ACUSTICA · acta acustica
Vol. 83 (1997) 659–669

Detection of Tones in Low-noise Noise: Further Evidence for the Role of Envelope Fluctuations

Armin Kohlrausch, Ralf Fassel, Marcel van der Heijden*, Reinier Kortekaas, Steven van de Par, Andrew J. Oxenham
Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Dirk Püschel
Akustikbüro Göttingen, Kornmarkt 2, 37073 Göttingen, Germany

Dedicated to Prof. Manfred R. Schroeder on the occasion of his 70th birthday

Summary
This paper investigates the role of envelope fluctuations in simultaneous masking conditions. Thresholds for tones in noise with a flat temporal envelope (low-noise noise, LNN) were compared with those in Gaussian noise. All measurements were performed with a running-noise presentation of 500-ms maskers. The sinusoidal signal was spectrally and temporally centered in the masker. The main findings were: (a) The 5.5-dB threshold difference between 100-Hz-wide Gaussian and LNN maskers at 1 kHz that was previously observed using frozen noise (cf. Hartmann and Pumplin [J. Acoust. Soc. Am. **83**, 2277-2289 (1988)]) is also apparant for running noise, although thresholds are generally higher in the latter condition. (b) The threshold difference between Gaussian and LNN maskers at 1 kHz reaches a maximum of 9.4 dB at a masker bandwidth of 25 Hz, while at 10 kHz, the difference reaches a maximum of 15 dB at bandwidths of 50 and 100 Hz. For a 100-Hz-wide masker presented at different center frequencies, there is no advantage for LNN maskers below 1 kHz. Towards higher frequencies, the difference between the two noises increases and reaches about 15 dB at 10 kHz. (c) At 1 kHz with a 100-Hz bandwidth, decreasing the signal duration fromm 500 to 20 ms increases the threshold difference to 7.6 dB. (d) Thresholds in a dichotic condition, in which the masker is in phase and the signal is out of phase, lie within 2 dB for the two noise types, and are nearly constant for masker bandwidths between 5 and 100 Hz. It is argued that the primary detection cue in LNN is not an increase in energy, but rather an increase in envelope fluctuations due to the addition of the signal. This hypotheses is supported by simulations with an auditory-filterbank model. The simulations further suggest that, for a large LNN advantage, it is not sufficient that the LNN envelope is flat at the output of the *on-frequency* filter. In addition, it is crucial that *off-frequency* filters also yield a flat temporal envelope.

PACS no. 43.66Dc, 43.66Ba, 43.60.Cg

1. Introduction

In an extension of earlier work by Schroeder [1] on signals with a low peak factor, Hartmann and Pumplin addressed the question of how a noise signal with a given bandpass power spectrum can be manipulated to have minimal envelope fluctuations [2, 3, 4]. For the resulting signal the term low-noise noise (LNN) was introduced. Envelope fluctuations for a noise signal were minimized by optimizing the phase spectrum using a gradient search procedure (for details see [2]). This procedure is quite time consuming and not easily implemented, which might be one reason why LNN has been used by only a few psychoacoustic laboratories. The only perception experiment published so far using LNN maskers is the one reported in [3], where maskers with bandwidths of 100 and 300 Hz, centered around 1 kHz, were used.[1] Due to the restricted availability of independent LNN waveforms, the same 500-ms masker waveform was presented in all intervals of the adaptive procedure, i.e., the experiment was performed as a frozen-noise measurement. In order to average out peculiarities of the fixed masker-signal interaction, measurements were repeated for six signal starting phases. A sinusoidal curve was fitted to the results and the mean value of this curve served as the threshold value. Using this procedure, thresholds for the 100-Hz-wide LNN masker were 5.5 dB lower than the average for two Gaussian-noise maskers (based on the data for five subjects in Table IV of [3]). A comparable measurement with 300-Hz-wide maskers at 1 kHz revealed no advantage for LNN over Gaussian noise. As the authors argue, this negative result for a bandwidth larger than the critical bandwidth at the signal frequency reveals the importance of peripheral filtering: once an LNN masker is substantially influenced by auditory filtering, its envelope will no longer be flat and its masking properties will be similar to that of Gaussian noise. In this vein, Hartmann and Pumplin speculated that even the 100-Hz-wide LNN masker might have been affected by the auditory filter at 1 kHz, given that the lower limit of filter bandwidth estimates is of the order of 100 Hz. In other words, the difference between the two noise types might have been even larger than the observed 5.5 dB, if a narrower bandwidth had been used.

One aspect of the data that was not discussed by Hartmann and Pumplin [3] was the general amount of masking:

Received 18 June 1996,
accepted 11 December 1996.

* Present address: Univ. of Connecticut Health Ctr, Farmington, CT 06030, USA

[1] While this article was being written, two other studies concerning low-noise noise were presented at the 131st meeting of the Acoustical Society of America [5, 6]. In both studies, LNN stimuli were calculated using the algorithm of Pumplin and Hartmann and the interaction between masker and signal was random.

if their thresholds are expressed as signal-to-overall-noise ratios, the average value for the 500-ms signal in the 100-Hz-wide Gaussian noises amounts to -11.2 dB. This value is much lower than the results from other studies, where thresholds are in the range of -3 to -5 dB (e.g. [7, 8, 9]). The most obvious reason for these rather low threshold values lies in stimulus differences: while in the other cited studies the interaction between signal and masker waveform was random in each observation interval, it was a frozen interaction in the study of Hartmann and Pumplin [3]. One of the questions that we want to address in the present investigation is, therefore, whether the observed LNN advantage is also seen in a running-noise experiment.

The results of Hartmann and Pumplin [3] have been cited by other authors as evidence for the importance of envelope cues in (monaural) signal detection: the addition of a sinusoidal signal to an LNN waveform introduces envelope variations, which might be a more potent detection cue than the energy increment [10]. Given the increasing interest in envelope-based cues, for instance in the context of modulation analysis and of comodulation masking release, it would be quite useful to have more experimental data about masking properties of LNN available.

The present paper is intended to provide masking data for LNN and Gaussian noise for a variety of experimental conditions. In order to do so, a time-efficient method to calculate LNN waveforms had to be found. This method will be explained in the following section. The first measurements addressed some of the questions that follow from the above discussion of the experiments by Hartmann and Pumplin [3]. Their experiment was first repeated as a running-noise experiment, and then narrower masker bandwidths and shorter signal durations were also tested. In order to test the influence of envelope-based cues in binaural processing, the experiment with different masker bandwidths was repeated for a dichotic condition with an in-phase masker and an out-of-phase signal (NoSπ). All these measurements were performed at 1 kHz. In order to learn more about the influence of auditory filtering on LNN masking behavior, the maskers in the two final experiments were presented at other center frequencies. First, a fixed masker bandwidth of 100 Hz was used at center frequencies between 250 Hz and 10 kHz, and secondly, masker bandwidths between 5 and 100 Hz were tested at a high center frequency of 10 kHz. In all experiments, thresholds were also obtained for Gaussian-noise maskers which served as a reference.

2. Method

2.1. Generation of low-noise noise stimuli

Initially, three different algorithms for generating LNN stimuli (termed LNN1, LNN2, LNN3) were used as described below. From experiment 2 onwards, all LNN measurements were performed using the first of these algorithms.

Method 1 (proposed by author SvdP) started with a Gaussian noise signal of a typical duration of 4 s and a rectangular power spectrum. In the two top panels of Figure 1, a section of the time function (left) and a section of the power spectrum (right) are shown for a 100-Hz-wide noise at 1 kHz. The following steps were iterated 10 times. The Hilbert envelope of the noise was calculated, the time waveform was divided by this envelope on a sample-by-sample basis and then restricted to its original bandwidth by zeroing the corresponding components in the power spectrum. The middle left panel in Figure 1 shows the waveform from the top panel after division by its Hilbert envelope and rescaling to the original rms value, the middle right panel shows the corresponding power spectrum. The two panels at the bottom show the waveform and power spectrum after the spectral limitation. It is obvious that flattening of the temporal envelope introduces spectral components outside the original bandpass spectrum, while restricting the bandwidth introduces envelope fluctuations. Fortunately, iteration of the procedure leads to a decreasing amount of envelope fluctuations and thus to a decreasing amount of spectral splatter after each division by the envelope. The result after 10 iterations is plotted in the middle panel of Figure 2, together with a Gaussian noise with the same rms value in the top panel. The bottom panel shows the envelope in dB for the low-noise noise (continuous curve) and the Gaussian noise (dashed curve) [2].

Method 2 (proposed by author MvdH) was very similar to the first method. Instead of dividing the waveform by its Hilbert envelope, the instantaneous waveform values were compressed by raising them to the power of 0.01. This was done separately for positive and negative waveform values. Again, the waveform was then restricted to its original bandwidth and the procedure was repeated 10 times.

In method 3 (proposed by author DP), first the spectral amplitude values for the noise signal were set to certain values. For the specific noise sample used in experiment 1, all spectral amplitudes between 950 and 1050 Hz had the same (nonzero) value and all other amplitudes were set to zero. Since the spectral components had a spacing of 0.5 Hz, the corresponding time signal had a duration of 2 s. The following steps were iterated approximately 20 times: first, the values of all time signal samples were limited (clipped) to a value that corresponded to the rms value of the noise signal (for samples with negative values this procedure was applied correspondingly). Then the amplitude values in the spectrum were reset to the original values, while the phases of the spectral components remained unchanged at the value they had after clipping.

Thus, all three methods resulted in a narrowband noise signal with a bandlimited power spectrum. In method 3,

[2] For interested readers, we can give an indication of the CPU times for this procedure on two different computers. Times are for a signal of 128000 samples, corresponding to 4 s duration at a sample rate of 32 kHz, and ten iterations. The time is 356 s on a Silicon Graphics with a CPU MIPS R2000A/R3000, FPU MIPS R2010A/R3010 running at 33 MHz (IP 12), and 78 s on a Silicon Graphics with a CPU MIPS R4400, FPU MIPS R4010 running at 174 MHz (IP 22). The computation time is linear with the number of iterations, and grows with signal duration in the same way as the FFT.

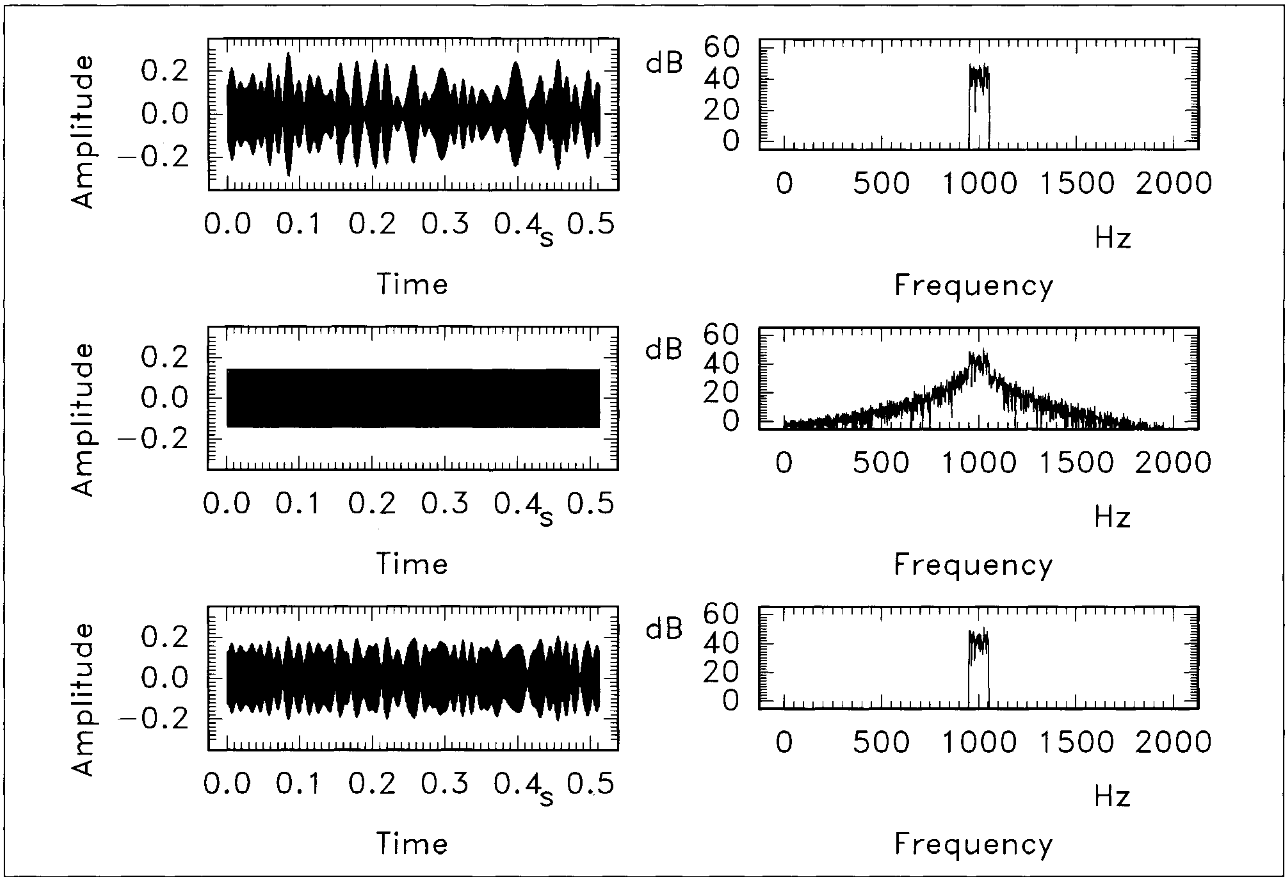

Figure 1. Time signals (left-hand panels) and corresponding power spectra (right-hand panels) illustrating method 1 for generating LNN. The two top panels illustrate a narrow-band Gaussian noise of about 500-ms duration. The two middle panels show the results after dividing the Gaussian noise by its Hilbert envelope. The two panels at the bottom show waveform and power spectrum after restricting the spectrum of the signal from the middle panel to the original bandwidth, ranging from 950 to 1050 Hz. The rms values of all time signals were adjusted to an equal value of 0.1.

the spectral amplitude values of the individual components within the passband were also maintained, while in methods 1 and 2 the spectral amplitude values after the iteration were different from those at the beginning of the iteration.

2.2. Measures of envelope fluctuations

In order to indicate the strength of noise fluctuations, we will use two different measures. One was used by Pumplin [2] and Hartmann and Pumplin [3] and is based on the normalized fourth moment of the waveform $x(t)$:

$$W = \overline{x^4}/(\overline{x^2})^2. \tag{1}$$

The values of the normalized fourth moment, W, range from 1.5 for a single sinusoid to 3 for a bandlimited Gaussian noise (of infinite duration). We are interested in signals having a value of W just above 1.5, indicating a flat temporal envelope. The LNN waveforms used by Hartmann and Pumplin [3] had values for W of 1.58 and 1.60.

The other measure expresses envelope fluctuation as the ratio between the standard deviation and the mean of the envelope, for which we will use the symbol V. This measure is useful if one wants to relate envelope fluctuations

Table I. Fluctuation measures for the 100-Hz wide LNN stimuli used in experiment 1. The three different methods are explained in the text. The measure W is based on the normalized fourth moment of the waveform, the measure V on the ratio between the standard deviation and the mean of the envelope.

Method	W	V [dB]
1	1.524	−24.1
2	1.535	−22.2
3	1.685	−13.2

to measures of modulation depth. For sinusoidal amplitude modulation with a modulation depth m (of an otherwise flat-envelope carrier), the measure V (in dB) is always 3 dB lower than the value of m (in dB). The lower limit of V is $-\infty$ dB for a flat-envelope signal, the theoretical value for a bandlimited Gaussian noise is -5.6 dB [3]. Table I shows the two measures, W and V, for the 100-Hz-wide LNN stimuli

[3] The mathematical basis for this number is that the envelope of a narrowband Gaussian noise has a Rayleigh distribution. For this distribution, the ratio between standard deviation and mean is given by: $\sqrt{\frac{4}{\pi} - 1}$, which is equal to 0.52, or -5.6 dB.

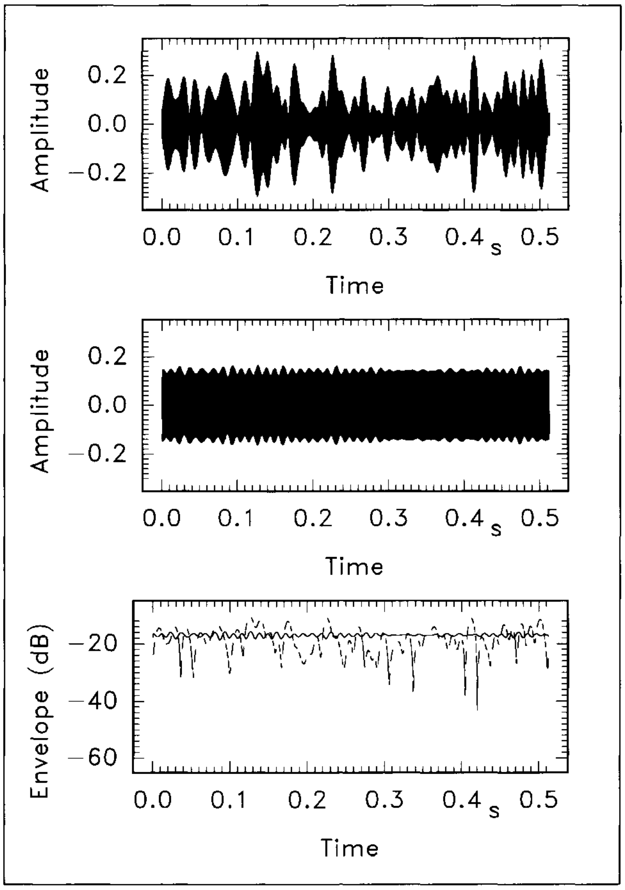

Figure 2. Waveform section for Gaussian noise (top panel) and LNN according to method 1 (middle panel). Both signals have a flat power spectrum between 950 and 1050 Hz and the same rms value of 0.1 (-20 dB). The bottom panels shows the envelopes in dB for the two noises. The continuous line indicates LNN, the dashed line Gaussian noise.

Table II. Fluctuation measures W and V for different numbers of iterations in generating LNN following method 1. The LNN was 100 Hz wide, centered at 1 kHz and had a duration of 4 s. The entries are the average values for 10 independent noise samples. The quantity V was averaged linearly. The entries in the first row are for Gaussian noise.

Number of iterations	W	V [dB]
0	3.030	−5.60
1	1.845	−11.1
2	1.701	−13.7
4	1.591	−17.8
6	1.552	−20.6
8	1.535	−22.6
10	1.526	−24.0

calculated with the three methods 1 to 3 that were used in the first experiment. The duration was 4 s for stimuli LNN1 and LNN2 and 2 s for stimulus LNN3.

Table II shows the values of W and V as a function of the number of iterations in applying method 1. The noise was 100 Hz wide, centered at 1 kHz and had a duration of 4 s. It is obvious that both measures decrease with the number of iterations. The decision to stop after 10 iterations for the noise stimuli used in the experiments was not based on a limit for the amount of fluctuations, but rather on convenience.

2.3. Threshold estimation procedure

Thresholds were measured with an adaptive three-interval forced-choice (3IFC) procedure that is described in detail in other recent publications (e.g. [11, 12]). The maskers were presented in three consecutive 500-ms intervals separated by 200-ms pauses. For all bandwidths, the masker level was held constant at 60 dB SPL. The signal was added to one randomly chosen interval, and its level was adjusted with a two-down one-up adaptive procedure. The step size for level changes was 4 dB in the beginning and 1 dB in the measuring phase of each run. The median value of the final 8 reversals of the signal level was taken as the threshold value. The 500-ms masker bursts for each interval were cut with random onset from a circular noise buffer of 4-s duration (2 s for the masker LNN3 in experiment 1). For Gaussian noise, a new buffer was calculated for each threshold run; for LNN, the same 4-s (2-s) buffer was used in all runs. Masker and signal were both gated with 10-ms raised-cosine ramps. With the exception of experiment 3, in which the signal duration was varied, the signal had the same duration as the masker and was added without any onset delay. In experiment 3, the signal was temporally centered in the masker interval.

All stimuli were generated digitally at a sample rate of 32 kHz. After D/A conversion with the built-in two-channel 16-bit converter of an Iris Indigo computer, and appropriate attenuation, the stimuli were presented to the subjects over headphones (Beyer DT 990). Presentation was diotic with the exception of experiment 4, in which the condition was $NoS\pi$.

Five subjects participated in the experiments. They were all members of the IPO hearing group, were experienced in this type of masking experiment and had no reported history of hearing impairment. For each parameter setting, each subject performed 4 adaptive runs. Individual data are given as the average across these four repeated measurements, and mean data are the average of the results for all subjects.

3. Results

3.1. Experiment 1: Running-noise presentation of LNN

The first experiment was designed to test whether the previously observed LNN advantage is also found in a running-noise presentation. In order to make our measurements comparable, the stimulus parameters were chosen in accordance with the experiment by Hartmann and Pumplin [3]. Besides a Gaussian noise, we used three different LNNs generated with methods 1 to 3. Four subjects participated in this experiment.

The results are shown in Table III. The four rows indicate the four masker types, the first four columns give the means and standard deviations of the individual data for the four subjects, and the last column gives the mean values across subjects, again with their standard deviations.

Table III. Masked thresholds obtained in experiment 1. Column 1 indicates the masker type, columns 2 to 5 give the individual data of four subjects (means and standard deviations based on four runs per noise type and subject) and column 6 gives the means and the standard deviations of the individual results.

Noise type	S1	S2	S3	S4	Mean
Gaussian	55.9	55.1	57.0	56.9	56.2
	(0.9)	(1.4)	(1.2)	(1.2)	(0.9)
LNN1	51.6	48.8	51.3	53.9	51.4
	(1.3)	(1.7)	(2.1)	(2.0)	(2.1)
LNN2	51.6	47.9	51.9	54.1	51.4
	(0.9)	(1.0)	(1.1)	(1.2)	(2.6)
LNN3	55.0	53.6	55.8	58.5	55.8
	(2.3)	(2.5)	(1.7)	(1.7)	(2.1)

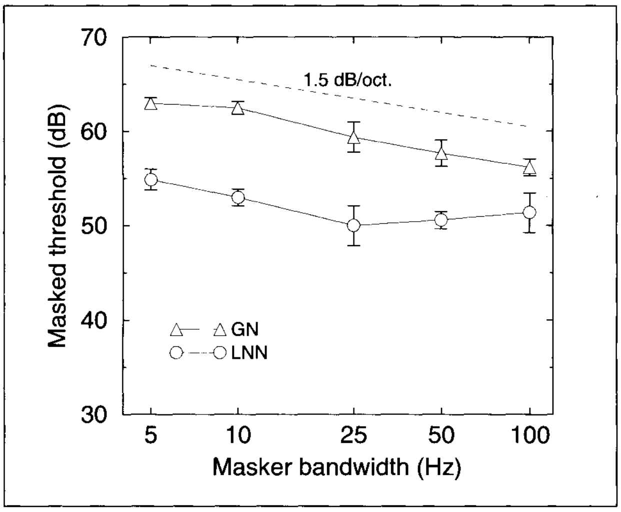

Figure 3. Masked thresholds (means ± 1 standard deviation) as a function of masker bandwidth from experiment 2. The triangles indicate Gaussian noise, the circles LNN. Masker and signal were both centered at 1 kHz and had a duration of 500 ms. Masker level was 60 dB SPL for all bandwdiths. The data are based on the results from four subjects.

If we look at the mean data in the last column, we see that Gaussian noise and the LNN with the largest amount of envelope fluctuations (LNN3) give similar threshold values. Thresholds for the two other LNN maskers (LNN1 and LNN2) are clearly lower, with a mean difference of 4.8 dB with respect to the Gaussian noise. This relation seen in the mean values also holds for the individual results of the four subjects. Thus we can conclude that for the flattest LNN stimuli, an advantage of about 5 dB with respect to a Gaussian-noise masker is observed if the maskers are presented as running noise. Based on these results, it was decided to perform all further LNN measurements in this paper with stimuli calculated using method 1 (LNN1). Also, only mean data are presented in the following experiments.

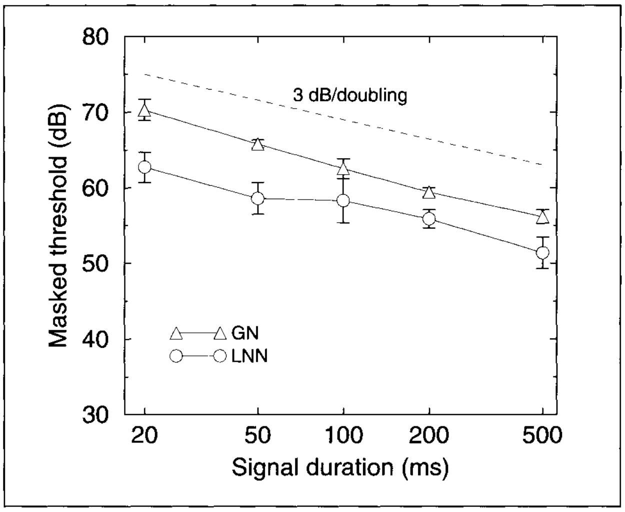

Figure 4. Masked thresholds as a function of signal duration from experiment 3. The masker had a duration of 500 ms, a bandwidth of 100 Hz and a level of 60 dB. Masker and signal were both centered at 1 kHz. The data are based on the results from four subjects.

3.2. Experiment 2: Variation of masker bandwidth

In this experiment, masked thresholds were obtained for Gaussian noise and LNN for bandwidths of 5, 10, 25, and 50 Hz. The same four subjects as in experiment 1 participated in this experiment. Figure 3 shows the mean data for Gaussian noise (triangles) and LNN (circles), where the 100-Hz data are those of experiment 1. Values for Gaussian noise decrease with increasing masker bandwidth at a rate of about 1.5 dB per doubling of the bandwidth (cf. the dashed line). This result agrees with the findings of other studies and is usually explained by the decreased variance of the energy estimate for noise samples with a larger bandwidth (e.g. [7]). Taking into account the constant noise level at all bandwidths, the signal-to-overall-noise ratio increases from -3.5 dB at 100 Hz to +3 dB at 5 Hz.

Thresholds for LNN initially decrease with increasing bandwidth and reach a minimum at a bandwidth of 25 Hz. At this bandwidth, the difference between the two noise types reaches a maximum of 9.4 dB. For larger bandwidths, thresholds increase slightly and the difference between the noise types becomes smaller.

This experimental result thus supports the speculation in [3] that the 100-Hz-wide masker at 1 kHz might already be affected by the 1-kHz auditory filter, rendering the internal temporal envelope of the masker less flat than the acoustic stimulus. However, as will be shown later in Sec. 3.5, this may not be the correct explanation for this observation.

3.3. Experiment 3: Variation of signal duration

In this experiment, masking by LNN and Gaussian-noise maskers was investigated for signal durations between 20 and 500 ms. The masker bandwidth for all measurements was 100 Hz.

The average results for four subjects are plotted in Figure 4. On and off-ramps of 10 ms are included in the signal duration.

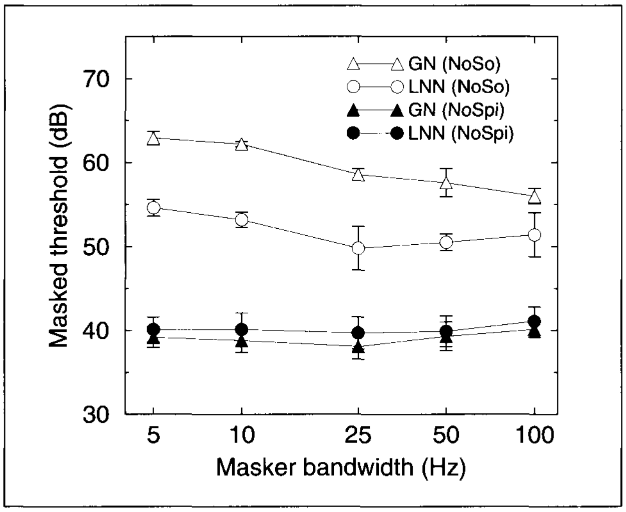

Figure 5. Masked thresholds for the conditions NoSo (open symbols) and NoSπ (filled symbols) from experiment 4. With the exception of the signal phase, all parameters were the same as in Figure 3.

The dashed line in the figure indicates a slope of -3 dB per doubling of the signal duration. Thresholds in Gaussian noise decrease with this slope, even for signal durations in excess of 200 ms. For LNN, this is only the case for signal durations between 100 and 500 ms. Averaged over the three shortest durations, the slope for the LNN data is less steep than for Gaussian noise, which leads to an increase of the LNN advantage from 4.8 dB (at 500 ms) to 7.6 dB (at 20 ms). This increase of the difference between LNN and Gaussian-noise maskers with decreasing signal duration does not support the prediction in [3] that, for frozen-noise masking experiments, the difference should decrease with decreasing signal duration[4]. That prediction was derived from an energy detection model in which the duration of signal and noise are so long that the listener can make multiple judgements, imagined to take place during subintervals of time.

3.4. Experiment 4: BMLDs

This experiment was a repetition of experiment 2 in a dichotic condition which creates a BMLD for Gaussian-noise maskers: the masker was presented in phase to both ears, the signal out of phase (NoSπ). The rationale for this experiment is based on the common notion that NoSπ thresholds can often be predicted by the change in the normalized interaural crosscorrelation that is introduced by adding the signal. For the crosscorrelation derived from the waveforms this change in crosscorrelation depends only on the signal-to-noise ratio (cf. [13, 14]). Based on this notion and the additional assumption that at 1 kHz the auditory system is capable of coding the fine structure of the stimulus waveform, the NoSπ thresholds are expected to be the same for both masker types.

The results of this experiment are represented in Figure 5 by the filled symbols. Triangles are used for Gaussian noise, circles for LNN. The open symbols indicate the NoSo data from experiment 2. The NoSπ thresholds for both conditions are virtually identical (average difference 1.1 dB) and they do not change much with the masker bandwidth. Both observations are in line with the above explanation using the normalized crosscorrelation value.[5]

3.5. Intermezzo: Simulations with a Gammatone filterbank

In this section, we will analyze the 'suboptimal' LNN advantage for the specific case of a 100-Hz-wide masker at 1 kHz. The data in Figure 3 show that the difference between Gaussian and LNN maskers increases with increasing bandwidth up to 25 Hz, but becomes smaller for wider bandwidths. Hartmann and Pumplin [3] had speculated that such an effect could occur due to the impact of the auditory filter at 1 kHz.

This idea was tested by filtering the 100-Hz-wide LNN masker with a Gammatone filter centered at 1 kHz. The equivalent rectangular bandwidth (ERB) of this filter was set to 132 Hz [15]. For a 3-s section of the LNN stimulus, the measure V was calculated before and after filtering. Filtering changed this value from -24.0 dB to -23.1 dB, a change of the same size as the difference in the quantity V for LNN1 and LNN2[6]. Since these two maskers led to the same threshold values, we can conclude that the on-frequency filter does *not* change the flatness of the LNN noise significantly. The same conclusion is derived from simulations with a linear basilar-membrane model [16].

An alternative explanation for the small LNN advantage also takes adjacent filters into account. Figure 6 shows the envelopes of filter outputs for five different filters: The on-frequency filter (top curve), the two filters with a center frequency 1 ERB above and below 1 kHz (two dashed curves in the middle), and the two filters 2 ERBs removed from the center frequency (two continuous curves at the bottom). Clearly, the four off-frequency filters have a highly modulated output.[7] By comparing the long- and short-dashed curves (filters 1 ERB removed), one can also recognize details of the LNN properties. LNN can approximately be considered as a frequency-modulated signal, with the lowest and highest values of the instantaneous frequency given by the

[4] In a personal communication, Hartmann stated that there would be no reason to limit this argument to frozen noise.

[5] It should be mentioned that two recent ASA meeting abstracts reported higher NoSπ thresholds for LNN than for Gaussian noise for some subjects [5, 6].

[6] In order to evaluate the influence of the filter bandwidth, we repeated this simulation with reduced bandwidths of the Gammatone filter. For ERB values of 100, 66, and 40 Hz, the measure V had values of -20, -15 and -9 dB, respectively.

[7] At the output of the linear basilar-membrane model, the amount of envelope fluctuations is different in channels tuned to frequencies below and above the masker frequencies. Channels tuned to frequencies above 1 kHz show the same small amount of fluctuations as the on-frequency channel. Channels tuned to frequencies below 1 kHz show a high amount of envelope fluctuations. This result reflects the asymmetric filter characteristic of a certain point along the basilar membrane. It supports the notion that the introduction of envelope fluctuations through filtering depends on the change in filter attenuation over the spectral range of the LNN stimulus.

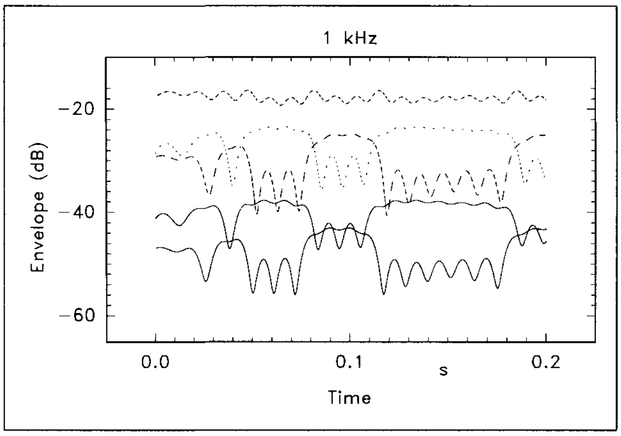

Figure 6. Sections of the Hilbert envelopes of the outputs of five Gammatone filters. The input was a 100-Hz-wide LNN (method 1), centered at 1 kHz. The five filters were centered on the noise band (top curve), ± 1 ERBs above or below the noise band (the two short- and long-dashed curves in the middle), and ± 2 ERBs above and below the noise band (the two continuous curves at the bottom). The bandwidths of the filters were chosen according to the formula in [15].

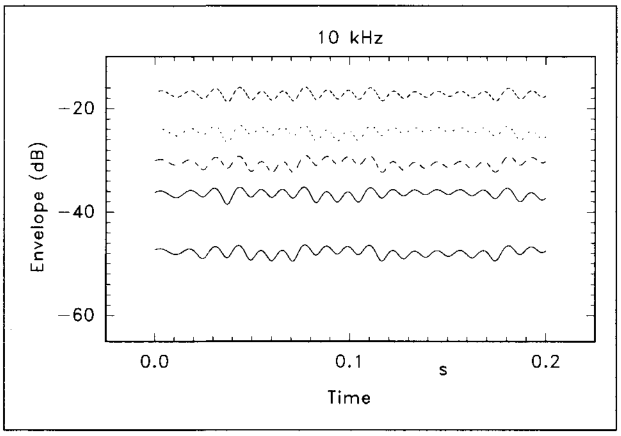

Figure 7. Same as Figure 6 for a 100-Hz-wide LNN at 10 kHz and five Gammatone filters centered at 10 kHz (top curve) and ± 1 (short- and long-dashed curves in the middle) and ± 2 ERBs above and below 10 kHz (two continuous curves at the bottom).

lower and upper spectral border (cf. [17]). If the instantaneous frequency equals the upper spectral limit, the filter above the noise will have a high output value, while the filter below the noise will have a low output value (and vice versa). The variation of the envelope is thus an indication of the filter attenuation over the spectral range of the noise band. The curves also suggest that the instantaneous frequency of the noise band rests mainly at either the upper or the lower spectral limit, and sweeps quickly (e.g., around 80 and 115 ms) through the noise spectrum. This analysis suggests that the disadvantage for the 100-Hz-wide LNN at 1 kHz may be caused by the fluctuations in off-frequency filters. This idea will be tested in the following experiments.

A direct prediction is that, for the same 100-Hz masker presented at a higher center frequency, we expect less modulation in the off-frequency filter outputs. This prediction is

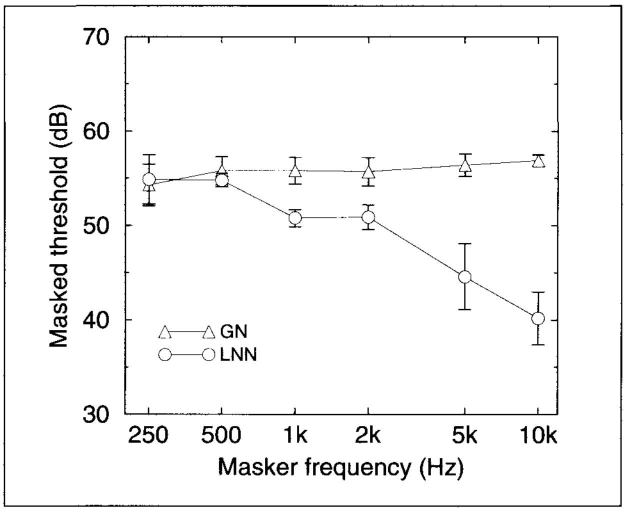

Figure 8. Masked thresholds obtained for 100-Hz-wide maskers as a function of their center frequency (experiment 5). The signal was spectrally centered in the masker. The data are based on the results from four subjects.

based on the knowledge that the slopes of the auditory filters are roughly constant on a Bark or ERB scale, and so they become shallower at high frequencies if expressed on a linear frequency scale. Figure 7 shows an analysis of a 100-Hz-wide LNN stimulus centered at 10 kHz. Again, the outputs of five Gammatone filters are shown, scanning the range from 2 ERB below to 2 ERB above the masker center frequency. In line with the above analysis, the filter outputs are now flat for on-frequency as well as for off-frequency filters. A prediction that will be tested in the final two experiments of this paper is that the full advantage for an LNN masker can only be obtained at a high frequency, and that this advantage should increase with increasing center frequency when an LNN with a constant absolute bandwidth is used.

3.6. Experiment 5: Variation of masker center frequency

This experiment is related to the last remark in the previous section. Gaussian-noise and LNN maskers with 100-Hz bandwidth were presented at 6 different frequencies between 250 Hz and 10 kHz. The measurements were performed by four subjects.

Figure 8 shows the mean results of the four subjects, with triangles indicating data for Gaussian noise and circles data for LNN. The threshold values for Gaussian noise depend very little on the frequency, except that the value at 250 Hz is 1.5 dB lower than all other values. Based on estimates of auditory filter bandwidths at different frequencies, and assuming a constant signal-to-noise ratio at the filter output as threshold criterion, one would expect constant thresholds at 1 kHz and above. Between 1 kHz and 250 Hz, thresholds should decrease by about 3 dB according to the ERB values [15], while they should remain constant according to the values published by Zwicker and Terhardt [18]. If one takes into account that the signal-to-noise ratio at the filter output increases with decreasing filter bandwidth (cf. the argument

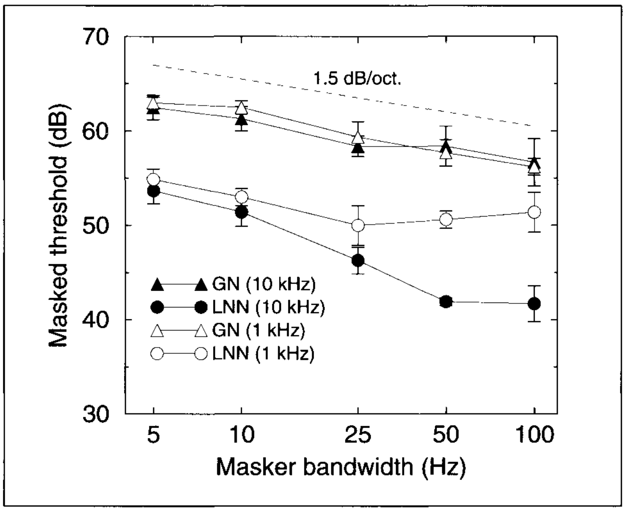

Figure 9. Comparison of masked thresholds obtained at 1 kHz (open symbols) and at 10 kHz (filled symbols). The data are based on the results from four subjects.

mentioned in section 3.2), the present data are in line with a prediction based on ERB values.

The LNN thresholds at the two lowest frequencies are the same as those for Gaussian noise. Towards higher frequencies, thresholds decrease continuously by about 15 dB. It may be that thresholds for LNN would continue to decrease beyond 10 kHz for a 100-Hz-wide LNN, but this was not tested. An analysis of the individual data reveals that for two subjects, thresholds at 10 kHz are no more than 1 dB below those at 5 kHz, and thus probably reveal the greatest possible advantage for LNN. For one subject, the decrease amounts to 5 dB and for the most sensitive subject, thresholds decrease by 10 dB between 5 and 10 kHz.

The absence of an LNN advantage at the two lowest frequencies corresponds to the observation by Hartmann and Pumplin [3] that a 300-Hz wide LNN masker at 1 kHz gives the same thresholds as a Gaussian noise. In both cases, the noise bandwidth is of the same order as, or even wider than, the critical bandwidth at the noise frequency and it is thus very likely, as Hartmann and Pumplin argued, that the noise waveform at the output of the auditory filter no longer has a flat temporal envelope.

3.7. Experiment 6: Variation of masker bandwidth at 10 kHz

This final experiment is a repetition of experiment 2, with the only difference that masker and signal were centered at 10 kHz. Figure 9 shows the data of both of these experiments, where the 10-kHz results are shown by the filled symbols and the 1-kHz results by the open symbols. The Gaussian-noise thresholds (triangles) are virtually identical for the two frequency regions. Such a result is to be expected, since the masker bandwidths always remain below the value of the corresponding critical bandwidths. For LNN (circles), the center frequency matters a lot. For the two narrowest bandwidths, thresholds for 1 and 10 kHz differ by no more than 1.5 dB. With further increasing bandwidth up to 50 Hz, the 10-kHz thresholds decrease strongly and level off at a value about 10 dB lower than the results at 1 kHz.

4. Discussion

The results of this study support the notion of Hartmann and Pumplin [3] that LNN maskers generally lead to lower masked thresholds in simultaneous, on-frequency conditions than Gaussian-noise maskers. To be more specific, we can conclude that, in order for such a difference to occur, the masker bandwidth has to be smaller than the auditory-filter bandwidth at the masker frequency (cf. Figure 8). The difference between LNN and Gaussian noise is maximal at a clearly subcritical bandwidth and it decreases at very narrow bandwidths (Figure 9). The value of the maximum difference increases with increasing center frequency (Figures 8 and 9).

What is the reason for the LNN advantage? This question can be answered in two ways, from psychoacoustic modelling and from reports of the subjects. When the sinusoidal test signal is added to an LNN, subjects can use different detection cues to distinguish this stimulus from the LNN masker alone: if the signal level is well above that of the masker, the best cue is the difference in level introduced by adding the signal. When the signal level is lowered and approaches the masker level, the cue changes to the detection of fluctuations in the signal interval. If the signal level is further lowered, the amount of introduced fluctuations decreases and the level at which the signal is still detectable depends on the amount of perceived fluctuations in the masker alone.

This "introspective" description agrees with the results from simulations. Figure 10 shows 5 different envelope functions, either for a 100-Hz-wide LNN masker alone or for a combination of an LNN masker plus an on-frequency sinusoid at four different signal-to-noise ratios. The conditions can best be indentified by looking at the envelopes at time 0.4 s. The lowest (continuous) curve indicates the envelope for the LNN masker alone and shows the degree of fluctuations present in the masker-alone interval. The four other curves show, in ascending order, masker-plus-signal combinations with an S/N ratio of -20, -10, 0, and +10 dB, respectively. The lowest S/N value corresponds to the thresholds for a 100-Hz-wide LNN at 10 kHz (see Figure 9). The level of the masker waveform was always -20 dB. Since the rms value of the envelope is always 3 dB higher than the rms value of the waveform, the average value of the masker-alone envelope is expected to be -17 dB.

At an S/N ratio of -20 dB, the addition of the signal hardly affects the envelope, while at an S/N ratio of 0 dB (long-dashed curve), the signal is able to increase the envelope by up to about 6 dB, or, at other time instants, to nearly cancel it. To describe these observations more quantitatively, we calculated the mean values of the five envelope functions (which are proportional to the waveform energies), and also the relation between the standard deviation and the mean, V, as defined earlier. These values are given in Table IV in dB as a function of the signal-to-noise ratio. The second column

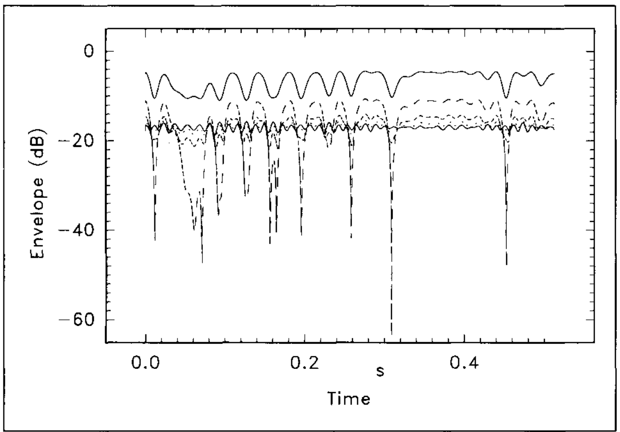

Figure 10. Envelopes of the 100-Hz-wide LNN (flat, continuous curve) and of combinations of the LNN and a sinusoid at signal-to-noise ratios of +10, 0, -10 and -20 dB. Around 0.4 s, the envelope functions increase monotonically with increasing S/N. The level of the masker waveform was always -20 dB.

Table IV. Analysis of the envelopes of combinations of a 100-Hz-wide LNN and an on-frequency sinusoid at different signal-to-noise ratios. Masker and signal had a duration of 4.096 s, the rms-value of the masker waveform was -20 dB. The first column gives the signal-to-noise ratio, where $-\infty$ indicates the masker-alone condition. The second column gives the mean values of the envelopes, and the third column the ratio between the standard deviation and the mean of the envelopes (= fluctuation measure V).

S/N [dB]	Mean [dB]	V [dB]
+10	−6.8	−13.3
0	−15.0	−6.2
−10	−16.9	−13.0
−20	−17.0	−20.5
$-\infty$	−17.0	−24.1

indicates the mean envelope value, the third indicates the fluctuation measure V.

The mean envelope value decreases monotonically with decreasing S/N ratio and is a valuable cue for the two highest S/N values only, but not for S/N ratios of -10 dB or lower. The amount of envelope fluctuations, on the other hand, is a nonmonotonic function of the S/N ratio. It is largest at an S/N ratio of 0 dB, and decreases for higher as well as for lower signal levels.

We can see that at the lowest experimental thresholds, i.e., for an S/N ratio of -20 dB, the change in fluctuations, expressed in the quantity V, amounts to about 3.5 dB. This value can be compared to predictions and data in other papers about the just noticeable change in envelope fluctuation or modulation depth. Maiwald [19] developed a model for the detection of amplitude modulations for carriers that have intrinsic envelope fluctuations with an effective modulation depth m. For a narrow-band Gaussian noise, m was assumed to have a value of 0.7 in his calculations. The sum of squares of the inherent fluctuations m and the external modulation, expressed as modulation depth, must exceed the squared value of m by 25 %, or 1 dB. Applied to our condition, the fluctuation measure V for the combination of LNN and a sinusoid could be as low as about -23 dB for the signal still to be detectable. Such a value for V corresponds to a signal-to-noise-ratio of about -27 dB, which is about 4 dB below the average thresholds for the most sensitive subject for the 100-Hz-wide LNN at 10 kHz.

Wakefield and Viemeister [20] measured the just-noticeable difference in modulation depth for broadband noise carriers. They expressed thresholds as $10\log(m_c^2 - m_s^2)/m_s^2)$, with m_s being the modulation depth of the standard and m_c the modulation depth of the comparison stimulus. For a standard modulation depth of -20 dB, this measure yielded values between about -1 and -3 dB (based on Figure 3 in [20]). For our stimuli, this measure results in a value of +0.9 dB for an S/N ratio of -20 dB. Thus, from both of these studies one would predict that the detection of a sinusoid added to an LNN could occur at even lower signal-to-noise ratios than those observed in experiments 5 and 6.

Finally, we discuss the dependence of the LNN thresholds on masker bandwidth. Here we have to explain two effects: first the decrease of LNN thresholds with increasing frequency for a constant-bandwidth masker (Figure 8), and secondly the increase in thresholds for decreasing masker bandwidths (Figure 9). The first effect was already mentioned in Section 3.5, where we stressed the importance of masker fluctuations in off-frequency filters. The simulations shown in Figures 6 and 7 have revealed that these fluctuations become smaller when the ratio between LNN bandwidth and the auditory filter bandwidth decreases. We can think of two ways in which fluctuations in off-frequency channels could increase signal thresholds in a 100-Hz-wide LNN at 1 kHz. It could be that the analysis of envelope fluctuations in the on-frequency channel is negatively influenced by the presence of masker fluctuations in off-frequency filters. In such a view, thresholds in a 100-Hz-wide LNN masker at 1 kHz are increased due to an effect resembling modulation detection interference (MDI, e.g. [21]). The interesting consequence of this view is that subjects are unable to ignore off-frequency channels, even if the relative stimulus bandwidth is as small as 5 or 10%.

Alternatively, we could assume that signal detectability in the on-frequency channel is not negatively influenced by the presence of masker fluctuations in off-frequency channels, but that *signal detectability in off-frequency channels* is non-optimal. In these filters, the masker alone has a high amount of envelope fluctuations and the increase in this quantity due to the signal is too small to be detectable. According to this view, subjects have to be able to detect the change in modulation depth caused by the addition of the signal in many peripheral filters in order to reach low thresholds. This distinction might look like an academic question, but it may be relevant for estimating the amount of internal noise per auditory channel in a quantitative model for modulation analysis (see, e.g., [22, 23]).

In order to understand the increase of LNN thresholds for very narrow bandwidths, we have to consider not only the amount of envelope fluctuations, but also the spectral distribution of these fluctuations. Figure 11 shows the power spectrum of the envelope of the 100-Hz-wide LNN wave-

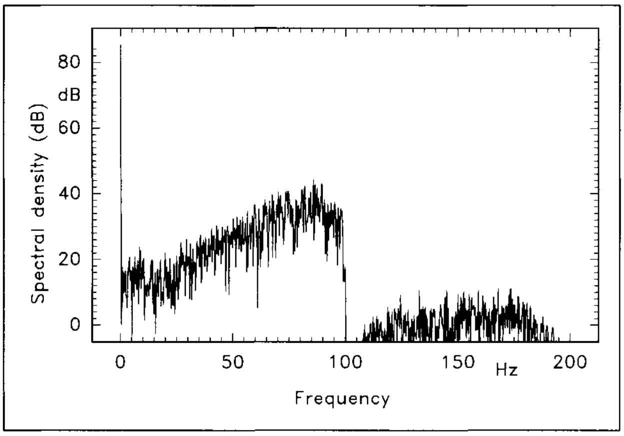

Figure 11. Power spectrum of the envelope of a 100-Hz-wide LNN of a total duration of about 4 s.

form in the region between 0 and 200 Hz. In contrast to Gaussian noise, for which the envelope spectrum continuously decreases, the envelope spectrum for LNN increases with frequency and reaches a maximum just below the LNN bandwidth. Beyond this bandwidth, it drops steeply [8]. For LNN with a narrower bandwidth, we get the same spectral shape on a compressed frequency scale, with the maximum always occurring just below the noise bandwidth.

Adding a sinusoidal signal to the masker waveform changes the envelope spectrum only in the low-frequency region up to half the LNN bandwidth. In this region, the spectrum becomes flat (with exception of the DC value) and the spectral level increases with the signal-to-noise ratio. Higher modulation frequencies (up to the LNN bandwidth) are not affected.[9] This can be seen from Fig 12, which shows envelope spectra for LNN plus a sinusoidal signal at two different signal-to-noise ratios. In the top panel, it is -10 dB, in the bottom panel, it is -20 dB.

The influence of masker bandwidth on the masked thresholds can now be understood, albeit qualitatively, by assuming that envelope fluctuations are analyzed through a modulation filterbank [24, 22, 23]. In this model, modulation filters with center frequencies below 10 Hz have a constant bandwidth of 5 Hz, while filters with higher frequencies have a constant relative bandwidth. In such a model, the detection of the signal added to an LNN would be possible by analyzing the outputs of filters tuned to low modulation frequencies: after all, these filters pass only a small amount of intrinsic masker fluctuations, and the change in modulation power caused by adding the signal is large. If, however, the masker bandwidth decreases, more and more power from the intrinsic masker fluctuations will pass through the (relatively widely tuned) low modulation filters and make the detection of the signal more difficult. At present, this explanation can only be given qualitatively and it remains to be tested in further simulations with the model proposed by Dau [24], whether the above argument is correct. Such work is presently in progress.

[8] One can note from this figure that the spectrum of the (linear) envelope of a bandlimited signal is not bandlimited, due to the square-root operation involved in calculating the Hilbert envelope from the analytical signal. Had we plotted the spectrum of the squared envelope instead, the spectrum would be strictly bandlimited to 100 Hz.

[9] Because envelope fluctuations reflect interactions between the spectral components in a stimulus, the addition of the sinusoidal signal to the LNN masker can only introduce extra fluctuations at those rates that correspond to spectral differences between sinusoid and masker. For a sinusoid spectrally centered in the masker, this is just half the masker bandwidth. If the added sinusoid has a frequency of, for instance, the upper spectral limit of the masker, changes in the modulation spectrum occur for modulation frequencies up to 100 Hz.

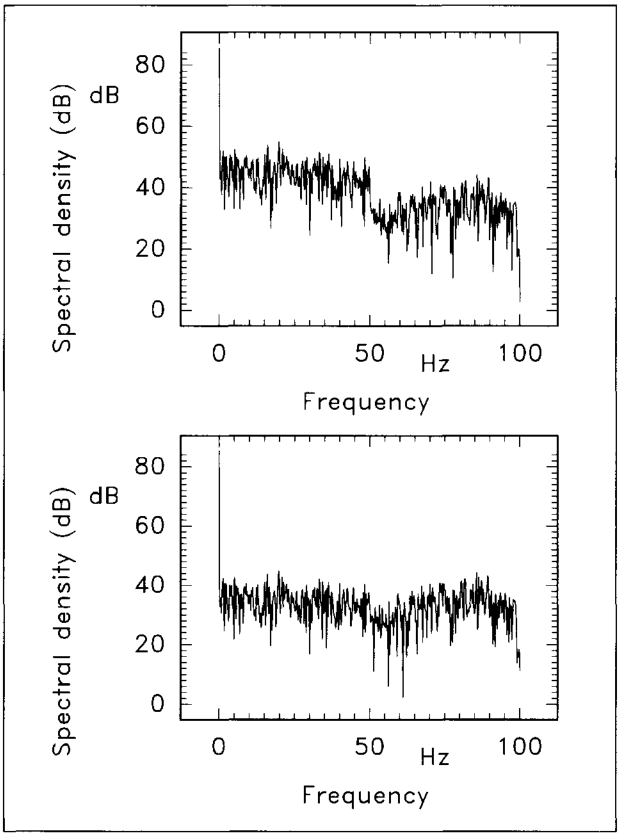

Figure 12. Power spectrum of the envelope of the sum of a 100-Hz-wide LNN and a sinusoid spectrally centered in the noise. The signal-to-noise ratio was -10 dB in the top panel and -20 dB in the bottom panel.

Acknowledgement

We would like to thank the three reviewers Torsten Dau, Christian Kaernbach and Brian C. J. Moore for their comments on an earlier version of this paper.

References

[1] M. R. Schroeder: Synthesis of low-peak-factor signals and binary sequences with low autocorrelation. IEEE Transact. Inf. Theor. **16** (1970) 85–89.

[2] J. Pumplin: Low-noise noise. J. Acoust. Soc. Am. **78** (1985) 100–104.

[3] W. M. Hartmann, J. Pumplin: Noise power fluctuation and the masking of sine signals. J. Acoust. Soc. Am. **83** (1988) 2277–2289.

[4] W. M. Hartmann, J. Pumplin: Periodic signals with minimal power fluctuations. J. Acoust. Soc. Am. **90** (1991) 1986–1999.

[5] D. A. Eddins: Temporal cues in monaural and binaural detection and discrimination. J. Acoust. Soc. Am. **99** (1996) 2471.

[6] J. W. Hall, III, J. H. Grose, W. M. Hartmann: The masking-level difference in low-noise noise. J. Acoust. Soc. Am. **99** (1996) 2595.

[7] C. E. Bos, E. de Boer: Masking and discrimination. J. Acoust. Soc. Am. **39** (1966) 708–715.

[8] G. Kidd, C. R. Mason, M. A. Brantley, G. A. Owen: Roving-level tone-in-noise detection. J. Acoust. Soc. Am. **86** (1989) 1310–1317.

[9] A. Kohlrausch, S. van de Par, A. J. M. Houtsma: A new approach to study binaural interaction at high frequencies. – In: Proceedings of the 10th international symposium on hearing, Irsee 94. G. Manley, G. Klump, G. Köppl, H. Fastl, H. Oeckinghaus (eds.). World Scientific, Singapore, New Jersey, London, Hong Kong, 1995, 343–353.

[10] B. C. J. Moore, J. W. Hall, III, J. H. Grose, G. P. Schooneveldt: Some factors influencing the magnitude of comodulation masking release. J. Acoust. Soc. Am. **88** (1990) 1694–1702.

[11] M. van der Heijden, A. Kohlrausch: The role of envelope fluctuations in spectral masking. J. Acoust. Soc. Am. **97** (1995) 1800–1807.

[12] T. Dau, D. Püschel, A. Kohlrausch: A quantitative model of the 'effective' signal processing in the auditory system: II. Simulations and measurements. J. Acoust. Soc. Am. **99** (1996) 3623–3631.

[13] N. I. Durlach, K. J. Gabriel, H. S. Colburn, C. Trahiotis: Interaural correlation discrimination: II. Relation to binaural unmasking. J. Acoust. Soc. Am. **79** (1986) 1548–1557.

[14] S. van de Par, A. Kohlrausch: Analytical expressions for the envelope correlation of certain narrow-band stimuli. J. Acoust. Soc. Am. **98** (1995) 3157–3169.

[15] B. R. Glasberg, B. C. J. Moore: Derivation of auditory filter shapes from notched-noise data. Hearing Research **47** (1990) 103–138.

[16] H. W. Strube: A computationally efficient basilar-membrane model. Acustica **58** (1985) 207–214.

[17] R. H. Margolis, A. M. Small: Masking with narrow-band FM noise. J. Acoust. Soc. Am. **56** (1974) 692–694.

[18] E. Zwicker, E. Terhardt: Analytical expression for critical-band rate and critical bandwidth as a function of frequency. J. Acoust. Soc. Am. **68** (1980) 1523–1525.

[19] D. Maiwald: Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas. Acustica **18** (1967) 193–207.

[20] G. H. Wakefield, N. F. Viemeister: Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise. J. Acoust. Soc. Am. **88** (1990) 1367–1373.

[21] W. A. Yost, S. Sheft, J. Opie: Modulation interference in detection and discrimination of amplitude modulation. J. Acoust. Soc. Am. **86** (1989) 2138–2147.

[22] T. Dau, B. Kollmeier, A. Kohlrausch: Modeling auditory processing of amplitude modulation: I. Detection and masking with narrowband carriers. J. Acoust. Soc. Am. (1997) submitted for publication.

[23] T. Dau, B. Kollmeier, A. Kohlrausch: Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration. J. Acoust. Soc. Am. (1997) submitted for publication.

[24] T. Dau: Modeling auditory processing of amplitude modulation. Dissertation. Universität Oldenburg, Germany, 1996.