Modelling across–and within–channel mechanisms in comodulation masking release

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

[1] Centre for Applied Hearing Research, Danish Technical University, 2800 Kgs. Lyngby, Danmark,[tpi,, tda]@elektro.dtu.dk

[2] Medizinische Physik, Universität Oldenburg, 26111 Oldenburg, Germany, stephan.ewert@uni-oldenburg.de

Introduction

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) [1] and is observed when the remote sound and the masker share coherent patterns of amplitude modulation. While a large body of data has been presented, the mechanisms underlying CMR are not clear. Neuronal suppression at a cochlear level, the detection of modulation beatings within auditory channels, and across-channel comparisons of temporal envelope information have been suggested to contribute to CMR. The present study extends an earlier model that includes an equalization-cancellation (EC) stage for the processing of modulations across the audio-frequency channels by a non-linear peripheral filtering stage. In the framework of the model, the combination and interaction of three main mechanisms were assessed: (i) suppression, (ii) within-auditory-channel cues related to amplitude modulations, and (iii) across-auditory-channel processes at higher, retro-cochlear stages. Experiments are presented to examine the relative role of these mechanisms. In particular, the influence of level and effects of auditory grouping on CMR were investigated.

Model

The present study is based on an earlier model including an equalization–cancellation (EC) stage that accounts for within-channel as well as across-channel CMR [2]. The lower part of Fig. 1 illustrates the EC-type mechanism in the monaural modulation domain. The outputs of the modulation filters tuned to the same centre frequency (except the DC-component) are further processed by subtracting the flanking band is from the signal or on-frequency band (cancellation-process) leading to a "noise reduction" in the signal band. In this study, the model was extended by a non-linear peripheral processing stage indicated in the upper part of Fig. 1. The peripheral filtering is performed by a dual resonance non-linear (DRNL) filterbank according to Meddis *et al* [3]. The DRNL filter has been shown to account for suppression effects [4] and CMR in specific situations [5]. The output of a combined outer- and middle-ear filter is passed into the linear and nonlinear path of the DRNL filter. In the linear path, the signal is passed through a gammatone filter and a lowpass–filter whereas in the nonlinear path, the signal is processed by a gammatone filter, a compressive nonlinearity, a second gammatone filter and a linear amplification. Finally, the combined signal is expanded by squaring the output. This model is referred to as the CASP (Computational Auditory Signal Processing and Perception) model [6] in the following. Here the performance of the CASP model is validated by two paradigms that show the necessity of suppression as well as an EC stage.

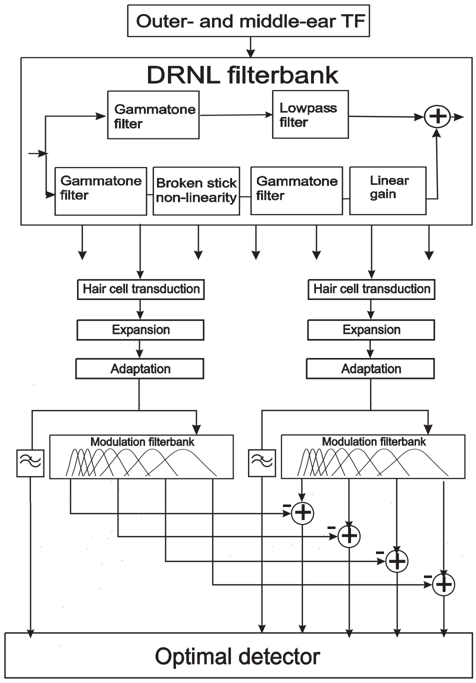

Figure 1: Block diagram of the non-linear across-channel perceptual model (CASP). The EC mechanism is incorporated for two auditory channels at the outputs of the modulation filterbank

Method

In the first experiment [5] the signal was a 2-kHz tone with a duration of 250 ms. The masker consisted of two 25-Hz wide multiplied noise bands of 500 ms duration. The on-frequency band (OFB) was centred at the signal frequency. The flanking band (FB) was placed at -3, -2, -1. and + 0.6 octaves relative to the signal frequency. Both masker and signal were gated with 50-ms raised-cosine ramps The level of the OFB was 20 dB SPL. The level of the flanking band was varied. The signal was temporally centred in the masker.

In the second experiment, the signal was a 1000-Hz pure tone with a duration of 187.5 ms. The masker consisted of 5 bands which were centred at 250, 500, 1000, 2000, and 4000 Hz covering a range of 4 octaves. The masker bands had the same duration as the signal. Masker and signal were gated with 20-ms raised-cosine ramps In the reference condition, the masker bands had random envelopes whereas in the comodulated condition, the maskers were frequency shifted versions of the masker centred at the signal frequency of 1000 Hz. The bandwidth of the maskers was 25 Hz. The level of each individual masker band (flanking bands and on-frequency band) was 35, 55, or 75 dB SPL.

CMR was calculated as the difference in threshold between the random and the comodulated condition. An adaptive, three-interval, 3-AFC procedure was used in conjunction with a 2-down 1-up tracking rule to estimate masked thresholds throughout all simulations and measurements.

Results

Figure 3 shows the measurement and model simulations for the first experimental condition [5]. The simulations here were carried out with a single-channel model**,** i.e. that no EC mechanism was assumed in the processing.

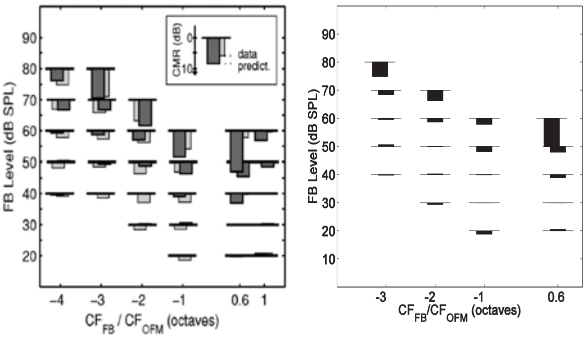

Figure 2: Left panel: Measured and predicted CMR for a temporal window model taken from [4]. Light gray shading denotes data, dark gray shading model predictions. Right panel: Predicted CMR for the current model shown in Fig. 1.

The tendency of higher CMR values occurring at larger differences in the level of the FB and the on-frequency band (increasing FB level) is shown by the temporal window model (left panel of Fig. 2) as well as by the CASP model (right panel of Fig.2). The CASP model in general shows slightly smaller CMR than the temporal window model. Significant differences between simulations and data are observed when the level of the flanking bands and the on-frequency band are identical. Figure 4 shows the results for the second experiment averaged over 5 subjects (open symbols).

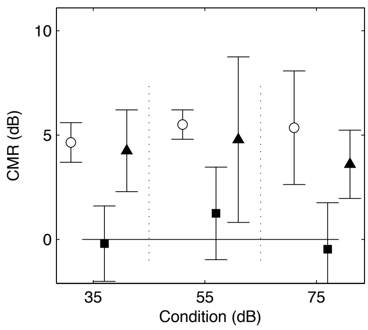

Figure 3: CMR for five subjects (open symbols) and for two model configurations at different masker levels. Closed squares are the CMR for the model without EC mechanism and the closed triangles are for the model with EC mechanism.

The subjects show a constant amount of CMR of around 5 dB for all masker levels. The model without an EC mechanism does not account for this behaviour. The predicted CMR does not significantly deviate from zero. Predictions for the model including an EC mechanism are in good agreement with the data for all levels.

Discussion

It has been shown [2] that the model described in Fig. 1 with a linear peripheral processing stage is able to account for within– and across–channel CMR in basic paradigms. When the levels of the on-frequency and flanking band differ significantly from each other, the suppression observed in the filter tuned to the signal (within-channel) seems to contribute to the overall amount of CMR. A level difference of 60 dB between the on-frequency and the flanking band leads to a CMR of about 9 dB (Figure 3). The CASP model, that includes a compressive nonlinearity in a peripheral processing stage accounts for the main findings in the data. However, the outcome of the second experiment suggests that a constant and robust amount of CMR of around 5 dB is observed over all masker levels as long as the levels among the bands (flanking and on-frequency) are equal. This cannot be explained within the model by a single frequency-channel processing even with a nonlinear peripheral processing. Therefore an across–channel processing remains a necessity.

Summary and conclusions

At least three different mechanisms contribute to the overall amount of CMR measured in various experimental settings:

- Within–channel cues for a near spectral separation of masker energy.
- Within–channel suppression for near and remote separation of flanking bands having different (higher) levels.
- Across–channel cues for remote separation of flanking bands having equal level.

References

[1] Buus, S.: *Release from masking caused by envelope fluctuations* J. Acoust. Soc. Am. **78** 1958-1965

[2] Piechowiak, T., Ewert, S. D., Dau, T.: *Modeling comodulation masking release using an equalization – cancellation mechanism* J. Acoust. Soc. Am. **121** 2111-2126

[3] Meddis, R. J., and O'Mard, L. P., Lopez-Poveda E. A.: *A computational algorithm for computing nonlinear auditory frequency selectivity* J.Acoust. Soc. Am. **109** 2852-2861

[4] Plack C. J., Oxenham, A. J.: *Linear and nonlinear processes in temporal masking* Acta Acoustic united with Acoustica **88** 348-358

[5] Ernst E. M. A., and Verhey J.L.: *Role of suppression and retro-cochlear processes in comodulation masking release* J. Acoust. Soc. Am. **120** 3843-3852

[6] Jepsen M. L., Ewert, S. D., Dau, T.: *A computational model of human auditory signal processing and perception* J. Acoust. Soc. Am, in review