Mechanisms of within- and across-channel processing in comodulation masking release

Tobias Piechowiak[1], Stephan D. Ewert[2] , Torsten Dau[1]

[1] *Centre for Applied Hearing Research, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark*

[2] *Carl von Ossietsky Universität Oldenburg, Medizinische Physik D-26111 Oldenburg, Germany*

The audibility of a target sound embedded in another masking sound can be improved by adding sound energy that is remote in frequency from both the masker and the target. This effect is known as comodulation masking release (CMR) and is observed when the remote sound and the masker share coherent patterns of amplitude modulation. Most ecologically relevant sounds, such as speech and animal vocalizations, have coherent amplitude modulation patterns across different frequency regions, suggesting that the detection and recognition advantages conveyed by such coherent modulations may play a fundamental role in our ability to deal with natural complex acoustic environments. While a large body of data has been presented, the mechanisms underlying CMR are not clear. This study proposes an auditory processing model that accounts for various aspects of CMR. The model includes an equalization-cancellation (EC) stage for the processing of stimulus information across the audio-frequency axis. The EC process, which is conceptually similar to the across-ear processing in binaural models, is assumed in the model to take place at the output of a modulation filterbank stage for each audio-frequency channel. This approach has been proven successful in several basic conditions of CMR (Piechowiak *et al*., 2007). In the present study, a modified version of the model is tested that includes a non-linear cochlear filtering stage, the dual resonance nonlinear filterbank (DRNL). It is investigated to what extent the within and across-frequency processes contributing to CMR depend on cochlear nonlinear processing.

INTRODUCTION

The detection of a sinusoid masked by a noise masker becomes easier when a noise masker with coherent amplitude modulation is placed remotely from the on-frequency masker. This enhancement in detectability is a phenomenon called comodulation masking release (CMR). CMR has initially been demonstrated by Hall *et al*. (1984a). It can be obtained in two classes of paradigms. The first is to use a single band of noise as the masker, centred around the signal frequency, and to compare thresholds for modulated and unmodulated maskers as a function of the masker bandwidth ("bandwidening paradigm"; e.g., Hall *et al*., 1984a; Haggard *et al*., 1990; Schooneveldt and Moore 1989; Carlyon *et al*., 1989). In the unmodulated case, the signal threshold increases until the critical bandwidth is reached while the threshold decreases in the modulated

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.).
ISBN: 87-990013-1-4. Print: Centertryk A/S.

case. The second method is to use a masker consisting of several narrow noise bands, one at the signal frequency (on-frequency band) and one or more flanking bands spectrally separated from the on-frequency band (e.g., Hall *et al.* 1984a, 1990; Schooneveldt and Moore 1987). Although a variety of experiments have shown that the auditory system is able to exploit coherent fluctuations effectively and that a reduction in signal threshold can be achieved, the processes underlying CMR are still not fully understood. CMR has commonly been assumed to be based on comparisons of the outputs of different auditory filters. However, Schooneveldt and Moore (1987, 1989) showed that even when the bandwidth of the masker is smaller than the critical bandwidth of the corresponding filter, a large amount of CMR can be observed. It results from a change of the envelope statistic of the stimuli due to the addition of the signal to the masker (e.g., Schooneveldt *et al*., 1989; Verhey *et al*., 1999). The CMR obtained in the bandwidening-paradigm can most likely be explained by such within-channel processing. These cues are not involved when the on-frequency and flanking bands in the second paradigm are widely separated in frequency. In such a condition, an across-channel CMR of about 4 – 5 dB has been found (e.g., Schooneveldt and Moore, 1987). Piechowiak *et al.* (2007) introduced an across-channel mechanism based on the binaural equalization–cancellation (EC) process (Durlach, 1963) in order to account for this type of CMR (see Model section).

The first part of the present study investigates “across-channel” CMR in conditions with largely separated masker bands. The second part investigates conditions where contributions from across-channel as well as within-channel processing are expected. The relative contribution of both effects depends on the shape of the auditory filters, which is associated with the level-dependent nonlinear processing on the basilar membrane. For example, assuming broadened auditory filters, more of the “remote” masker activity would contribute to “within–channel” processing than in the case of narrower filters. In terms of modelling, the second part of the study considers the impact of the non-linear, level-dependent dual resonance nonlinear (DRNL) filtering stage on the predicted amount of CMR.

MODEL

The overall model structure is based on the detection model by Dau *et al.* (1997). It consists of a basilar-membrane (BM) stage, halfwave rectification and lowpass filtering (at 1 kHz), an adaptation stage, a modulation filterbank, and an optimal detector as the decision device. As in Piechowiak *et al.* (2007), an extended version of the original model was used in order to be able to account for across–channel CMR. This model is illustrated in Fig.1. The across–(peripheral) channel mechanism is assumed to take place at the output of the modulation filterbank (for all modulation filters except the lowest one that includes the dc component). The outputs of the individual modulation filters at the flanking band frequencies are subtracted from the corresponding outputs at the signal frequency. This process is denoted as cancellation in Fig 1. The outputs of the low-pass filters in the different peripheral channels remain unaffected. When more than two peripheral filters are considered, the weighted sum of the activity of the flanking filters is computed and subtracted from the on-frequency channel. Weighting here means that the output of the referring channel is scaled by its rms value. Calculating the weighted sum can be considered as the equalization process. The activity is then transferred to the optimal detector.

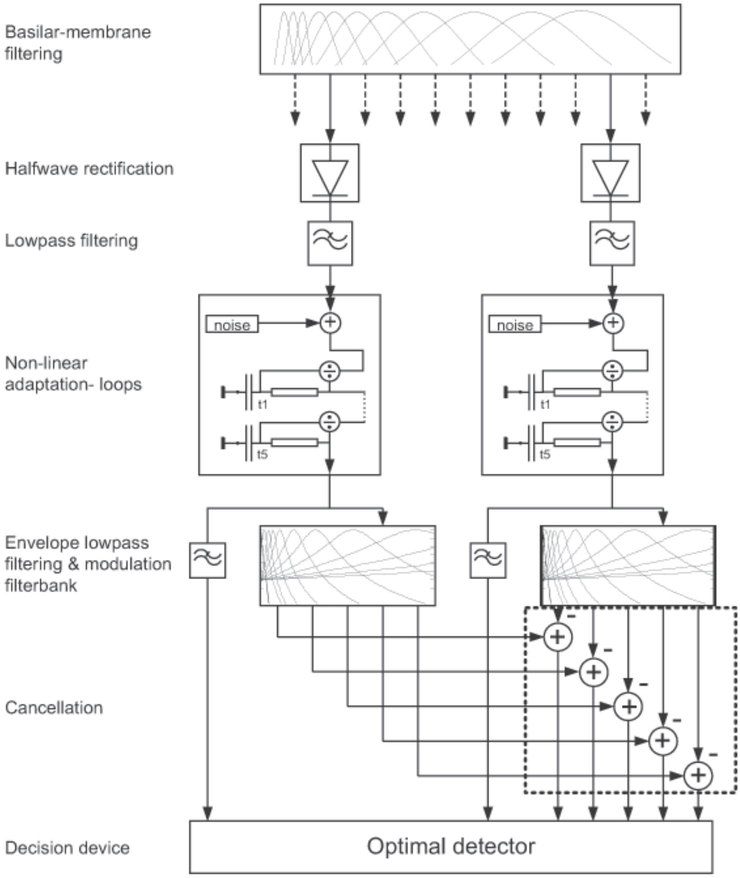

Fig. 1: Block diagram of the monaural across-channel model for CMR. The EC-mechanism is here exemplarily indicated for two auditory channels and is implemented at the output of the modulation filterbank. The activity at the flanking bands is subtracted from the corresponding activity at the on-frequency band (C-process).

In the present study, the effect of level dependent BM filtering is simulated using the DRNL filterbank (Meddis *et al.*, 2001). The results are compared with those obtained with the linear gammatone filterbank assumed in the original model implementation. The DRNL consists of two parallel processing pathways, one linear and the other one compressive non-linear. The output of the filter represents the sum of the outputs of the nonlinear and linear part. For illustration, the frequency response of a DRNL filter at 2 kHz to isointensity pure tones is shown in Fig. 2 (dashed line) together with the corresponding gammatone transfer function (solid line). At the chosen intensity of

60 dB SPL, the response of the DRNL filter is strongly asymmetric while the gammatone filter is symmetric.

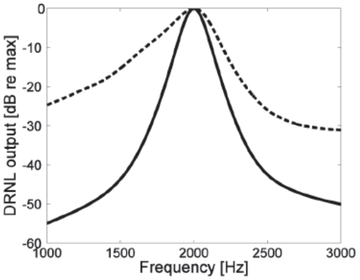

Fig. 2: Intensity transfer function in response to an isointensity pure tone of a gammatone filter (solid line) and a DRNL filter (dashed line) with a centre frequency of 2000 Hz at a level of 60 dB. At this intensity the response for the DRNL filter is asymmetric compared to the symmetric response of the gammatone filter.

EXPERIMENT 1: COMODULATION MASKING RELEASE (CMR) WITH FOUR BANDS OF NOISE

Method

Four normal hearing listeners from 24 to 40 years participated in the test. All had extensive prior training in psychoacoustical listening tests. An adaptive, three-interval, three-alternative forced-choice (AFC) procedure was used in conjunction with a 2-down 1-up tracking rule to estimate the 70.7% correct point of the psychometric function. The initial step size was 4 dB, which was reduced to 2 and 1 dB after the second and fourth reversals, respectively. Threshold was defined as the mean of the levels at the last six reversals of a threshold run. Four threshold estimates were obtained from each listener in each condition, final thresholds were calculated as the mean of these four estimates per subject.

Stimuli

The signal was a 1000-Hz tone with a duration of 187 ms. The masker consisted of five bands of noise centred at 250, 500, 1000, 2000 and 4000 Hz. Signal thresholds were determined for masker bandwidths of 25, 50, 100 and 200 Hz. The bands were generated in the time domain and restricted to the desired bandwidth in the frequency domain. The masker as well as the signal were windowed with 20 ms raised-cosine ramps and then presented diotically over Sennheiser HD580 headphones.

Results and discussion

Figure 3 shows the results of Experiment 1. Signal detection thresholds are plotted as a function of the masker bandwidth (left panel). Circles and squares show the thresholds for the uncorrelated and comodulated conditions, respectively. Open symbols represent the measured data and filled symbols show the model predictions using gammatone filters (gray filled) or DRNL filters (black filled). The right panel shows the amount of CMR, defined as the difference between the uncorrelated and comodulated conditions. Significant CMR of 4-5 dB is observed only for bandwidths of 25 Hz and 50 Hz [F1,22 = 38.59, p<0.001 and F1,22 = 32.18, p<0.001]. No significant CMR was observed for the larger bandwidths of 100 Hz and 200 Hz [F1,22 = 1.67, p=0.21 and F1,22 = 0.02, p=0.89] . Thus, CMR in this type of paradigm is restricted to bandwidths <= 50 Hz which is consistent with previous studies (e.g., Moore and Emmerich, 1990). This indicates that across-channel CMR is a phenomenon where the masker is dominated by slow envelope fluctuations. In the simulations, both types of filtering produce essentially the same results in this experimental condition. Significant CMR is observed for the bandwidth of 25 Hz and 50 Hz [F1,18 = 15.38, p<0.001 and F1,18 = 16.91, 0.001, respectively]. No significant CMR is obtained for 100 Hz and 200 Hz bandwidth [F1,18 = 6.48, p=0.02 and F1,18 = 6.29, p=0.02]. There is a good accordance between measured data and simulations although the model generates slightly higher thresholds than the data thresholds. No CMR at all is predicted when the EC process is not applied. This supports the hypothesis that the CMR in this type of paradigm results from "true across-channel" processing

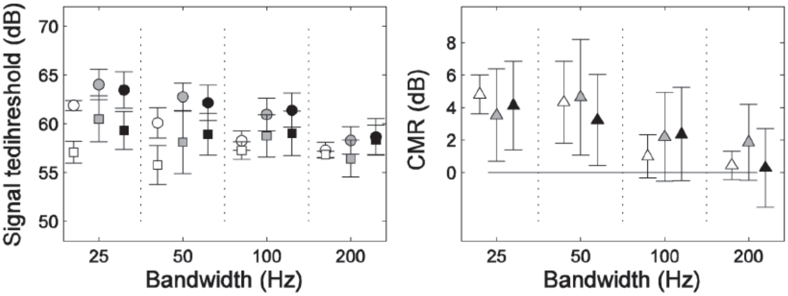

Fig. 3: Left panel: Detection thresholds for the 1-kHz tone in the presence of five noise bands as a function of the bandwidth of the noises. Open symbols indicate average experimental data and filled symbols show simulation results for gammatone filtering (light shaded symbols) and DRNL filtering (filled symbols). Circles and squares represent results for the uncorrelated and comodulated conditions, respectively. Right panel: CMR effect for the conditions of the left panel.

EXPERIMENT 2: COMODULATION MASKING RELEASE (CMR) WITH ONE FLANKING BAND VARYING IN FREQUENCY

Method

The setup and procedure were the same as in Experiment 1. The same subjects participated in this experiment.

Stimuli

The signal was a 2000-Hz tone with a duration of 400 ms. The masker consisted of one band of noise with a varying centre frequency of 1000, 1400, 1800, 1900, 2100, 2200, 2600 and 3000 Hz. The bandwidth of the masker was 25 Hz. Each level had an overall level of 67 dB. The bands were generated in the time domain and restricted to the desired bandwidth in the frequency domain. The masker as well as the signal were windowed with 10 ms raised-cosine ramps and then presented diotically over Sennheiser HD580 headphones.

Results and Discussion

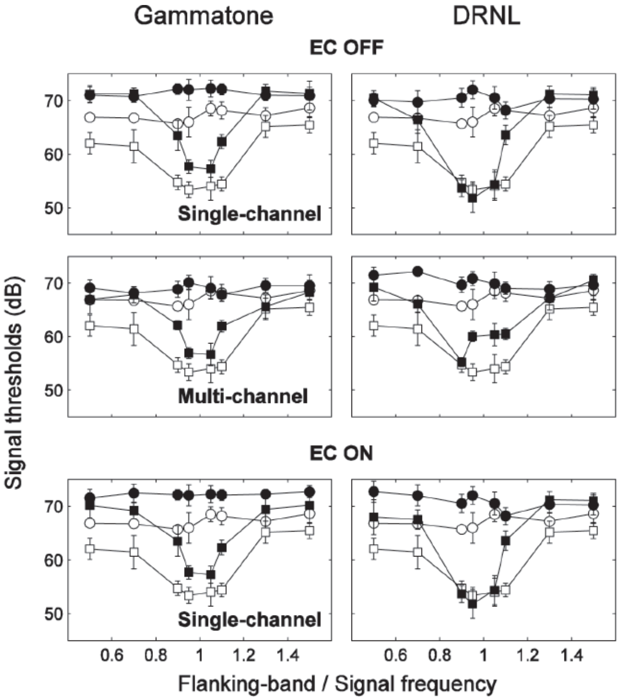

Fig. 4: Detection thresholds for the 2-kHz tone in the presence of two noise bands as the function of the spectral distance to the tone. Circles and squares represent results for the random and comodulated conditions, respectively. Open symbols show averaged experimental data while the left panels show the corresponding simulations for the gammatone filters (filled symbols left panel) and the right panels the simulations for the DRNL filters (filled symbols right panel). In the upper row, only the filter tuned to the on-frequency band was used with the EC process switched off. The middle panel shows simulations for a range of filters (one half octave below to half an octave above on-frequency band) with the EC process switched off. In the lower panel, the EC process was switched on and only the on-frequency band was used.

Figure 4 shows the results of the experiment. Signal thresholds are plotted as a function of the ratio between flanking-band and signal frequency. Squares denote thresholds when the masker bands are comodulated and circles when they are uncorrelated. In all panels, open symbols indicate average measured data. CMR is determined as the difference between uncorrelated and comodulated conditions. It reaches 12 – 14 dB when flanker and signal frequency are close to each other (ratios between 0.9 and 1.1). For large spectral separations between on-frequency and flanking band, the data show an asymmetry. CMR amounts to 3-4 dB at higher flanking-band frequencies and 5-6 dB for lower flanking - band frequencies. The data agree quite well with the results of Schooneveldt and Moore (1987). Filled symbols illustrate simulated thresholds. The left panels show the simulation thresholds using gammatone peripheral filtering. The right panels show the simulated thresholds for the case when the DRNL filters perform the peripheral filtering. (no new paragraph after this).

The model predicts slightly elevated thresholds as in the first experiment. The panels in the first row show simulations performed with only one peripheral filter centred at the signal frequency that was used for detecting the signal. For small spectral separations the model predicts a large amount of CMR that corresponds to that found in the experimental data. For the most remote spectral separations, no CMR is predicted. The model predicts lower thresholds in the comodulated case when the DRNL filters are used. This is caused by the larger bandwidth of the DRNL filters compared to the bandwidth of the gammatone filters. In the second row simulations were performed with five peripheral filters that were used for signal detection in order to introduce “off-frequency” listening into the model. They were centred at 1567, 1772, 2000, 2254 and 2537 Hz. Their centre frequencies were chosen in such a way that the gammatone filters had a spacing of one equivalent rectangular bandwidth. This resulted in a higher filter density of the DRNL filters because the DRNL filters’ magnitude responses were not overlapping at the 3 dB points as in the case for the gammatone filters. The simulated thresholds in this configuration are similar to those obtained in the single-channel case; a larger amount of CMR is predicted for small spectral separations and no CMR for relatively large spectral separations. The thresholds for the uncorrelated condition are slightly lower than in the single-channel case. When multiple DRNL filters are used, an asymmetry in the amount of CMR is observed where a larger amount of CMR is predicted at lower than at higher flanking band frequencies. This is consistent with the measured data; for example, the thresholds measured for a flanking-band frequency of 2600 Hz (rel. frequency 1.3) are not significantly different for the uncorrelated and comodulated condition whereas a CMR of around 6 dB is measured at a flanking band frequency of 1400 Hz (relative frequency 0.7). Since all filter outputs contribute independently to the detection process the observed asymmetry leads to the assumption that not necessarily the filter centred at the signal frequency is used for signal detection but a filter shifted towards lower frequencies where the within-channel cues are more salient. The bottom panels shows the simulations when the EC process was turned on. The filters that are contributing to the EC process should be statistically independent from the on-frequency filter channel. To achieve this, the correlation of the output of the respective off-frequency channel and the on-frequency filter channel was calculated, using broadband noise as the input signal. The limit of correlation, at which the filters were included for the EC process, was chosen to be 5%. In this configuration, only the filters with the relative frequencies of 0.5, 0.7, 1.3 and 1.5 contributed when gammatone filters were used. The thresholds that were not included in the across-channel process were the same as in the single-channel simulation. In the case of DRNL filters, the outcome was different. Due to the shallower slope of the DRNL magnitude transfer function towards lower frequencies, only the filters at relative frequencies of 0.5 and 0.7 were considered in the EC process. In this configuration, the model produced the largest agreement with the experimental data. Here, in contrast to the multi-channel configuration, the asymmetry is created by the weighted contribution of the filters to the EC process.

CONCLUSIONS

Following conclusions can be drawn from the results:

- A monaural auditory processing model was presented that accounts for comodulation masking release (CMR) in perceptual listening tests. The model distinguishes between contributions from within-channel processing and contributions resulting from across-channel processing. For the across-channel process, an equalization-cancellation stage was assumed, conceptually motivated by models on binaural processing.
- The model accounts for the main findings in the presented paradigms of CMR: (i) CMR with widely spaced flanking bands (where only across-channel processing does contribute), and (ii) CMR with one flanking band varying in frequency where within-channel processing dominates at close separations while across-channel processes determine thresholds at large separations.
- The simulation results support the earlier hypothesis that at least two different processes can contribute to CMR. Within-channel contributions can be as large as 15 dB. CMR resulting from across-channel process is robust but small (3-5 dB) and only observable at small bandwidths (below about 50 Hz) of the flankers.
- The current implementation of the model does include a nonlinear, level-dependent cochlear filtering stage which broadens the applicability of the model. The effect of a level-dependent frequency selectivity was investigated using DRNL filters instead of gammatone filters. It shows that an explicit across-channel process is needed to account for the outcome in some of the experimental conditions.

REFERENCES

Carlyon, R., and Stubbs, R. (**1989**). “Detecting single-cycle frequency modulation imposed on sinusoidal, harmonic, and inharmonic carriers,” J. Acoust. Soc. Am. **85**, 2563–2574.

Dau, T., Kollmeier B., and Kohlrausch, A. (**1997a**). “Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,” J. Acoust. Soc. Am. **102**, 2893–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). “Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,” J. Acoust. Soc. Am. **102**, 2906–2919

Durlach, N. I. (**1963**). “Equalization and cancellation theory of binaural masking-level differences,” J. Acoust. Soc. Am. **35**, 1206–1218.

Haggard, M. P., Hall, J. W., and Grose, J. H. (**1990**). “Comodulation masking release as a function of bandwidth and test frequency,” J. Acoust. Soc. Am. **88**, 113–118.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984a**). “Detection in noise by spectrotemporal pattern analysis,” J. Acoust. Soc. Am. **76**, 50–56.

Hall, J. W., and Grose, J. H. (**1990**). “Comodulation masking release and auditory grouping,” J. Acoust. Soc. Am. **88**, 119–125.

Meddis, R., O’Mard, L. P., and Lopez-Poveda, E. A. (**2001**). “A computational algorithm for computing nonlinear auditory frequency selectivity,” J. Acoust. Soc. Am. **109**, 2852–2861.

Schooneveldt, G. P., and Moore, B. C. (**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**, 1944–1956.

Schooneveldt, G. P., and Moore, B. C. (**1989**). “Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth and signal duration”, J. Acoust. Soc. Am. **85**, 273-281.

Verhey, J., Dau, T., and 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**, 2733–2745.