Spectral integration in binaural signal detection

D.J. Breebaart, S.L.J.D.E. van de Par and A. Kohlrausch
e-mail: breebaar@ipo.tue.nl

Abstract

For both monaural and binaural masking, the spectral content of the masker and of the signal to be detected are important stimulus properties influencing the detection process. It is generally accepted that the auditory system separates the incoming signals in several frequency bands. It is not clear a priori, however, how the auditory system combines these separate band-limited signals. In this paper, a spectral integration model is presented that weights the binaural information of various auditory filters adaptively to minimize the errors due to internal noise. The model is tested for several binaural signal detection conditions. It can account for several phenomena concerning binaural signal detection, including the level-dependent wider binaural critical bandwidth in band-widening experiments and the effect of transitions of the interaural masker correlation at a certain frequency. These results suggest that the auditory system combines information across auditory filters under conditions where such processing enhances the detection performance.

Introduction

When a tone is presented together with another signal, like broadband noise for example, the threshold of audibility of the tone is usually higher than the threshold of audibility of the tone by itself. The amount of masking depends on the spectral content of the masking noise and the centre frequency of the test tone. This observation can be explained 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) (Zurek & Durlach, 1987). This difference is generally referred to as binaural masking level difference (BMLD) and can amount up to 25 dB. The increased sensitivity for the antiphasic signal in the NoSπ condition stems from the generation of interaural differences between the signals arriving in both ears.

One unresolved issue concerning the band-filter concept is that for certain binaural conditions, the bandwidth of the auditory filters seems wider than the bandwidth reported from monaural experiments (cf. Glasberg & Moore, 1990). Hall, Tyler and Fernandes (1983) measured the detection threshold for an antiphasic sinusoid added to an interaurally in-phase noise masker with a constant spectral level as a function of the bandwidth of the noise. They found that for masker levels well above the absolute threshold, the detection threshold still increases for masker bandwidths that amount up to 2 to 3 times the monaural bandwidth. In contrast, when they estimated the auditory filter bandwidth from experiments with broadband noise containing a notch of variable bandwidth, the estimat-

ed bandwidths were in line with estimates from monaural experiments. The discrepancy between these different values of the critical bandwidth is still unresolved. We propose, however, that these phenomena are related to the fact that the auditory system not only processes the signals in the on-frequency filter, but can also use off-frequency filters. Clearly, for a broadband masker, the interaural differences are largest for the auditory filter that is centred on the test-signal frequency. For a narrowband masker or for a wideband test signal in a broadband masker, however, it is not clear a priori which auditory filter results in the highest change in interaural differences. Moreover, if several auditory filters have similar interaural cues, an important issue is whether the auditory system can combine the information across these filters and in what way this would occur.

Several modelling efforts have been made to predict binaural masked thresholds. Most models that have emerged for binaural signal processing in recent decades rely on a cross-correlation approach (e.g., Colburn, 1977; Lindemann, 1986; Stern & Shear, 1996). However, the scope of each of these cross-correlation models separately has been rather limited till now. In a recent paper (Breebaart, Van de Par & Kohlrausch, 1997), for example, we noted that most current binaural models cannot account for the experimental data with narrowband maskers. Furthermore, most of these models rely on the processing of one auditory filter only.

To resolve some of these failures, we developed a model based on a binaural interaction mechanism different from the one used in cross-correlation models. This model is based on the processing of the difference between the signals arriving in both ears. From neurophysiological studies, it is known that a subclass of neurons found in the auditory pathway are found to be excited by the signals from one ear and inhibited by the signals from the other ear (e.g., Joris, 1996). The opposite influence of the two ears makes these cells sensitive to interaural differences; effectively, it results in a difference signal for the sounds arriving in the left and right ear. This difference signal can be rendered for each auditory filter. To continue processing of these separate signals, a spectral integration model is proposed that combines the binaural difference signals to enhance the detection performance of the model.

Model description

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 2nd order gammatone filterbank with filters of equivalent rectangular bandwidth (ERB) (Glasberg & Moore, 1990) and a spectral spacing of 1 filter 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 & Rose, 1988), followed by a chain of 5 adaptation loops (Dau, Püschel & Kohlrausch, 1996). The almost logarithmic input-output characteristic, for stationary signals of the chain of adaptation loops, ensures that interaural differences are rendered relative to the total signal intensity within each auditory filter.

The second stage simulates the binaural interaction found in the auditory system. For each auditory filter, the difference signal between inputs from the left and right side is computed and half-wave rectified. To incorporate a finite detection threshold, an inde-

pendent Gaussian-noise signal with a constant RMS value is added to the difference signal of each auditory filter.

The third stage, which combines the contralateral inhibited signals, 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 linear weighted combination of the difference signals for each auditory filter, and that the weights are learned in such a way that under certain assumptions (see Green & Swets, 1966), the weighting optimally reduces the internal error. Assume that the difference output, integrated over time for auditory filter *n* is denoted by E_n. Then for E_n we have:

$$E_n = \varepsilon_n + \int_0^T \lceil L_n(t) - R_n(t) \rceil dt. \quad (1)$$

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. The brackets denote the half-wave rectification process.

We will now use the model as an 'artificial observer' using a 3-interval, forced choice procedure. In this procedure, two stimuli contain a masker alone, while the third stimulus contains a masker plus test signal. The model's task is to identify which interval contains the test signal. This is done in the following way. Say we have a template, that is stored in memory, consisting of the mean output signal $\bar{E}_n$ of several masker-alone realizations. Then the task would be to determine which stimulus induces an internal representation that differs most from the template. The distance U_n between template $\bar{E}_n$ and the actual signal E_n for filter *n* can be described as:

$$U_n = \left| E_n - \bar{E}_n \right|. \quad (2)$$

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:

$$U = \sum_n \frac{\mu_n}{\sigma_n^2} U_n, \quad (3)$$

which is, under certain assumptions (Green & Swets, 1966), the least-squares estimator for the presence of a test signal.

The procedure that was used to determine the model predictions was similar to the procedure used with human observers. The model tries to determine the interval that contains the test signal. The level of the test signal was varied adaptively and the above process was repeated until a detection threshold was determined. Furthermore, the internal representations of the processed stimuli are used to update the estimates of $\bar{E}_n$, σ_n^2 and μ_n.

Experiments and model predictions

Maskers with phase transitions in the frequency domain

One way to determine the spectral acuity of the model is by using an NoπSπ stimulus con-

figuration (Kohlrausch, 1988). Here, a 500-ms broadband masker is used with a constant spectral energy density (No=41 dB/Hz), of which the interaural correlation of the masker is +1 below 500 Hz and -1 above 500 Hz. If an out-of-phase sinusoid (Sπ) is used as a test signal, the effective stimulus configuration is NoSπ for a sinusoid below 500 Hz (i.e., a BMLD is expected). Above 500 Hz, the condition corresponds to NπSπ and no BMLD should occur. On the other hand, if the interaural correlation relation of the masker is reversed, an NπoSπ condition is obtained, and the BMLD should occur for test frequencies above 500 Hz. Due to the limited spectral resolution of the binaural auditory system, a gradual change in the detection thresholds is expected for signal frequencies around 500 Hz. Figure 1 shows the experimental data with human subjects (white symbols, 4 subjects) as a function of the frequency of the test signal. The black symbols represent the model predictions. The left panel represents the NoπSπ condition, the right panel represents the NπoSπ condition. Clearly, the thresholds change gradually if the test signal frequency is increased from 300 to 600 Hz, indicating that the system has a limited frequency resolution. This resolution is well represented by the model.

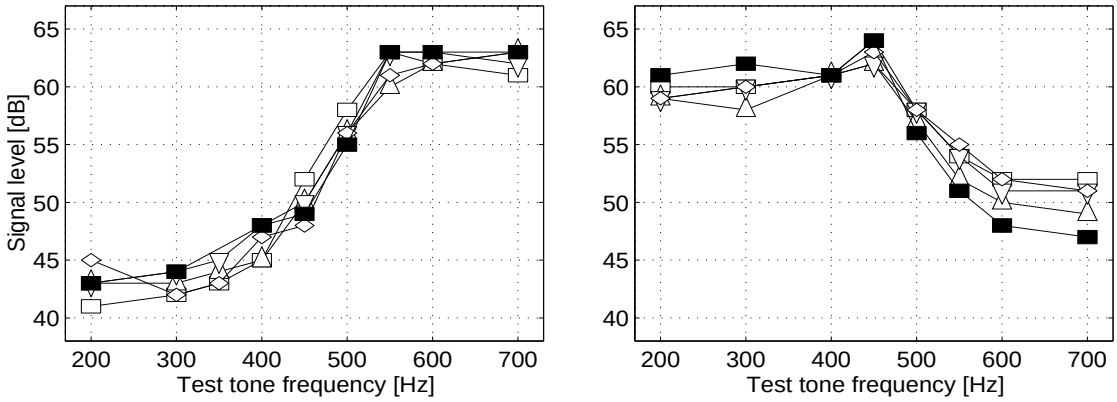

Figure 1: Experimental masked thresholds for four subjects (white symbols) and model predictions (black squares) for the NoπSπ condition (left panel) and for the NπoSπ condition (right panel) as a function of the frequency of the test signal. Experimental data adapted from Kohlrausch (1988).

Combined level and bandwidth dependence

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, Tyler and Fernandes (1983). We determined the detection threshold for a 500-Hz, out-of-phase sinusoid added to a band-limited in-phase noise masker (NoSπ) with a spectral 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 2 by the black symbols; the white symbols denote the experimental data from Hall, Tyler and Fernandes. 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 78 Hz (Glasberg & Moore, 1990). In this respect, the model predictions agree perfectly with the experimental data. However, there is a systematic difference between the model predic-

tions 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 difference does not reflect model artefacts, but is related to differences between data sets.

The right panel shows thresholds as a function of the notch bandwidth centred 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 threshold in wideband noise) as the measure of 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 that 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.

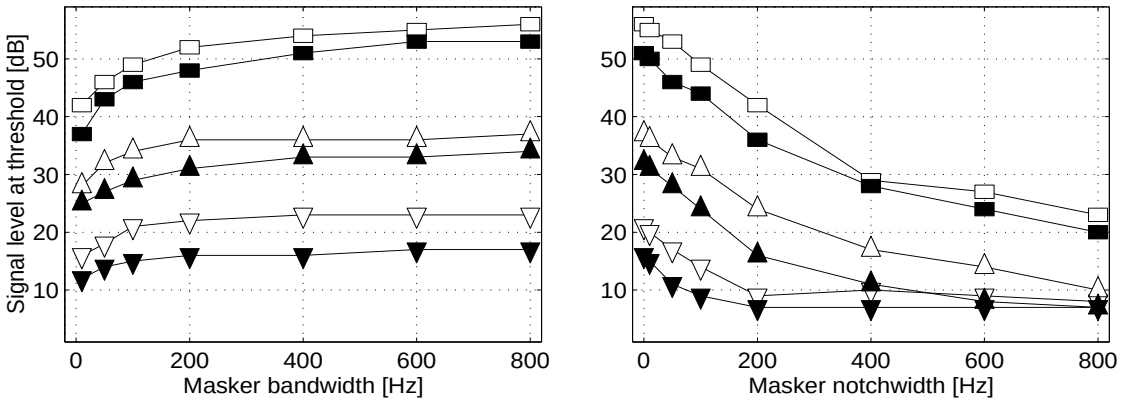

Figure 2: Detection thresholds for an NoSπ condition as a function of the bandwidth of a constant-spectral-level noise (left panel) and as a function of the bandwidth of a notch centred around the signal frequency (right panel). The white symbols represent experimental data adapted from Hall, Tyler and Fernandes (1983), the black symbols are model predictions. Legend: squares: No=50 dB/Hz; upward triangles: No=30 dB/Hz; downward triangles: No=10 dB/Hz.

Multicomponent test signal

In the experiments described above, the spectral integration of the model was determined by the spectral properties of the *masker*. The experiments of Langhans and Kohlrausch (1992) were also simulated to investigate the effect of spectral properties of the *test signal*. Here, the masker consists of a broadband in-phase noise (No=47 dB/Hz), while the test signal consists of an antiphasic harmonic complex with a flat amplitude spectrum, a spectral spacing between adjacent components of 10 Hz and a centre frequency of 400 Hz. The number of harmonics was varied to generate test signals of different bandwidths. The experimental data and model predictions are shown in Figure 3. The thresholds are expressed as a masking level per component. Clearly, for bandwidths beyond the critical band (>5 components), the thresholds still decrease with an increasing number of components for both experimental data and model predictions. Furthermore, the slope relating thresholds to the number of harmonics is very similar, indicating that both the model and the human observers use a similar across-frequency integration method. For this condition, however, the model predictions are 3 dB higher than the thresholds from human observers as a result of differences between data sets.

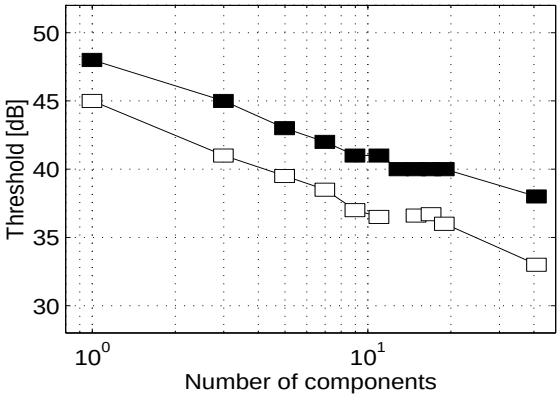

Figure 3: Masked threshold indicated in terms of level per component as a function of the number of components. White symbols represent experimental data from Langhans and Kohlrausch (1992), the black symbols are model predictions.

Discussion

The linear weighted addition of separate channel information reduces the internal error for a narrowband masker, due to the fact that the internal error variables ε_n are uncorrelated. With increasing masker bandwidth, the error in the decision variable increases due to masking of more and more off-frequency filters, disabling the internal error reduction through spectral integration. This process even continues for masker bandwidths that exceed the auditory filter bandwidth, resulting in a larger binaural effective filter bandwidth for band-widening conditions. These results suggest that the binaural system analyses more than one auditory filter if such a spectral integration enhances the detection performance. It also implies that the sensitivity to interaural cues for each auditory filter and the determination of binaural critical bandwidths can best be estimated from a wideband noise masking a narrowband signal, disabling the across-frequency integration.

References

Breebaart, J., Par, S. van de & Kohlrausch, A. (1997). Binaural signal detection with phase-shifted and time-delayed noise maskers. *Journal of the Acoustical Society of America*. In press.

Colburn, H.S. (1977). Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *Journal of the Acoustical Society of America, 64,* 95-106.

Dau, T., Püschel, D. & Kohlrausch, A. (1996). A quantitative model of the 'effective' signal processing in the auditory system. I. Model structure. *Journal of the Acoustical Society of America, 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. *Hearing Research, 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. *Journal of the Acoustical Society of America, 73,* 894-898

Joris, P.X. (1996). Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *Journal of Neurophysiology, 76,* 2137-2156.

Kohlrausch, A. (1988). Auditory filter shape derived from binaural masking experiments. *Journal of the Acoustical Society of America, 84,* 573-583.

Langhans, A. & Kohlrausch, A. (1992). Spectral integration of broadband signals in diotic and dichotic masking experiments. *Journal of the Acoustical Society of America, 91,* 317-326.

Lindemann, W. (1986). Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *Journal of the Acoustical Society of America, 80,* 1608-1622.

Stern, R.M. & Shear, G.D. (1996). Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay. *Journal of the Acoustical Society of America, 100,* 2278-2288.

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

Zurek, P.M. & Durlach, N.I. (1987). Masker-bandwidth dependence in homophasic and antiphasic tone detection. *Journal of the Acoustical Society of America, 81,* 459-463.