An explanation for the apparantly wider critical bandwidth in binaural experiments

J. Breebaart[1], S. van de Par[1,2] and A. Kohlrausch[1,2]

[1]*IPO Center for User-System Interaction*
P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands
d.j.breebaart@tue.nl

[2]*Philips Research Laboratories Eindhoven*
Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands

1. Introduction

When a tone is presented together with another signal, like broadband noise, the threshold of audibility of the tone is usually higher than the threshold of audibility of the tone alone. The amount of masking depends, among other aspects, on the spectral extent of the masking noise and the center frequency of the tone. The masked thresholds can be predicted by assuming that the auditory system splits the incoming sounds into several band-limited signals (Fletcher, 1940).

Besides the spectral content, the interaural relationship of both masker and test signal influences the masked threshold. When a broadband noise is presented in phase to both ears, and pure tones are presented out of phase to each ear simultaneously (NoSπ condition), masked thresholds are generally lower than for the case when both the noise and the tone are presented in phase (NoSo condition). This difference is generally referred to as binaural masking level difference (BMLD) and can amount up to 25 dB (Zurek and Durlach, 1987). The increased sensitivity for the out-of-phase signal in the NoSπ condition stems from the occurence of interaural differences between the signals arriving in both ears.

One paradoxical issue concerning the band-filter concept is that certain binaural experiments suggest that the bandwidth of the auditory filters is wider than the bandwidth typically estimated from monaural experiments (eg. Glasberg and Moore, 1990). Hall *et al*. (1983) measured detection thresholds for an NoSπ condition as a function of the bandwidth of the masking noise keeping the spectral level of the masker constant. They found that the detection threshold continues to increase for masker bandwidths up to 2 to 3 times the monaural bandwidth, for the conditions where the masking spectral level is high. In contrast, when they estimated the auditory filter bandwidth from experiments with broadband noise containing a notch of variable bandwidth, the estimated bandwidths were in line with estimates from monaural experiments. We think that these different estimates are related to the fact that the auditory system not only processes the signals in the on-frequency filter, but also uses off-frequency filters in conditions in which such a process enhances the detection performance. For a broadband masker, the interaural differences are largest for the auditory filter that is centered on the test-signal frequency. For a narrowband masker, however, it is not a single auditory filter that contains interaural differences. In fact, several auditory filters have similar interaural cues and we will discuss that it is an important issue to consider that the auditory system may integrate binaural information across various filters. In this paper it will be explained how the integration of binaural information across frequency can lead to apparently wider

critical bandwidths even when the underlying auditory filters have bandwidths consistent with monaural estimates.

2. Binaural signal detection model

To demonstrate the role of across frequency integration, a signal detection model is developed which satisfies several boundary conditions. Most importantly, the model includes a range of bandpass filters to simulate the peripheral filtering of the cochlea which have bandwidths which match the widths of monaural estimates obtained by Glasberg and Moore (1990). The model that is developed here is able to generate quantitative threshold predictions, both for monaural and binaural conditions and it is a simplified version of a more elaborate model by Breebaart *et al* (2000).

The model is divided into three stages. The first stage simulates the peripheral preprocessing of the outer, middle and inner ear. The outer and middle ear processing is simulated by a time-invariant bandpass filter with a roll-off of 6 dB/oct. below 1 kHz and -6 dB/oct. above 4 kHz. The processing of the inner ear is simulated by a 3rd order gammatone filterbank with equivalent rectangular bandwidths (ERBs) corresponding to the estimates published by Glasberg and Moore, (1990) and a spectral spacing of 2 filters per ERB. The effective signal processing of the inner haircells is modelled by a half-wave rectifier, a 5th order low-pass filter at 650 Hz (Weiss and Rose, 1988), followed by a chain of 5 adaptation loops (Dau *et al*., 1996). The almost logarithmic input-output characteristic of the chain of adaptation loops reduces the dependence of thresholds on the overall stimulus level.

The second stage simulates the binaural interaction found in the auditory system. This process is implemented as an excitation-inhibition (EI) interaction: the waveform from one ear has an excitatory effect on the activity of a neural unit, the other ear has an inhibitory effect. To generate the EI interaction, for each auditory filter, a difference signal between inputs from the left and right side is computed. This difference is squared and integrated over time. To incorporate a finite detection threshold, an independent Gaussian-noise signal with a constant RMS value is added to the difference signal of each auditory filter. Thus the difference output for auditory filter n, integrated over time denoted by E_n can be written as:

$$E_n = \varepsilon_n + \frac{1}{T}\int_0^T \left(L_n(t) - R_n(t)\right)^2 dt \,.$$

Here, $L_n(t)$ and $R_n(t)$ represent the preprocessed signal for filter n from the left ear and the right ear, respectively, ε_n represents independent error variables (i.e., internal noise) which follow a Gaussian distribution with zero mean and fixed RMS, and T is the integration time (i.e., stimulus duration).

The third stage, which combines the difference signals in the frequency domain, will be described in more detail, since this stage is of great importance for the spectral integration. The idea is that the model can derive a linearly weighted combination of the difference signals for each auditory filter, E_n, and that the weights are learned in such a way that under certain assumptions (see Green and Swets, 1966), the weighting optimally reduces the internal error.

We will now use the model as an `artificial observer' in a 3-interval, forced choice procedure. In this procedure, two intervals contain only the masker, while the third interval contains the masker plus signal. The model's task is to identify which interval contains the signal. This is done in the following way. Assume that within the

model a template is stored which consists of the mean output $\underline{E}_n$ of several masker-alone realizations. Then the task would be to determine which interval induces an internal representation that differs most from the template. The distance U_n between template $\underline{E}_n$ and the actual signal E_n for filter n can be described as:

$$U_n = E_n - \underline{E}_n .$$

The variance of U_n resulting from the internal noise and masker uncertainty is denoted by σ_n^2, while the mean difference in U_n between masker plus test signal and masker alone near threshold level is denoted by μ_n. The following linear combination U of the signals U_n is used as a decision variable:

$$U = \sum_n \frac{\mu_n}{\sigma_n^2} U_n ,$$

which is, under certain assumptions (Green and Swets, 1966), the least-squares estimator for the presence of the signal in a detection task.

The procedure that was used to determine the model predictions was similar to the procedure used with human observers. After each trial (consisting of three intervals) the model determines the interval which, according to the distance measure U, is most likely to contain the signal. Depending on the correctness of the response, the signal level for the next trial is adjusted according to a two-down one-up procedure, and a threshold estimate is derived from a certain number of reversals of the signal level track. Furthermore, the internal representations of the processed stimuli are used to update the estimates of $\underline{E}_n$, σ_n^2 and μ_n.

3. Simulation of band-limited and notched NoSπ data

As briefly discussed in the introduction, binaural band-widening masking conditions result in effective auditory filter bandwidths that are larger than for notch-widening masking conditions. To examine whether the model can account for this phenomenon, we determined the thresholds for the experiments originally performed by Hall *et al*. (1983). We determined the detection threshold for a 500-Hz NoSπ condition with a spectral masker level of 10, 30, and 50 dB/Hz as a function of the bandwidth of the noise. The results are shown in the left panel of Figure 1 by the black symbols; the white symbols denote the experimental data from Hall *et al*. For all three noise levels, the detection threshold first increases with increasing bandwidth up to a few hundred Hertz. However, the bandwidth at which the thresholds stop increasing is larger for higher noise levels. Moreover, the effective filter bandwidth is larger than the reported monaural critical bandwidth of 79 Hz (Glasberg and Moore, 1990). In this respect, the model predictions agree perfectly with the experimental data. However, there is a systematic difference between the model predictions and the experimental data. For signal levels above the absolute threshold, the thresholds of the model are 4 dB lower than the thresholds of human observers. This indicates that the model is slightly more sensitive to binaural cues than the average listener.

The right panel shows thresholds as a function of the notchwidth centered around the test signal frequency (500 Hz), for the same three spectral noise levels. For all spectral levels, the thresholds first decrease with increasing notchwidth. If we take the 3-dB threshold improvement (relative to the threshold in wideband noise) as an

estimate of the frequency resolution, the bandwidth of the auditory filter does agree with the monaural critical bandwidth for all three noise levels. The black symbols represent the model predictions. The predictions agree with the experimental data with respect to the frequency selectivity; the point of 3-dB improvement is very similar for the experimental data and model predictions.

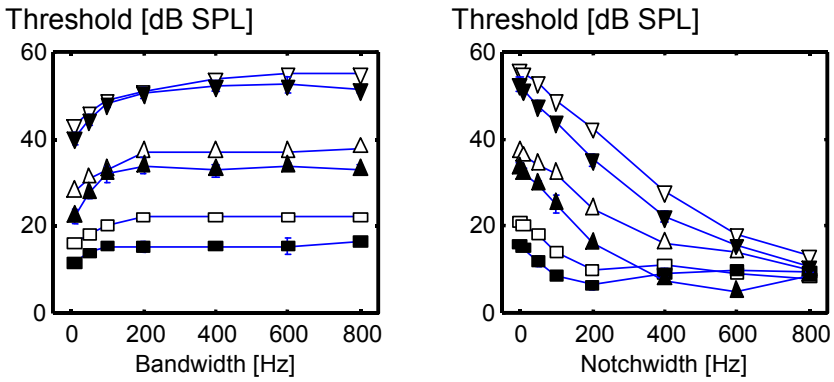

Figure 1. Detection thresholds for an NoSπ condition as a function of the bandwidth of a constant-spectral-level noise (left panel) and as a function of the notchwidth of a notch centered around the signal frequency (right panel). The black symbols are model predictions, the white symbols are experimental data adapted from Hall et al. (1983). The squares denotea masker level of 10 dB/Hz, the upward triangles 30 dB/Hz and the downward triangles 50 dB/Hz.

The explanation for this result in this specific experimental paradigm relies on the across-frequency integration of information, as described by van de Par and Kohlrausch (1999). For a narrowband masker (i.e., below the monaural critical bandwidth), the complete stimulus energy lies within the on-frequency channel. For off-frequency channels, the entire stimulus resides at the skirt of the filters and is therefore reduced in its power. However, the *relative* amount of masker and signal energy is hardly changed. Since the signal-to-masker ratio within an auditory channel determines detectability, information about the presence of the signal is not only available in the on-frequency channel, but also in several off-frequency channels. The only limitation for this extended availability is the absolute threshold: if the stimulus is attenuated too strongly it becomes undetectable in that channel. This extended availability of information is depicted in the left panel of Fig. 2. In this figure, the model output E_n for a 500-Hz NoSπ stimulus is shown as a function of the masker bandwidth and of the filter number. The masker had a spectral level of 50 dB/Hz and the signal level was equal to the corresponding model predictions as shown in the left panel of Fig. 1. The middle and right panels of Fig. 2 show cross-sections of the left panel along the filternumber and masker bandwidth axes, respectively.

We remind the reader that if a masker alone is presented, the output E_n will be zero (neglecting the internal noise) since the masker can be canceled completely. Thus, any increase in the activity E_n can be used as a cue for the presence of the signal. This increase by the presence of the Sπ signal is shown in Fig. 2. If the masker bandwidth is very small (10-Hz wide), a whole range of filters shows activity (see solid line in the middle panel of Fig. 2). The fact that in this condition the cue for detection is available in several channels enables reduction of the internal error in the following way. The internal error which is added to the different filter outputs is *independent* across filters. Thus, individual noise sources add up by their intensities.

On the other hand, the output in different filters by the addition of the signal is available in several filters and is *correlated* and therefore adds up linearly. Thus, if the model uses the sum of activities E_n across several filters instead of using the output of only one filter, the internal signal-to-noise ratio is increased. Therefore, the magnitude of the activities per filter can be relatively small (see solid lines in the middle and right panels of Fig. 2). If the masker bandwidth is increased, the activities in the off-frequency channels start to be *reduced*. This can be observed by the dashed and dash-dotted lines in the right panel of Fig. 2. Hence the reduction process of the internal error by across-frequency integration is less effective. This must be compensated by an increase in the signal level to achieve equal detectability. Thus, the output for the on frequency channel increases with increasing bandwidth to compensate for the less-effective across-frequency integration process. This increase is shown by the solid line in the right panel of Fig. 2.

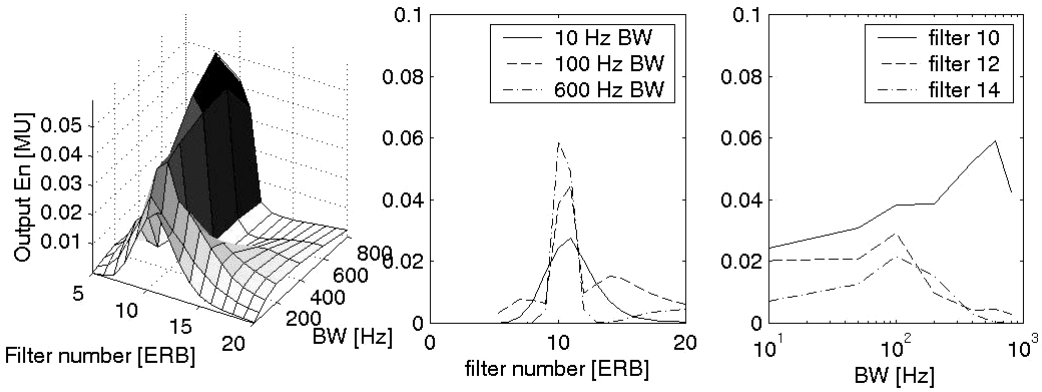

Figure 2. EI output as a function of the auditory filter number and masker bandwidth(left panel) in an NoSπ condition with a fixed spectral masker level and a signal level corresponding to the model predictions shown in the left panel of Fig. 1.The middle panel shows the output for masker bandwidths of 10, 100 and 600 Hz. The right panel shows the output of filter 10, 12 and 14 as a function of the masker bandwidth.

The across-frequency integration process as described above only works at high masker levels. At low masker levels (i.e., 10 dB/Hz), the stimulus energy in the off-frequency channels is below the absolute threshold and hence cannot contribute to the detection process. The explanation of across-frequency integration also accounts for the fact that the notched-noise stimuli do not show a wider effective critical band, as shown in the right panel of Fig. 1. In this condition, all off-frequency channels are masked and hence do not give any valuable information regarding the detection process. Therefore the wider binaural critical bandwidth is not observed in this condition.

4. Across-frequency integration in monaural detection

If we consider a monaural bandwidening experiment, the thresholds resemble those of an energy detector following a peripheral bandpass filter. The limiting factor for detection is the external variability of the masker waveform. Since this external variability is highly correlated across auditory filters there is no gain to be expected from across frequency integration of monaural information and thus we donot expect a wider effective critical bandwidth, such as is the case when detection is limited by internal errors instead of external stimulus variability. An interesting experiment was presented by Kohlrausch *et al.*(1998). They performed a band-limited NoSo detection experiment with frozen-noise samples instead of running noise in which case *internal* errors are determining detectability. As can be expected on the basis of our reasoning, the critical bandwidths estimated from such an experiment also appears to be wider than what is commonly found in monaural experiments.

The spectral level of the noise was either 10 or 50 dB/Hz. The noise had a duration of 400 ms. The 300-ms signal was centered in the noise burst. It was a 10-Hz-wide frozen noise, which was constructed from the central 10 Hz of the masker band by shifting all phases by 90 degrees. Both masker and signal were gated with 50-ms Hanning ramps. The data are shown in Fig. 3 (open symbols) as means of four subjects. The open squares denote the 10 dB/Hz condition, the open triangles the 50 dB/Hz condition. At subcritical bandwidths, thresholds increase with a slope of 3 dB/oct., which is equivalent to a constant signal-to-masker ratio. This slope agrees with the one found in binaural conditions and differs from the slope in monaural experiments with running noise. Beyond a certain bandwidth, thresholds remain constant. This bandwidth is different for the two noise levels; it amounts about 70 Hz for the 10 dB/Hz condition and 150 Hz for the 50 dB/Hz condition. Furthermore, the difference between the 10 dB/Hz thresholds and the 50 dB/Hz thresholds is smaller at

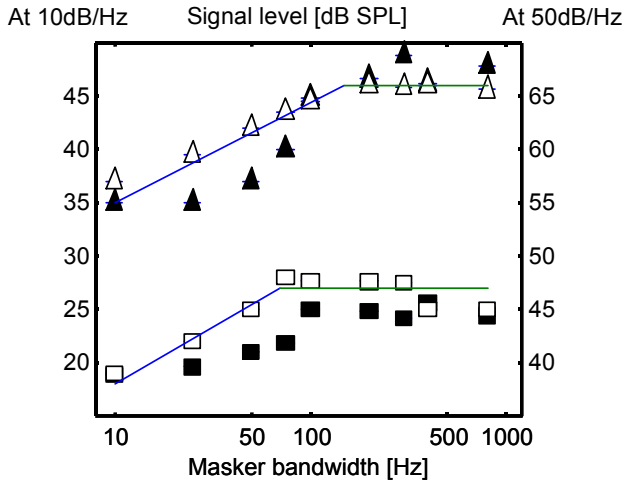

Figure 3. Thresholds of a 10-Hz frozen noise masked by a frozen noise of variable bandwidth with a spectral level of 10 dB/Hz (squares, axis on left side of figure) and 50 dB/Hz (triangles, axis on right side of figure). The open symbols are adapted from Kohlrausch et al. (1998), the filled symbols are model predictions. The solid lines are fits to the experimental data.

10 Hz bandwidth than at 1000 Hz bandwidth. At 1000 Hz bandwidth, the difference is about 40 dB, which is expected based on the difference in masker energy. At 10 Hz bandwidth, this difference amounts 36 dB. This difference is smaller because at 50 dB/Hz and 10 Hz bandwidth, the across-frequency integration results in lower thresholds, while at 10 dB/Hz it does not. These observations are in line with the observation from binaural masking experiments: if the internal error limits the detection and the cue is available in several filters, the detection is enhanced by integration of information across frequency.

The experiment by Kohlrausch *et al.* can easily be modelled by making a small modification to the model presented here. Instead of using the difference signal between the left and right channels, we now only use one channel, resulting in:

$$E_n = \varepsilon_n + \frac{1}{T}\int_0^T L_n(t)\ dt$$

All other processing by the model remains unchanged. All simulations were repeated 5 times to reduce the effect of specific masker waveforms. The results of the simulations are shown in Fig. 3 by the filled symbols. Although the 3 dB/oct is not so nicely present in the model predictions as in the experimental data, the model predicts the wider effective critical band at the higher masker level: at 10 dB/Hz the model has constant thresholds for bandwidths beyond 100 Hz, while at 50 dB/Hz thresholds remain constant above 200 Hz.

5. Conclusions

The apparently wider critical bandwidth that is sometimes observed in binaural but also in monaural conditions can be explained by assuming that information is integrated across frequency. This integration reduces the uncorrelated internal errors for narrowband maskers. With increasing masker bandwidth, the error in the decision variable increases due to a gradual masking of the off-frequency filters, disabling the internal error reduction through spectral integration. This process occurs for masker bandwidths that exceed the auditory filter bandwidth, resulting in an apparently larger effective filter bandwidth for band-widening conditions.

6. References

Breebaart, J., van de Par, S. and Kohlrausch, A. (2000). A model for the effective signal processing in the binaural auditory system based on contralateral inhibition. I. Model setup. *J. Acoust. Soc. Am.* In preparation.

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.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, 12, 47-65.

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

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

Hall, J.W., Tyler, R.S. & Fernandes, M.A. (1983). Monaural and binaural auditory frequency resolution measured using bandlimited noise and notched-noise masking. *J. Acoust. Soc. Am.*, 73, 894-898

Kohlrausch, A., Breebaart, J. and van de Par, S. (1998). The influence of masker variability on estimates of monaural and binaural critical bandwidths. *J. Acoust. Soc. Am.* 103 (5, Pt.2), 3082.

Van de Par, S. and Kohlrausch, A. (1999). Dependence of binaural masking level differences on center frequency, masker bandwidth and interaural parameters. *J. Acoust. Soc. Am*. 106, 1940-1947.

Weiss, T.F. & Rose, C. (1988). A comparison of synchronization filters in different auditory receptor organs. *Hear. Res.*, 33, 175-180.

Zurek, P.M. & Durlach, N.I. (1987). Masker-bandwidth dependence in homophasic and antiphasic tone detection. *J. Acoust. Soc. Am.*, 81, 459-463.