Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Estimating individual listeners' auditory-filter bandwidth in simultaneous and non-simultaneous masking

Jörg M. Buchholz, Sabine Caminade, Olaf Strelcyk, and Torsten Dau

Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, Ørsteds Plads, 2800 Kgs. Lyngby, Denmark.

PACS: 43.66.DC, 43.66.BA

ABSTRACT

Frequency selectivity in the human auditory system is often measured using simultaneous masking of tones presented in notched noise. Based on such masking data, the equivalent rectangular bandwidth (ERB) of the auditory filters can be derived by applying the power spectrum model of masking and assuming a rounded-exponential filter shape. If a forward masking paradigm is used instead of simultaneous masking, filter estimates typically show significantly sharper tuning. This difference in frequency selectivity has commonly been related to spectral suppression mechanisms observed in the cochlea. Considering bandwidth estimates from previous studies based on forward masking, only average data across a number of subjects have been considered. The present study is concerned with bandwidth estimates in simultaneous and forward masking in individual normal-hearing subjects. In order to investigate the reliability of the individual estimates, a statistical resampling method is applied. It is demonstrated that a rather large set of experimental data is required to reliably estimate auditory filter bandwidth, particularly in the case of simultaneous masking. The poor overall reliability of the filter estimates was found to be mainly related to the very short tone duration (i.e., 10 ms) that was chosen. Applying 300-ms long tones in simultaneous masking drastically improved the reliability of the filter estimates. The tone duration in forward masking had to be very short to elicit a sufficient amount of masking. Based on extensive data for three subjects, the difference between forward and simultaneous masking estimates of auditory filter bandwidth was found to be even larger than previously reported, with a bandwidth decrease by a factor of about 1.8 rather than 1.4. The results of the study can be used to optimize the measures of frequency selectivity which is particularly useful when studying consequences of (individual) hearing impairment.

INTRODUCTION

Frequency selectivity in the human auditory system is often measured using simultaneous masking of tones presented in notched-noise. Based on such masking data, the equivalent rectangular bandwidth (ERB) of the auditory filters can be derived by applying the power spectrum model of masking and assuming a rounded-exponential (roex) filter shape (e.g., Patterson and Moore, 1986; Glasberg and Moore, 1990). If a forward masking paradigm is used instead of simultaneous masking, filter estimates typically show significantly sharper tuning than observed in simultaneous masking (e.g., Moore, and O'Loughlin, 1986; Oxenham and Shera, 2003; Oxenham and Simonson, 2006). This difference in frequency selectivity has typically been related to (instantaneously acting) suppression mechanisms on the basilar membrane, although the underlying mechanisms are still poorly understood. Evidence has been provided that frequency resolution measured in forward masking conditions are in better agreement with physiological data based on otoacoustic emissions (e.g., Shera et al., 2002; Pickles, 2008). This suggests that forward masking is more appropriate to derive auditory frequency resolution. However, Ruggero and Temchin (2005), for example, argued that physiologically measured filters are broader than those observed in forward masking. Despite this controversy, the comparison between auditory filter bandwidth measured in forward and simultaneous masking provides some measure of the amount of suppression in the cochlear, and thus might allow conclusions regarding the involved active mechanisms. This is of particular importance for hearing impaired listeners, because they often show a reduction in these active mechanisms due to a loss (or dysfunction) of outer hair-cells. Since every hearing impaired person has a unique hearing loss, auditory frequency resolution needs to be measured in individual subjects. A large number of studies exists that measured frequency resolution in individual hearing-impaired listeners using notched-noise simultaneous masking (e.g., see Moore, 2003, for an overview). However, studies that directly compared filter bandwidth in (notched-noise) forward and simultaneous masking solely considered masked thresholds averaged across a number of (normal-hearing) subjects (e.g., Oxenham and Shera, 2003; Oxenham and Simonson, 2006). The critical difference of these measurements to the standard simultaneous notched-noise measurements is that, in order to allow the comparison between forward and simultaneous masking data, very short test signals need to be applied (i.e., test signals with a duration of about 10 ms are used rather than several hundreds of ms). Using such short test signals to measure individual auditory filter shapes, it was found in a pilot study that, even for normal-hearing listeners, filter estimates can be unreliable.

In the present study, auditory filter bandwidth is estimated in individual normal-hearing subjects using notched-noise forward and simultaneous masking. In order to assess the accuracy of auditory filter estimates when applying very short test signal, a parametric bootstrap method was used (Efron and Tibshirani, 1993). This statistical resampling method allowed (quantitative) conclusions on the reliability of a given auditory filter estimate. The methods and results described in the present study are important for future comparisons of auditory filter shapes measured in (notched-noise) simultaneous and forward masking in individual hearing-impaired listeners. This in turn is important for a better understanding of suppression effects (or loss of suppression) in the impaired cochlear.

The present manuscript is organized as follows: First the experimental methods are described which were applied to measure thresholds for three normal-hearing subjects in notched-noise forward and simultaneous masking. Based on the resulting threshold data, individual auditory filters are derived. In order to investigate the influence of the number of notch-conditions as well as measurement repetitions on the filter estimates, a subset of the threshold data is analysed in addition to the complete data set. The reliability of the filter estimates is then analysed using the resampling method.

METHODS

Subjects

One female and two male listeners, aged between 21 and 25 years, served as subjects. They were all experienced listeners in notched-noise masking experiments and had audiometric thresholds within 15 dB hearing level. Given the difficulty of the task and the high level of concentration required, the listener's were asked to take regular breaks during the measurement sessions.

Stimuli

Masked thresholds were measured in notched-noise simultaneous and forward masking, mainly following the procedures described by Oxenham and Shera (2003). For both conditions, signal and masker were the same, only the temporal position of the signal was changed. An illustration of the temporal characteristics of the signal and the masker is given in Fig. 1. The signal was a 10-ms long 2-kHz tone, and the masker was a 400-ms long notched noise. Both were gated with 5-ms raised-cosine ramps, so that the signal had no steady-state portion. In the simultaneous-masking condition (Fig. 1a), the onset of the signal occurred 380 ms after the onset of the masker. In the forward-masking condition (Fig. 1b), the silent interval between masker offset and signal onset was 5 ms.

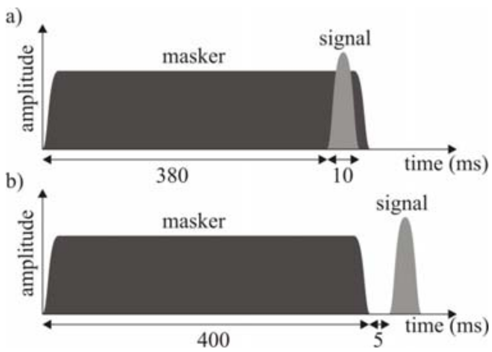

Fig. 1: Illustration of the temporal characteristics of the masking noise (dark grey) and the test tone (light grey).

Before the actual masked threshold measurements were carried out, the test signal's threshold in quiet was measured for each listener. For all subsequent measurements, the signal level was fixed at 10 dB above the individual threshold in quiet. The masker consisted of a white-noise band with a spectral notch around the signal frequency f_0 (see Fig. 2). Masked thresholds were measured by varying the masker level (see Rosen et al., 1998) for six symmetrical notch conditions ($\Delta f/f_0$: 0.0, 0.05, 0.1, 0.2, 0.3 and 0.4) and four asymmetrical notch conditions ($\Delta f/f_0$: 0.1|0.3, 0.3|0.1, 0.2|0.4, and 0.4|0.2), where Δf denotes the spacing between the inner noise edges and the signal frequency f_0. The outside edges of the noise masker were fixed at $\pm 0.8 f_0$.

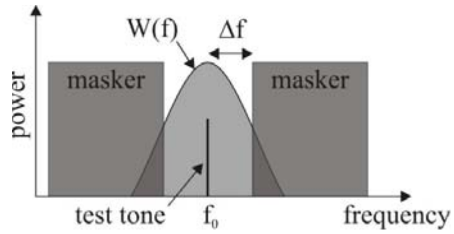

Fig. 2: Illustration of the power spectra of the notched-noise masker, the test tone, and the on-frequency auditory filter *W*(*f*).

Procedures

Masked thresholds were measured using a three-interval, three-alternative forced-choice (3I-3AFC) weighted up-down method (Kaernbach, 1991) tracking the 75% correct point on the psychometric function. Three intervals were successively presented, two of them only containing the masker, and one of them containing the masker and the test signal. The subject had to select the interval containing the test signal. The 10 notch conditions were randomly presented in a block, and a block was repeated six times. The listeners got visual feedback on the correctness of their response after each trial.

A run was terminated after 12 reversals. The mean value (i.e., the masked threshold) and the standard deviation (std) were determined using the masker level at the last 8 reversals. If the standard deviation exceeded 2.5 dB, the run was excluded and the threshold was measured again. When all the measurements were completed, the mean threshold and the standard deviation across runs were calculated. If the first measurement was 1.5 dB lower than any of the last three ones, this run was considered training and was re-measured. This operation was repeated until no training effect remained. A total of 26 of the 360 runs were discarded in this way. Moreover, thresholds which lay outside of a 90% normal-based confidence interval were considered as outliers and were also re-measured. The final thresholds were determined from at least 6 repetitions which were considered to be free from training effects and outliers.

Auditory filter derivation

Auditory filters were derived from the measured individual thresholds, and their estimated equivalent rectangular bandwidth (ERB) was analyzed using a resampling method. In order to investigate the effect of a limited data-set on the estimated bandwidth of auditory filters derived in notched-noise forward and simultaneous masking, the following two data-sets were considered:

- 10 notch conditions repeated 6 times (data-set: 10/6)
- 7 notch conditions repeated 3 times (data-set: 7/3)

The first (and complete) data-set was measured according to the procedures and criteria described above. For the second data-set the same criteria and stimulus conditions as applied in Oxenham and Shera (2003) were used, meaning that only some of the measured data were taken into account and previously excluded data were considered again. The first run of each condition was considered as training and was discarded. The standard deviations within track and across tracks were limited at 4 dB. One symmetric notch condition ($\Delta f/f_0$: 0.05) and two asymmetric notch conditions ($\Delta f/f_0$: 0.1|0.3, 0.3|0.1) were removed.

Auditory filter shapes were estimated from the individual notched-noise forward and simultaneous masking data by applying the methods described by Glasberg and Moore (1990), i.e. utilizing the power spectrum model of masking. Similar to Oxenham and Shera (2003), the shape of the auditory filters, $W(g)$ (see Fig. 2), was approximated by a variation of the rounded-exponential filter, roex(p_l,ω_l,t_l,p_u), as described by Patterson et al. (1982). The lower side of this filter (i.e., $f \le f_0$) is given by:

$$W(g) = (1-\omega)(1+p_l g)\exp(-p_l g) + \omega(1+p_l g/t)\exp(1+p_l g/t) \quad \text{(Eq. 1)}$$

where p_l, ω, and t are free parameters used for optimizing the fitted filter and g is the normalized deviation from the filter's centre frequency f_c and defined by $g = |f-f_c|/f_c$. Equation 1 defines a filter with two slopes, a steep one that mainly determines the shape of the filter's tip (defined by p_l) and a shallow one that realizes the filter's tail (defined by p_l/t). The knee-point between these two slopes is controlled by the parameter ω. The upper side of the roex(p_l,ω_l,t_l,p_u) filter (i.e., $f > f_0$) was realized by a single slope given by:

$$W(g) = (1+p_u g)\exp(-p_u g) \quad \text{(Eq. 2)}$$

A nonlinear minimization routine was implemented in MATLAB to find the best-fitting roex filter in the least-squares sense, assuming that the signal was detected using the filter with the maximum signal-to-noise ratio at its output. The shift of the filter's centre frequency f_c from the signal frequency f_0 (i.e., the amount of off-frequency listening) was limited to 10%. Finally, the ERB of the derived auditory filter was calculated.

Statistical analysis

Common procedures for estimating auditory filter shape take only the mean value of a limited set of masking data into account and thereby disregard that this mean value is only an estimate of a true threshold X_t. However, assuming that the masked thresholds are normally distributed around the true threshold X_t, the following holds: For a sample of n measured thresholds X, the ratio $(\overline{X}_n - X_t)/s_{\overline{X}}$ has a Student's t distribution with $n-1$ degrees of freedom, where $\overline{X}_n$ represents the mean of the measured thresholds and $s_{\overline{X}}$ the sample standard deviation divided by $\sqrt{n}$. In Fig. 3, the t-student distributions for $n = 2$, 3, and 6 measurements are shown together with the standard normal distribution. It can be seen that the t distribution for 3 measurements (dashed line) is much wider than a normal distribution (solid line), but for 6 measurements (dashed-dotted line) comes already very close to the normal distribution.

In order to determine the reliability of the estimated ERBs, bootstrap percentile intervals were calculated. The method consists of three steps:

1. For each notched-noise threshold, a bootstrap replica X* is generated: $X^* = \overline{X}_n + r_t \cdot s_{\overline{X}}$, with r_t randomly drawn from a t distribution. This results in a new simulated notched-noise threshold curve.
2. A filter is derived for this resampled threshold curve.
3. This process is iterated N times.

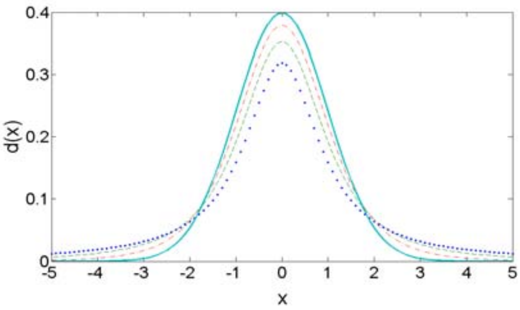

Fig. 3: T-student distribution for samples of 2 (dotted line), 3 (dashed line), and 6 measurements (dashed-dotted line), as well as the normal distribution (solid line), which represents an infinite number of measurements.

From this sample of N filters a distribution of the filter's parameter of interest was derived, which describes the uncertainly inherent in the filter parameter estimation. Here, only the filter's ERB distribution was considered. A number of 1500 replicas was found sufficient to obtain stable results. Filters with a root-mean-square (rms) error larger than 2.5 dB were discarded in order to reduce bias due to poor fits (resulting from unrealistic threshold curves).

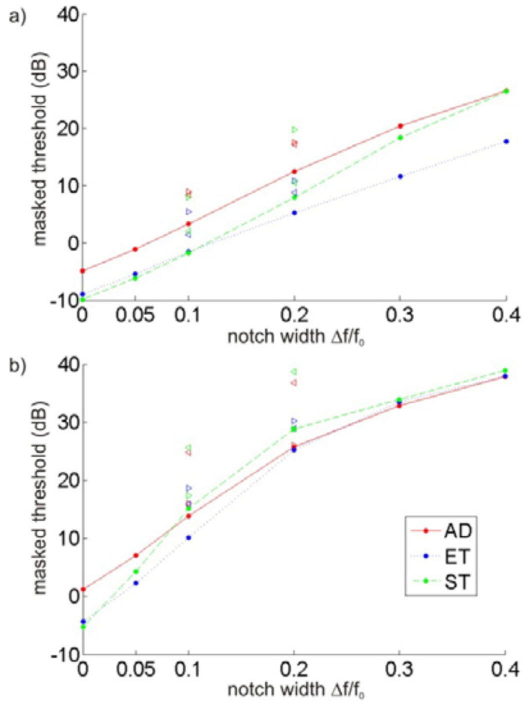

Fig. 4: Individual mean data for the simultaneous (panel a) and forward masking condition (panel b). Connected circles denote symmetric and triangles asymmetric notch conditions (left pointing triangle, $\Delta f/f_0$: 0.3|0.1, 0.4|0.2; right pointing triangle, $\Delta f/f_0$: 0.1|0.3, 0.2|0.4).

RESULTS

Experimental data

Figure 4 shows the individual thresholds in simultaneous masking (panel a) and in forward masking (panel b). Thresholds are represented by circles for the six symmetric conditions and by triangles for the four asymmetric conditions. For both simultaneous and forward masking, the masked threshold increases with increasing symmetrical notch width. Comparing the data in Figure panels 4a and 4b, two general trends can be observed. First, the increase in threshold with increasing notch width is steeper in forward than in simultaneous masking. Second, the values for the four asymmetric notch conditions are closer to each other in simultaneous than in forward masking. The fist trend indicates better frequency selectivity in forward than in simultaneous masking, and the second trend indicates a more asymmetric auditory filter shape in forward masking.

Table 1: Forward masking ERB data.

		ERB (Hz)	rms	med (Hz)	5% (Hz)	95% (Hz)	width (Hz)
AD	10/6	188	1.1	186	153	215	62
	7/3	177	0.3	174	120	242	122
ET	10/6	176	1.1	175	158	189	31
	7/3	141	0.7	139	114	193	79
ST	10/6	126	1.0	125	115	137	22
	7/3	134	0.6	130	108	153	45

Derived auditory filters

Table 1 and 2 provide a comparison between the derived ERBs of the auditory filters estimated from the mean forward and simultaneous masking thresholds measured for the two data-sets (the small and complete data-sets 7/3 and 10/6, respectively). In forward masking (table 1), both data-sets result in a similar mean ERB value of around 160 Hz across subjects. Hence, the small data-set would have been sufficient for deriving auditory filter bandwidth. In simultaneous masking (table 2), the complete data-set results in a mean ERB value of about 290 Hz. For the small data-set, the ERB values vary significantly around a mean value of 216 Hz. This large difference in ERB values between the two data-sets indicates that, in contrast to the forward masking case, the small data-set is not sufficient for estimating the individual auditory filter's bandwidth in the case of simultaneous masking. However, for both data-sets, the filter estimates are broader in simultaneous masking than in forward masking and which supports the conclusion of Oxenham and Shera (2003) that frequency selectivity observed in forward masking is higher than in simultaneous masking. The ERB ratio between filter estimates in simultaneous and forward masking is 1.8 for the complete data-set and 1.4 for the small data-set.

Table 2: Simultaneous masking ERB data.

		ERB (Hz)	rms	med (Hz)	5% (Hz)	95% (Hz)	width (Hz)
AD	10/6	290	0.5	290	252	319	67
	7/3	250	1.1	261	202	360	158
ET	10/6	293	1.9	298	261	360	99
	7/3	207	0.9	214	170	294	124
ST	10/6	284	1.3	284	273	294	21
	7/3	191	0.8	209	163	309	146

Table 1 and 2 also present the root-mean-square (rms) error of the filter fit. The overall rms error for both data-sets and all subjects was larger in simultaneous masking (1.1 dB) than in forward masking (0.8 dB).

Statistical analysis

The ERB distributions that resulted from the bootstrap analysis are shown in Fig. 5 for all subjects, the two data-sets, and both masking conditions. These ERB distributions do not follow a normal distribution, which was confirmed by Jarque-Bera tests (Jarque and Bera, 1987). The median values, the lower (5%) and upper (95%) percentiles, as well as the distance between the two percentiles (as a measure of the width of the distributions) were calculated from the derived ERB distributions and are presented in table 1 for the forward masking condition, and in table 2, for the simultaneous masking condition. Generally, the ERB distributions are narrower (by a factor of about 1.5) in forward masking than in simultaneous masking. Moreover, the complete data-set leads to narrower ERB distributions (with a mean width of 38 Hz in forward masking and 62 Hz in simultaneous masking, averaged across all subjects), than for the small data-set (with a mean width of 82 Hz and 134 Hz, respectively).

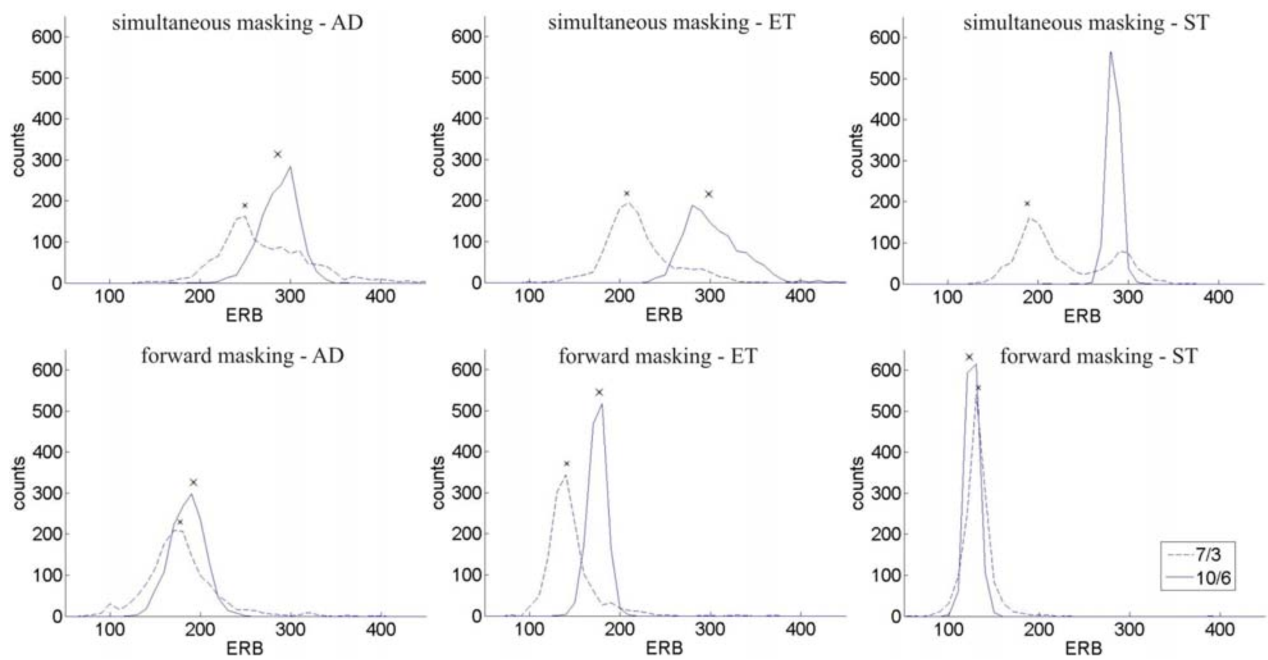

Fig. 5: Distribution of estimated individual auditory bandwidth (in ERBs) resulting from the statistical resampling method for the three subjects AD, ET, and ST for simultaneous (upper row) and forward masking (lower row). The complete data-set (condition 10/6) is indicated by the solid lines and the small data-set (7/3 condition) by the dashed lines. The small and large crosses indicate the individual ERB estimates resulting from actually measured thresholds for the 7/3 condition and 10/6 condition, respectively.

Hence, the filter bandwidth estimates are more reliable in the forward masking case and the complete data-set than in the simultaneous masking case and the small data-set, respectively.

The median values of the ERB distributions are consistent with the ERB values derived from the measured notched-noise thresholds (see table 1 and 2). However, when analyzing the ERB distributions derived from the small data-set in simultaneous masking, either a two-peak distribution (subject ST) or considerably skewed one-peak distributions (subjects AD and ET) were found. In the forward masking condition, such two-peak or skewed behavior was not observed for subjects AD and ST, and a slight skewness was observed for subject ET. Interestingly, those ERB estimates that differ significantly between the small and complete data-set also show skewed (or two-peak) ERB distributions. This observation might suggest that the proposed statistical analysis, combined with a skewness calculation, can be used as a tool for evaluating the reliability (or accuracy) of an auditory filter (bandwidth) estimate. Thereby, the skewness of the ERB distributions might be calculated by the third standardized moment, but further research is needed to determine the relationship between skewness and accuracy (or reliability) of the filter estimate.

DICUSSION

Simultaneous versus forward masking

Auditory filter estimates were derived using short tones in simultaneous and forward masking experiments and their reliability was assessed using a statistical resampling method. Oxenham and Shera (2003) observed that on average, auditory filters measured in forward masking were about 1.4 times narrower than when measured in simultaneous masking. Applying the same experimental procedures as used by Oxenham and Shera (i.e., using 7 notch conditions and 3 repetitions), roughly the same ratio between filter bandwidths was observed in the present study. However, for filters derived by using 10 notch conditions with more measurement repetitions, the estimated bandwidths were found to be about 1.8 times narrower in forward masking than in simultaneous masking. As these later filters are more precise than the previous ones (see Fig. 5), the factor of 1.8 probably represents a better estimate of the difference in frequency resolution measured in notched-noise forward and simultaneous masking (at least at 2 kHz).

Additionally, despite the fact that the forward masking thresholds generally exhibited a larger standard deviation (1.8 dB on average) than the simultaneous masking thresholds (1.4 dB), it was generally observed that ERBs are better defined in forward masking than in simultaneous masking. This observation was based on: (i) the results for the two data-sets, which provide similar ERB estimates in the 10/6 and 7/3 conditions and (ii) the statistical analysis results showing narrower ERB distributions for the forward masking condition (Fig. 5).

The larger overall rms fitting errors for the simultaneous masking condition than for the forward masking condition may indicate that the applied roex filter model is less successful in accounting for masked thresholds obtained for short signals in simultaneous masking.

Signal duration

Since difficulties with auditory filter estimation in simultaneous masking conditions with long test signals have not been reported in the literature (e.g., Glasberg and Moore, 1990), auditory filters were also determined using a long test signal. In this additional experiment, the masker was the same as for the short signal measurements (see methods section), but the test signal was 300 ms long and started 50 ms after the masker onset. The resulting experimental data and the corresponding ERB estimates and distributions can be found in Table 3. Comparing the width of the ERB distributions for the short signal (table 2) and for the long signal (table 3) shows that ERBs are significantly better defined when using long test signals. Moreover, since the ERB estimates derived for the 10/6 and 7/3 conditions are very similar, a small number of measurements seems to be sufficient to reliably estimate filter bandwidth for long test tones. Hence, the difficulty in deriving auditory filters in simultaneous masking can be mainly attributed to the short test signal duration. This observation might suggest that, when comparing auditory filter bandwidth in forward and simultaneous masking, the difficulties observed in simultaneous masking with the short test signal could be solved by applying a long test signal. Unfortunately, this is problematic because the filter bandwidths derived for the long test tones (table 3: about 230 Hz for the 10/6 condition) are slightly narrower than for the short test tones (table 2: about 290 Hz for the 10/6 condition). Further research is required to understand why this difference in auditory filter bandwidth exists between short and long test signals.

Table 3: Simultaneous masking ERB data for a 300-ms long test tone.

		ERB (Hz)	rms	med (Hz)	5% (Hz)	95% (Hz)	width (Hz)
AD	10/6	258	1.0	258	245	271	26
	7/3	267	0.7	267	233	308	75
ET	10/6	250	1.2	250	239	266	27
	7/3	275	0.7	272	239	294	55
ST	10/6	188	0.6	188	171	206	35
	7/3	165	0.2	163	111	230	119

Filter model

Throughout literature a large number of different variations of the roex filter have been proposed and no optimal general solution has been identified. Hence, in order not to limit the present findings to one specific filter model, an additional roex filter version was considered. Similar to Oxenham and Shera (2003), the additional filter model was the roex(p_l,p_u,ω,t) filter proposed by Glasberg et al. (1984). In contrast to the roex(p_l,ω_l,t_l,p_u) filter used throughout the previous sections, the roex(p_l,p_u,ω,t) filter employs two slopes at either side of the filter, and both are described by Eq. 1. In order to limit the number of free parameters (i.e., 4), the parameters ω and p are assumed to be the same for both sides, whereas the parameter p can differ between the lower side (p_l) and the upper side (p_u).

When the roex(p_l,p_u,ω,t) model is applied to derive auditory filter bandwidth, basically the same results were found as given in table 1 and 2 for the roex(p_l,ω_l,t_l,p_u) model. The only difference was observed in simultaneous masking for the short signal, where the roex(p_l,p_u,ω,t) model resulted in consistently narrower auditory filters than the roex(p_l,ω_l,t_l,p_u) model (i.e., mean ERB across listeners is 230 Hz rather than 290 Hz), while no consistent differences were observed in terms of goodness of fit. Hence, for the simultaneous masking condition with a short signal, the two roex filters are not equivalent, which is in contrast to Oxenham and Shera (2003) who found no significant differences in ERB estimates for these two roex filters. The difference in ERB estimate is here mainly due to the filter slopes close to the center frequency, where the slope for notch widths of $0f_c$ to $0.1f_c$ is steeper for the roex(p_l,p_u,ω,t) filter. This difference in slope can be seen in Fig 6, where the simultaneous masking data for the short signal are shown exemplarily for subject ET together with the two corresponding model predictions. Hence, the estimated ERB is overly sensitive to the measurement points very close to the filter tip, which might suggest that the roex family is not optimal for modeling auditory filters for short signals in simultaneous masking.

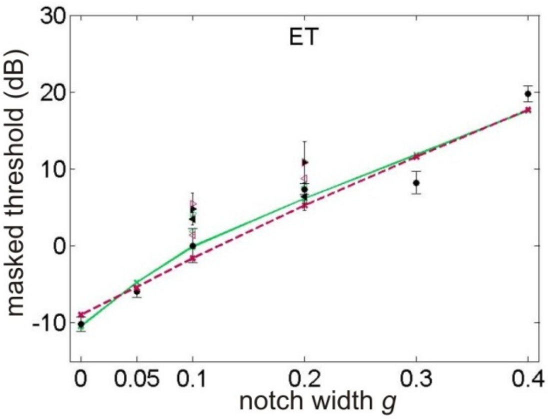

Fig. 6: Measured masked thresholds and corresponding model predictions for the roex(p_l,p_u,ω,t) (solid line) and roex(p_l,ω_l,t_l,p_u) filters (dashed line) for subject ET in simultaneous masking.

Data-sets and repetitions

In table 2 it is shown that, in simultaneous masking, the individual filters derived from the small data-sets are narrower than the ones derived from the complete sets. When reorganizing the existing complete data set to derive data sets of 7 notch conditions and 6 repetitions (7/6 condition) as well as 10 notch conditions and 3 repetitions (10/3 condition), it was found that the narrower ERB is mainly influenced by the additional notch conditions and not the number of repetitions. Since these additional notch conditions are all very close to the filter tip, it seems to be important that the filter tip is well defined by an appropriate number of measurement points around the tip to accurately derive auditory filter bandwidth in terms of the ERB. Furthermore, if the entire shape of the filter is of interest even 10 conditions are probably not enough for such very short test signals.

The statistical analysis (table 1-3) showed that, for the same number of notch conditions, the ERB distribution is much narrower for 6 than for 3 repetitions This improvement is due to the reduced variability of the threshold estimates (both in terms of a smaller standard error of the mean and lower tails of the t distribution).

In conclusion, it is necessary to ensure that (i) enough repetitions are made to precisely define the masked thresholds and (ii) enough conditions are measured to determine accurately the shape of the filter, in particular around the tip.

Interpretation

While the width of the ERB distribution provides a rough estimate of the accuracy of the derived ERB values, a closer study of the shape of the distributions provides important further information. For instance, the ERB distribution of subject ST in simultaneous masking (Fig. 5) shows a single narrow peak for the complete data-set, and two peaks for the small data-set. The main peak is located where the ERB of the filter estimated from the mean thresholds lies while the second peak is located at the same ERB as the one predicted by the complete data-set. The second peak is another possible filter width given by the data, and since it is consistent with the ERB of the filter derived from the complete set, it may be interpreted as the “correct” ERB. Thus, the statistical analysis provides a tool for judging the reliability of an ERB estimate. The statistical analysis should be of particular importance when unrealistic ERB values are found, such as in Oxenham and Shera (2003), where an ERB value of 10 Hz was observed at 2 kHz. In this case, the proposed statistical analysis might have revealed another peak with a more realistic value.

As suggested above, the implemented statistical analysis provides much more information about the reliability of a filter estimate than the simple rms value used throughout the literature. It takes into account the error introduced by the measurement variability, and is influenced by the number of notch conditions and repetitions as well as the goodness of the fit between measured and predicted thresholds. Further experiments are required to develop a quantitative method that allows the “online” control of psychoacoustical experiments to derive auditory filter estimates with a predefined accuracy.

SUMMARY AND CONCLUSIONS

Individual auditory filter bandwidths were estimated using a notched-noise paradigm in simultaneous and forward masking. The influence of the number of notch conditions and repetitions of threshold measurements on the accuracy of the filter estimates was investigated using a statistical resampling method. It was found that the uncertainty of bandwidth estimates using a 10-ms long test signal was significantly larger for filters derived in simultaneous masking than in forward masking. This increased uncertainty in simultaneous masking was mainly due to the short duration of the signal and was not observed when the signal duration was increased to 300 ms (as typically employed in auditory filter estimation). Moreover, it was shown that the accuracy of filter bandwidth estimates depends not only on the standard deviation of the thresholds but also on the number of repetitions of each condition as well as the number of notch conditions. In view of these results, it is suggested that three repetitions should be considered in simultaneous masking with long signals as well as in forward masking. This number of repetitions seems to be insufficient for simultaneous masking measurements with short test signals, where at least six repetitions are required to reliably determine auditory filter bandwidth. Moreover, it seems to be particularly important to include conditions with very narrow spectral notches to better define the tip of the auditory filter. However, the exact number of measurements depends on the specific signal configuration, individual variability as well as on the required accuracy of the estimates. Here, resampling methods provide a useful tool for assessing the reliability of filter estimates and thus, may help to efficiently estimate filter properties with predefined accuracy. Furthermore, the methods and results described in the present study may be of particular importance for estimation of auditory filter shapes using short test signals in hearing-impaired individuals. Resampling could help delineating if a larger-than-normal across-subject variability was due to increased measurement error or genuinely larger variability of auditory filter shapes in the hearing impaired.

REFERENCES

[1] Efron, B., and Tibshirani, R. J. (1993), An Introduction to the Bootstrap. New York: Chapman and Hall.

[2] Glasberg, B. R, Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I. (1984). “Dynamic range and asymmetry of the auditory filter.” J. Acoust. Soc. Am. **76**, 419-427.

[3] Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data." Hearing Research **47**, 103-138.
[4] Jarque, C. M., and Bera, A. K. (1987). "A test for normality of observations and regression residuals." International Statistical Review **55**, 1-10.
[5] Kaernbach, C. (1991). "Simple adaptive testing with the weighted up-down method." Percept. Psychophys. **49**, 227–229.
[6] Moore, B. C. J. and O'Loughlin, B. J. (1986). "The use of nonsimultaneous masking to measure frequency selectivity and suppression," in "Frequency selectivity in hearing", edited by B. C. J. Moore, Academic Press.
[7] Moore, B. C. J. (2003). "An introduction to the psychology of hearing." Fifth edition, Academic Press.
[8] Oxenham, A. J., and Shera, C. (2003). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking." JARO **4**, 541-554.
[9] Oxenham, A. J., and Simonson, A. M. (2006). "Level dependence of auditory filters in nonsimultaneous masking as a function of frequency." J. Acoust. Soc. Am. **119**, 445-453.
[10] Patterson; R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982), "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold." J. Acoust. Soc. Am. **72**, 1788--1803
[11] Patterson, R. D. and Moore, B. C. J. (1986). "Auditory filters and excitation patterns as representation of frequency resolution," in "Frequency selectivity in hearing", edited by B. C. J. Moore, Academic Press.
[12] Patterson, R. D., Holdsworth, J., Nimmo-Smith, I., and Rice, P. (1988). "Svos final report: The auditory filterbank." Tech. Rep. APU report 2341.
[13] Pickles, J. O. (2008). "An Introduction to the Physiology of Hearing," third edition, Emerald.
[14] Rosen, S., Baker, R. J., and Darling, A. (1998). "Auditory filter nonlinearity at 2 kHz in normal hearing listeners." J. Acoust. Soc. Am. **103**, 2539-2550.
[15] Ruggero, M. A. and Temchin, A. N. (2005). "Unexceptional sharpness of frequency tuning in the human cochlea." PNAS **102**, 18614-18619.
[16] Shera, C. A., Guian Jr., J. J., and Oxenham, A. J. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements." Proc. Natl. Acad. Sci. **99**, 3318-3323.