Biol Cybern (2007) 97:397–411
DOI 10.1007/s00422-007-0179-8

ORIGINAL PAPER

Biological Cybernetics

# The role of the auditory periphery in comodulation detection difference and comodulation masking release

**Michael Buschermöhle · Jesko L. Verhey · Ulrike Feudel · Jan A. Freund**

Received: 22 December 2006 / Accepted: 30 August 2007 / Published online: 9 October 2007


**Abstract** Natural sounds often exhibit correlated amplitude modulations at different frequency regions, so-called comodulation. Therefore, the ear might be especially adapted to these kinds of sounds. Two effects have been related to the sensitivity of the auditory system to common modulations across frequency: comodulation detection difference (CDD) and comodulation masking release (CMR). Research on these effects has been done on the psychophysical and on the neurophysiological level in humans and other animals. Until now, models have focused only on one of the effects. In the present study, a simple model based on data from neuronal recordings obtained during CDD experiments with starlings is discussed. This model demonstrates that simple peripheral processing in the ear can go a substantial way to explaining psychophysical signal detection thresholds in response to CDD and CMR stimuli. Moreover, it is largely analytically tractable. The model is based on peripheral processing and incorporates the basic steps frequency filtering, envelope extraction, and compression. Signal detection is performed based on changes in the mean compressed envelope of the filtered stimulus. Comparing the results of the model with data from the literature, the scope of this unifying approach to CDD and CMR is discussed.

## 1 Introduction

One of the prominent aspects of auditory scene analysis is the formation and discrimination of auditory objects (Griffiths and Warren 2004). Having identified an object, listeners can follow it even under adverse circumstances. Object formation may be guided by many different cues, one of which is amplitude modulation. It has been shown that the detectability of signals can depend on the correlation structure of modulations in an auditory stimulus (e.g., Nelken et al. 1999; Singh and Theunissen 2003). As amplitude modulations due to sound production and sound propagation occur frequently in nature, this cue might be especially important for animals and humans (e.g., Langemann and Klump 2001; Hofer and Klump 2003; Verhey et al. 2003).

Several experiments in auditory object formation are concerned with correlated amplitude modulations and their effect on signal detection. Two such experimental paradigms are comodulation detection difference (CDD) and comodulation masking release (CMR, see Table 1 for a summary of all abbreviations used in the paper). In the first of these paradigms (e.g., Cohen and Schubert 1987; Fantini and Moore 1994; Wright 1990; Hall et al. 2006), the signal to be detected is a narrow noise band (signal band) in the frequency domain masked by one or several additional noise bands (flanking bands). If the envelope of the signal band fluctuates in the same way as that of the flanking bands, detection thresholds are found to be higher than if the flanking bands share a common envelope while the signal band has a differing envelope. The difference in these thresholds is called the comodulation detection difference (CDD). For the case of more than one flanking band, three main correlation conditions may be distinguished: (i) all correlated (AC), if all envelopes are the same, (ii) all uncorrelated (AU) for mutually different envelopes, and (iii) co-uncorrelated (CU) if the flanking band envelopes are the same while the signal band has a different envelope (see Fig. 1, top row). Generally, signal detection thresholds in CDD experiments are lowest in the CU condition. With this nomenclature, CDD is often defined by subtracting the AC threshold from the CU

M. Buschermöhle (✉) · J. L. Verhey · U. Feudel · J. A. Freund
International Graduate School for Neurosensory Sciences,
Carl-von-Ossietzky Universität Oldenburg,
26111 Oldenburg, Germany
e-mail: buschermoehle@icbm.de

**Table 1** Abbreviations used in the paper

| Abbreviation | Meaning |
|---|---|
| AC | All correlated |
| AU | All uncorrelated |
| CDD | Comodulation detection difference |
| CMR | Comodulation masking release |
| CU | Co-uncorrelated |
| ERB | Equivalent rectangular bandwidth |
| FB | Flanking band |
| OFM | On-frequency masker |
| SB | Signal band |
| SPL | Sound pressure level |

threshold, and therefore it is usually negative (McFadden 1987).

Stimuli commonly used in CMR experiments consist of a narrow noise band (the on-frequency masker, OFM) masking a pure tone signal and one or several flanking noise bands serving as additional maskers. If the flanking bands are modulated in the same way as the OFM (which will be referred to as the AC condition according to the above terminology), then signal detection thresholds are lower than if flanking bands and OFM have differing amplitude modulations (in the following referred to as the AU condition). This threshold difference has been termed comodulation masking release (CMR, Hall et al. 1984). Example stimuli for this kind of CMR experiments are sketched in the bottom row of Fig. 1. A second kind of CMR experiments exists in which there is only one on-frequency masker centered on the signal sinusoid. Common amplitude fluctuations in this case may be obtained by modulating the whole on-frequency masker with a lowpass noise. This condition can also be called the AC condition, while the unmodulated case can be termed the AU condition. With increasing OFM bandwidth, threshold differences between AC and AU condition tend to increase in these experiments, even if the masker bandwidth exceeds the width of a typical auditory filter. This finding has led to the hypothesis that across-channel processes may underlie the CMR effect (Hall et al. 1984), however, see Verhey et al. (1999). In the present work, CMR is defined as threshold difference between the AU and AC conditions (AU–AC, usually positive).

Several mechanisms have been hypothesized to account for CMR (see Verhey et al. 2003 for a review) and CDD (see Moore and Borrill 2002 and references therein for further information). Generally, the explanations can be divided into within-channel and across-channel accounts of the observed effects. This refers to the number of auditory channels involved in the processing of the stimuli. (The inner ear can be

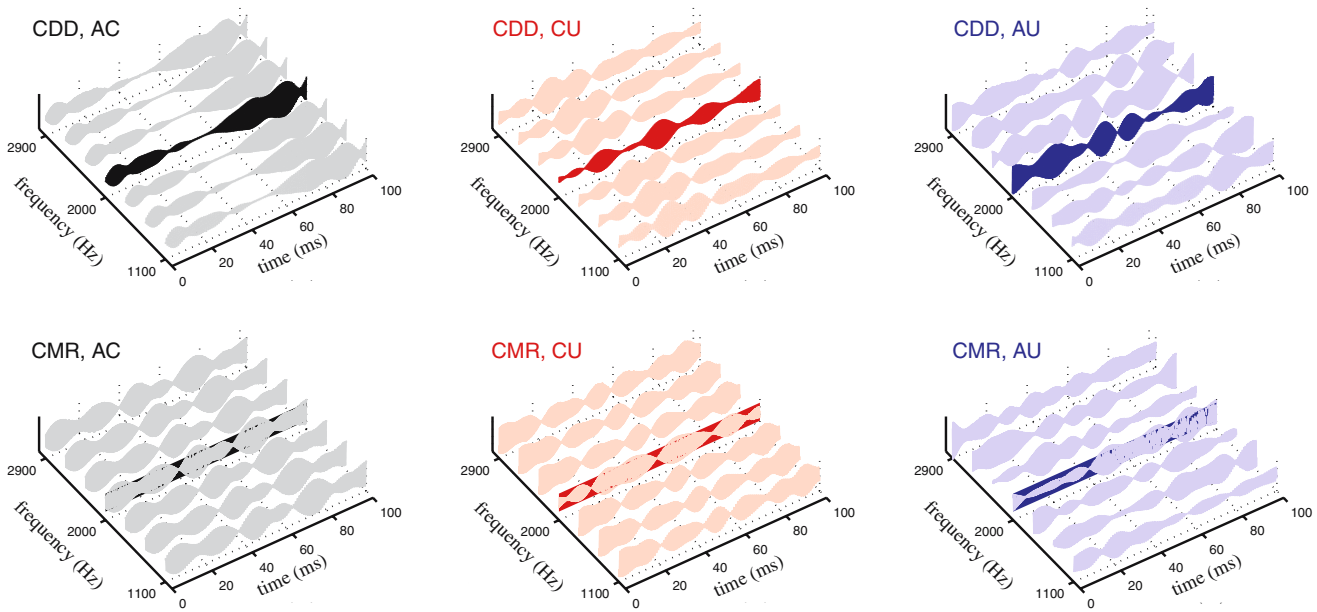


**Fig. 1** Schematic illustration of typical correlation conditions for the noise bands used in CDD (*top row*) and CMR (*bottom row*) experiments. The maskers are *light-colored* while the signal is *dark colored*. In the CDD case, the masker consists only of flanking bands (*FB*). In the CMR case, the masker consists of the FBs and an additional on-frequency masker (*OFM*). The signal in the CMR case is a sine tone with a flat envelope, which is illustrated here by a *dark-colored bar* which is partly covered by the OFM. *Left* FBs and SB/OFM have correlated amplitude modulations (all correlated, AC). *Middle* FBs are comodulated, SB/OFM has uncorrelated amplitude fluctuations (co-uncorrelated, CU). *Right* amplitude fluctuations of FBs and SB/OFM are uncorrelated (all uncorrelated, AU). (Color online)

viewed as a filter bank that analyzes sounds according to a row of auditory filters or channels.)

The across-channel explanation for the psychophysical CMR experiments is that the auditory system compares the output of a channel that is centered on the signal to those centered on the flanking bands. Model realizations of such across-channel comparisons may be (i) correlation models, in which the outputs of several channels are cross-correlated with one another, (ii) equalization–cancellation models, in which the outputs of several channels are first equalized in overall level and then subtracted from each other, or (iii) dip listening models, in which the auditory system is assumed to be able to detect times of low amplitudes of the flanking band channels in order to improve the signal-to-noise ratio.

Dip listening, however, does not necessarily have to rely on across-channel processing. Mechanical suppression on the level of the cochlea and other mechanisms have been proposed to allow for dip listening also within one auditory channel (see Moore and Borrill 2002; Ernst and Verhey 2006). Another possibility of a within-channel explanation is that changes in the temporal waveforms may be registered by the auditory system and lead to different signal detection thresholds in the various correlation conditions. A quantitative within-channel model for CMR is proposed in Verhey et al. (1999). In this model, the most important stage is a spectral decomposition of the envelope within the auditory filter by means of a modulation filter bank located after the inner ear's frequency filter bank on the auditory pathway (Dau et al. 1997), the output of which is compared to a stored "image" for a suprathreshold signal by using cross correlation. This within-channel model can explain the experimental data in the accompanying experiments by an effective reduction of the modulation depth.

Possible explanations for psychophysical CDD findings are very similar to those used for CMR. Different qualitative across-channel and within-channel mechanisms have been proposed. Borrill and Moore (2002) and Moore and Borrill (2002) conclude that CDD is most likely mainly a within-channel effect. They introduce a quantitative within-channel model in which the auditory system can detect times during which the signal-to-masker ratio in the output of one channel is temporally above a certain threshold. These times are combined to the cumulative detection time over the total duration of the stimulus. If this cumulative time is sufficiently large, then the signal is detected.

These effective models are quite elaborate and use technical approaches to explain psychoacoustical data. Their neuronal realization is, however, still not fully understood. Wide-band inhibition in the cochlear nucleus was proposed as the neural mechanism underlying CMR, where the response of neurons that are excited by the signal frequency of the stimulus is inhibited by neurons that react to a wider range of frequencies around the signal frequency (Pressnitzer et al. 2001; Meddis et al. 2002; Neuert et al. 2004). This mechanism leads to a stronger firing rate response to the signal during masker dips in AC conditions than in AU or CU conditions. Another proposed neuronal explanation for CMR is suppression of firing rate locking to the masker envelope (Nelken et al. 1999; Las et al. 2005), which means that the sinusoidal signal prevents the stimulated neurons from locking to the masker envelope and thus allows for a better signal detectability in comodulated noise (corresponding to the AC condition) than in unmodulated noise (corresponding to the AU condition). The neuronal networks necessary for these explanations of CMR require at least a few excitatory and inhibitory synaptic connections to enable the proposed responses (see e.g., Meddis et al. 2002).

For CDD, a much simpler neuronal model based on neurophysiological recordings was proposed by Buschermoehle et al. (2006). The model involves only an excitatory stage and uses the mean value of the compressed envelope of a stimulus filtered by a single auditory channel as a detection cue. Although all of the above-mentioned neuronal models can account for general aspects of the respective psychophysical experiments, their predictions have been compared only to a very limited set of experiments.

To our knowledge, none of the models for CMR or CDD has aimed at explaining both effects at the same time. The aim of the present work is to show that a very simple neuronal model can be applied to various psychophysical CDD and CMR experiments. The physiology-based within-channel model introduced by Buschermoehle et al. (2006) will be extended in order to account for both effects. For this purpose, first the basic steps of the model will be described, then its main mathematical expressions will be introduced, and subsequently its applicability to CDD and CMR experiments will be demonstrated. Finally, there will be a comparison of the model's predictions with data from the literature.

In the model, it is assumed that the mean firing rate of neuronal populations covaries with the mean compressed envelope of the filtered stimulus. Such a locking of neuronal firing rates to the envelopes of auditory stimuli is described e.g., in Schreiner and Urbas (1988) and Joris et al. (2004). The presence of a compressive nonlinearity in rate-intensity functions of auditory nerve fibers has been demonstrated (Koeppl and Yates 1999; Saunders et al. 2002). This compression reflects a nonlinear transformation in the peripheral auditory system: A relatively large range of input sound levels is transduced into a smaller range of basilar membrane vibrations, where the slope of the input/output function on a double logarithmic scale is less than one dB/dB. A very simple model of this compression is used here: the instantaneous value of the sound envelope after filtering is raised to the power of $\alpha$ with $0 < \alpha < 1$. In Buschermoehle et al. (2006), an expression for the time and trial expectation value of the compressed envelope of the filtered stimulus is derived

for those CDD stimuli that were used for the corresponding physiological experiments. With this expression, the characteristics of neuronal firing rates could be simulated (see Bee et al. 2007 for details on the experimental setup). Here, the calculations are generalized to psychophysics and to describe CMR as well.

The central model quantity which is used in this paper is the mean compressed envelope of the filtered stimulus. The extraction of this quantity by the auditory system may be realized by a very simple neuronal mechanism: At least to a first approximation, the firing rate of neurons in the auditory nerve codes the envelope of the filtered stimulus directly. The properties of the basilar membrane result in filtering and compression at the same time. The effective half wave rectification of the signal and the limited ability to phase lock to the stimulus results in an effective extraction of the envelope.

## 2 Model structure

The model is organized in several steps. First, the incoming stimulus is filtered by applying a bandpass filter that is centered on the signal frequency. In order to allow for analytical calculations, each noise band is attenuated as a whole according to the response of a filter centered on the signal using the magnitude transfer function of the gammatone filter bank described in Hohmann (2002).

This procedure means that a phase preserving filter is assumed. For the model results presented here, the difference between this form of frequency-dependent attenuation and filtering may be neglected: numerical results with the original filters do not differ qualitatively from the analytical approximations (not shown here). Quantitative differences may be compensated by choosing adequate parameters.

After frequency filtering, the trial averaged envelope of the resulting stimulus is calculated. Then a compressive nonlinearity is applied to the envelope by raising the envelope to the power of $\alpha$ with $0 < \alpha < 1$. To finally get an expression for the mean of this compressed envelope across time and trials, the statistics of the analytical signal (Gabor 1946) is analyzed.

The stimuli in CDD as well as in flanking band CMR experiments consist of a number of noise bands centered at certain frequencies and in the case of CMR an additional single sinusoid. Numbering the flanking noise bands from 1 to $K$ and giving the on-frequency masker (or accordingly the signal band) the number 0, the analytical signal can be written as:

$$s(t) = d_0 b \mathrm{e}^{2\pi i f_0 t} + \sum_{k=0}^{K} a_k d_k \sum_{n=-N}^{N} \mathrm{e}^{i[2\pi(f_k + n\Delta\nu)t + \phi_{k,n}]}, \tag{1}$$

where each noise band is regarded as a sum of $2N+1$ pure tones with random phases. The factors $d_k$ produce the attenuation due to the filtering process. The envelope of the filtered stimulus is $|s(t)|$, while the filtered stimulus itself is $\mathrm{Re}(s(t))$. Each individual noise band is centered at $f_k$ and has a bandwidth of $2N\Delta\nu$. The noise bands are composed of individual sinusoids with amplitudes $a_k$, while the sinusoidal signal at $f_0$ for CMR-experiments has the amplitude $b$ and is attenuated by the filter factor $d_0$ (for CDD-experiments, $b = 0$). The phases $\phi_{k,n}$ distinguish the different correlation conditions: In the AU condition, the $\phi_{k,n}$ are independent and uniformly distributed in $[0; 2\pi]$. In the AC condition, all phases within one band (constant $k$, varying $n$) are random, while the same set of phases is used for all the different bands. And finally, in the CU condition, the phases of the flanking bands (i.e., $k \in \{1, \ldots, K\}$) are random within one band but the same for different bands, while the phases of the signal band (i.e., $k = 0$) are independent of the phases of the FBs and randomly distributed.

The trial average of the squared envelope across phases (i.e., the ensemble average) can be calculated and simplified by splitting the absolute square $|s(t)|^2$ into a sum of $\mathrm{Re}^2(s(t))$ and $\mathrm{Im}^2(s(t))$ and by using addition theorems. Straightforward calculations lead to:

$$\langle |s(t)|^2 \rangle_\phi = d_0^2 b^2 + (2N+1) \sum_{k,k'=0}^{K} a_k d_k a_{k'} d_{k'} \cdot \langle \delta(\phi_{k,0}, \phi_{k',0}) \rangle_\phi \cos(2\pi |f_k - f_{k'}| t). \tag{2}$$

This expression can be split into the three correlation conditions. The simplest case is the AU condition, where one can write $\langle \delta(\phi_{k,0}, \phi_{k',0}) \rangle_\phi = \delta(k, k')$:

$$\langle |s(t)|^2_{\mathrm{AU}} \rangle_\phi = d_0^2 b^2 + (2N+1) \sum_{k=0}^{K} a_k^2 d_k^2. \tag{3}$$

In the AC condition, the $\delta$-term is always one, such that

$$\langle |s(t)|^2_{\mathrm{AC}} \rangle_\phi = d_0^2 b^2 + (2N+1) \sum_{k,k'=0}^{K} a_k a_{k'} d_k d_{k'} \cdot \cos(2\pi |f_k - f_{k'}| t). \tag{4}$$

Finally, in the CU condition, the $\delta$-term does not vanish for all $k, k' \neq 0$ and if $k = k' = 0$. Thus, one gets

$$\langle |s(t)|^2_{\mathrm{CU}} \rangle_\phi = d_0^2 b^2 + (2N+1) d_0^2 a_0^2 + (2N+1) \sum_{k,k'=1}^{K} a_k a_{k'} d_k d_{k'} \cdot \cos(2\pi |f_k - f_{k'}| t). \tag{5}$$

In Appendix A, Eqs. 3–5 are rewritten in a notation using the overall levels of the noise bands. The differences between the

three correlation conditions can be seen in the interference terms $\cos(2\pi|f_k - f_{k'}|t)$.

Averaging the above equations over time will yield the same result in all three correlation conditions. The important step for getting quantitative differences in the temporal averages is to not consider the mean squared envelope but the mean compressed envelope (see also van de Par and Kohlrausch 1998; Verhey et al. 2007), which will be done in the following.

In order to proceed, we first consider the case of $b = 0$ (i.e., CDD). Here, the stimulus consists of noise bands only. This means that for large $N$, the distribution of squared envelope values may be approximated by an exponential distribution (Lawson and Uhlenbeck 1950). The mean of this distribution at time $t$ is given by $\mu^2(t) := \langle |s(t)|^2 \rangle_\phi$. (We define the term $\mu^2(t)$ for $b = 0$.) Now let $Y$ denote the random variable describing the value of the stimulus envelope. Then for any $\alpha > 0$ the time-dependent expectation value $E_t(Y^\alpha)$ can be computed from the exponential distribution of $Y^\alpha$ as

$$E_t(Y^\alpha) = \Gamma\left(\frac{\alpha+2}{2}\right)|\mu(t)|^\alpha \tag{6}$$

Here, $\Gamma(.)$ denotes the complete gamma function (e.g., Weisstein 2002). In a last step, the time- and ensemble-averaged value of the compressed envelope is given by integrating $E_t(Y^\alpha)$ over time. If the $f_k$ share a common integer divisor $\tilde{f}$, the integration only needs to be done over one period $T = 1/\tilde{f}$. Otherwise, one needs to take the limit of $T \to \infty$ for getting the expectation value:

$$E(Y^\alpha_{\mathrm{CDD}}) = \Gamma\left(\frac{\alpha+2}{2}\right)\frac{1}{T}\int_0^T |\mu(t)|^\alpha \mathrm{d}t. \tag{7}$$

Equations 6 and 7 are derived for the CDD case ($b = 0$). In the case of CMR the equations will still be a reasonable approximation for small $b$ (i.e., $b \ll a_0$). But with increasing amplitude of the sinusoidal signal in the CMR stimuli, the distribution of the squared envelope values will not be exponential anymore. For $b \gg a_0$ the sinusoid will dominate the envelope, which means that the distribution of $Y^2$ can be approximated by a $\delta$-peak at $d_0^2 b^2$, such that $E(Y^\alpha) \approx d_0^\alpha b^\alpha$. As the calculations leading to Eq. 7 cannot be carried out without making an assumption about the distribution of squared envelope values, we choose to approximate $E(Y^\alpha)$ by getting the right asymptotic behavior for small $b$ and for large $b$ within one expression. The simplest way to approximate this asymptotic behavior is to use the prefactor $\Gamma((\alpha+2)/2)$ only for $\mu^2(t)$, but not as a prefactor for the term describing the signal sinusoid (i.e., the term including $b$):

$$E(Y^\alpha_{\mathrm{CMR}}) = \frac{1}{T}\int_0^T \left(d_0^2 b^2 + \Gamma\left(\tfrac{\alpha+2}{2}\right)^{\frac{2}{\alpha}}|\mu^2(t)|\right)^{\frac{\alpha}{2}} \mathrm{d}t. \tag{8}$$

This expression has the desired asymptotic behavior.

A comparison of the predictions made by Eqs. 7 and 8 with realizations of stimuli from simulations is shown in Fig. 2a, b. Although for simplicity constant amplitudes of the sinusoids making up the stimuli have been assumed, the calculations hold with good accuracy also for stimuli with Rayleigh-distributed amplitudes, which are frequently used in experimental setups. They are also reasonably accurate for describing stimuli where the noise bands are generated by multiplying lowpass noise with sinusoids centered at the noise band center frequencies. This means that for explaining the general effects the model calculations can also be applied to CDD and CMR stimuli from different authors that have been generated in slightly different ways. A remarkable feature of the model curves shown in Fig. 2a is that their general shape is very similar to the characteristics of neuronal firing rates measured in the avian auditory forebrain during presentation of CDD stimuli (Buschermoehle et al. 2006; Bee et al. 2007). Differences of the curves for different correlation conditions, including the dip in the AC curve for CDD, result from interference between the components for correlated phases (see the cosine terms in Eqs. 4 and 5). The occurrence of the dip in the AC curve for CDD can be understood by noting that interference due to correlated phases results in a reduction of mean compressed envelopes already in the AC and CU conditions relative to the AU condition if no signal band is present. In the AC condition, the addition of the signal band introduces further interference and therefore results in a reduction of the mean compressed envelope for intermediate signal levels. This reduction is counterbalanced as soon as the signal band dominates the stimulus and therefore also the stimulus envelope. See Buschermöhle et al. (2006) for further details.

### 2.1 Mechanism of signal detection

A basic signal detection scheme can be envisaged by assuming that the time averaged compressed envelope of a filtered sample stimulus is represented somewhere in the auditory system, possibly by the firing rate of a population of neurons. This estimate $y$ is a random variable and will have an error $\sigma$ associated with it which is due to variability within the stimulus as well as variability in its neural representation. In the following, this error $\sigma$ is assumed to be constant (in particular, independent of signal and masker level), which allows for using the $d'$-measure from signal detection theory (Green and Swets 1966):

$$d'(L_S) = \frac{y(L_S) - y(-\infty)}{\sigma}, \tag{9}$$

where $y(L_S)$ denotes the estimate of the mean compressed envelope value of the filtered stimulus when the signal level is at $L_S$ and $y(-\infty)$ denotes the same estimate when the signal

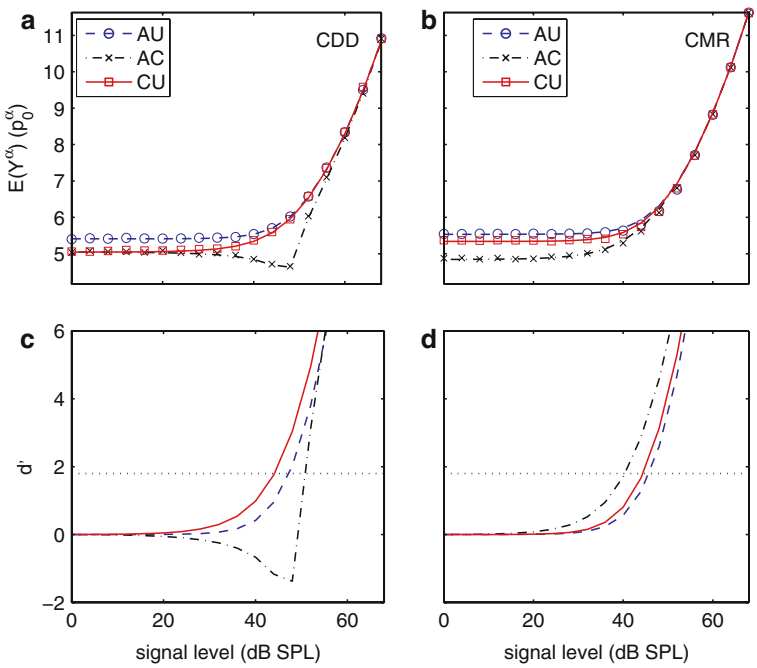


**Fig. 2** Mean compressed envelope values for CDD (**a**) and CMR (**b**) stimuli as well as the corresponding index of discriminability ($d'$) curves (**c** for CDD and **d** for CMR). *Markers* indicate simulated data, while *lines* are derived from Eqs. 7, 8 and 9. The overall level is set to 50 dB SPL for each flanking band, and the overall level of the on-frequency masker is 40 dB SPL. Compression $\alpha = 0.3$. Center frequency $f_0 = 2.0$ kHz. Flanking bands (and on-frequency masker in case of CMR) have a bandwidth of 100 Hz and are centered at 1.7, 1.85, 2.0, 2.15, and 2.3 kHz. Each band is attenuated as a whole according to the magnitude transfer function of a gammatone filter centered at 2.0 kHz (bandwidth parameter $\gamma = 1$). The *dotted horizontal line* in **c** and **d** indicates the decision criterion $D = 1.8$. The signal levels at which the $d'$-curves cross this criterion are the signal detection thresholds. (Internal error $\sigma = 0.3$, color online)

is absent. The larger the value of $d'$, the further apart are the two distributions of $y$ with and without signal. If $d'$ exceeds a predefined criterion $D = 1.8$, then the signal is said to have been detected, while for $d' < D$, the signal is not detected. The signal detection threshold can be defined as the signal level at which $d'$ reaches $D$. Note that the parameters $\sigma$ and $D$ together only comprise one free parameter: a change in the decision criterion $D$ can be equivalently expressed as a change in the internal error $\sigma$, which essentially rescales the $d'$ axis.

With this signal detection scheme, one can identify the reason for CDD and CMR in Fig. 2. For CDD, consider the AC and CU curves in Fig. 2a. Calculating $d'$ means shifting them to zero and rescaling them by $\frac{1}{\sigma}$ (see Fig. 2c). Because the AC curve is always beneath the CU curve, the former will cross the detection criterion $d' = D$ at a higher signal level than the latter, which directly corresponds to the CDD effect. For CMR (Fig. 2b), the AU and AC curves are both monotonically increasing and have the same asymptotic behavior for $L_S \to \infty$, but values for vanishing signal are not the same. Calculating $d'$ corresponds to shifting both curves vertically relative to each other, and therefore, the AC curve will always cross the detection criterion at lower signal levels than the AU curve (see Fig. 2d), i.e., a CMR is predicted.

## 2.2 Model parameters

Up to now, two free model parameters have been introduced: the compression $\alpha$ and the internal error $\sigma$. The third and final model parameter $\gamma$ describes the width of the auditory gammatone filter in ERB. The ERB (equivalent rectangular bandwidth) of a bandpass filter with arbitrary amplitude response is the bandwidth of a corresponding filter with rectangular amplitude response that has the same peak response and passes the same total amount of power. The relation between center frequency $f_c$ and ERB bandwidth for the human auditory system is given by the following empirical formula from Glasberg and Moore (1990) for $\gamma = 1$:

$$\mathrm{ERB}(f_c) = \gamma \cdot (24.7\ \mathrm{Hz} + 0.1079\, f_c). \tag{10}$$

The prefactor $\gamma$ is introduced here to be able to change the filter width. For exploring the model's ability to explain different CDD and CMR experiments, these three parameters will be adjusted within reasonable ranges in the following. The parameter meanings and their ranges are summarized in Table 2. The influence of the different parameters is discussed in Sect. 4. The compression in the human auditory system can vary roughly between $\alpha = 0.1$ and $\alpha = 0.8$ (for a review, see Bacon et al. 2003). The internal error $\sigma$ is a quantity that

**Table 2** Model parameters used in this paper

| Parameter | Range | Description |
|---|---|---|
| $\alpha$ | 0.1–0.6 | Compressive exponent (no units) |
| $\sigma$ | 0.1–0.35 | Internal error (in $p_0^\alpha$, compressed reference pressure) |
| $\gamma$ | 0.8–5.0 | Filter width (in ERB) |

is hard to measure, and therefore, its reasonable range for the human auditory system is not known. We regard $\sigma$ as an adjustable parameter which should be positive and smaller than the dynamic range of mean compressed envelopes that arises at sound pressure levels between 0 and 100 dB. The filter bandwidth parameter $\gamma$ should be close to one for normal hearing humans. Values up to about 1.5 may be realistic for within-channel processing of stimuli (see Glasberg and Moore 1990). If $\gamma$ is larger than that, then the processing must be regarded as across-channel processing.

## 3 Results

The very simple model introduced here has only three free parameters and has been derived from a model for neuronal firing rates obtained in CDD experiments. The basic modeling steps frequency filtering, envelope extraction, and compression are directly motivated by peripheral auditory processing. Only temporal averaging and decision making are presumably central processes. Despite its simplicity, the model will now be applied to different CDD and CMR experiments described in the literature. Model and experimental results are compared in order to determine the scope of the proposed model. All relevant parameters for the experiments modeled in the present study are summarized in Table 3.

### 3.1 Frequency spacing between signal and flanking bands for CDD

Several experimental studies discuss the effect of frequency distance between flanking bands and the signal band for CDD (e.g., Borrill and Moore 2002; Cohen and Schubert 1987; McFadden 1987). In the present study, the model is applied to stimulus setups comparable to the ones described in these three studies. The leftmost column in Fig. 3 compares model predictions with experimental data from Borrill and Moore (2002), where the signal band was flanked symmetrically by two noise bands, at a distance $\Delta f$ above and below the signal frequency, respectively. In the central column of Fig. 3, model and experiment are compared for data published in Cohen and Schubert (1987). There, the stimulus consisted of a flanking band at a constant center frequency and signal band at varying distances $\Delta f$ from the flanking band. In the third column of Fig. 3, signal detection thresholds for an experiment described in McFadden (1987) are plotted. Here, the stimulus consisted of one signal band at a constant center frequency and a flanking band with varying center frequency.

The model predictions are in qualitative agreement with the experimental data. In agreement with the data in all three experiments, the difference between AC and CU thresholds (the CDD) decreases for increasing frequency spacing but does not reach zero for the largest $\Delta f$. The thresholds are determined by the form of the auditory filter since the filter used for signal detection is always the one centered on the signal band. The signal band will be detected as soon as its influence on the total filter output dominates the mean compressed envelope, which happens at signal levels close to the

**Table 3** An overview of the experimental setups and respective modeling parameters discussed in the present paper

| Reference | Paradigm | Signal frequency (kHz) | FB frequencies (kHz) | Bandwidth (Hz) | Noise band level | $\alpha$ | $\sigma$ | $\gamma$ | Remarks |
|---|---|---|---|---|---|---|---|---|---|
| Borrill and Moore (2002) | CDD | 1.5 | $\Delta f = \pm$**0.1 ... ±1.4** | 20 | 78 dB SPL | 0.3 | 0.1 | 1 | – |
| Cohen and Schubert (1987) | CDD | **0.2…6.0** | 1.0 | 100 | 73 dB SPL | 0.6 | 0.1 | 1 | – |
| McFadden (1987) | CDD | 2.5 | **1.5…3.5** | 100 | 70 dB SPL | 0.6 | 0.1 | 1 | – |
| Schooneveldt and Moore (1987) | CMR | 2.0 | **1.0…3.0** | 25 | 67 dB SPL | 0.2 | 0.1 | 5 | – |
| McFadden (1987) | CMR | 2.5 | **1.5…3.5** | 100 | 70 dB SPL | 0.2 | 0.1 | 5 | – |
| Hall et al. (1984) | CMR | 1.0 | **0.7…1.3** | 100 | 60 dB SPL | 0.2 | 0.1 | 5 | – |
| McFadden (1987) | CDD | 2.5 | **2 or 4 bands at 1.5…3.5** | 100 | 70 dB SPL | 0.2 | 0.1 | 1 | – |
| McFadden (1987) | CMR | 2.5 | **2 or 4 bands at 1.5…3.5** | 100 | 70 dB SPL | 0.2 | 0.1 | 1.4 | – |
| Borrill and Moore (2002) | CDD | 1.5 | 0.9, 2.1 | 20 | **53…83 dB SPL** | 0.3 | 0.2 | 0.8 | – |
| Ernst and Verhey (2005) | CMR | 8.0 | 1.0 | 100 | **OFM level: 20…60 dB SPL** | 0.3 | 0.1 | 1 | Off-frequency listening |
| Hall et al. (1984) | CMR | 1.0 | 1.0 | **100…700** | 40 dB SPL spectrum level | 0.25 | 0.35 | 1.2 | – |

The experimentally varied quantities are printed in bold face

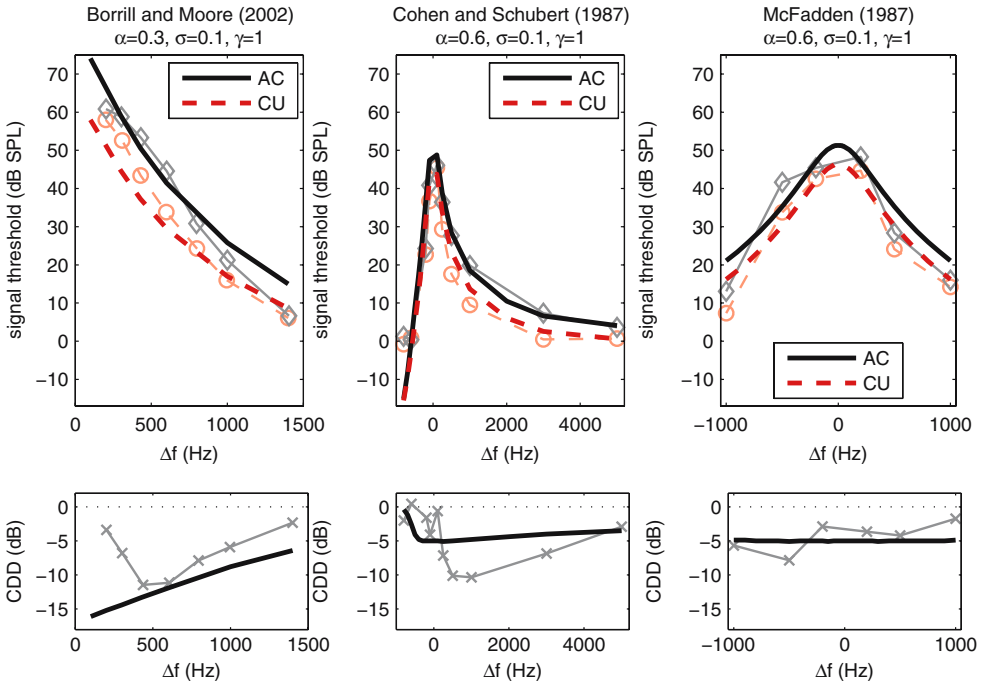


**Fig. 3** Effect of frequency spacing between signal and flanking band on CDD. *Dark thick lines* mark model results. *Light thin lines with symbols* indicate experimental data. The *three columns* show data from three different CDD experiments. References and model parameters used are indicated in the *figure titles*. The *top row* shows signal detection thresholds, the *bottom row* shows the corresponding threshold differences (CDDs). Experimental parameters are summarized in Table 3. (Color online)

level of the attenuated flanking band. The clearest deviations between model and experiment can be found for $\Delta f$ close to 0 Hz. There, the model predicts quite large CDDs while in experiments the CDDs are relatively small.[1] The plot for the experiments by Cohen and Schubert (1987) is not symmetrical to $\Delta f = 0$ Hz. This reflects the fact that for each signal frequency a different filter is picked, and that filters with increasing signal frequencies have larger bandwidths (cf. Eq. 10). In contrast, the plot for data from McFadden (1987) is largely symmetrical with respect to $\Delta f = 0$ Hz because the signal frequency and therefore the auditory filter does not change. Here, the thresholds reflect the shape of the symmetrical filter.

All three experiments could be modeled reasonably well with very similar parameters. It is noteworthy that the filter bandwidth is at a value that is typical for the human ear, indicating that the whole CDD effect may be due to peripheral within-channel processes.

### 3.2 Frequency spacing between the signal and one flanking band for CMR

The effect of frequency distance between one flanking band and the on-frequency masker for CMR is discussed amongst others in Schooneveldt and Moore (1987), McFadden (1987), and Hall et al. (1984). The model is compared to the results reported in these studies in Fig. 4. In that figure, the left column shows data from Schooneveldt and Moore (1987), where the signal frequency was kept constant and one flanking band was varied in its center frequency. The frequency distance between the flanking band's center frequency and the signal frequency is denoted as $\Delta f$. The second column in Fig. 4 shows data from McFadden (1987), where the on-frequency masker and the signal had a constant center frequency while the flanking band was centered at varying distances $\Delta f$ from the signal. In the rightmost column of Fig. 4, experimental results from Hall et al. (1984) are compared to model predictions. The stimuli here consisted of an on-frequency masker and a signal at a constant frequency and one flanking band at a frequency distance $\Delta f$. The case $\Delta f = 0$ Hz was treated as a special case in that experiment: for that case there was only the on-frequency masker present (without the flanking band) which means that there was no difference between the AC and AU correlation conditions and therefore both signal detection thresholds are the same in the model as well as in the experiments.

A general finding for all three experiments is that thresholds in the AU case increase for decreasing $\Delta f$, while thresholds in the AC case behave in the opposite way. For large frequency separations, the model predicts a decrease in the threshold difference (i.e., the CMR), which can also be seen in the experimental data. The general shape of the depicted model curves can be understood by considering that for large frequency separations, the flanking band is completely filtered out by the auditory filter and therefore only the on-frequency masker determines the signal threshold. The closer the flanking band is to the on-frequency masker,

[1] Note that for $\Delta f = 0$ Hz, the AC condition needs to be treated separately (see Appendix A), because here signal band and masker band are exactly the same except for level differences. We do not explicitly treat this case here, because it is not important for the general model results.

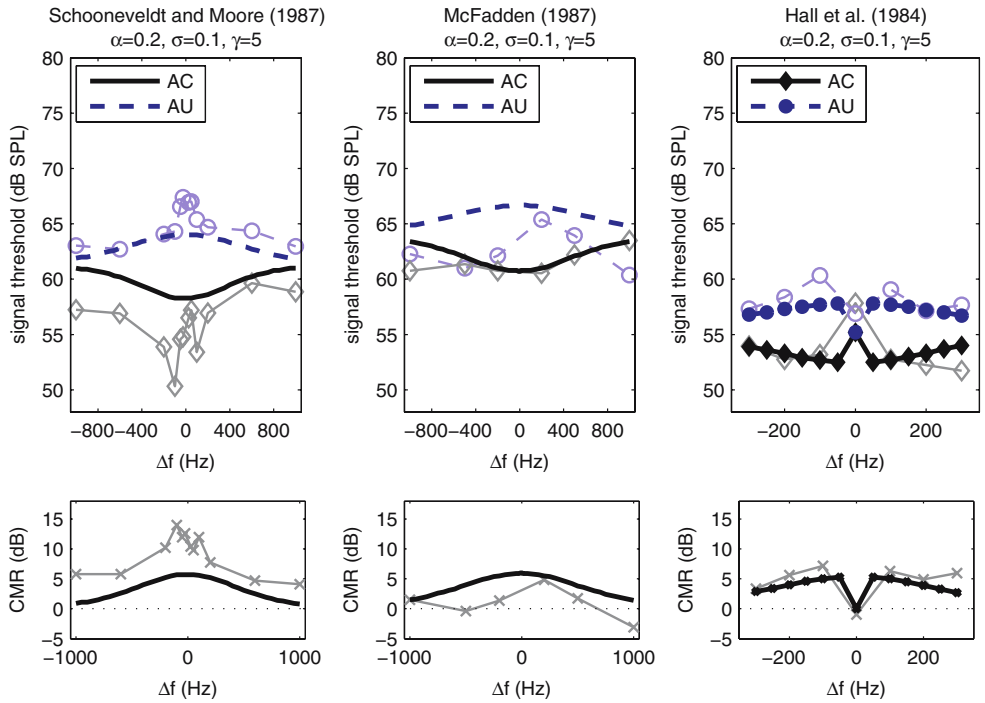


**Fig. 4** Effect of frequency spacing on CMR. *Dark thick lines* mark model results. *Light thin lines with symbols* indicate experimental data. The *three columns* show data from three different CMR experiments. References and model parameters used are indicated in the *figure titles*. The *top row* shows signal detection thresholds, the *bottom row* shows the corresponding threshold differences (CMRs). In all experiments, the signal was a pure sine tone at a constant signal frequency, and there were two noise bands present: the on-frequency masker centered on the signal frequency and the flanking band with varying frequency distance from the signal $\Delta f$. The signal frequency, noise bandwidth and noise band levels were different for the different experiments (see Table 3). Model predictions for the experiment by Hall et al. (1984) were obtained only for certain frequency distances indicated by the *symbols*. (Color online)

the more it influences the on-frequency masker. In the AU case, it increases the power that falls into the auditory filter and therefore rises thresholds, while in the AC case, the flanking band causes beating with the on-frequency masker and therefore leads to reduced mean compressed envelopes and facilitates signal detection.

The parameters used for modeling the three experiments are exactly the same, which means that in spite of slightly varying experimental conditions, the model can account for the general shape of the experimental data without varying the model parameters. An important note needs, however, to be made concerning the parameter $\gamma$ determining the bandwidth of the auditory filter. The value $\gamma = 5$ is well above the realistic range for typical auditory filters. A model with a realistic auditory filter would underestimate the CMR for medium to large spectral separations between signal and flanking bands. This means that CMR is presumably not solely due to within-channel processes as it was the case for CDD in Sect. 3.1. A model with a filter width larger than that observed in the auditory periphery can be interpreted as an effective realization of an across-channel process. It should be noted that it does not rely on the classical across-channel processes as they were explained in Sect. 1. This point will be discussed further in Sect. 4.1.

### 3.3 Number of flanking bands for CDD and CMR

It has been pointed out that the amount of CDD as well as CMR depends on the number of flanking bands (McFadden 1987; Wright 1990, for CDD, and Hall et al. 1984; McFadden 1987 for CMR). This effect is investigated for the model by setting up the stimuli similar to those used in McFadden (1987), where the consequences of adding further flanking bands are investigated for both CDD and CMR. The model results are compared to experimental data in Fig. 5. In the experiments, only the AC and CU conditions were considered, which is why the model results are restricted to these two correlation conditions.

For CDD (Fig. 5a, c), it is observed in the model and in the experiments that both thresholds are higher if the flanking bands are closer to the signal band. This is the case because by shifting the flanking bands closer to the signal band, they are less attenuated due to filtering and therefore contribute to increased thresholds. The amount of CDD in the model grows the closer the flanking bands come to the signal band. This is not reflected in the experimental data. Also the experimental finding that by adding two further noise bands, the amount of CDD increases is not reflected in the model. These discrepancies between model and experiment are consistent

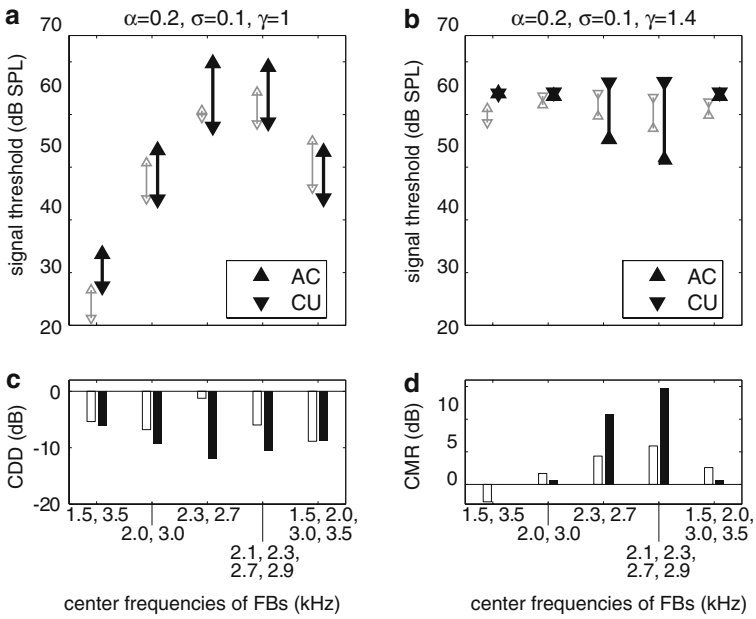


**Fig. 5** Effect of number of bands on CDD (**a, c**) and CMR (**b, d**). *Shaded triangles* mark signal detection thresholds as determined from the model. *Light open triangles* indicate experimental data from McFadden (1987). **c** CDD, *open bars* experimental data, *filled bars* model. **d** CMR. Model parameters are specified in the *titles*. Experimental parameters are summarized in Table 3

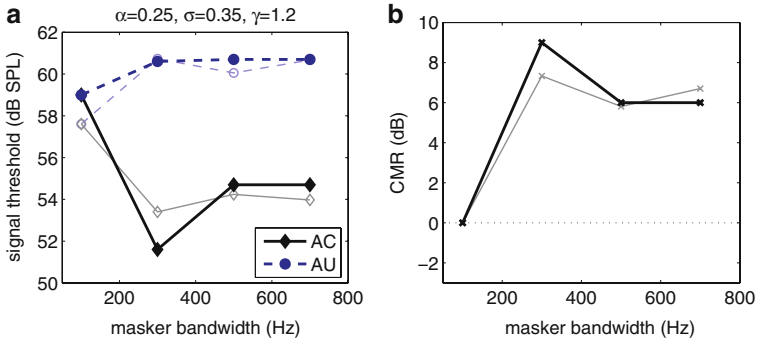


**Fig. 6** Effect of masker bandwidth on CMR. *Dark thick lines with shaded symbols* mark model results. *Light thin lines with open symbols* indicate experimental data from Hall et al. (1984). The on-frequency masker is composed of up to seven noise bands with a bandwidth of 100 Hz each, and the bandwidth is increased by symmetrically adding further noise bands to the OFM. **a** Signal thresholds, **b** threshold differences. (Color online)

with those found in Sect. 3.1, where the CDDs for very small $\Delta f$ increased in the model while they decreased in the experiments. Still, the general characteristics of model predictions are reasonably close to those of the experimental data.

Comparison of model predictions and experimental results for CMR is shown in Fig. 5b, d. Here, the model predicts considerable CMR only for the cases where the flanking bands are closest to the on-frequency masker. This generally agrees with the experimental data. The model thresholds also are in the same range as the experimental thresholds. It should be noted that the predicted CMR is even larger than in the experimental data despite the fact that the model uses a filter width that is reasonably close to that of the auditory filters. Thus, for this experiment across-channel processing seems not be necessary to account for the CMR.

A further experiment, described in Hall et al. (1984), investigates the influence of noise bandwidth on CMR by symmetrically adding 100 Hz wide noise bands to the right and left of the on-frequency masker. The additional noise bands can either have the same envelope fluctuations as the first on-frequency masker (AC) or independent envelope fluctuations (AU). This experiment can, therefore, also be seen as an experiment concerning the number of noise bands. Comparison of the data from Hall et al. (1984) with model predictions is shown in Fig. 6a, b.

There, the on-frequency masker bandwidth is increased from 100 over 300 and 500 to 700 Hz by repeatedly adding two noise bands of a bandwidth of 100 Hz above and below the original on-frequency masker. The model thresholds show a behavior which resembles that of the experimental CMR

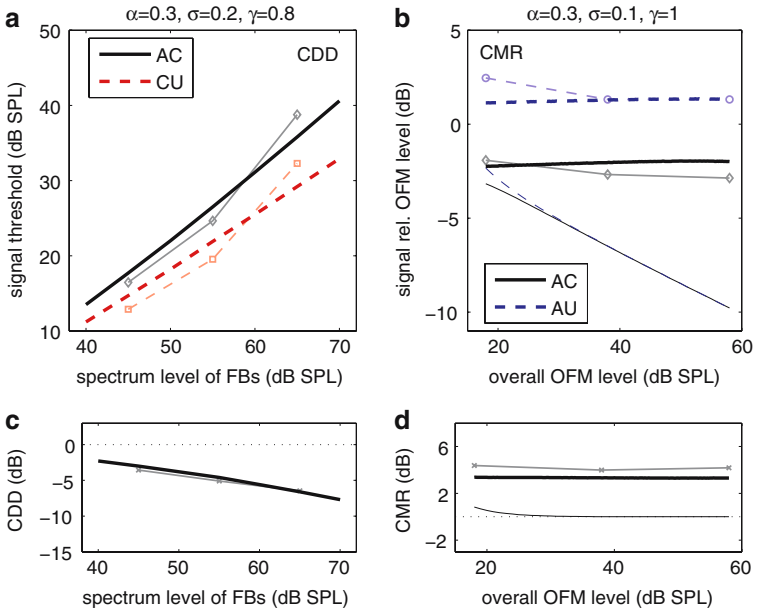


**Fig. 7** Effect of level of noise bands on CDD (**a, c**) and CMR (**b, d**). **a** CDD; *dark thick lines* mark signal detection thresholds as determined from the model, *light lines with symbols* indicate mean experimental data from Borrill and Moore (2002). The signal band is centered between two flanking bands. **b** CMR; *light lines with symbols* indicate data from Ernst and Verhey (2005). *Dark thick lines* represent model results for off-frequency listening. *Dark thin lines* show model results for on-frequency listening (filter centered on the signal frequency). The flanking band is located 7 kHz below the on-frequency masker at 8 kHz. **c, d** Threshold differences. (Color online)

thresholds closely. In agreement with the data, the predicted CMR does not change strongly for on-frequency masker bandwidths larger than 500 Hz because further masker power is only added at frequencies outside the filter, which does not influence the envelope statistics significantly.

The parameters used for modeling the CDD and CMR experiments in this section are similar. The most important difference is the filter bandwidth $\gamma$. For CMR, slightly larger filters were needed than for CDD. However, all parameter values can still be viewed as modeling within-channel processes.

### 3.4 Influence of noise band level

A further experimental parameter which may affect signal detection thresholds is the level of the noise bands. For CDD, the dependence of signal thresholds on flanking band level was explored in Borrill and Moore (2002). In that study, the signal band was centered between two flanking bands. Experimental and model data for this setup are compared in Fig. 7a, c. One finds that signal thresholds generally increase with rising flanking band level. As in the data, predicted AC thresholds are generally higher than the CU thresholds. The model quantitatively predicts the CDD in this experiment.

There are several studies involving the dependence of CMR on noise band level. The dependence of CMR on the level of one flanking band, keeping the on-frequency masker level constant, is investigated in Schooneveldt and Moore (1987). In Moore and Shailer (1991), the dependence of CMR on the level of several flanking bands is examined keeping the on-frequency masker level constant or varying it with the flanking bands. In Cohen (1991) and Ernst and Verhey (2005) the influence of on-frequency masker level on the detection of the signal sinusoid is analyzed while the overall level of one flanking band is kept constant. As the last two studies report a CMR over a range of several octaves, we choose to compare the model with the data of the latter study, Ernst and Verhey (2005).

As can be seen from Fig. 4, the basic model setup yields significant CMRs only for frequency distances of up to 1,000 Hz between on-frequency masker and flanking band, which is considerably less than the 7 kHz separation between the two bands here. Therefore, the model needs to be changed from using the filter centered on the signal frequency (on-frequency listening) to using the filter which attenuates flanking band and on-frequency masker such that their levels *after filtering* are the same. This corresponds to an off-frequency listening strategy (Patterson and Nimmo-Smith 1980; O'Loughlin and Moore 1981) and means that a different filter is picked from the filter bank than the one centered on the signal. If the filter used in the model is chosen in such a way, then beating between on-frequency masker and flanking band after filtering has a big effect on the envelope statistics in the AC condition and thus facilitates signal detection in the AC condition. Now, the central frequency of the filter used for signal detection depends on the levels of on-frequency masker and flanking band. The modeling results for on- and off-frequency listening are plotted in Fig. 7b, d. One finds a good correspondence between the off-frequency listening model and the experimental data: The amount of CMR is nearly independent of on-frequency masker level.

Although the thresholds determined by the on-frequency listening model are lower than those predicted by the off-frequency listening model and therefore represent the better signal detection strategy, the performance of the auditory system in this experiment is better described by the

off-frequency listening model. This is not the case for the previously modeled experiments.

## 4 Discussion

In the present paper, it is proposed that main features of CDD and CMR experiments can be understood by registering changes in the mean compressed envelope of a stimulus filtered by a single auditory channel. The model consists of the five stages frequency filtering, envelope extraction, compression, averaging, and signal detection, which may be realized at very early stages of the auditory pathway. There are three free parameters for the model: the filter bandwidth $\gamma$, the compressive exponent $\alpha$, and the internal error $\sigma$. When applying the model to experiments from the literature, these model parameters are adjusted within reasonable ranges to demonstrate the scope of the model.

### 4.1 Filter bandwidth $\gamma$

The bandwidth of the auditory filters is usually kept close to the values given in Glasberg and Moore (1990) (corresponding to $\gamma \approx 1$ in Eq. 10). This choice is valid for most normal hearing subjects. The good correspondence between data and model predictions indicates that CDD and CMR effects in most of the experiments are largely due to within-channel cues. Only for the CMR experiments investigating the effect of frequency spacing between on-frequency and flanking bands, a standard auditory filter bandwidth cannot account for the experimental thresholds. By increasing $\gamma$, the present model can also explain across-channel contributions to CMR. Indeed, the general discussion of comodulation experiments in the literature indicates that CDD experiments can be accounted for by within-channel processes (see Borrill and Moore 2002), while CMR also needs across-channel processing (e.g., Verhey et al. 2003). This is in agreement with the findings in the present study. Compared to other models, the presented model has the advantage that it does not involve suppression (Ernst and Verhey 2006), inhibition (Pressnitzer et al. 2001), or analysis of temporal features (Verhey et al. 1999; Nelken et al. 1999) but only sequential excitatory feed-forward processing to account for within- as well as across-channel processes. Using the model with realistic filter widths can help to quantify the contribution of within-channel cues in comodulation experiments.

### 4.2 Compression $\alpha$

The values of the parameter $\alpha$ describing compression in the auditory system are within the realistic range for humans (e.g., Bacon et al. 2003). For model simplicity, the compression is assumed to be independent of the stimulus level. This suffices as an explanation for the general trends in the data presented in this study. A more realistic level dependent compression would have been possible but at the cost of a more complex model which could not have been treated analytically. Due to the fact that the different experiments described in the present study were performed with different subjects and at varying overall stimulus levels, adjusting the compressive exponent to the different experiments is a reasonable assumption.

### 4.3 Signal detection criterion $D$ and internal error $\sigma$

The criterion for signal detection $D = 1.8$ has been chosen because of the fact that two Gaussian distributions with their means separated by 1.8 standard deviations can be distinguished quite well and because this criterion has been used in other studies as well (e.g., Klump and Nieder 2001; Langemann and Klump 2001). As discussed in Sect. 2.1, this parameter can be scaled by choosing a different value for $\sigma$ and may, therefore, be regarded as a constant rather than a free parameter of the model.

The parameter $\sigma$ for the supposedly constant internal error that the auditory system faces when detecting changes in the mean compressed envelope has been chosen by fitting the model thresholds to the data. The variability of the estimate of the compressed envelope by the auditory system is affected by at least two contributions: on the one hand, the signal statistics yield a time-varying envelope value, and on the other hand, the partly random firing of auditory nerve fibers makes the number of spikes in a certain time interval a random number. In Buschermöhle et al. (2006) it is assumed that the standard deviation of firing rates of small populations of neurons in the starling's auditory forebrain only marginally depends on the level of the signal in CDD experiments. This may be taken as a hint that the biggest contribution to $\sigma$ is the firing rate variability due to random spiking and not due to the stimulus variability, as the variability of the stimulus envelope in CDD experiments changes with signal level. For simplicity, the parameter $\sigma$ is regarded as a constant although the firing rate variability of auditory nerve neurons may depend on the overall level of a stimulus. Surely there is also an intersubject variability concerning the exact value of $\sigma$. The different psychophysical procedures and stimulus or signal durations used in the experiments can affect this parameter as well. This may be an explanation for the fact that $\sigma$ needs to be adjusted slightly in order to model the discussed experiments.

### 4.4 Parameter dependence of the model results

Even if the model has only three free parameters, its results may depend critically on their specific choice. To get an idea of the influence of a certain parameter set on the model, we

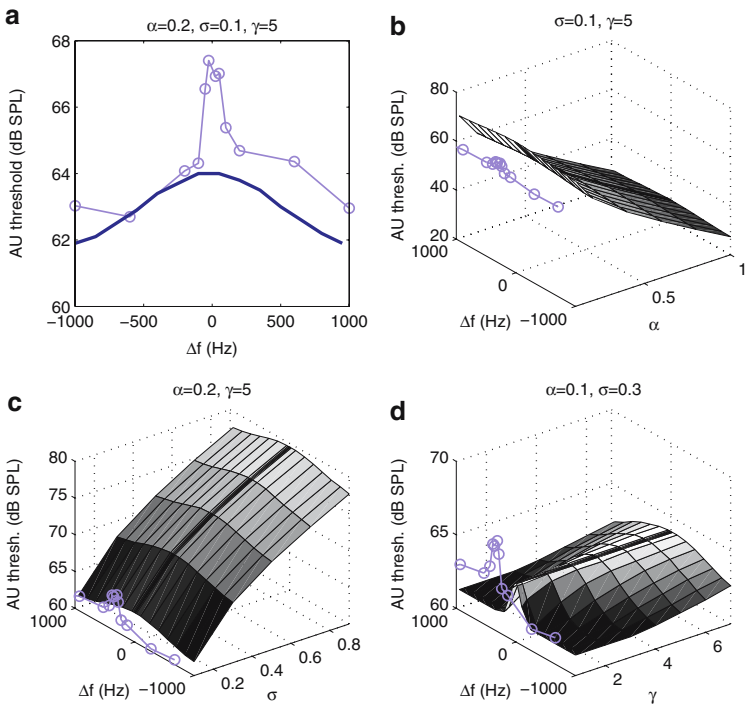


**Fig. 8** Influence of parameter choice on model results for the exemplary case of data from Schooneveldt and Moore (1987). **a** Replot of the AU signal detection thresholds shown in the *top left* subplot of Fig. 4. **b, c, d** Dependence of the AU threshold on $\alpha$, $\sigma$, and $\gamma$, respectively. Each surface plot shows the thresholds in the AU condition predicted by the model while the *light colored connected symbols* show the experimental data. (Color online)

discuss the results for an example stimulus setup for the range of parameters used in the present publication (see Fig. 8, Table 2). The example stimulus setup is the first frequency spacing CMR experiment discussed in Sect. 3.2. The dependence of thresholds on the compression $\alpha$ for given $\sigma$ and $\gamma$ (Fig. 8b) is easily understood: the larger $\alpha$, the quicker do the mean compressed envelopes increase with signal level and the lower is the signal level at which the signal is detected. Also the general dependence of thresholds on $\sigma$ (Fig. 8c) is easy to explain: larger values of $\sigma$ cause the $d'$-curves to reach the detection criterion $D$ at higher signal levels, which means that the signal is harder to detect and the thresholds increase to higher levels. The influence of the filter bandwidth $\gamma$ on the model thresholds (Fig. 8d) is to increase the frequency region in which the noise band centered on the signal frequency and the flanking bands can interact to influence the mean compressed envelopes. In the example shown here, a broader filter (larger $\gamma$) leads to a broader region of increased thresholds around $\Delta f = 0\,\mathrm{Hz}$.

One can generally say that by increasing $\alpha$ or by decreasing $\sigma$ the absolute value of the thresholds can be adjusted while by reducing $\gamma$, the size of the affected frequency region can be reduced.

### 4.5 Influence of temporal processing

For determining the mean value of compressed envelopes in the model, the average with respect to infinite time and the whole ensemble of possible stimulus realizations is computed. This means that the duration of intervals in the experiments is not explicitly taken into account. There are two possible ways of extending the model by considering the stimulus duration: on the one hand, the parameter $\sigma$ may be changed accordingly (larger $\sigma$ for shorter signal intervals). On the other hand, one can perform Monte Carlo simulations of the experiments and use the realizations of stimuli to compute mean compressed envelope values. These kinds of simulations may also be used for determining signal detection thresholds in cases where the stimulus statistics is not calculated as easily as in the cases discussed here.

A further aspect that indicates a possible oversimplification in the proposed model is that for human listeners it is possible to perceive the temporal fine structure of the stimuli. This is not included in the model, as the model averages over time. Different perceptual impressions of maskers alone and maskers plus signal in the model are only possible for one stimulus feature, which might correspond to the overall loudness. But subjects do not only perceive loudness differences in stimuli with and without signal. Therefore, the model might be improved by taking into account temporal aspects of the stimuli (such as it is e.g., assumed in Verhey et al. 1999).

### 4.6 Comparison to other low-level models

Besides our approach to model the influence of the auditory periphery on CDD and CMR, there are basically two other models discussing the contributions of low-level processing on comodulation effects. These contributions are either suppression due to nonlinear filtering on the basilar membrane

(Ernst and Verhey 2006) or lateral inhibition on a neuronal level in the cochlear nucleus (Meddis et al. 2002).

The model by Ernst and Verhey (2006) is in its general setup comparable to the model proposed here: only the output of one auditory filter is analyzed, which can be done in principle at the level of the auditory nerve. The filtering however is nonlinear, such that suppression effects are considered. The filter output is not converted into an envelope and then averaged but convolved with a temporal window. The decision variable in that model is the quotient of the maximum intensity during the presentation of masker and signal and the maximum intensity during the presentation of the masker alone. This means that the model complexity is comparable to that of the model discussed in the present paper. However, the model cannot be treated analytically due to the nonlinear filtering. Additionally, the model is only applied to the stimuli used in the accompanying experiments and a general applicability in comodulation experiments is not discussed.

The model by Meddis et al. (2002) can quantitatively explain neuronal recordings made in accompanying experiments as well as general trends in psychophysical experiments. One drawback is that it is not clear how this model can explain other comodulation experiments (the set of stimuli used in the accompanying experiments was rather limited and did not consist of noise bands but amplitude modulated pure tones). Another drawback is that the neuronal circuitry necessary for the model by Meddis et al. (2002) is relatively complex compared to the fact that the model presented here relies only on neuronal firing rates at the level of the auditory nerve. One aspect which has been clearly shown by the model in the present paper is that a reduction in neuronal firing rate with increasing stimulus level does not necessarily have to involve inhibition or suppression, as the dip in the AC condition for CDD stimuli (see Fig. 2) is caused by destructive interference within a purely excitatory model.

## 5 Conclusions

In this paper, a very simple model of the auditory periphery essentially based on peripheral processing is introduced that is able to reproduce several aspects of psychophysical CDD and CMR experiments while being derived from physiological investigations. Deviations between the model predictions and experimental results which have been discussed above still leave room for higher level processes. The model may provide insights into the causes of psychophysical thresholds and their dependence on different experimental parameters. One notable feature of the model is the possibility of performing analytical calculations which gives a basic understanding of the importance of the parameters and how they may change the model's behavior. The proposed model constitutes a unifying approach to CDD and CMR and may be instrumental in developing more sophisticated simulation models for experiments with comodulated stimuli.

**Acknowledgments** We thank Mark A. Bee and Georg M. Klump for their comments and support. This work was supported by the International Graduate School for Neurosensory Science, Systems, and Applications (DFG GRK 591).

## A Envelope statistics with levels

Denoting the overall level of noise band $k$ by $L_k$ and the overall level of the signal sinusoid for CMR by $L_b$, Eqs. 3–5 can be written in terms of levels instead of sine amplitudes. The RMS-value of a sum of $2N+1$ sinusoids of amplitude $a$ with independent random phases is $a\sqrt{(2N+1)/2}$. Setting the reference sound pressure to an RMS value of $p_0$ (for dB SPL, $p_0 = 20\,\mu\text{Pa}$), has the consequence that the level of the mentioned sum of sinusoids is $L = 20\log_{10}\left(\frac{a\sqrt{(2N+1)/2}}{p_0}\right)$. This expression can be solved for $a$ to give

$$a = p_0 10^{\frac{L}{20}}\sqrt{\frac{2}{2N+1}}. \tag{11}$$

Inserting Eq. 11 into Eqs. 3–5 and remembering that $L_b = 20\log_{10}\left(\frac{b}{\sqrt{2}p_0}\right)$ when normalizing to an RMS value of $p_0$, one gets:

$$\mu_{\text{AU}}^2 = 2p_0^2 d_0^2 10^{\frac{L_b}{10}} + 2p_0^2 \sum_{k=0}^{K} 10^{\frac{L_k}{10}} \cdot d_k^2 \tag{12}$$

$$\mu_{\text{AC}}^2 = 2p_0^2 d_0^2 10^{\frac{L_b}{10}} + 2p_0^2 \sum_{k,k'=0}^{K} 10^{\frac{L_k}{20}} 10^{\frac{L_{k'}}{20}} \cdot d_k d_{k'} \cos(2\pi |f_k - f_{k'}| t) \tag{13}$$

$$\mu_{\text{CU}}^2 = 2p_0^2 d_0^2 10^{\frac{L_b}{10}} + 2p_0^2 d_0^2 10^{\frac{L_0}{10}} + 2p_0^2 \sum_{k,k'=1}^{K} 10^{\frac{L_k}{20}} 10^{\frac{L_{k'}}{20}} \cdot d_k d_{k'} \cos(2\pi |f_k - f_{k'}| t) \tag{14}$$

Eqs. 12–14 are independent of the bandwidth of the noise bands and the frequency spacing $\Delta\nu$ of the component sinusoids. They rely on the assumptions that all phases of the summed sinusoids within one band are independent and random and that the number of added sinusoids is large (i.e., $N \gg 1$).

The equations do not hold if two identical bands are superimposed at the same frequency (this happens for band spacing experiments in the AC condition for $\Delta f = 0\,\text{Hz}$). In that case, the two superimposed bands with their sine amplitudes $a_1$ and $a_2$ can be viewed as one band with the sine amplitude $a_1 + a_2$ resulting in an RMS-value of $(a_1 + a_2)\sqrt{(2N+1)/2}$, which changes Eqs. 12–14 accordingly.

## References


Bacon SP, Fay RR, Popper AN (Eds) (2003) Compression: from Cochlea to Cochlear Implants. Springer, New York

Bee MA, Buschermöhle M, Klump GM (2007) Detecting modulated signals in modulated noise: II. Neural thresholds in the songbird forebrain. Eur J Neurosci 26(7):1979–1994

Borrill SJ, Moore BCJ (2002) Evidence that comodulation detection differences depend on within-channel mechanisms. J Acoust Soc Am 111(1):309–319

Buschermöhle M, Feudel U, Klump GM, Bee MA, Freund JA (2006) Signal detection enhanced by comodulated noise. Fluct Noise Lett 6(4):L339–L347

Cohen MF (1991) Comodulation masking release over a three octave range. J Acoust Soc Am 90(3):1381–1384

Cohen MF, Schubert ED (1987) The effect of cross-spectrum correlation on the detectability of a noise band. J Acoust Soc Am 81:721–723

Dau T, Kollmeier B, Kohlrausch A (1997) Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. J Acoust Soc Am 102(5):2892–2905

Ernst S, Verhey JL (2005) Comodulation masking release over a three octave range. Acta Acust United Ac 91:998–1006

Ernst SMA, Verhey JL (2006) Role of suppression and retro-cochlear processes in comodulation masking release. J Acoust Soc Am 120(6):3843–3852

Fantini DA, Moore BCJ (1994) Profile analysis and comodulation detection differences using narrow bands of noise and their relation to comodulation masking release. J Acoust Soc Am 95(4):2180–2191

Gabor D (1946) Theory of communication. J IEE (London) 93:429–457

Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. Hear Res 47:103–138

Green DM, Swets JA (1966) Signal detection theory and psychophysics. Wiley, New York

Griffiths TD, Warren JD (2004) What is an auditory object?. Nat Rev Neurosci 5:887–892

Hall JW, Haggard MP, Fernandes MA (1984) Detection in noise by spectro-temporal pattern analysis. J Acoust Soc Am 76(1):50–56

Hall JW, Buss E, Grose JH (2006) Comodulation detection differences for fixed-frequency and roved-frequency maskers. J Acoust Soc Am 119(2):1021–1028

Hofer SB, Klump GM (2003) Within- and across-channel processing in auditory masking: a physiological study in the songbird forebrain. J Neurosci 23(13):5732–5739

Hohmann V (2002) Frequency analysis and synthesis using a gammatone filterbank. Acta Acust United Ac 88:433–442

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84(2):541–577

Klump GM, Nieder A (2001) Release from masking in fluctuating background noise in a songbird's auditory forebrain. NeuroReport 12(9):1825–1829

Köppl C, Yates G (1999) Coding of sound pressure level in the barn owl's auditory nerve. J Neurosci 19(21):9674–9686

Langemann U, Klump GM (2001) Signal detection in amplitude-modulated maskers. I. Behavioural auditory thresholds in a songbird. Eur J Neurosci 13:1025–1032

Las L, Stern EA, Nelken I (2005) Representation of tone in fluctuating maskers in the ascending auditory system. J Neurosci 25(6):1503–1513

Lawson JL, Uhlenbeck GE (1950) Threshold signals. McGraw-Hill, New York

McFadden D (1987) Comodulation detection differences using noise-band signals. J Acoust Soc Am 81(5):1519–1527

Meddis R, Delahaye R, O'Mard L, Sumner C, Fantini DA, Winter I, Pressnitzer D (2002) A model of signal processing in the cochlear nucleus: comodulation masking release. Acta Acust United Ac 88(3):387–398

Moore BCJ, Borrill SJ (2002) Tests of a within-channel account of comodulation detection differences. J Acoust Soc Am 112(5):2099–2109

Moore BCJ, Shailer MJ (1991) Comodulation masking release as a function of level. J Acoust Soc Am 90(2):829–835

Nelken I, Rotman Y, Yosef OB (1999) Response of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157

Neuert V, Verhey JL, Winter IM (2004) Responses of dorsal cochlear nucleus neurons to signals in the presence of modulated maskers. J Neurosci 24(25):5789–5797

O'Loughlin BJ, Moore BCJ (1981) Off-frequency listening: effects on psychoacoustical tuning curves obtained in simultaneous and forward masking. J Acoust Soc Am 69(4):1119–1125

Patterson RD, Nimmo-Smith I (1980) Off-frequency listening and auditory-filter asymmetry. J Acoust Soc Am 67(1):229–245

Pressnitzer D, Meddis R, Delahaye R, Winter IM (2001) Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. J Neurosci 21(16):6377–6386

Saunders JC, Ventetuolo CE, Plontke SKR, Weiss BA (2002) Coding of sound intensity in the chick cochlear nerve. J Neurophysiol 88:2887–2898

Schooneveldt GP, Moore BCJ (1987) Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. J Acoust Soc Am 82(6):1944–1956

Schreiner CE, Urbas JV (1988) Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hear Res 32(1):49–63

Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114(6):3394–3411

van de Par S, Kohlrausch A (1998) Diotic and dichotic detection using multiplied noise maskers. J Acoust Soc Am 103(4):2100–2110

Verhey JL, Dau T, Kollmeier B (1999) Within-channel cues in comodulation masking release (CMR): experiments and model predictions using a modulation-filterbank model. J Acoust Soc Am 106(5):2733–2745

Verhey JL, Pressnitzer D, Winter IM (2003) The psychophysics and physiology of comodulation masking release. Exp Brain Res 153:405–417

Verhey JL, Rennies J, Ernst SMA (2007) Influence of envelope distributions on signal detection. Acta Acust United Ac 93(1):115–121

Weisstein EW (2002) CRC concise encyclopedia of mathematics. Chapman & Hall/CRC, London

Wright BA (1990) Comodulation detection differences with multiple signal bands. J Acoust Soc Am 87(1):292–303