The procedure predicts the same thresholds for two maskers independent of their bandwidth and *rate* of envelope fluctuations, as long as their envelope *statistics* coincide. The results should, therefore, be compared with data obtained with small masker bandwidths, where limitations in temporal resolution should have the smallest effect on detection. In Fastl (1975) the model predictions agreed reasonably well with data for a 32-Hz-wide masker (for this bandwidth, the masker fluctuations are too slow to be smeared out by forward and backward masking). In any event, the just proposed procedure should not be regarded as a general model to predict masking by fluctuating maskers; we only introduce it here in order to demonstrate the effect of nonlinear scaling when applied to fluctuating stimuli.

Results of the prediction are shown in Fig. 5, where the difference between calculated thresholds for noise maskers and tonal maskers (masking release) is plotted (cf Fig. 2). The model predictions agree with the data of Fig. 2 in that a release from masking by the fluctuations is predicted, which is greater for multiplication noise than for Gaussian noise, and which shows a "saturation" at masker levels above 80 dB. But the predicted amount of release is much smaller than in the data of Fig. 2, both for the 20- and the 100-Hz bandwidth data: For Gaussian noise the predicted maximum release is 8 dB and for multiplication noise it is 13 dB. The experiments for a 20-Hz bandwidth show releases of 18 and 25 dB, respectively. The 100-Hz bandwidth data show releases of 13 and 21 dB, respectively. Despite this quantitative disagreement, the calculations demonstrate that a simple transform-and-average model is able to describe a release from masking due to envelope fluctuations. If the transforming part is compressive (i.e., if it disproportionally favors the envelope minima), the release will be greatest for maskers such as multiplication noise, which show distinct envelope minima. Apparently, the amplitude-to-dB transformation followed by the masker-to-target-level substitution described above, does not appreciate the envelope minima sufficiently. It is an exercise in curve fitting to find an even stronger compressive transformation that would "repredict" the 20-Hz data of Fig. 2.

In the light of the results shown in Fig. 5, Buus' (1985) remark that "it seems clear that no model which bases detection solely on the statistics of energy (or some transformation thereof) within a single channel can explain [his results]" is slightly misleading. If energy refers to the total energy, it is of course correct. But transforming the instantaneous power *before* averaging it will generally lead to different predictions for different masker statistics. The average power of a fluctuating masker is just one choice of conveniently describing its strength; it is by no means clear that the average power of a stimulus is a sensible measure of its masking effectiveness. In fact, we feel that the very effect of "masking release caused by envelope fluctuations" suggests that the masker's average power is not a good measure, and that the effect by itself does not justify Buus' (1985) conclusion that "...the auditory system applies some rather sophisticated multi-channel process to optimize detection."

Since masker envelope fluctuations can play an important role in the amount of masking produced by narrow-band maskers, this aspect cannot be ignored in "addition-of-masking" studies, where masking by a combination of different narrow-band maskers is measured. Combining two 50-Hz-wide noise maskers, Lutfi (1983) observed an "excess masking" of 10–17 dB over the amount of masking which would be expected from a simple power addition. In a reaction to this study, Moore (1985) argued that the excess masking should partly be ascribed to the influence of masker fluctuations. It is interesting to note that in the original study by Lutfi (1983) the maskers were Gaussian bands of noise, whereas Moore (1985) in his replication used multiplication noise, thereby ignoring possible differences in masking effect by these two types of noises. Moore's (1985) statement that "when two independent narrow-band noises are added together, the minima in the combined waveform are less pronounced than for either alone," is correct for multiplication noise as used in his study, but incorrect for Gaussian noise as applied in the original study by Lutfi (1983). This is so, because any sum of Gaussian processes is again a Gaussian process, and because any bandlimited Gaussian signal has a Rayleigh-distributed envelope. Thus a combination of Gaussian noise bands has the same envelope distribution as either component. It is of course true that addition of spectrally distant bands will introduce *faster* envelope fluctuations, but this is a bandwidth effect and not a matter of envelope distribution.

Another example that illustrates the role of masker statistics of combined maskers can be based on experiment I of the present study. According to the averaged data presented in Fig. 1, both a 100-Hz-wide Gaussian noise G of 78 dB SPL and a 100-Hz-wide multiplication noise M of 82 dB SPL produce the same masked threshold of 40 dB SPL. Addition of two independent samples of G results in the same kind of Gaussian noise, with a level of 78+3=81 dB SPL. From Fig. 1 we read that this "$G+G$" masker produces a threshold of 46 dB SPL. On the other hand, incoherent addition of two independent samples of M (with a 90-deg carrier phase difference) results in a Gaussian noise; this is the so called "quadrature" method of producing Gaussian noise. The level of this Gaussian noise is 82+3=85 dB SPL, and from Fig. 1 we find a masked threshold of 54 dB SPL. The difference of 8 dB is entirely due to the *type* of noise of the masker constituents. Models in which the masking effectiveness of a complex is predicted from the masking produced by the individual components (e.g., Humes and Jesteadt, 1989; Humes and Lee, 1992), will predict the same thresholds for the "$G+G$" and the "$M+M$" masker.

In summary, our data show that, particularly with high masker levels, attention should be paid to the exact nature of narrow-band stimuli used to mask a target at a frequency above the masker frequency. Statistics and rate of the masker fluctuations have a considerable effect on masked thresholds. These differences are by no means restricted to extremely slowly fluctuating maskers. On the contrary, masker fluctuations are relevant under many common experimental conditions, such as the "low-frequency tail" of psychoacoustic tuning curves. Unfortunately, in many published papers the influence of the type of noise used as masker (particularly multiplication noise versus Gaussian noise) has generally been neglected.

ACKNOWLEDGMENTS

The authors wish to thank Adrian Houtsma, Larry Humes, and an anonymous reviewer for valuable comments on earlier drafts of the manuscript.

---

[1] The choice of cutting off the envelope spectrum above 100 Hz is a compromise between affecting the waveform spectrum of the noise and affecting its envelope distribution. Generally, it is not possible to generate a stimulus that unifies a predescribed spectrum with arbitrary envelope characteristics. The problems encountered here are analogous to the problem of generating spectrally flat noise with a flat temporal envelope ("low-noise noise," Hartmann and Pumplin, 1988).

Bos, C. E., and de Boer, E. (1966). "Masking and discrimination," J. Acoust. Soc. Am. 39, 708–715.

Buus, S. (1985). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. 78, 1958–1965.

Egan, J. P., and Hake, H. W. (1950). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. 22, 622–630.

Fastl, H. (1975). "Loudness and masking patterns of narrow noise bands," Acustica 33, 266–271.

Fletcher, H. (1940). "Auditory patterns," Rev. Mod. Phys. 12, 47–65.

Glasberg, B. R., and Moore, B. C. J. (1990). " Derivation of auditory filter shapes from notched-noise data," Hear. Res. 47, 103–138.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Hartmann, W. M., and Pumplin, J. (1988). "Noise power fluctuations and the masking of sine signals," J. Acoust. Soc. Am. 83, 2277–2289.

Heijden, M. van der, and Kohlrausch, A. (1994). "Using an excitation-pattern model to predict auditory masking," Hear. Res. 80, 38–52.

Humes, L. E., and Jesteadt, W. (1989). "Models of the additivity of masking," J. Acoust. Soc. Am. 85, 1285–1294.

Humes, L. E., and Lee, L. W. (1992). "Two experiments on the spectral boundary conditions for nonlinear addition of masking," J. Acoust. Soc. Am. 92, 2598–2606.

Johnson-Davies, D. B. (1981). "The analogy between psychophysical and physiological tuning curves," unpublished Ph.D. thesis, University of Cambridge, England.

Lawson, J. L., and Uhlenbeck, G. E. (1950). *Threshold Signals* (McGraw-Hill, New York).

Levitt, H. (1971). "Transformed up–down methods in psychoacoustics," J. Acoust. Soc. Am. 49, 467–477.

Lutfi, R. A. (1983). "Additivity of simultaneous masking," J. Acoust. Soc. Am. 73, 262–267.

Moore, B. C. J. (1985). "Additivity of simultaneous masking revisited," J. Acoust. Soc. Am. 78, 488–494.

Moore, B. C. J., and Glasberg, B. R. (1987). "Factors affecting thresholds for sinusoidal signals in narrow-band noise maskers with fluctuating envelopes," J. Acoust. Soc. Am. 71, 69–79.

Mott, J. B., and Feth, L. L. (1986). "Effects of the temporal properties of a masker upon simultaneous-masking patterns," in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York), pp. 381–386.

Patterson, R. D. (1974). "Auditory filter shape," J. Acoust. Soc. Am. 55, 802–809.

Schafer, T. H., Gales, R. S., Shewmaker, C. A., and Thompson, P. O. (1950). "The frequency selectivity of the ear as determined by masking experiments," J. Acoust. Soc. Am. 22, 490–496.

Schöne, P. (1979). "Mithörschwellen-Tonheitsmuster maskierender Sinustöne," Acustica 43, 197–204.

Stelmachowicz, P. G., and Jesteadt, W. (1984). "Psychophysical tuning curves in normal-hearing listeners: Test reliability and probe level effects," J. Speech Hear. Res. 27, 396–402.

Wegel, R. L., and Lane, C. E. (1924). "The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear," Phys. Rev. 23, 266–285.

Wright, B. A. (1992). "The role of envelope fluctuations in an apparent demonstration of suppression in simultaneous masking," J. Acoust. Soc. Am. 91, 3436–3442.

Zwicker, E. (1970). "Masking and psychological excitation as consequences of the ear's frequency analysis," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Springer, Berlin), pp. 376–396.

Zwicker, E. (1976). "Mithörschwellen-Periodenmuster amplitudenmodulierter Töne," Acustica 36, 113–120.

Zwicker, E. (1979). "Zur Nichtlinearität ungerader Ordnung des Gehörs," Acustica 42, 149–157.