Appendix A

Zwicker's loudness model

In Fig. A.1 the four stages of the model are shown. Firstly, the transfer characteristics of outer and middle ear are taken into account. The effect of outer and middle ear can be considered as a frequency specific attenuation. Glasberg and Moore (1990) proposed a correction which is based on the assumtion that the transfer characteristics of outer and middle ear are reflected by the 100–phon equal loudness counter (ELC) rather than by the absolut threshold curve on which the MAF–correction is based (minimum audible field). Moore and Glasberg assume that the absolute threshold curve below 1 kHz is mainly determined by internal noise in the cochlea rather than by transfer characteristics of outer and middle ear. For frequencies above 1 kHz the ELC–correction and the MAF–correction become equivalent. However, since the stimuli employed in this study were selected from the spectral range between 1300 and 3000 Hz, it is not crucial which of both corrections is applied.

In the second step the excitation level per critical band is calculated. This could be done in either of two ways. They differ mainly in the way they calculate the excitation pattern across critical bands.

Firstly, in the way originally proposed by Zwicker and collegues (Zwicker, 1960; Zwicker and Scharf, 1965; Paulus and Zwicker, 1972; Zwicker and Fastl, 1990). Spread of excitation across critical bands is determined from masking patterns of pure tones masked by narrowband noises. Zwicker and collegues assumed that these masking pattern reflect the neural exitation evoked by such tones. If the slope of the skirts of the excitation pattern are plotted on a Bark[1] rather than a linear frequency scale, then they are approximately independent of the center frequency of the critical band under consideration. The slope of

[1]The Bark or the ERB scale both are scales which are more closely related to how sound is represented in the auditory system. Both scales assume that one unit, ERB or Bark, corresponds to a constant length along basilar membrane. While the Bark scale assumes this length to be 1.3 mm, it corresponds to 0.86 mm for the ERB scale. They yield about the same filter bandwidth for frequencies between 1 – 5 kHz, but differ markedly below and above. For a comparison of both scales see Fig. B.1, p. 96.

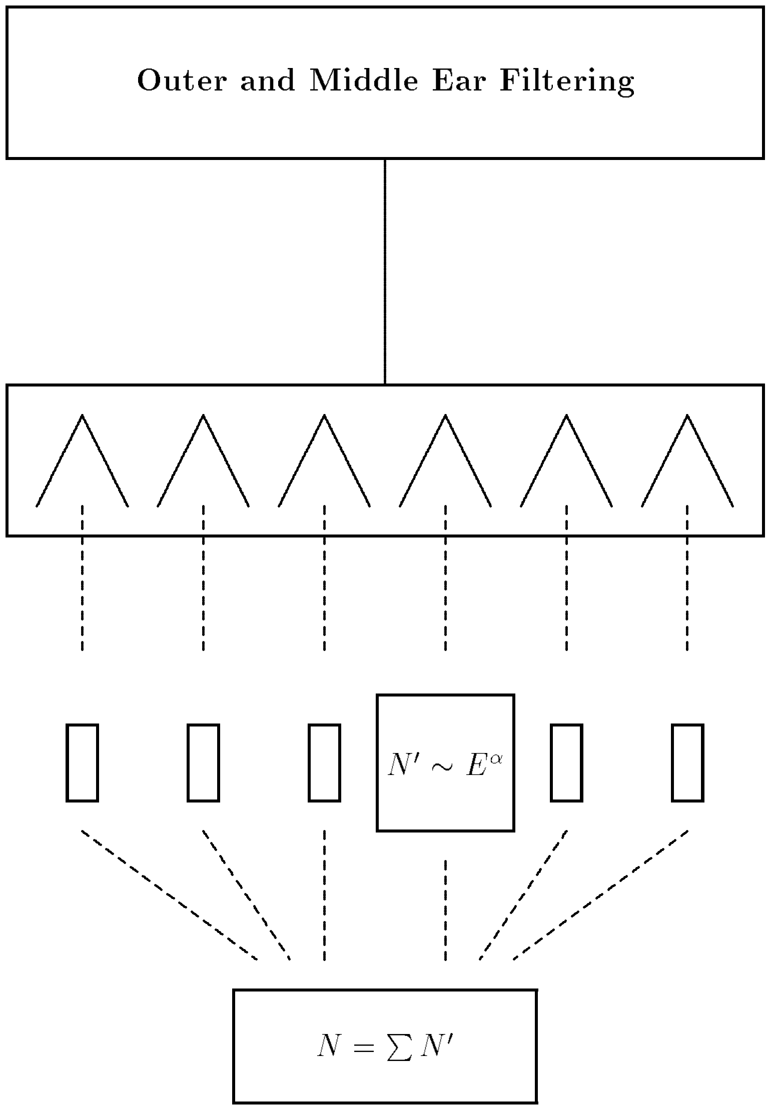

Fig. A.1: *Schematic representation of the different stages of Zwicker's model. A fixed filter representing transfer through outer and middle ear is followed by an auditory filter bank from which excitation patterns E are calculated. In the third step these excitation patterns are transformed into specific loudness N' by a power law relationship. Summing the specific loudness across bands yields the overall loudness N.*

these skirts depend on level being about 20 dB/Bark for the low frequency skirt and about 10 dB/Bark for the high frequency skirt.[2]

Masking patterns deduced from narrowband noise maskers may be influenced by several factors such as beat detection, combination products and off–frequency listening (Moore, 1993). In the second method excitation patterns are calculated from auditory filters. The auditory filter shapes have been determined using a notched–noise method (Moore, 1989; Moore, 1993). This method appears to be less influenced by the detection of beats or combination products. Auditory filters are thought to be symmetrical on a linear frequency scale. On a linear frequency scale their bandwidth increases, while on the ERB–scale their bandwidth is constant of 1 ERB. The shape of auditory filters is described using the empirical formula (Patterson et al., 1982):

$$W(g) = (1 + g \cdot p) \cdot exp(-p \cdot g) \quad (A.1)$$

where g represents the normalized deviation from the center of the filter and p the slopes of the filter skirts. This filter shape has been called "rounded exponential filter" or "ROEX filter".

Moore and Glasberg (Moore and Glasberg, 1983; Moore and Glasberg, 1987; Glasberg and Moore, 1990; Moore, 1993) have described a way of deriving excitation patterns using the concept of auditory filters.

Following Moore and Glasberg, excitation patterns represent the output of auditory filters as a function of their center frequency. The excitation evoked by a tone can then be constructed from the amount of energy falling in auditory filters with different center frequencies, as is shown in the upper panel of Fig. A.2, points a – e. This energy represents the excitation evoked by a tone at the center frequency of the respective filter (see lower panel of Fig. A.2). Thus, the level dependence of excitation patterns is determined by the level dependence of auditory filter shapes. Glasberg and Moore (1990) proposed the following empirical formulae for the relation between level L and the slopes of upper and lower frequency filter skirts, p_u and p_l respectively:

$$p_l(L) = p_l(51) - 0.38 \cdot \left(\frac{p_l(51)}{p_l(51, 1kHz)} (L - 51.0) \right) \quad (A.2)$$

$$p_u(L) = p_l(51) \quad (A.3)$$

$p_l(51)$ and $p_l(51, 1\,\text{kHz})$ are parameters accounting for the bandwidth of the respective filter.

There is still some controversy about whether or not the upper frequency skirt $p_u(L)$ also depends on level. However, this level dependence is not crucial for modeling hearing impairment since it yields only small changes in calculated loudness and can thus be neglected.

In this study, the latter method of calculating excitation patterns was chosen. It is much easier to broaden auditory filters using this approach, which could be important for

[2] For intermediate sound pressure levels, i.e., 60 dB SPL.

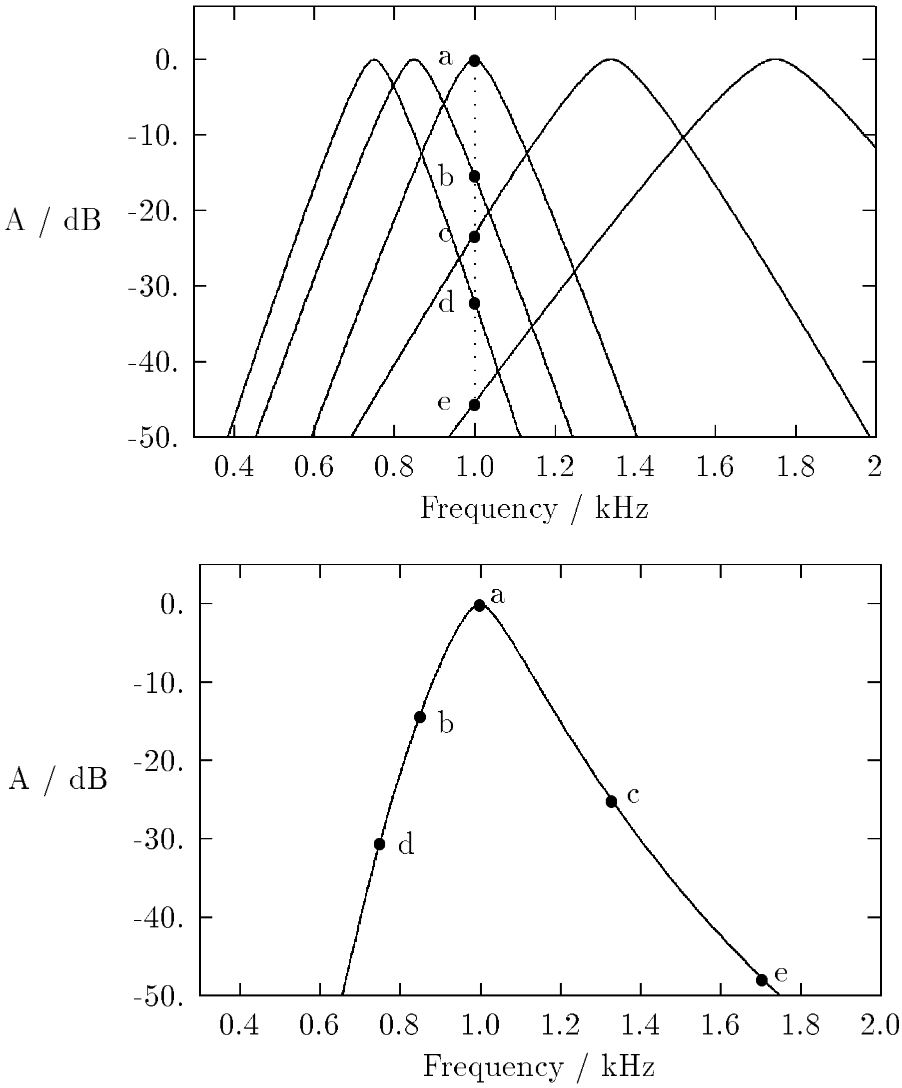

Fig. A.2: *Constructing excitation patterns from auditory filters. The upper panel shows the auditory filters and the lower panel shows the derived excitation patterns. The excitation patterns evoked by a 1kHz sinusoid can be derived by calculating the output of different auditory filters (a – e). It is assumed that the output of different auditory filters determine the excitation at the center frequency of the respective filter. Note that although the filters are assumed to be symmetrical on a linear frequency scale, the excitation patterns are not. This is due to the increase of filter bandwidth with increasing center frequency. Adapted from Moore and Glasberg (1987).*

modeling hearing impairment. Broadening of filters can simply be achieved by multiplying the above mentioned formulae A.3 and A.2 with a factor (which could also depend on frequency).

The third step in the loudness model is to calculate specific loudness N' from excitation E. Two different formulae have been proposed in the literature. Both account for threshold in quiet E_{ThQ} by an internal noise. Equation A.5 was proposed by Zwicker, while eq. A.4 was applied by Moore and Glasberg (1995)

$$N' = C \cdot \left(E^{\alpha} - E_{ThQ}^{\alpha}\right) \tag{A.4}$$

$$N' = C \cdot \left(\frac{E_0}{E_{ThQ}}\right)^{\alpha} \cdot \left(\left(0.5 + \frac{E}{E_{ThQ}}\right)^{\alpha} - 1\right) \tag{A.5}$$

For the exponent α a value of 0.23 is used in normal–hearing subjects. Thus, there is a nonlinear, compressive relationship between stimulus level and loudness. One might think of this power law as reflecting the nonlinear processing in the auditory system. Both attempts to account for threshold (thereby accounting for a steeper loudness function near threshold) yield similar results at medium to high levels. Thus, it is not crucial which of both is used for modeling loudness perception in hearing impaired listeners.

The overall loudness N is calculated from specific loudness by integrating specific loudness N' across frequencies on an ERB or Bark scale.

$$N = \int N'(z)dz \tag{A.6}$$

Thus, the total area under the specific loudness function is calculated. The physiological correlate of the summation of loudness across critical bands might be the summation of neural activity, i. e. loudness may be related to the total neural activity evoked by the sound, summed across the whole activity pattern. Thus, the area under the specific loudness function could be interpreted as approximately proportional to the total neural output evoked by the stimulus.

Appendix B

Comparison of Bark and ERB scale

Both the Bark or the ERB scale are scales which are more closely related to how sound is represented in the auditory system than is a linear frequency scale. Both scales assume that the bandwidth of an auditory filter ("critical band") corresponds to a constant length

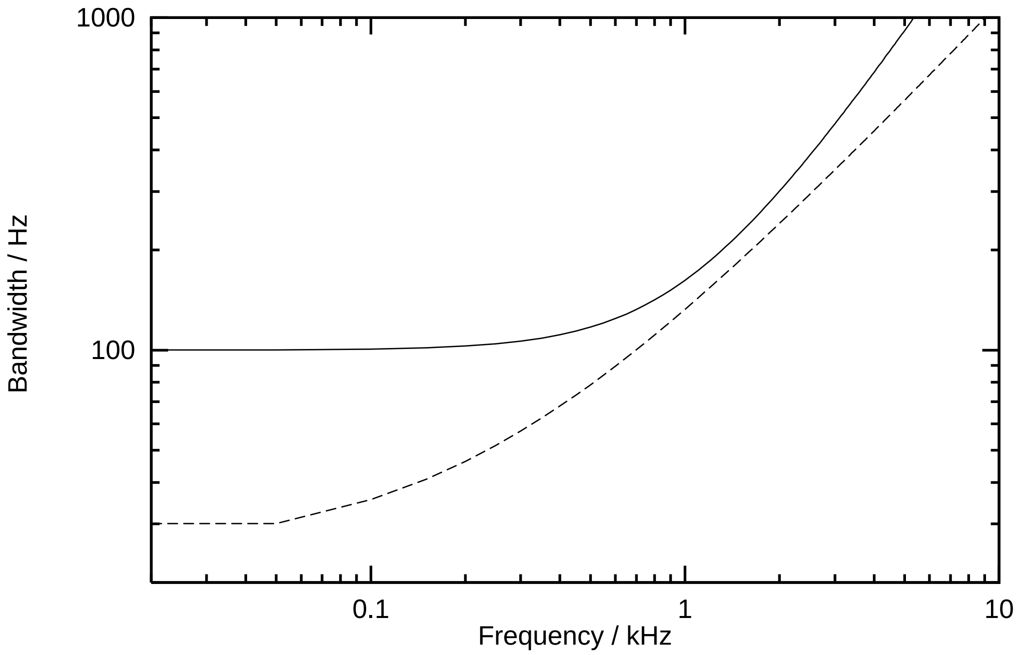

Fig. B.1: *Comparison of filter bandwidths obtained using the ERB (dashed) and the Bark (solid) scales. The filter bandwidth in units of Hz is plotted versus center frequency of the respective filter. Between 1 and 5 kHz both scales yield similar filter bandwidths, while they differ markedly at lower and higher frequencies.*

along the basilar membrane. While the Bark scale suggests this length to be 1.3 mm, it corresponds to 0.86 mm when using the ERB scale. Thus, on a linear frequency scale, the filter bandwidth increases with increasing center frequency for frequencies greater than a "cut–off" frequency. This cut–off frequency is assumed to be about 500 Hz for the Bark scale and about 100 Hz for the ERB scale. For frequencies smaller than this cut–off frequency, the filter bandwidth is constant on a linear scale. However, both scales yield about the same filter bandwidth for frequencies between 1 and 5 kHz, and differ markedly below and above. The following empirical equations have been proposed to calculate the auditory filter bandwidth Δf (Hz) as a function of frequency f (kHz):

$$\Delta f = 25 + 75 \cdot \left(1 + 1.4 \cdot f^2\right)^{0.69}, \tag{B.1}$$

for the Bark scale (Zwicker and Fastl, 1990), and

$$\Delta f = 24.7 \cdot (4.37 \cdot f + 1), \tag{B.2}$$

for the ERB scale (Moore, 1993). In Fig. B.1 the filter bandwidths calculated from these two equations are plotted against center frequency of the respective filter.

Appendix C

Audiometric data of the hearing–impaired subjects

Subject\Frequencies (kHz)	0.02	0.125	0.25	0.5	0.75	1.0	1.5	2.0	3.0	4.0	6.0	8.0	10.	12.0
AP	58.	58.	58.	58.	58.	58.	58.	58.	58.	58.	58.	58.	58.	58.
AW	5.	5.	5.	20.	25.	30.	55.	60.	62.	50.	60.	60.	60.	60.
CS	40.	40.	40.	40.	40.	50.	60.	60.	65.	60.	60.	40.	65.	65.
EH	40.	40.	40.	40.	40.	40.	43.	48.	55.	60.	70.	70.	70.	70.
HF	10.	10.	10.	10.	20.	40.	56.	60.	78.	78.	78.	78.	78.	78.
HK	15.	15.	15.	25.	30.	50.	50.	58.	58.	60.	60.	70.	70.	70.
JC	15.	15.	15.	15.	20.	35.	70.	83.	90.	90.	90.	95.	95.	95.
JKl	15.	15.	15.	30.	45.	55.	60.	70.	86.	90.	91.	92.	93.	94.
JKn	45.	45.	45.	49.	50.	50.	55.	69.	71.	72.	73.	74.	76.	78
MU	80.	80.	80.	80.	80.	70.	74.	80.	80.	80.	80.	80.	80.	80.
RB	25.	25.	25.	30.	30.	35.	48.	56.	62.	68.	74.	76.	78.	79.
RS	60.	60.	60.	60.	60.	52.	57.	60.	60.	60.	60.	60.	60.	60.
RW	35.	35	40.	45.	50.	55.	60	65.	70.	85.	85.	85.	85.	85.
UH	20.	20.	25.	35.	45.	55.	60.	65.	65.	65.	65.	65.	65.	65.

Table C.1: *Individual audiometric data (air conduction thresholds) of the 14 hearing–impaired listeners. The thresholds are taken from the loudness scaling experiments. Thresholds for frequencies below 250 Hz and above 6 kHz are estimates based on the shape of the pure tone audiogram.*

Appendix D

Instructions for the loudness scaling procedure

The purpose of this experiment is to determine how loud some noise bursts sound to you. After each stimulus presentation (a series of two short noise bursts) a categorical scale (Fig. D.1) with the 11 categories:

too loud – loud – intermediate – soft – very soft – inaudible

with intermediate values is presented to you on the display of this touch sensitive screen. Please rate the loudness of the sound by selecting one of these categories and touching the screen. All that matters is how loud the bursts sound to you. There are no “wrong” or “right” answers. The noise bursts will be presented in random order. The experiments take about 15 – 20 minutes.

Further informations and instructions such as the begining and the end of the experiment will be given to you on the display of the touch screen.

Do you have any questions?

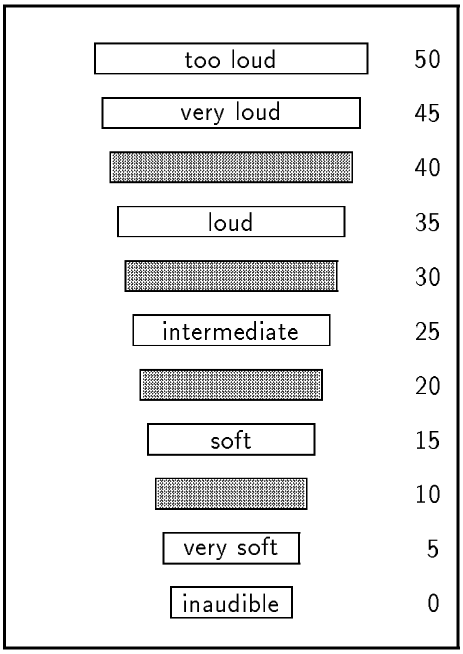

Fig. D.1: *Schematic plot of the different categories from which subjects had to select.*