A quantitative model of the 'effective' signal processing in the auditory system. II. Simulations and measurements

Citation for published version (APA):

Dau, T., Püschel, D., & Kohlrausch, A. G. (1996). A quantitative model of the 'effective' signal processing in the auditory system. II. Simulations and measurements. *Journal of the Acoustical Society of America*, 99(6), 3623-3631. <https://doi.org/10.1121/1.414960>

DOI:

[10.1121/1.414960](https://doi.org/10.1121/1.414960)

Document status and date:

Published: 01/01/1996

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.

A quantitative model of the “effective” signal processing in the auditory system. II. Simulations and measurements

Torsten Dau^{a)} and Dirk Püschel^{b)}

Drittes Physikalisches Institut, Universität Göttingen, Bürgerstr. 42-44, 37073 Göttingen, Germany

Armin Kohlrausch^{c)}

Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 20 June 1994; revised 24 January 1996; accepted 18 February 1996)

This and the accompanying paper [Dau *et al.*, J. Acoust. Soc. Am. **99**, 3615–3622 (1996)] describe a quantitative model for signal processing in the auditory system. The model combines several stages of preprocessing with a decision device that has the properties of an optimal detector. The present paper compares model predictions for a variety of experimental conditions with the performance of human observers. Simulated and psychophysically determined thresholds were estimated with a three-interval forced-choice adaptive procedure. All model parameters were kept constant for all simulations discussed in this paper. For frozen-noise maskers, the effects of the following stimulus parameters were examined: signal frequency, signal phase, temporal position and duration of the signal within the masker under conditions of simultaneous masking, masker level, and masker duration under conditions of forward masking, and backward masking. The influence of signal phase and the temporal position of the signal, including positions at masker onset, was determined for a random-noise masker and compared with corresponding results obtained for a frozen noise. The model describes all the experimental data with an accuracy of a few dB with the following exceptions: forward-masked thresholds obtained with brief maskers are too high and the change in threshold with a change in signal duration is too small. Both discrepancies have their origin in the adaptation stages in the preprocessing part of the model. On the basis of the wide range of simulated conditions we conclude that the present model is a successful approach to describing the detection process in the human auditory system. © 1996 Acoustical Society of America.

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk

INTRODUCTION

This is the second paper about a quantitative model for the internal representation of time-varying signals in the auditory system. While the first paper (Dau *et al.*, 1996) concentrated on details of the implementation of the model, the present paper compares model predictions with the performance of human observers.

In developing the model we followed the following philosophy. The model should allow the prediction of masked thresholds in a great variety of simultaneous and nonsimultaneous conditions. No restrictions are included as to the duration, spectral shape and waveform statistics of masker and signal. Not covered are conditions that require across-channel comparisons like comodulation masking release (CMR) and binaural masking level differences (BMLDs). The model should consist of preprocessing stages that represent the “hard wired” part of signal processing in the human hearing system. The term “hard wired” is used in the sense that this part is common for normal-hearing subjects, that it is not subject to learning effects and that it should be appli-

cable to a wide variety of different experimental tasks. The output of the preprocessing stages can be interpreted as a multidimensional, time-varying activity pattern, in a similar way as, for instance, the stabilized auditory image in the model developed by Patterson and colleagues (e.g., Patterson *et al.*, 1992). Limitations of resolution are simulated by adding internal noise with a constant variance value to the activity pattern.

The activity pattern is further analyzed by a decision device. In the description of an earlier version of the model (Kohlrausch *et al.*, 1992) this detector was simply realized as a temporal integrator with a threshold device behind it. In the implementation described in the present paper it is realized as an optimal detector, which is realized as a cross correlator between the internal representation of the “expected” temporal signal (template) and the representation of the actually “received” waveform (see accompanying paper).

Our implementation of the optimal detector implies that a template of the signal (in a masking experiment) which has to be detected is derived at the beginning of the experiments. This template is used throughout a simulated experimental run and is compared with the various intervals to determine which one contains the stimulus. Thus in the same way that, in a psychoacoustic experiment, the subject has to get an idea of what she or he has to listen for, the optimal detector must derive its template at the beginning of each simulated measurement.

^{a)}Current address: Carl-von-Ossietzky Universität Oldenburg, Graduiertenkolleg Psychoakustik, Postfach 2503, D-26129 Oldenburg, Germany.

^{b)}Current address: Akustikbüro Göttingen, Kornmarkt 2, D-37073 Göttingen, Germany.

^{c)}All correspondence should be sent to Armin Kohlrausch at this address. E-mail: kohlraus@natlab.research.philips.com

Simulations were performed for a variety of masking conditions that are particularly suited to test the influence of the various stages of the model. As for the implementation of the nonlinear adaptation stage (feedback loops), several forward-masking conditions are simulated which always include the *transition* between simultaneous and forward masking. These conditions investigate the influence of masker duration and masker level. In order to test the choice for an internal noise with constant rms value in combination with the optimal detector, a great number of simultaneous masking conditions with *frozen* noise are investigated. In order to allow a direct comparison between model predictions and human performance, thresholds for most simulated conditions were also measured for a number of subjects using exactly the same acoustic stimuli and the same psychophysical procedure as in the simulations.

We want to emphasize that all simulations presented in this paper were performed with the same values of the model parameters.

I. METHOD

A. Procedure and subjects

Masked thresholds were measured and simulated using an adaptive three-interval forced-choice (3IFC) procedure. The masker was presented in three consecutive intervals, separated by silent intervals of 200 ms. In one randomly chosen interval the test signal was added to the masker. The subject's task was to specify the interval containing the test signal. The threshold was adjusted by a two-down one-up algorithm (Levitt, 1971), which converges at a signal level corresponding to a probability of being correct of 70.7%. The step size was 4 dB at the start of a run and was divided by 2 after every two reversals of the signal level until the step size reached a minimum of 1 dB, at which time it was fixed. Using this 1-dB step size, ten reversals were obtained and the median value of the signal levels at these ten reversals was used as the threshold value. In all the simulations, an algorithm was used to derive the mean values as described in the accompanying paper. The algorithm leads to the level at which the model yields 70.7% correct responses in a simulated 3IFC procedure. The subjects received visual feedback during the measurements. The procedure was repeated four times for each signal configuration and subject. All figures show the median and interquartile ranges based on four single measurements. All five subjects had experience in psychoacoustic measurements and had normal hearing. They were between 23 and 29 years old.

B. Apparatus and stimuli

All acoustic stimuli were digitally generated at a sampling frequency of 30 kHz. The stimuli were transformed to analog signals with the aid of a two-channel 16-bit D/A converter, attenuated, low-pass filtered at 10 kHz and diotically presented via headphones (Beyer DT880 monitor with diffuse-field equalizer) in a soundproof booth. Signal generation and presentation of trials were controlled by a computer using a signal-processing software package developed at the psychoacoustic laboratory in Göttingen.

FIG. 1. Masked thresholds of a short signal presented in a frozen-noise masker as a function of its temporal position relative to the masker onset. The filled symbols (●) indicate the simulated thresholds. The measured thresholds of the three subjects are represented by open symbols. Signal: 5-ms Hanning pulse at 1 kHz; masker bandwidth: 20–5000 Hz, masker duration: 300 ms, masker level: 77 dB SPL. Subjects: AS (◇) SM (△), TD (○).

The maskers used in the different experiments had a flat spectrum between 20 and 5000 Hz. The starting phase of each spectral component was uniformly distributed in the range of 0° – 360° . Unless explicitly stated, a masker level of 77 dB SPL was used. This level corresponds to a spectrum level of 40 dB. In all the experiments with a deterministic noise masker (frozen noise), the masker waveform was presented in the same manner in all intervals. During the experiments with a running-noise masker, an independent sample of the noise was presented in each interval. The signals were sinusoids of different durations—depending on the particular experiment. In each case they were windowed over their entire length with a Hanning window.

II. DETERMINISTIC NOISE MASKERS

A. Simultaneous masking

In a reproducible frozen-noise masker, the masked threshold of a simultaneously presented short signal depends on the spectral properties of the stimuli and also on the temporal position and phase of the signal relative to the masker (Gilkey *et al.*, 1985; Hanna and Robinson, 1985; Langhans and Kohlrausch, 1992; von Klitzing and Kohlrausch, 1994).

1. Temporal position of short signals in frozen-noise maskers

In the first experiment the thresholds of a 5-ms 1-kHz sinusoidal signal were simulated and measured at several temporal positions in a fixed 300-ms section of a broadband noise masker. Figure 1 shows the results of the model (filled symbols) and those of three subjects (open symbols). The abscissa indicates signal onset relative to masker onset. The ordinate indicates the signal level at the threshold (masked threshold) in dB SPL. The thresholds vary markedly with the temporal position. These variations reflect the interaction between the frozen-noise waveform and the signal that is typical of frozen-noise measurements (Hanna and Robinson, 1985). The simulated and the measured thresholds of all sub-

FIG. 2. Masked thresholds for a fixed temporal position ($t=115$ ms) of the signal in frozen noise as a function of the starting phase of the signal. Signal and masker parameters as in Fig. 1. Subjects: MH (Δ), AE ($\square$), TD ($\circ$), optimal detector ($\bullet$).

jects show a similar pattern up to the signal position with $t=155$ ms, with a difference of 10 dB between the highest and the lowest thresholds. In this temporal range the simulated values are close to the values of the most sensitive human observer. In the case of later temporal positions, the agreement between the simulated and measured thresholds is poorer, but only in one temporal position ($t=165$ ms) is the predicted threshold different from measured thresholds by more than 3 dB. However, the variability in the experimental data of the three subjects also increases for later signal positions.

2. Relative phase at a fixed temporal position of the signal in the masker

As pointed out above, the threshold of a short signal presented in frozen noise depends on its temporal position and its phase relative to the masker. In Fig. 2 simulated and measured thresholds are plotted for a fixed temporal position ($t=115$ ms) of the signal in the noise. The relative signal phase is indicated on the abscissa. The ordinate indicates the signal level at the threshold. Masker and signal were the same as in the first experiment. The phase “0” corresponds to the configuration of the previous subsection. There is a sinusoidal dependence of simulated and measured thresholds on the relative phase (cf. Langhans and Kohlrausch, 1992; von Klitzing and Kohlrausch, 1994). The shape of the function is captured by the simulation, whereas the absolute values differ somewhat from those of the subjects. Interestingly the thresholds of the most sensitive subject are always within 1 to 2 dB of the predicted values. This suggests that the higher threshold values of the other subjects could be precisely simulated if the model were made less sensitive by increasing the variance of the internal noise.

3. Signal integration

In this experiment we investigated the influence of signal duration on the measured and simulated thresholds. In experiments with signals presented in noise maskers the thresholds were found to decrease with an increasing signal duration up to about 200 ms (e.g., Plomp and Bouman, 1959). The auditory system seems to be able to integrate

FIG. 3. Masked thresholds of a signal presented in frozen noise for several signal durations. Signal: 1 kHz; masker bandwidth: 20–5000 Hz, masker duration: 600 ms, masker level: 77 dB SPL. Subjects: AS ($\diamond$), MH ($\square$), TD ($\circ$), optimal detector ($\bullet$).

signal intensity in detection tasks over at least 200 ms. That is why a time-integrating process with a time constant of about 200 ms is often postulated (Zwicker and Feldtkeller, 1967; Florentine, 1986). We investigated whether the model is able to stimulate this effect using a signal presented in a 600-ms section of a frozen-noise masker with a level of 77 dB SPL.

Figure 3 shows the results of the model and those of three subjects for signal durations of 10, 20, 40, 80, 160, 320, and 400 ms. The onset for all signals occurred 100 ms after masker onset. The signal duration is indicated on the abscissa and the ordinate indicates the signal level at the threshold. As expected, there is a 3-dB decay per doubling of signal duration in the measured thresholds up to a duration of 160 ms. However, the model does not show a similar decay. The difference between the lowest and highest thresholds is only 8 dB in contrast to 17 dB in the measurements. How can this discrepancy be explained? As remarked in Sec. I B of the accompanying paper, the feedback loops of the adaptation model (Püschel, 1988) are very sensitive to fast changes in the input. This sensitivity causes very strong oscillations at the output of this stage, when the signal is switched on. To illustrate this, the internal representation of the template was computed and plotted in Fig. 4 for different signal durations. In each case the signal started 100 ms after the masker had been switched on. The ordinate indicates the amplitude of the nonlinear processed waveform and is scaled in model units (MU). There is a considerable overshoot at the beginning of the test tones. The overshoot is the dominant part of the representation of each template. Therefore, the model calculations do not lead to the expected 3-dB decay per doubling of signal duration as would be the case for a “perfect” integration. To obtain better agreement between model predictions and the data in terms of the effect of signal duration, it would be necessary to reduce the considerable overshoot at the adaptation stage (see the discussion).

4. Signal frequency

An important parameter affecting simultaneous masking is the signal frequency. In the next experiment the thresholds of a 200-ms sinusoidal signal presented in a 600-ms section

FIG. 4. Representations of the normalized template for different signal durations. The signal started 100 ms after masker onset. The ordinate is scaled in model units (MU).

of a broadband noise masker were measured and simulated. The signal started 200 ms after masker onset. In this experiment the masker was generated with a flat spectrum between 0 and 15 kHz. The level of the masker was 87 dB SPL. The signal frequency fell in the range from 200 to 5000 Hz. Only the channel of the basilar membrane that was tuned to the signal frequency was considered in the model. With increasing frequency, the bandwidth of the bandpass filters on the basilar membrane increases. Consequently, with increasing frequency of the signal more energy of the masker falls within the filter bandwidth, and the masked thresholds increase. In addition, because of the deterministic masker waveform, one would expect a fine structure in the thresholds that results from the direct phase-sensitive interaction between the masker and the signal. Figure 5 shows the results of the model and the measured thresholds. The abscissa indicates the signal frequency and the ordinate indicates the signal level at the threshold. As the figure demonstrates, the model is able to simulate the fine structure of the plot of the measured thresholds as well as the overall increase with frequency.

FIG. 5. Masked thresholds of a 200-ms signal presented in a 600-ms section of a frozen-noise masker as a function of the signal frequency. Subjects: AE ($\square$), AS ($\diamond$), TD ($\circ$), optimal detector ($\bullet$).

FIG. 6. Masked thresholds of the 200-ms signal presented with a phase shift of $\pi/2$ in the same frozen noise as in Fig. 5. Subjects: AE ($\square$), AS ($\diamond$), TD ($\circ$), optimal detector ($\bullet$).

In a further experiment the starting phase of each signal relative to the masker was shifted by $\pi/2$. Figure 6 shows the simulated and measured thresholds. Once again the thresholds tend to increase with an increasing frequency, but this time the fine structure of the thresholds is quite different, especially at low frequencies. Again, simulation and measurement are in good agreement.

B. Forward masking

The threshold of a signal in forward masking decreases with an increasing temporal separation between the masker and the signal. The signal typically reaches its absolute threshold 200 ms after masker offset (e.g., Stein, 1960; Elliot, 1971; Duifhuis, 1973).

A major feature of forward masking is that the threshold decreases rapidly for small masker-signal delays and more slowly for large delays (Zwicker, 1984). Another typical feature of forward-masking curves is the nonlinear dependence of the initial slope on masker level and masker duration. The rate of decay of masking is greater for short maskers and high levels than for long maskers and low levels (Zwicker, 1984; Kohlrausch and Fassel, 1996).

In addition, the initial slope of forward masking depends on the signal frequency. Due to the mechanical filtering process in the inner ear, the offset of the masker is prolonged and this ringing takes a longer time at low frequencies where the mechanical filters have a narrower bandwidth. Therefore, the first parts of forward-masking curves are shallower for low signal frequencies than for high frequencies. In addition, a phase-sensitive interaction between the signal and a frozen-noise masker can be observed at the beginning of the forward-masking curve. Measurements and simulations of these aspects of forward masking can be found in Langhans (1991). In the present study we simulated the dependence of the threshold function on the level and duration of the masker.

Figure 7 shows the simulated and measured thresholds of a 10-ms 1-kHz signal presented at the end and after a fixed 200-ms section of a frozen-noise masker. It shows the threshold level in forward masking as a function of the temporal separation between masker offset and signal offset. The filled

FIG. 7. Forward-masked thresholds of a short signal as a function of its temporal position relative to the offset of a broadband frozen-noise masker. The filled circles (●) indicate the simulated thresholds, the measured thresholds of the subject TD being represented by the open symbols (○). Signal: 10 ms, 1 kHz; masker: 200 ms, 20–5000 Hz, 77 dB SPL.

symbols represent the simulated thresholds while the open circles indicate the measured thresholds of one subject (TD). The two left most data points correspond to a simultaneous masking condition.

As described in Dau *et al.* (1996) with regard to the adaptation model of Püschel, five feedback elements were used in series with different time constants for the individual feedback loops. In Püschel's original simulations the time constants of the five feedback stages were spaced at equal increments between 5 and 500 ms. That led to the time constants $\tau_1=5$ ms, $\tau_2=129$ ms, $\tau_3=253$ ms, $\tau_4=376$ ms, and $\tau_5=500$ ms. For the present simulation we applied a modified version of the adaptation model. One of the larger time constants (376 ms) was replaced by a smaller one (50 ms). This caused a steeper threshold function for short delays, which led to a better agreement between the results of the measurement and simulation in the case of both the present results and the previous data (Langhans, 1991). The threshold for the signal ending 40 ms after masker offset ($t=40$ ms) was still about 15 dB above the absolute threshold (19 dB SPL under this condition).

1. Dependence on masker level

Next we examined the sensitivity of the model to changes in the forward-masker level. Thresholds were measured for three subjects at three masker levels, 77, 57, and 37 dB SPL. Figure 8 shows the individual results and, in the bottom right panel compares the median data with the simulated thresholds. The uppermost curve in the top-left panel (subject TD) is replotted from Fig. 7. For simultaneous masking ($t=-5, 0$ ms), in the simulation as well as in the measured data, each 20-dB change in masker level leads to a 20-dB change in signal threshold. For forward masking ($t>0$ ms) the rate of decay of masking increases with increasing masker level for all subjects and the simulated data. Although there are small differences across subjects with respect to the steepness of the threshold functions, the results of the simulations are in good agreement with the median

FIG. 8. Simulated and measured forward-masked thresholds of a short signal as a function of its temporal position relative to the offset of a broadband frozen-noise masker. The two upper panels and the lower left panel show the measured thresholds of three subjects (TD, AS, JV) for different masker levels: 77 dB SPL (○), 57 dB SPL (◇), and 37 dB SPL (△). The lower right panel shows the median measured thresholds of all subjects (open symbols) together with the simulated thresholds (filled symbols). Signal: 10 ms, 1 kHz; masker: 200 ms, 20–5000 Hz.

values of the measurement, and are also consistent with data reported in the literature (Penner, 1974; Jesteadt *et al.*, 1982; Moore and Glasberg, 1983).

2. Dependence on masker duration

In the next experiment we analyzed forward-masking curves for two different masker durations. Figure 9 shows the simulated and measured thresholds of a 5-ms 2-kHz signal presented after a 200-ms section of the masker at 77 dB SPL. The measured thresholds of two subjects were in very

FIG. 9. Forward-masked thresholds of a short signal as a function of its temporal position relative to the offset of a 200-ms broadband noise masker. Signal: 5 ms, 2 kHz; masker duration: 200 ms, masker bandwidth: 20–5000 Hz, masker level: 77 dB SPL. Subjects: AS (◇), SM (△), TD (○), optimal detector (●).

FIG. 10. Forward-masked thresholds of a short signal presented after a 25-ms section of a broadband noise masker. Signal: as in Fig. 9; Masker duration: 25 ms, masker bandwidth: 20–5000 Hz, masker level: 77 dB SPL. Subjects: AS ($\diamond$), SM ($\triangle$), TD ($\circ$), optimal detector ($\bullet$).

good agreement with the model. The third subject (SM) performed better than the model, especially at short masker-signal delays (nearly 9 dB). But it should be emphasized that this subject was a highly trained observer who had gained much experience in similar experiments. In addition, the time constants were adjusted to fit the data of listener TD in those experiments. Nevertheless, the model covers the range of the three observers quite well.

The use of shorter masker durations at the same masker level as above leads to some differences between the measured and calculated data. Figure 10 shows the signal level at the threshold as a function of the temporal separation between the masker and the signal. The masker had a duration of 25 ms. The simultaneous masked thresholds of all subjects and those of the model are almost identical at $t=0$ ms and are also similar to the values for the long masker. At greater delay times, differences of about 10 dB between the results of the simulation and the measurements appear. The simulated threshold function has a shallower decay than the measured thresholds. There was a steeper decay at a masker duration of 25 ms than at a duration of 200 ms in both the measurements and in the simulations, but this effect was not as strong in the simulations as in the measurements. This means that the calculated thresholds are too high for short maskers.

The reason why the model overestimates the thresholds lies in the pronounced overshoot of the feedback loops that was already mentioned in connection with the signal integration. The slope of forward-masking curves directly reflects the discharging of the capacitors of the five feedback loops. The discharge will take a short time if only the capacitors with the short time constants have been charged; it will take a longer time if the capacitors with long time constants have also been charged. The former will be the case for a short masker, the latter for a long masker.

However, the pronounced overshoot at the onset of every masker will produce a significant charge of all capacitors. Thus irrespective of the duration of the masker, the feedback with the longest time constant will contribute a little to forward masking and therefore the curve predicted for a short

FIG. 11. Simulated ($\bullet$) and measured (open symbols) backward-masked thresholds of a short signal presented before a broadband frozen-noise masker was switched on. Signal: 10 ms, 1 kHz; masker bandwidth: 20–5000 Hz, masker level 77 dB. Subjects: AE ($\square$), AS ($\diamond$), TD ($\circ$).

masker will be too shallow compared with the experimental data. This discrepancy again points out that the amount of overshoot response in the model should be limited in an improved version.

C. Backward masking

Under backward-masking conditions, the signal threshold is typically raised for signal-masker intervals, Δt , up to 20 ms (e.g., Duifhuis, 1973). For our backward-masking measurements and simulations we used the same 200-ms section of a broadband masker as in our first forward-masking experiment. The signal was a 10-ms 1-kHz sinusoid that was presented before and at the beginning of the masker. Figure 11 shows the results of the simulation (filled circles) and of three subjects (open symbols). The temporal position of the signal onset relative to the masker onset is indicated on the abscissa. All signal positions with $t \geq 0$ ms correspond to simultaneous masking, those with $t < -10$ ms correspond to backward masking and the positions in between indicate the transition between these two conditions. The ordinate indicates the signal level at the threshold.

The data measured for the three subjects differ considerably. Subject AE showed strong backward masking up to a signal masker interval of 60 ms. In contrast, the two other subjects showed only a small backward-masking effect. For these two subjects the signal threshold equals the absolute threshold for $t \geq -20$ ms. At later signal positions the threshold increases abruptly.

The simulated data show even smaller backward masking effects. Thresholds are higher than the absolute threshold only for $t > -15$ ms. The backward masking in the model is caused by the properties of the basilar-membrane filtering. Because the linear model of the basilar membrane used (Strube, 1985) has no sharp resonances, the impulse responses of those filters are relatively short. Therefore backward masking occurs only for very small temporal separations between the signal and the masker. If, in contrast, use were to be made of a nonlinear basilar-membrane model, the initial bandwidth would be narrower and this should lead to

a somewhat stronger backward-masking effect in the simulation. However, to the extent that backward masking reflects more “central” processes, the model in its current version does not simulate the effect.

III. SIMULATIONS AND MEASUREMENTS USING STOCHASTIC NOISE MASKERS

Thresholds of sinusoidal signals presented in random noise can be calculated as a function of the spectral properties and the duration of the stimuli (Zwicker and Fastl, 1990). The temporal position of the signal in the masker does not influence the threshold, except when the signal is presented shortly after the masker has been switched on (overshoot effect).

In this section some simulations of simultaneous masking of short signals in random noise will be discussed. The level and spectral properties of the masker were chosen to match those in the corresponding experiments with frozen noise.

In all experiments discussed so far we have considered deterministic noise maskers. In each interval the difference between the current representation and the “stored” representation of the deterministic masker was calculated. This led to two intervals containing only interval noise and one interval containing the nonlinearly transformed signal plus internal noise. To apply the model to random noise, a different sample of the masker was presented in each interval. To model the “reference” of the masker, we calculated the mean internal representation of several masker samples. This averaged reference was calculated once before the adaptive procedure was started. During the adaptive procedure we computed the difference between the actual representation and the reference in each interval; this gave three different representations that were affected by the statistical properties of the *external noise* in addition to the internal noise.

In order to be able to generate the template, we computed an averaged suprathreshold internal representation of several signal masker configurations. The top panel of Fig. 12 shows an example of a mean suprathreshold representation of the sum of masker and signal. The signal is a 5-ms 1-kHz sinusoid that starts 130 ms after the onset of a 300-ms section of a random-noise masker. The middle panel of Fig. 12 shows the averaged internal representation of the masker alone. The bottom panel shows the normalized difference between the two representations, i.e., the template.

The algorithm for calculating the threshold was similar to that used in the simulations with deterministic maskers. Again, the internal noise set a limit to the performance. But unlike in the simulations with frozen noise, the “criterion” d' was given by the difference between the largest “noise-correlation value” (i.e., the largest correlation coefficient of the two noise-alone representations with the template) and the correlation value between the signal interval and the template. When this difference was smaller than the limit of resolution determined by the internal noise the signal was not detected.

Because the model behaves essentially stochastically, the median thresholds of four “runs” were computed for each simulated condition. Figure 13 shows the simulated and

FIG. 12. Generation of the template representation (at the bottom) of a short signal in a random-noise masker. The signal started 130 ms after the masker had been switched on. The template is the normalized difference between the mean representation of the masker plus the suprathreshold signal (at the top) and the mean representation of the masker alone (in the middle). The ordinate is scaled in model units (MU).

measured thresholds of a 5-ms 1-kHz signal presented in random noise (filled and open squares, respectively). The phase of the signal relative to the masker at a fixed temporal position ($t=115$ ms) is indicated on the abscissa. The ordinate indicates the signal level at the threshold. As expected, neither in the simulation nor in the measurement does the threshold depend on the signal phase relative to the masker. To enable comparison of these results with those obtained in deterministic noise the corresponding simulated and mea-

FIG. 13. Simulated (filled symbols) and measured thresholds (open symbols) of a 5-ms 1-kHz signal in frozen noise (circles) and in random noise (squares), plotted as a function of the starting phase of the signal. Subject: TD.

FIG. 14. Simulated (filled symbols) and measured thresholds (open symbols) of a 5-ms 1-kHz signal in frozen noise (circles) and in random noise (squares), plotted as a function of the temporal position of the signal relative to masker onset. Subject: TD.

sured results (of the same subject) with frozen noise are plotted in the same figure. The thresholds in random noise lie in the upper third part of the range of the thresholds obtained in frozen noise. This is in good agreement with the evidence reported in the literature (Langhans and Kohlrausch, 1992).

Similar results were obtained when the temporal position of the signal was varied. Figure 14 shows the simulated and measured thresholds of the signal presented in random noise (squares) and in frozen noise (circles). The temporal position of the signal is indicated on the abscissa as the time of the signal onset relative to the masker onset. The random noise thresholds again lie in the upper third part of the range of the thresholds obtained with frozen noise. For only two of the eleven delays (105, 145 ms) are the frozen-noise thresholds higher than the random-noise thresholds. This holds for both the model and the human observer.

We also simulated the effect of presenting the signal shortly after the masker onset. Figure 15 shows the simulated thresholds of the signal in random and frozen noise for “earlier” temporal positions of the signal. The random-noise thresholds (squares) remained at an almost constant level of

FIG. 15. Simulated thresholds for several signal positions shortly after the frozen-noise masker (circles) and the random-noise masker (squares) had been switched on.

79 dB SPL, whereas the corresponding frozen-noise thresholds (circles) were as low as 67 dB SPL when the signal was presented just after the masker onset.

How can these thresholds at masker onset be explained? In the simulations we calculated the difference between the actual representation of the masker plus signal and the “stored” representation of the masker itself. In the case of frozen-noise maskers the difference contains the nonlinearly transformed signal only embedded in internal noise. Because the adaptation model responds almost linearly to the onsets of the stimuli, the computed signal-to-noise ratio is larger at signal positions shortly after the beginning of the masker than at later temporal positions of the signal.

In the case of random-noise maskers, on the other hand, the signal-to-noise ratio remains almost constant because the representations of the different samples of the masker have strong “oscillations” that are not correlated with each other. The difference between masker-plus-signal and masker fluctuates substantially at the beginning of the masker owing to the external noise, so the “energy” of the masker and the “energy” of the signal both increase proportionally. Therefore the threshold in a random-noise masker does not depend on the signal’s temporal position.

These results also indicate that the present model does not predict overshoot in random or frozen noise. As discussed in a recent paper by von Klitzing and Kohlrausch (1994), one reason could be that we have so far not included a correct model for basilar-membrane nonlinearities.

IV. SUMMARY AND DISCUSSION

In this study some results of simulations concerning backward, simultaneous and forward masking have been described and compared with measured thresholds. Deterministic (frozen) and stochastic (random) broadband noise maskers were used and only diotic configurations were considered.

In all the simulations the temporal fine structure of the representation of the stimuli at the output of the auditory filter tuned to the signal frequency was analyzed. The standard deviation of the internal noise was determined in a basic experiment of level discrimination so as to satisfy Weber’s law. The time constants of the nonlinear adaptation model were determined in an appropriate simulation of forward masking in order to adapt them to the results of a “normal” listener. To get a consistent and complete model, the same parameters were applied in all further simulations.

In most experiments the simulated and the measured thresholds showed a very good agreement. It was possible to simulate the range of thresholds as well as the fluctuations in the thresholds. In the simulations concerning signal integration and forward masking using short maskers, some discrepancies between the measurements and the results of the simulations were observed. These discrepancies were probably caused by the extremely high sensitivity of the adaptation model at the onset of signals. For a better prediction of these thresholds within the framework of the present model, it is necessary to limit the strong onset response.

We conclude that the preprocessing model can describe the main features of temporal masking in the auditory sys-

tem. Furthermore, it seems to be reasonable to assume that a subject is able to “store” a representation of the suprathreshold stimulus and compare it with the current representation of the sound. The subject apparently bases decisions on the similarity of the two compared “patterns.”

The optimal detector seems to be a successful concept for describing detection processes in the human hearing system. On the basis of the wide range of conditions for which we can predict thresholds reasonably well, we expect to be able to simulate more complex experiments too, like gap detection and detection tasks with complex maskers other than noise. Such work is currently being done in Oldenburg and Göttingen.

ACKNOWLEDGMENTS

We would like to thank all our colleagues of the psychoacoustic research groups at the Universities of Göttingen and Oldenburg for their fruitful discussions on the content of this paper. We also thank Dr. Brian Moore for his comments and suggestions concerning this study and Martin Hansen and Andres Sander for their critical reading of earlier versions of this paper. The two anonymous reviewers also provided very constructive criticism. This study was supported by a grant from the Deutsche Forschungsgemeinschaft (Ko 1033/3-1).

- Dau, T., Püschel, D., and Kohlrausch, A. (1996). “A quantitative model of the ‘effective’ signal processing in the auditory system: I. Model structure,” *J. Acoust. Soc. Am.* **99**, 3615–3622.
- Duifhuis, H. (1973). “Consequences of peripheral frequency selectivity for nonsimultaneous masking,” *J. Acoust. Soc. Am.* **54**, 1471–1488.
- Elliot, L. L. (1962). “Backward and forward masking of probe-tones of different frequencies,” *J. Acoust. Soc. Am.* **34**, 1116–1117.
- Florentine, M. (1986). “Level discrimination of tones as a function of duration,” *J. Acoust. Soc. Am.* **79**, 792–798.
- Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (1985). “Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise,” *J. Acoust. Soc. Am.* **78**, 1207–1219.
- Hanna, T. E., and Robinson, D. E. (1985). “Phase effects for a sine wave masked by reproducible noise,” *J. Acoust. Soc. Am.* **77**, 1129–1140.

- Jesteadt, W., Bacon S. P., and Lehmann, J. R. (1982). “Forward masking as a function of frequency, masker level, and signal delay,” *J. Acoust. Soc. Am.* **71**, 950–962.
- Klitzing, R. von, and Kohlrausch, A. (1994). “Effects of masker level on overshoot in running- and frozen-noise maskers,” *J. Acoust. Soc. Am.* **95**, 2192–2201.
- Kohlrausch, A., and Fassel, R. (1996). “Binaural masking level differences in nonsimultaneous masking,” in *Binaural and Spatial Hearing*, edited by R. Gilkey and T. R. Anderson (Erlbaum, Hillsdale, NJ, in press).
- Kohlrausch, A., Püschel, D., and Alpehi, H. (1992). “Temporal resolution and modulation analysis in models of the auditory system,” in *The Auditory Processing of Speech*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin, New York), pp. 85–98.
- Langhans, A. (1991). “Psychoakustische Messungen und Modellvorstellungen zur Wahrnehmung reproduzierbarer und statistisch fluktuierender Signale,” Ph.D. thesis, University of Göttingen.
- Langhans, A., and Kohlrausch, A. (1992). “Differences in auditory performance between monaural and diotic conditions. I: Masked thresholds in frozen noise,” *J. Acoust. Soc. Am.* **91**, 3456–3470.
- Levitt, H. (1971). “Transformed up–down procedures in psychoacoustics,” *J. Acoust. Soc. Am.* **49**, 467–477.
- Moore, B. C. J., and Glasberg, B. R. (1983). “Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise,” *J. Acoust. Soc. Am.* **73**, 1249–1259.
- Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (1992). “Complex sounds and auditory images,” in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 429–443.
- Penner, M. J. (1974). “Effects of masker duration and masker level on forward and backward masking,” *J. Acoust. Soc. Am.* **56**, 179–182.
- Plomp, R., and Boumann, M. A. (1959). “Relation between hearing threshold and duration for tone pulses,” *J. Acoust. Soc. Am.* **31**, 749–758.
- Püschel, D. (1988). “Prinzipien der zeitlichen Analyse beim Hören,” Ph.D. thesis, University of Göttingen.
- Stein, H. J. (1960). “Das Absinken der Mithörschwelle nach dem Abschalten von weißem Rauschen,” *Acustica* **10**, 116–119.
- Strube, H. W. (1985). “A computationally efficient basilar-membrane model,” *Acustica* **58**, 207–214.
- Zwicker, E., and Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger* (Hirzel, Stuttgart).
- Zwicker, E. (1984). “Dependence of post-masking on masker duration and its relation to temporal effects in loudness,” *J. Acoust. Soc. Am.* **75**, 219–223.
- Zwicker, E., and Fastl, H. (1990). *Psychoacoustics* (Springer-Verlag, Berlin).