

# Within- and Across-Channel Processing in Auditory Masking: A Physiological Study in the Songbird Forebrain

Sonja B. Hofer<sup>1</sup> and Georg M. Klump<sup>2</sup>

<sup>1</sup>Technische Universität München, Lehrstuhl für Zoologie, 85748 Garching, Germany, and <sup>2</sup>Carl von Ossietzky Universität Oldenburg, FK 5, Arbeitsgruppe Zoophysiology und Verhalten, 26111 Oldenburg, Germany

Synchronous envelope fluctuations in different frequency ranges of an acoustic background enhance the detection of signals in background noise. This effect, termed comodulation masking release (CMR), is attributed to both processing within one frequency channel of the auditory system and comparisons across separate frequency channels. Here we present data on CMR from a study in field L2 of the auditory forebrain of the European starling (*Sturnus vulgaris*) using two 25-Hz-wide bands of masking noise that provide the opportunity to distinguish between within-channel and across-channel effects. Acoustically evoked responses were recorded from unrestrained birds via radio telemetry. The signal was a 800 msec pure tone presented at the most sensitive frequency of the units in a previously determined frequency-tuning curve (FTC). One band of masking noise was centered on the signal frequency while the flanking band of noise was presented either within the limits of the excitatory FTC (i.e., within the same frequency channel as the on-frequency masker) or in the suppression area of the FTC (i.e., in a separate channel). For flanking bands inside the excitatory FTC, signal detection thresholds based on the rate code were lower in noise maskers with identical envelope fluctuations (comodulated) than in maskers with uncorrelated envelopes resulting in a neural CMR of  $\sim 4$ –7 dB. For flanking bands inside the suppression areas, the neural CMR was reduced. Although the average neural CMR was below the behaviorally determined CMR, a subsample of between 11 and 26% of the recording sites resembled the behavioral performance.

**Key words:** auditory forebrain; bird; masking release; auditory scene analysis; envelope correlation; narrow-band noise

## Introduction

In everyday life, we perceive our acoustic environment as being composed of many interfering sounds originating from separate sources. In analogy to visual image analysis, Bregman (1990) has introduced the concept of auditory scene analysis to summarize the processes involved in our perception of separate sources in acoustic scenes. These processes are also active in the detection of a signal in a masking background of sounds from other sources. In the natural environment, the acoustic background is often characterized by synchronous envelope fluctuations over a range of frequencies (Langemann and Klump, 1994; Nelken et al., 1999). This coherent modulation of the envelope (comodulation) is exploited by the auditory system to improve the detection of signals in acoustic scenes. Because signal detection thresholds are lower in comodulated background sounds than in masking sounds having no correlated envelope fluctuations, the effect has been termed “comodulation masking release” (CMR) (Hall et al., 1984). On the basis of psychophysical studies in humans, CMR has been attributed both to cues that are provided within one

frequency channel of the auditory system (e.g., as described by the auditory filter or critical band) and to cues that rely on the comparison across separate frequency channels (Hall et al., 1984; Buus, 1985; Moore, 1992). One type of experiment that allowed a clear distinction between the role of within- and across-channel cues used two separate narrow bands of masking noise (McFadden, 1986; Cohen and Schubert, 1987; Schooneveldt and Moore, 1987; Moore and Schooneveldt, 1990). One band was centered on the signal frequency (on-frequency band) and the second band (flanking band) could be freely positioned at any other frequency to stimulate the same auditory filter as the on-frequency band or a different auditory filter. If the on-frequency and the flanking band had identical envelope fluctuations, a masking release was observed in comparison with a stimulation with two bands of noise with uncorrelated envelopes. This effect was especially large if the two bands of noise were presented in the same auditory filter, but masking release could also be found if the noise bands stimulated well separated auditory filters. CMR is not limited to the human auditory system. Using two narrow bands of noise as the masker and a tonal signal, CMR has been demonstrated in psychoacoustic experiments in a small mammal, the Mongolian gerbil (*Meriones unguiculatus*) (Wagner and Klump, 2001), and in a songbird, the European starling (*Sturnus vulgaris*) (Klump et al., 2001). The evidence for CMR in behavioral studies in animals provides the unique opportunity to study its neurophysiological basis. Here we present results from a neurophysiological study in the auditory forebrain of awake Euro-

Received Oct. 30, 2002; revised March 24, 2003; accepted April 18, 2003.

The study was funded by a grant from the Deutsche Forschungsgemeinschaft (FOR 306). We thank Mark Bee for his comments on a previous version of this manuscript.

Correspondence should be addressed to Georg M. Klump, Carl von Ossietzky Universität Oldenburg, FB 7, AG Zoophysiology und Verhalten, Postfach 2503, 26111 Oldenburg, Germany. E-mail: Georg.Klump@uni-oldenburg.de.

S. B. Hofer's present address: Max-Planck-Institute for Neurobiology, Am Klopferspitz 18a, 82152 München-Martinsried, Germany.

Copyright © 2003 Society for Neuroscience 0270-6474/03/235732-08\$15.00/0

pean starlings using the described flanking-band paradigm. The direct comparison of neural response to previously published behavioral data in the same species provides the opportunity to determine whether the CMR observed in the neural response is sufficient to explain the size of the psychophysical effect, and in turn may provide for a better understanding of the mechanisms underlying CMR.

## Materials and Methods

**Preparation and recording.** The data were recorded from eight wild-caught adult European starlings, *Sturnus vulgaris*, of both sexes (four females and four males). Multiunit activity was recorded using two different types of electrodes: (1) epoxide-insulated tungsten electrodes with core diameters of 75  $\mu\text{m}$  and impedances ranging from 6 to 10 M $\Omega$  (FHC, Bowdoinham, ME) and (2) custom-built electrodes from Teflon-coated platinum-iridium wires (core diameter 25  $\mu\text{m}$ ; AM-Systems, Everett, WA). The latter were etched electrochemically (6 V AC) in an aqueous solution of  $\text{NaNO}_3$  (15%) and  $\text{NaOH}$  (5%) after exposing the tip by burning off the insulation. During the etching process the wire broke at a constriction leaving a sharp deinsulated tip. The wire was then carefully heated to partly reseal the insulation at the tip yielding electrodes with impedances ranging from 4 to 10 M $\Omega$  measured in a 0.9% NaCl solution. Amphenol integrated circuit sockets were used as connectors. Up to four electrodes were placed on a small head-mounted microdrive that was used to lower the electrodes to a maximal depth of 5 mm into the brain. Stainless-steel wires served as indifferent electrodes.

Surgery was performed under general anesthesia with 0.8–3% halothane after a subcutaneous injection of atropine (0.05 ml). The bird's head was fixed in a stereotaxic holder by ear bars with the beak inclined  $\sim 45^\circ$  below the horizontal plane. The skull was exposed and the electrodes were implanted into the right hemisphere through a small opening in the skull and dura positioned 1.8–2.3 mm rostral and 0.7–0.9 mm lateral to the caudal bifurcation of the sinus sagittalis. These stereotaxic coordinates were chosen to penetrate the input layer of the auditory forebrain (field L2; analogous to the mammalian primary auditory cortex). The indifferent electrodes were inserted through a second opening in the skull into the left rostral hemisphere. The microdrive, indifferent electrodes, and a small socket for attaching a radio transmitter were fixed to the skull with dental cement. Recordings started 4 d after surgery at the earliest.

Multiunit activity was recorded via radio telemetry (FM radio transmitter FHC type 40-71-1, 5.3 gm including battery). During the search for recording sites by lowering the electrodes via the microdrive, the bird was wrapped in a small jacket to prevent wing and leg movements. Once a recording site was identified by auditory-evoked neural activity, the bird was released into the experimental cage (35  $\times$  27  $\times$  52 cm) inside a sound-proof booth (IAC 403A, Industrial Acoustics, Niederkrüchten, Germany). The neural activity transmitted from the freely moving bird was received by a commercial FM tuner (Pioneer TX 970, Willich, Germany) via an antenna placed inside the booth. The signals were bandpass filtered (600–3600 Hz), amplified, digitized (Sound Blaster PCI128; 16 bit analog-to-digital, 44.1 kHz), and stored on the disc of a Linux workstation (500 MHz Celeron PC). Recordings were synchronized with the presentation of acoustic signals via the sound card (16 bit digital-to-analog, 44.1 kHz).

After termination of the experiments the birds were killed with sodium pentobarbital, and their brains were fixed with Zamboni's Reagent. Lesions created by the electrode tracts were identified in 50  $\mu\text{m}$  slices to determine the electrode position.

The care and treatment of the animals were in accordance with the procedures of animal experimentation approved by the Government of Upper Bavaria, Germany. All procedures were performed in compliance with the NIH Guide for the Care and Use of Laboratory Animals.

**Stimuli.** All stimuli were synthesized digitally by the soundcard in the workstation. Pure-tone signals and maskers were produced in two separate channels that were adjusted independently in level by computer-controlled attenuators (PA4; Tucker-Davis Technologies, Gainesville, FL). The channels were mixed in a high-fidelity amplifier (NAD 3100; NAD Electronics, Pickering, Ontario, Canada) and presented through a loudspeaker (Canton Twin 700, Canton Elektronik, Weilrod, Germany)

mounted at the ceiling of the booth  $\sim 50$  cm above the position of the bird's head. Additional sound-absorbing material (Plano50 plus Pyramide 100/100, absorption coefficient  $>0.99$  in the frequency range used; Illbruck, Leverkusen, Germany) lined the sound-proof booth. The sound field was calibrated using a sound-level meter (General Radio 1982 precision sound-level meter; GenRad, Concord, MA) and a condenser microphone (General Radio  $\frac{1}{2}$  inch microphone type 1982–9611) placed at 5 possible locations of the bird's head during the experiments.

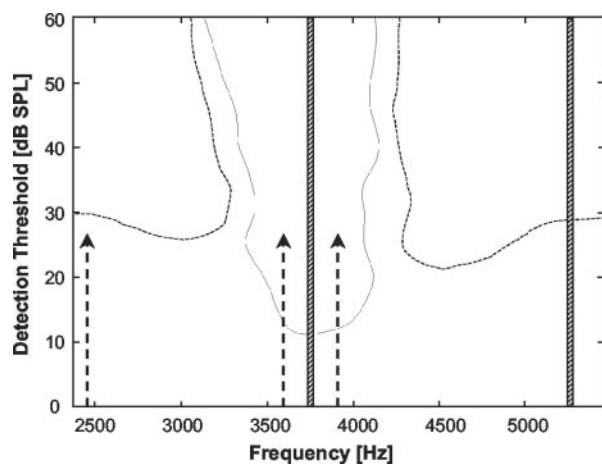
Before the masking experiment, the spectral tuning properties of each recording site were determined by analyzing responses to 200 msec tone bursts (including 10 msec Hanning ramps) of 165 different frequency-level combinations (11 frequencies at 15 sound pressure levels, ranging from 0 to 70 dB SPL in 5 dB steps). Signals were repeated at a rate of one per second.

The stimuli used for the masking experiment consisted of 800 msec tone signals (including 50 msec Hanning ramps, repeated every 2 sec) presented in 5 dB increments from 0 to 70 dB SPL in a continuous masker. The signal frequency was always specified as the most sensitive frequency [characteristic frequency (CF)] of the tuning curve of the recording site. The continuous masker consisted of either one or two noise bands with a bandwidth of 25 Hz. One noise band (the on-frequency band) was always centered on the signal frequency, whereas the other band (flanking band) had different frequency positions. In the first condition, the center frequency of the flanking-band was positioned at the lower or upper border of the auditory filter centered on the signal frequency [alternating between recording sites; bandwidth calculated after Buus et al. (1995)], which was always inside the excitatory frequency-tuning curve (FTC) of the recording site. In the second and third condition, the flanking band was centered on a frequency lying four auditory filter bandwidths above or below the signal frequency, respectively. For each condition the envelopes of the on-frequency and the flanking band were either correlated (i.e., comodulated) or uncorrelated. In a fourth condition, the reference condition, only the on-frequency band was presented as the masker. Each masker band was computed by multiplying low-pass-filtered noise with a cutoff at 12.5 Hz (successive 10 sec segments drawn randomly from a 30 sec noise computed *de novo* by an inverse Fourier transform for every repeated tone presentation) by a sinusoid with the desired center frequency. For the correlated condition, the sinusoids were multiplied by identical noise resulting in identical envelopes. In the condition with uncorrelated envelopes, a different low-pass noise was used for each band. The spectrum level of the noise bands was adjusted to the response characteristics of the recording site (i.e., 10–20 dB above the threshold of neurons driven by a 25-Hz-wide band of noise at the CF) and ranged from 0 to 42 dB SPL (average: 13.4 dB SPL). This level of stimulation by the masker alone was sufficient to provide masking and allowed detection of a further increase of the response when adding the tone signal (because the masker alone did not saturate the neural response).

**Data analysis.** Each tone signal was repeated 20 times, and the first 10 artifact-free presentations were included in the analysis. The computer registered an artifact in the recording if the level of the neural signal was more than four times above the level used to discriminate spikes. The multiunit responses were detected using a software window discriminator. Multiunit recordings from the seven recording sites with the highest signal-to-noise ratio ( $\geq 3:1$ ) were subjected to spike sorting based on Bayesian methods (Lewicki, 1994). A total of 27 single units could be isolated for further analysis using the same amplitude threshold used to analyze the multiunit recordings. There was good correspondence between the impulse rate of each multiunit recording and the added impulse rates of the sorted single units derived from it. Single-unit and multiunit data were analyzed separately using the same procedures.

The response to tones played to determine the tuning characteristics of the recording site was determined by the count of impulses detected during a 200 msec window delayed by the latency typical for the recording site (8–14 msec) relative to the onset of the tone presentation. The excitatory threshold was defined as the response rate lying  $1.5\times$  above the mean spontaneous rate of the recording site (determined in a 200 msec window preceding the tone). A decrease in the mean response rate below the spontaneous rate divided by 1.5 was defined as suppression of neural activity. Using these criteria, excitatory FTCs and suppressive flanking areas for each recording site were constructed (Fig. 1).

Activity elicited by tones in narrow-band maskers was determined by



**Figure 1.** Typical FTC of a recording site in area L2 with a central excitatory area (solid line) and a suppressive area (dashed line) at each side. The hatched bars show the spectral positions of one set of maskers used in the experiment: an on-frequency band positioned at the CF in the excitatory part of the FTC and a flanking band positioned in a suppressive area. The arrows indicate the alternative positions of the flanking band used in the other conditions. The test tone was always presented at the CF.

counting impulses in three different time windows of 100, 400, and 800 msec. The responses of the neurons during the first 100 msec of the signal presentation reflected mostly the phasic response characteristic, whereas the 400 msec time window beginning 200 msec after the start of the signal comprised only the tonic response to the tone signal. The 800 msec window included the whole signal-driven response. All time windows were shifted by the response latency of the unit. To determine the activity evoked by the masker alone, the level of the tone was attenuated maximally to a value well below the threshold in quiet, and identical analysis windows were used. The neural detection threshold for the signal in noise was reached when the mean response exceeded the mean spike rate elicited by the masker alone by 1.8 SDs. If no threshold could be calculated for a specific time window, the recording was excluded from the corresponding analysis, resulting in different sample sizes for the various paradigms. The magnitude of neural masking release was determined either as the difference between the detection thresholds in correlated and uncorrelated noise bands [CMR(U–C)] (Schooneveldt and Moore, 1987) or relative to the reference condition in which only the on-frequency band acted as the masker [CMR(R–C)]. Positive values of CMR indicate lower detection thresholds for tones in correlated noise bands.

## Results

Because the multiunit analysis was possible for a larger number of recording sites representing a bigger sample of auditory neurons, we first present multiunit data that we subsequently compare with the single-unit data under a special subheading.

### Tuning characteristics

For the current study, tuning characteristics of 24 multiunit clusters were analyzed. The knowledge of the spectral tuning properties was a prerequisite for the masking experiments, because we used the CF of each recording site as the signal frequency and placed the flanking bands in the frequency range of either the excitatory tuning curve or the suppressive sidebands. The primary-like response pattern and the histological verification indicated that all recording sites were located in the area of the input layer L2 of the bird auditory forebrain. The CFs ranged from 0.8 to 6.0 kHz, and the response threshold at CF (best threshold) varied between 2.0 and 17.2 dB SPL ( $8.2 \pm 4.1$  dB SPL, mean  $\pm$  SD; always provided in this format). The bandwidth of the excitatory tuning curve 10 and 40 dB above the threshold of

the neurons was 807 Hz ( $\pm 382$  Hz) and 1186 Hz ( $\pm 493$  Hz), respectively, corresponding to  $3.0 \pm 1.9$  and  $4.3 \pm 2.2$  auditory-filter bandwidths in the starling (Buus et al., 1995). Almost all recording sites (92%) had suppressive sidebands on one or both sides of the excitatory FTC. The best response threshold of the suppression areas was on average 14.3 dB ( $\pm 14.4$  dB) above the best threshold at CF. On average, suppression was observed 602 Hz ( $\pm 301$  Hz) below the CF and 874 Hz ( $\pm 534$  Hz) above the CF (frequencies measured 10 dB above the best suppressive threshold), corresponding to  $2.2 \pm 1.1$  and  $2.8 \pm 1.7$  auditory-filter bandwidths, respectively. After these analyses, we could select the appropriate stimulus paradigms for each recording site. On-frequency band and signal were positioned at the CF. Flanking bands positioned on the borders of the auditory filter centered at the CF were always inside the excitatory FTC. Because the flanking bands that differed by four auditory-filter bandwidths from the CF were supposed to lie inside suppression areas, recording sites without suppressive sidebands or with sidebands that were separated from the CF by more than four auditory-filter bandwidths were excluded from the masking experiments involving stimulation of suppressive areas.

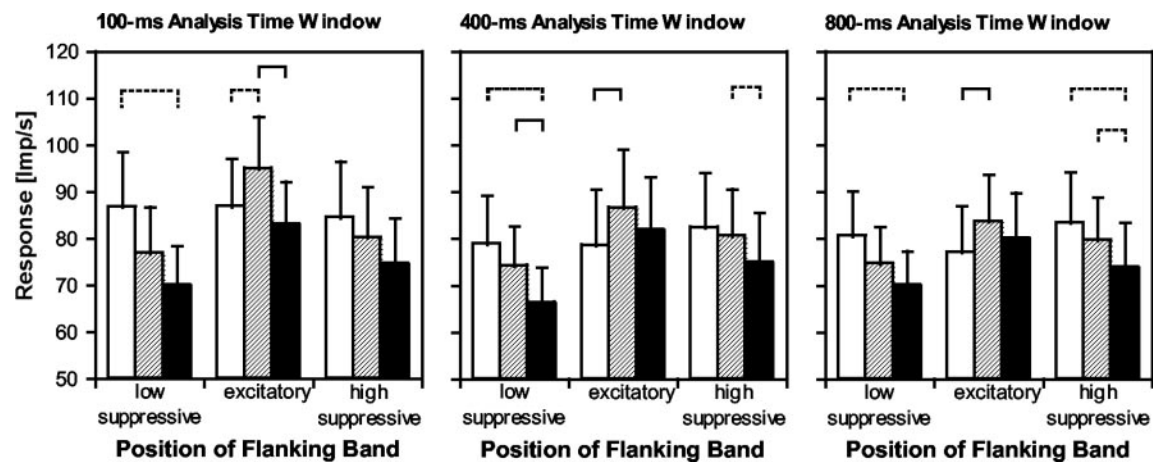
### Responses to the masker alone

The mean neural response rates to the different maskers, i.e., on-frequency band alone (reference condition, white bars), on-frequency band plus uncorrelated flanking band (hatched bars), and on-frequency band plus correlated flanking band (black bars) are shown in Figure 2 together with the results of a Wilcoxon test comparing the responses with the different maskers. The response in the reference condition is evaluated separately for each type of flanking band (positioned in the low suppressive, excitatory, or high suppressive area of the FTC), because the samples of recording sites differed. For all analysis-time windows, additional flanking bands in the excitatory area evoked significantly higher impulse rates than the on-frequency band alone if they were uncorrelated but not if they were correlated. There was no significant difference in any time window between the response to the on-frequency band alone and the slightly weaker response observed when adding uncorrelated flanking bands in the suppressive areas. Adding flanking bands with correlated envelopes resulted in significantly lower response rates than in the reference condition in four of the six possible conditions (i.e., upper and lower suppression area each for three time windows). The correlated-noise response was always weaker than the uncorrelated-noise response. For flanking bands in the excitatory FTC, this difference was significant only for the 100 msec analysis window. For flanking bands in the suppressive areas, however, a significantly lower impulse rate or a trend toward a reduced response to correlated versus uncorrelated noise bands was found only in the two longer 400 and 800 msec time windows and not in the 100 msec time window.

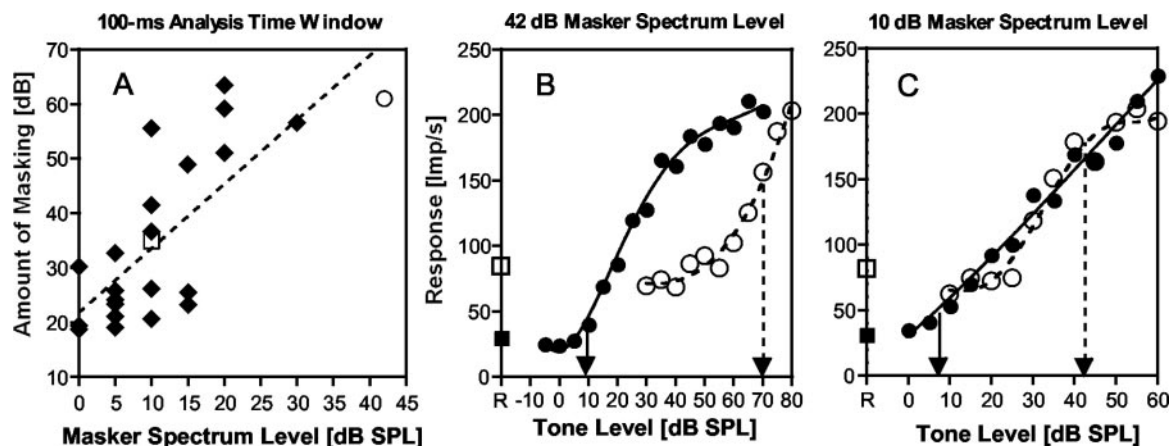
### Amount of masking

The detection of the signal was impaired when presented inside the continuous noise bands. The difference between the detection threshold of the CF tone in noise and the response threshold of the FTC at the CF provides a measure for the strength of masking (because we did not use identical stimulus durations, it is only a measure for the strength of masking and not the exact amount of masking). For all conditions, masking grew significantly with the increasing spectrum level of the masker, which varied between recording sites. Figure 3A demonstrates this correlation between amount of masking and level of the noise for the condition in which the signal was masked by two uncorrelated noise bands inside the excitatory FTC. The mean amount of





**Figure 2.** Mean neural spike rates elicited by the masker alone. Responses are shown for the on-frequency band alone (reference, white bar) and the on-frequency band presented together with an uncorrelated (hatched bar) or correlated flanking band of noise (black bar). Flanking bands were presented either within the excitatory frequency-tuning curve or in the upper or lower suppressive areas of the frequency-tuning curve. Because sample sizes differed among the various conditions (flanking bands in the low suppressive, excitatory, and high suppressive areas), the appropriate data set describing the response in the reference condition is presented along with the data obtained when stimulating with the additional flanking band. The size of the analysis-time window is indicated above each graph. The error bars indicate SEM. The brackets above the bars indicate significant differences (solid line,  $p < 0.01$ ; dotted line,  $p < 0.05$ ; Wilcoxon test).



**Figure 3.** Amount of masking (i.e., difference between signal threshold in noise and threshold at CF in the FTC determined in quiet) in relation to the level of the masker. *A*, Diamonds show multiunit data from all recording sites for the condition in which two uncorrelated noise bands inside the excitatory FTC served as the masker and a 100 msec integration-time window adjusted by the latency of the neuron was used to determine thresholds in the masked condition. The open circle and quadrangle show the data from recording sites for which rate-intensity functions are shown in *B* and *C*, respectively. The dashed line shows the linear regression. Parameters describing the regression for different integration-time windows are listed in Table 1. *B*, *C*, The response rate of the neurons in relation to the level of the tone. Filled circles show responses to the tone alone, and the solid line indicates a fit used for threshold estimation; the arrow points to the threshold in quiet. The filled quadrangle indicates the activity of the neurons without presentation of the tone (reference for threshold estimation *R*); the open circles show responses to the tone in noise determined in the first 100 msec of the response, and the dashed line indicates a fit used for threshold estimation; the arrow points to the tone threshold in noise. The open quadrangle indicates the activity of the neurons when stimulated with noise alone (reference for threshold estimation *R*).

masking and parameters describing the correlations are shown in Table 1 for all masking conditions and time windows. The slope of the regression line indicates that on average masking grows by  $1.0 \pm 0.2$  dB per 1 dB increase in spectrum level. The mean amount of masking in the different conditions was at least 20 dB, indicating a sufficient signal masking. As is demonstrated by the quadrangles in Figure 3, *B* and *C*, the masking noise increased the firing rate of the neurons compared with the spontaneous activity without any stimulation. Because of this increase in baseline activity, the tone threshold of the neurons in noise is higher than the tone threshold in quiet.

#### Detection thresholds and CMR

For the two types of maskers (correlated and uncorrelated), different signal-detection thresholds are found. Figure 4 provides an example of single-unit data. The raster plots show that the tones presented in

the continuous noise masker elicit a phasic-tonic response pattern with an offset suppression being typical for neurons in the input area L2 of the starling's auditory forebrain. A response to the tone can be observed at a lower SPL of the tone in the correlated condition (Fig. 4*B*) than in the uncorrelated condition (Fig. 4*A*). The rate-intensity function for the cell (Fig. 4*C*) indicates a masking release of 10 dB. The multiunit data show similar patterns, but response rates are higher because of the fact that the spikes of the individual units add up. All detection thresholds of the CF tones in the different masking paradigms are presented in Figure 5 for the multiunit data. The pattern of masked signal-detection thresholds resembled the pattern of activation by the different maskers alone. When both the on-frequency band and the uncorrelated flanking band excited the recording site, the detection thresholds for the signal were significantly higher than in the reference condition with the on-frequency band alone (Wilcoxon test;  $p < 0.02$ ; all analysis-time windows). The ad-

**Table 1.** Amount of signal masking (difference between the masked threshold and the best threshold at the CF in the FTC) and parameters describing the correlation between the amount of masking and the spectrum level of the masker

Analysis-time window	Masker stimuli	Amount of masking (mean $\pm$ SD)	Slope of regression line	$R^2$	$p$	$n$
100 msec	Reference	29.8 $\pm$ 15.4	1.1	0.49	<0.001	23
	U excitatory	35.5 $\pm$ 15.5	1.2	0.58	<0.001	23
	C excitatory	30.2 $\pm$ 15.9	1.0	0.37	0.002	23
	U high suppressive	33.0 $\pm$ 16.1	1.3	0.61	<0.001	16
	C high suppressive	27.9 $\pm$ 14.6	1.2	0.66	<0.001	16
	U low suppressive	30.2 $\pm$ 16.1	1.2	0.51	0.001	18
	C low suppressive	28.0 $\pm$ 15.6	1.3	0.57	<0.001	18
	Reference	23.2 $\pm$ 15.7	0.7	0.20	0.062	20
400 msec	U excitatory	36.2 $\pm$ 15.8	0.8	0.31	0.011	20
	C excitatory	29.0 $\pm$ 17.7	0.8	0.21	0.043	20
	U high suppressive	32.5 $\pm$ 16.3	0.9	0.25	0.047	16
	C high suppressive	27.7 $\pm$ 16.5	1.2	0.47	0.003	16
	U low suppressive	31.8 $\pm$ 15.7	0.9	0.42	0.003	19
	C low suppressive	28.4 $\pm$ 14.2	0.7	0.31	0.014	19
	Reference	20.3 $\pm$ 13.6	0.9	0.37	0.003	24
	U excitatory	29.7 $\pm$ 15.4	1.0	0.48	<0.001	24
800 msec	C excitatory	25.8 $\pm$ 15.2	0.7	0.22	0.021	24
	U high suppressive	24.7 $\pm$ 13.7	0.8	0.34	0.014	17
	C high suppressive	21.4 $\pm$ 15.7	1.2	0.58	<0.001	17
	U low suppressive	27.7 $\pm$ 18.1	1.2	0.45	0.001	21
	C low suppressive	27.2 $\pm$ 17.1	1.0	0.38	0.003	21
	Reference	20.3 $\pm$ 13.6	0.9	0.37	0.003	24
	U excitatory	29.7 $\pm$ 15.4	1.0	0.48	<0.001	24
	C excitatory	25.8 $\pm$ 15.2	0.7	0.22	0.021	24

U, uncorrelated noise bands; C, correlated noise bands

dition of the uncorrelated flanking band in the suppressive areas on either side of the CF, however, caused no significant change in the masked signal thresholds. When comparing the detection thresholds in two correlated noise bands with the reference condition, no significant differences could be found in any of the flanking-band positions and analysis-time windows. Thus, no significant CMR(R–C) can be found in the L2 forebrain neurons. Figure 6 shows the amount of CMR(R–C) for the different positions of the flanking band. There is a small negative CMR(R–C) when the flanking band provides additional excitation and mostly a positive CMR(R–C) when the flanking band in the high or low suppression area does not lead to additional excitation, but even the larger mean amounts of CMR(R–C) determined with 400 and 800 msec analysis-time windows in the condition with the flanking band in the upper suppression area are far from being significant (all  $p > 0.1$ ; Wilcoxon test). The distribution of the neural CMR(R–C) was very broad, however, and between 13 and 25% of the recording sites showed a CMR(R–C) resembling or exceeding the behavioral results (Klump et al., 2001) for the different conditions and analysis-time windows.

If the two noise bands were presented inside the excitatory FTC, for all analysis windows the detection thresholds of the tone were significantly lower in correlated noise bands than in uncorrelated noise bands. This demonstrates a significant neural release from masking resulting from comodulation [CMR(U–C)]. The average amounts of CMR for the different flanking-band positions and time windows are shown in Figure 7, and detailed information describing the variation of the data is provided in Table 2. The mean amount of CMR(U–C) for flanking-band positions inside the excitatory FTC ranged from 4.0 to 7.2 dB for the different time windows. In the condition with the flanking band in the upper spectral suppression area, the mean amount of CMR(U–C) was similar or slightly smaller than in the condition with an excitatory flanking band. However, the difference between the detection thresholds in the uncorrelated and correlated noise bands was significant only for the shortest analysis-time window (Table 2). When the flanking band was placed in the suppression area below the CF, the CM-

R(U–C) was considerably smaller than in the condition with an excitatory flanking band, and none of the differences in the thresholds between uncorrelated and correlated noise bands was significant.

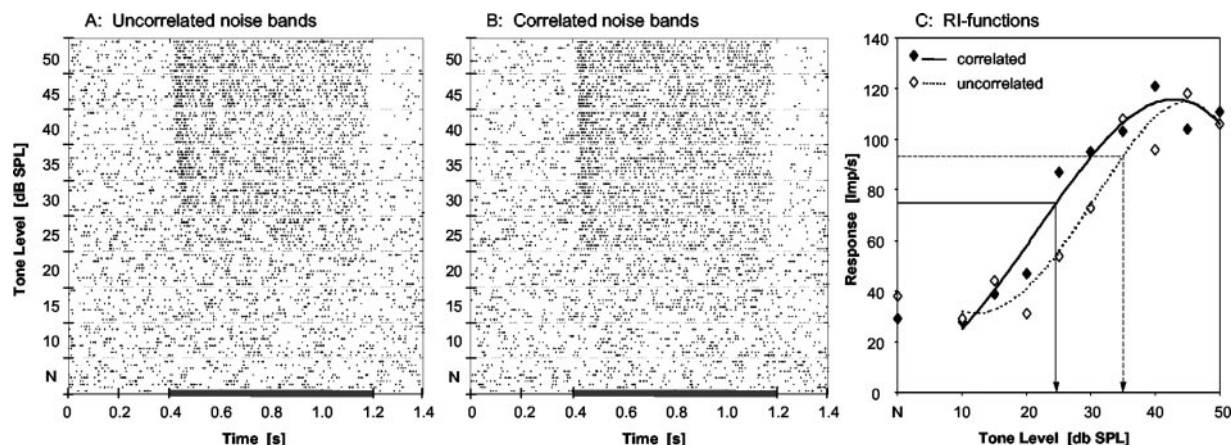
Although the different analysis-time windows possibly represent different components of the response (phasic versus tonic response components or the total response), we observed no significant effect of the size of the analysis-time window on the amount of CMR(U–C). Furthermore, we also found no significant interactions between the size of the time window and the condition on the amount of CMR(U–C) (two-way repeated-measures ANOVA). With two exceptions in which the results were nearly significant, the detection thresholds in the different conditions depended on the size of the analysis-time window (data are shown in Fig. 5; all  $p < 0.05$  in a Friedman test with the exception of two  $p$  values of 0.059 for the thresholds obtained with the on-frequency band alone and two correlated bands in the excitatory area). Detection thresholds tended to decrease with increasing anal-

ysis-time windows (Fig. 5).

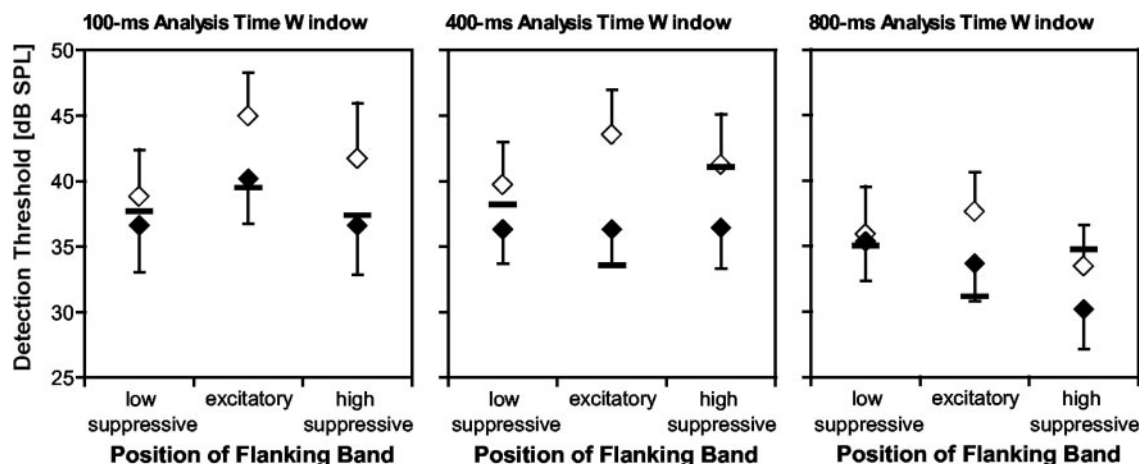
To evaluate the relationship between CMR and masker-driven activity, we correlated the CMR(U–C) with the difference in activity elicited by uncorrelated and correlated noise bands. For the condition with the flanking band in the lower suppression area, the CMR(U–C) was positively correlated with the difference in activity (Spearman rank correlation; 100 msec:  $r_s = 0.50$ ,  $p = 0.034$ ; 400 msec:  $r_s = 0.54$ ,  $p = 0.026$ ; 800 msec:  $r_s = 0.54$ ,  $p = 0.011$ ). A similar positive correlation was found only for the 800 msec time window if the flanking band was presented in the upper suppression area (Spearman rank correlation; 100 msec:  $r_s = 0.33$ ,  $p = 0.21$ ; 400 msec:  $r_s = 0.06$ ,  $p = 0.83$ ; 800 msec:  $r_s = 0.54$ ,  $p = 0.024$ ) and if both the on-frequency band and the flanking band were located inside the excitatory FTC (Spearman rank correlation; 100 msec:  $r_s = 0.31$ ,  $p = 0.16$ ; 400 msec:  $r_s = 0.32$ ,  $p = 0.18$ ; 800 msec:  $r_s = 0.50$ ,  $p = 0.013$ ).

#### Comparison of single-unit and multiunit data

The tuning properties of the multiunit data closely resembled the underlying single units. On average, the CF of single units differed from that of the corresponding multiunit recording by only 0.8%. The deviations for the best threshold and bandwidth of the excitatory tuning curve 10 and 40 dB above the threshold of the neurons were on average 9.3, 4.6, and 5.5%, respectively. The differences for the characteristics of the suppressive areas (frequency borders of suppression areas and most sensitive suppression thresholds) ranged from 2.3 to 8.0%. When studied in the masking paradigm, single units and multiunits did not differ with respect to the masked thresholds and CMR. When comparing the sample of single-unit data ( $n$  ranging from 19 to 27) with the sample of multiunit data ( $n = 17$ ), excluding the multiunit recordings from which the single units were derived with a statistical test for independent samples (Mann–Whitney  $U$  test), we found no significant differences in threshold or CMR (only one  $p$  value  $< 0.05$  was found in 45 comparisons, which was not significant when applying the appropriate Bonferroni correction). The single-unit CMR(U–C) ranged from



**Figure 4.** Example of the response of a single unit to tones presented in uncorrelated (*A*) and correlated (*B*) bands of masking noise. Both narrow-band maskers were placed in the excitatory frequency-tuning curve of the cell. In the raster plots (dots indicate individual action potentials), 10 repetitions of the stimulus are shown at each tone level. The block marked "N" shows the response to the masker alone. The black bar at the lower axis indicates the time of tone presentation. *C* shows rate-intensity functions for the response in a 100 msec time window adjusted for the response latency of the neuron. Curves are fitted to the raw data, and thresholds are shown by the arrows that are computed as the sound pressure level at which the response of the cell is 1.8 SDs above the average response observed when stimulating with the masker alone. The masking release of the neuron, i.e., the difference between the tone thresholds in the uncorrelated and the correlated conditions, was 10 dB.



**Figure 5.** Mean detection thresholds of the signal in uncorrelated (open diamond) and correlated (filled diamond) noise bands in relation to the different positions of the flanking band. Detection thresholds for the reference condition for each corresponding sample of recording sites are indicated by the horizontal line. The error bars indicate SEM. The size of the analysis-time window is indicated above each graph.

4.2 to 7.3 dB for a flanking band in the excitatory area of the tuning curve, being significant for the 100 and 400 msec analysis-time window (100 msec:  $p = 0.017$ ; 400 msec:  $p = 0.03$ ; 800 msec:  $p = 0.33$ ; Wilcoxon test comparing masked thresholds in the uncorrelated and the correlated condition). The CMR(U–C) was smaller and never significant for flanking bands in the suppressive areas, ranging from  $-2.3$  to  $4.0$  dB. Similar to the multiunit data, the CMR(R–C) showed no consistent pattern and was never significantly different from zero.

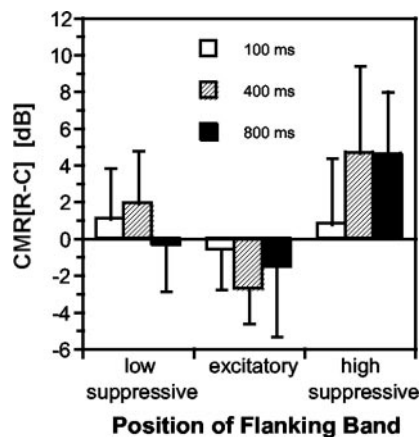
## Discussion

Here we demonstrate physiological CMR in the auditory forebrain of awake starlings using the flanking-band paradigm from psychophysical studies (McFadden, 1986; Cohen and Schubert, 1987; Schooneveldt and Moore, 1987). This study provides the unique opportunity to directly compare neural and behavioral performance from the same species (Klump et al., 2001). Because single-unit and multiunit data on CMR were very similar, we concentrate here on the multiunit data, which are based on a larger sample of recording sites.

Because psychophysical studies suggest that processing of the

acoustic input both within and across different auditory frequency channels contributes to CMR, it is necessary to relate the neural tuning properties to the CMR effect. The frequency-tuning characteristics (range of CFs, bandwidths, suppression) and phasic-tonic response patterns of the units that we recorded in layer L2 of the bird forebrain, an area that is analogous to the mammalian primary auditory cortex, were very similar to those reported by Nieder and Klump (1999) from the same area. The frequency tuning at the level of layer L2 resembles the tuning of peripheral auditory neurons and is closely related to psychophysical auditory filters (Buus et al., 1995; Nieder and Klump, 1999). The flanking bands that were presented 0.5 critical-band units (auditory filters) above or below the CF were always within the excitatory FTC, and therefore the masker components and the signal excited the same frequency channel providing within-channel cues. The flanking bands that were presented four critical-band units above or below the CF did not lie within the same excitatory FTC as the on-frequency band and the signal and thus provided only across-channel cues. All maskers clearly increased the thresholds of the CF signal compared with the best





**Figure 6.** Mean amount of CMR(R–C) in relation to the position of the flanking band. CMR(R–C) is the difference between the signal-detection threshold in the reference condition and in correlated noise bands. Error bars indicate SEM.

threshold in quiet, indicating masking even at the lowest noise level (0 dB spectrum level).

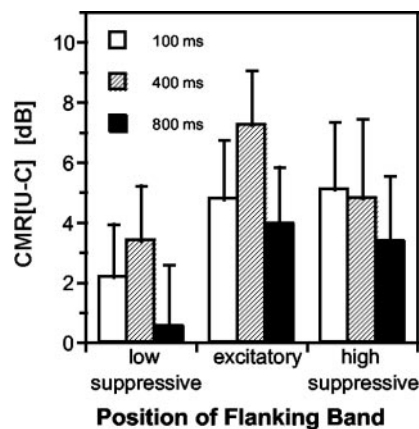
We used three different time windows to analyze the neural response to the signal in case different components of the response pattern would yield differing results for masking or masking release. Detection thresholds for signals in noise generally declined for longer analysis-time windows. This may be attributed to the fact that the ratio between the variance and the mean of the masker-driven activity was significantly decreased for longer time windows (all  $p < 0.002$ ; Wilcoxon test); however, the masking release did not depend on the analysis-time window.

A substantial neural CMR(U–C) of between 4.0 and 7.2 dB could be demonstrated for signal detection when both noise bands were placed inside the excitatory tuning curve, i.e., for a condition providing within-channel cues. Starlings showed a CMR of  $\sim 11$  dB with the identical stimulus paradigm in a psychophysical study (2 kHz signal, flanking band placed at the border of the auditory filter) (Klump et al., 2001). A corresponding difference between physiology and behavior can also be found in other studies of CMR in starlings and gerbils (Klump and Nieder, 2001; Nieder and Klump, 2001; Foeller, 2001). Parker and Newsome (1998) suggest that the behavioral performance could be represented by the response of the most sensitive neurons, which would require that the brain has selective access to these cells. On average,  $>20\%$  of the recording sites stimulated with excitatory maskers showed a CMR that was comparable with or even exceeded the behavioral release from masking (Table 2). These neurons could determine the behavioral performance.

**Table 2.** Several parameters describing CMR(U–C) and  $p$  values of the comparison between detection thresholds in correlated and uncorrelated noise (Wilcoxon test)

Analysis-time window	Masker stimulus	Mean $\pm$ SD	CMR(U–C) (dB)		% Units with CMR more than or equal to behavioral CMR <sup>a</sup>	$p$	$n$
			Maximum	Minimum			
100 msec	Excitatory	$4.8 \pm 9.6$	35.2	–8.7	21	0.034	24
	High suppressive	$5.1 \pm 9.0$	26.7	–14.1	19	0.030	16
	Low suppressive	$2.2 \pm 7.5$	25.2	–9.6	11	0.372	18
400 msec	Excitatory	$7.2 \pm 8.1$	26.1	–4.1	25	0.001	20
	High suppressive	$4.8 \pm 10.6$	25.0	–12.8	19	0.121	16
	Low suppressive	$3.4 \pm 7.5$	20.5	–7.5	26	0.147	19
800 msec	Excitatory	$4.0 \pm 9.3$	31.6	–15.8	21	0.049	24
	High suppressive	$3.4 \pm 9.0$	28.3	–6.7	18	0.287	17
	Low suppressive	$0.6 \pm 9.1$	13.3	–21.6	19	0.414	21

<sup>a</sup>Percentage of the recording sites showing an amount of CMR(U–C) that is equal to or exceeds the mean behavioral CMR(U–C).



**Figure 7.** Mean amount of CMR(U–C) in relation to the position of the flanking band. CMR(U–C) is the difference between the signal-detection threshold in uncorrelated and correlated noise bands. Error bars indicate SEM.

The condition with the flanking bands inside the suppression areas of the FTC represents a setting possibly providing across-channel cues. The mean amount of CMR(U–C) at the level of L2 was generally smaller for this condition and was significant in only one case (Table 2). Psychophysically, starlings show no decline in CMR(R(U–C)) for flanking bands placed outside the auditory filter centered on the signal (Klump et al., 2001). The psychophysical masking release was on average 9.4 and 13.1 dB for a flanking band centered four auditory filter bandwidths below and above the signal frequency, respectively. Nevertheless, also in this condition, between 11 and 26% of the recording sites were at least as sensitive as observed in psychophysics, thus providing a substrate for explaining the performance of the animal (Parker and Newsome, 1998). A similar relation between psychophysical and neurophysiological amounts of masking release has also been found in other studies of masking release in the starling (Klump and Nieder, 2001; Nieder and Klump, 2001). These studies also indicate that across-channel processing is less important on this level of the auditory system.

The observation that physiological patterns of excitation closely resemble masking patterns suggests that the amount of excitation provided by the masker determines masked thresholds. According to a study of cat auditory nerve fibers by Delgutte (1990), excitation produced by the masker dominates masking for signal frequencies near or below the masker frequency, and suppression affects the detection thresholds for signals with frequencies above the masker frequency, but only for high masker levels. Our results are consistent with the observations of Delgutte (1990), because we used very low masker levels and found no influence of suppression on the masking of the signal. One could argue that the noise level was too low to affect neural suppression, because the suppressive response thresholds in the FTC are higher than the excitation thresholds at CF, but also at moderate noise levels that were certainly above the suppression threshold we could find no influence of the flanking band in the suppressive area on the detection threshold of the signal and no increased CMR. In addition in the study by Nieder and Klump (2001), in which the level of the flanking bands was increased to account for the

higher suppression thresholds, maskers in suppression areas did not result in a substantial CMR.

Because uncorrelated maskers generally excited the neurons more than correlated maskers of the same acoustic energy, this difference in excitation could contribute to CMR(U–C). Although the correlations between the amount of CMR(U–C) and the difference in response strength support this view, the amount of variance explained by these correlations is negatively related to the mean CMR(U–C) (Spearman rank correlation;  $r_s = -0.755$ ,  $p < 0.02$ ). That is, the correlations were greater in the conditions in which only a small CMR was observed. Thus, CMR can be caused only partly by differences in activation by the correlated and uncorrelated maskers. Using a different stimulus paradigm, Klump and Nieder (2001) observed a release from masking without an underlying difference in activation.

Comodulation masking release can also be defined relative to the signal threshold measured only in the on-frequency band [CMR(R–C)]. Both starlings and humans show substantial CMR(R–C) in psychophysical studies (Schooneveldt and Moore, 1987; Klump et al., 2001). In the present neurophysiological study, in which the on-frequency band alone had the same spectrum level as the two correlated noise bands, no significant CMR(R–C) could be demonstrated. It appears that for the overall sample of neurons, CMR mainly compensates for an increased masking caused by the added flanking band [see also McFadden (1986) for a psychophysical parallel]. Nevertheless, for every condition a considerable subsample of recording sites showed a CMR(R–C) [as shown before for CMR(U–C)] providing a basis for the behavioral performance.

Physiological correlates of comodulation masking release have also been demonstrated in other species. Using wide-band maskers with 10 Hz sinusoidal or trapezoid modulations and large depth of modulation, Nelken et al. (1999) demonstrated a release from masking of between 10 and  $>30$  dB in the cat's primary auditory cortex. Using less regularly modulated maskers, however, Foeller (2001) observed a release from masking of only 6 dB in the gerbil primary auditory cortex. Pressnitzer et al. (2001) found neural CMR(R–C) in the ventral cochlear nucleus of guinea pigs with flanking components presented remote from the signal frequency to minimize within-channel cues. In primary-like responding cells, the mean amount of masking release was 2.4 dB, showing in the brainstem an effect quite similar to our results. The two studies are barely comparable, however, because Pressnitzer et al. (2001) used seven sinusoidally amplitude-modulated tones as maskers and 50 msec tone signals presented in the masker-dips. Verhey and Winter (2002) showed neural CMR in the ventral cochlear nucleus of guinea pigs with stochastic maskers (20 Hz noise bands) and 200 msec tone signals, but they provide no data on the mean amount of masking release. Pressnitzer et al. (2001) and Meddis et al. (2002) suggested a neural circuit for the brainstem in which transient inhibition explains the results of their CMR experiment. Because we used long signals and continuous statistically fluctuating maskers, we do not think that the model is directly applicable.

Summarizing our results in area L2 of the starling, the physiological findings on the basis of the rate code can explain the behavioral performance if we assume that the most sensitive neurons can be exploited selectively. If the average population response determines perception, however, the neural data fail to explain some aspects of the behavioral performance such as CMR(R–C) or the across-channel contributions to CMR. It remains to be shown whether the currently unexplained CMR effects are caused by processing in auditory areas postsynaptic to L2. On the other hand, the spike rate is only one source of information that the auditory system could exploit. The study by Nelken et al. (1999) using disruption of phase

locking to the masker envelope in the presence of the signal as the detection cue suggests that spike timing may provide another source of information. For example, the synchronized neural activity that results from comodulation of the noise could enable the auditory system to improve signal detection in fluctuating background noise.

## References

- Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: MIT.
- Buus S (1985) Release from masking caused by envelope fluctuations. *J Acoust Soc Am* 76:1958–1965.
- Buus S, Klump GM, Gleich O, Langemann U (1995) An excitation-pattern model for the starling (*Sturnus vulgaris*). *J Acoust Soc Am* 98:112–124.
- Cohen MF, Schubert ED (1987) Influence of place synchrony on detection of a sinusoid. *J Acoust Soc Am* 81:452–458.
- Delgutte B (1990) Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. *J Acoust Soc Am* 87:791–809.
- Foeller ER (2001) Mechanisms of inhibition and neuronal integration for signal processing in the primary auditory cortex of the Mongolian gerbil (*Meriones unguiculatus*). Doctoral dissertation, Ludwig-Maximilians-Universität München.
- Hall JW, Haggard MP, Fernandes MA (1984) Detection in noise by spectrotemporal pattern analysis. *J Acoust Soc Am* 76:50–56.
- Klump GM, Nieder A (2001) Release from masking in fluctuating background noise in a songbird's auditory forebrain. *NeuroReport* 12:1825–1829.
- Klump GM, Langemann U, Friebe A, Hamann I (2001) An animal model for studying across-channel processes: CMR and MDI in the European starling. In: Physiological and psychophysical bases of auditory function (Breebart DJ, Houtsma AJM, Kohlrausch A, Prijs VF, Schoonhoven R, eds), pp 266–272. Maastricht, The Netherlands: Shaker.
- Langemann U, Klump GM (1994) Acoustic perception in birds: can they profit from atmospheric turbulences to improve signal detection? *J Ornithol* 135:422.
- Lewicki MS (1994) Bayesian modeling and classification of neural signals. *Neural Comput* 6:1005–1030.
- McFadden D (1986) Comodulation masking release: effects of varying the level, duration and time delay of the cue band. *J Acoust Soc Am* 80:1658–1667.
- 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 Acustica* 88:387–398.
- Moore BCJ (1992) Across-channel processes in auditory masking. *J Acoust Soc Jpn* 13:25–37.
- Moore BCJ, Schooneveldt GP (1990) Comodulation masking release as a function of bandwidth and time delay between on-frequency and flanking band maskers. *J Acoust Soc Am* 88:725–731.
- Nelken I, Rotman Y, Yosef OB (1999) Responses of auditory-cortex neurons to structural features of natural sound. *Nature* 397:154–157.
- Nieder A, Klump GM (1999) Adjustable frequency selectivity of auditory forebrain neurons recorded in freely moving songbird via radiotelemetry. *Hearing Res* 127:41–54.
- Nieder A, Klump GM (2001) Signal detection in amplitude-modulated maskers: II. Processing in the songbird's auditory forebrain. *Eur J Neurosci* 13:1033–1044.
- Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–277.
- Pressnitzer D, Meddis R, Delahaye R, Winter IM (2001) Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. *J Neurosci* 21:6377–6386.
- 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* 82:1944–1956.
- Verhey JL, Winter IM (2002) The effect of random maskers on comodulation masking release in the cochlear nucleus. Abstract presented at 25th Annual Midwinter Research Meeting of the Association for Research in Otolaryngology, St. Petersburg, FL, January.
- Wagner E, Klump GM (2001) Comodulation masking release in Mongolian gerbils (*Meriones unguiculatus*) studied with narrow-band maskers. In: Elsner N, Kreutzberg GW (eds) Göttingen Neurobiology Report 2001. Thieme-Verlag, Stuttgart New York, p 415.