Exp Brain Res (1999) 124:311–320 © Springer-Verlag 1999

RESEARCH ARTICLE

Andreas Nieder · Georg M. Klump

Time course of simultaneous masking in the starling's auditory forebrain

Received: 20 April 1998 / Accepted: 16 September 1998

Abstract Simultaneous masking of pure tones was studied in the primary auditory forebrain of a songbird species, the European starling (*Sturnus vulgaris*). The responses of 32 multi-unit clusters in the input layer of the auditory neostriatum (field L2a) were recorded via radiotelemetry from freely moving birds. The probe was a 10-ms tone burst at the units' characteristic-frequency (CF) presented 20 dB above the threshold. The masker was an 80-ms tone burst presented either at the units' CF (excitatory masker) or at a frequency located in inhibitory side-bands (inhibitory masker) of the units' tuning curves. The probe was presented either 3 ms or 63 ms after masker onset. Probes presented at a 3-ms delay were influenced at significantly lower levels of an excitatory masker than probes presented at a 63-ms delay. The mean difference in masker level at the detection thresholds for both probe delays was 8 dB. No difference in masker level was observed for inhibitory-frequency maskers. The observed neural masking effects may be explained by at least four mechanisms: (1) swamping of the probe response by the response to the masker, (2) a reduction of the probe response during neural adaptation of the response to the masker, (3) a reduction of the probe response during side-band inhibition in the central nervous system, and (4) suppression originating in the cochlea.

Key words Two-tone masking · Overshoot · Auditory telencephalon · Radiotelemetry · Songbird

A. Nieder (✉) · G.M. Klump
Institut für Zoologie, Technische Universität München,
Lichtenberg Str. 4, D-85747 Garching, Germany

Present address: A. Nieder
Lehrstuhl für Zoologie/Tierphysiologie,
Institut für Biologie II, RWTH Aachen, Kopernikusstr. 16,
D-52074 Aachen, Germany
e-mail: nieder@corvus.bio2.rwth-aachen.de
Tel.: +49-241-805859, Fax: +49-241-8888133

Introduction

The term "simultaneous masking" refers to acoustic events when a signal that ought to be detected (called a "probe") is presented at the same time as a disturbing sound, the "masker." As a consequence, the threshold of audibility for the probe is raised by the masker, i.e., the detection threshold for the signal is increased in the presence of a masker (see overview in Moore 1989). Traditionally, the perceptual masking of one sound by another has been thought to be influenced only by the auditory filter centered at the signal frequency (i.e., the critical band; see Scharf 1970). Fletcher (1940) already attributed the spectral filtering in the auditory system to the frequency-tuning properties of the cochlea. These spectral filter properties are reflected, with some modification, in the excitatory frequency-tuning curves of neurons throughout the auditory system. Indeed, excitatory filter bandwidths of auditory neurons that were measured applying a masking paradigm matched results from behavioral measurements quite well both in the cat's inferior colliculus (Ehret and Merzenich 1985) and cortex (Ehret and Schreiner 1997) and in the starling's forebrain (Nieder and Klump 1999).

There is considerable evidence, however, that observers routinely make comparisons across auditory filters rather than listening only through a single filter (see review in Moore 1992). In addition, the patterns of temporal fluctuations of the masking noise both within and across auditory filters substantially influence the amount of masking (e.g., in a phenomenon called "comodulation masking release" which has been observed in both humans and the European starling; see Hall et al. 1984; Klump and Langemann 1995). Temporal effects in simultaneous masking within and across auditory filters have been found to be of great importance in this context (see McFadden and Wright 1992). The psychophysical "overshoot" effect is another phenomenon demonstrating temporal effects in masking (Elliott 1965; Zwicker 1965a, b; Fastl 1979). In an overshoot experiment, the masking of a brief signal by a simulta-

neous long-duration masker is found to be greatest when the onset of signal and masker coincide and decreases as the signal onset is delayed relative to that of the masker. The magnitude of the overshoot effect increases with increasing bandwidth of the masker (Zwikker 1965a; Bacon and Viemeister 1985) providing further evidence that both processes within one auditory filter and computations across several auditory filters may contribute to temporal masking effects (see also Wright 1997).

Despite the significance of temporal and spectral effects in simultaneous masking for understanding auditory signal detection in the natural environment, with its ubiquitous background noise, only a few studies in the auditory periphery have so far investigated the underlying neural mechanisms (Smith and Zwislocki 1971). To facilitate the interpretation of the experimental results, the neurophysiological studies used gated two-tone stimuli (a tonal masker and a tonal probe) in the experiments rather than spectrally and temporally complex signals. Here we present data on temporal effects in simultaneous masking from a neurophysiological study in the forebrain of a songbird species that has proven to be an excellent model for various aspects of the perception of masked signals (Klump and Langemann 1995; Langemann et al. 1995). We recorded multi-unit activity in awake and freely moving European starlings (*Sturnus vulgaris*) while presenting a probe with two different delay times relative to the masker onset in a two-tone masking paradigm. To evaluate spectral effects, we presented maskers both within the excitatory tuning curve of the cell cluster (i.e., at the characteristic frequency) and at an adjacent inhibitory frequency (i.e., a frequency close to the center of an inhibitory side-band). Presenting masker and probe at the same frequency within the limits of the excitatory tuning curve is analogous to a perceptual paradigm in which masker and probe stimulate the same auditory filter. Presenting the probe in the center of the excitatory tuning curve and the masker at the center of an inhibitory side-band should be analogous to a perceptual paradigm in which information from two separate auditory filters interact.

In the auditory periphery, two mechanisms are thought to affect the analysis of a probe tone masked by a second tone. Firstly the swamping (or line-busy) hypothesis assumes that the excitatory response to the masker creates a level of activity such that the additional excitatory response to the probe tone cannot produce any further increment in the response (Pickles 1984). Secondly, suppression of the probe-tone response by the masker may occur even if the masker alone does not elicit any excitation (Sachs 1969). According to the suppression hypothesis, the activity elicited by the probe will be reduced to a level similar to the spontaneous activity such that the probe tone cannot be detected. We will discuss whether these mechanisms are sufficient for explaining the observed patterns in forebrain units or if additional mechanisms in the central auditory system affect temporal effects in simultaneous masking.

Materials and methods

Preparation and recording

Experiments were performed on eight adult starlings (*Sturnus vulgaris*) of both sexes caught in the wild. The birds weighed between 70 and 99 g. A detailed description of the manufacturing of electrodes and the surgical procedures is given elsewhere (Nieder and Klump 1999). Briefly, polymide-insulated nickel-chrome resistance wires with core diameters of 17 µm were sharpened at the tip in a procedure similar to that described by Jacob and Krüger (1991). The impedance of the microelectrodes measured in 0.9% NaCl ranged from 300 kΩ to 1 MΩ (1000 Hz, a.c.). Several microelectrodes were attached to the sockets of an Amphenol IC-socket connector.

The birds were given atropine (0.05 ml) subcutaneously and anesthetized with 0.8–3% halothane. They were fixed in an stereotaxic holder and body temperature was maintained. A small hole was made in the skull 1.5–2.5 mm rostral and 0.8–1.1 mm lateral of the bifurcation of the sinus sagittalis. Stereotaxic coordinates were chosen to reach the input layer of the field L complex, L2a. Electrodes were inserted into the auditory neostriatum as bundles. After implantation of two indifferent electrodes, the whole array was fixed to the skull with dental cement and the wound was closed using tissue glue. In addition, a socket for the transmitter was glued onto the electrode array. After the birds awoke from anesthesia, they were carried to their home cages. Recordings started after 2 days or more but not before the animals had eaten. The care and treatment of these birds were in accordance with the procedures of animal experimentation approved by the Government of Upper Bavaria, Germany.

All multiple-unit data presented here were recorded via radiotelemetry from a freely moving bird. During the recording sessions, the bird was sitting in a cage (25 cm×53 cm×35 cm) inside a custom-built soundproof booth. The birds were habituated to the stimuli and the setup and usually sat calmly on the only perch, in the center of the cage, during recording. Food and water were provided *ad libidum*. A miniature FM radiotransmitter (type 40-71-1; Frederick Haer, USA) with a high-impedance input stage was attached to the bird's skull. Transmitted signals were filtered (bandpass 500–5000 Hz) and amplified. The recorded signal was digitized with a sampling rate of 32 kHz and stored on the disk of a Silicon Graphics Indy Workstation. Recordings containing artifacts caused by movement of the bird (i.e., signals with amplitudes that were more than twice as large as the amplitude of the neuronal discharge) were automatically rejected and the stimulus was presented again. The multiple-unit activity ("impulses") was extracted using window-discriminator software representing an amplitude-threshold device with a constant nontriggering delay of 0.5 ms. The recordings were stable several days after implantation, since the floating electrodes were not moved once implanted. Constancy of the amount of background activity and the cluster's frequency tuning during the recording session was used as an indication that we recorded from the same site throughout the period of data collection. By comparing the spontaneous activity of the multi-unit recording with the mean spontaneous activity obtained in single-unit recordings from the starling's forebrain, we can estimate that on average we recorded the activity of five cells per cluster.

Stimulus generation

Prior to the experiments, the soundproof booth was calibrated with a sound-level meter (General Radio 1982 precision sound-level meter) and a condensor microphone (General Radio 0.5″ microphone type 1962-9611). The microphone was placed at about the location where the bird's head would be while it was sitting on the perch in the cage. The frequency response of the setup was corrected to vary within ±2 dB.

All stimuli were produced by an SGI Workstation, using its 16-bit digital-to-analog converter at a sampling rate of 32 kHz. They were adjusted in level by a computer-controlled attenuator (TDT

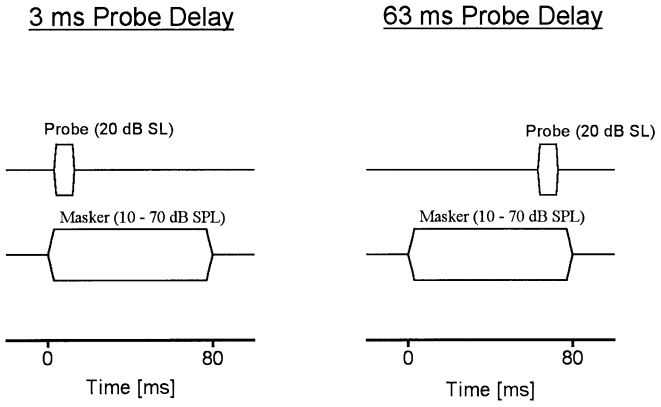

Fig. 1 The simultaneous-masking protocol used in this study. The probe frequency was always the recording site's characteristic-frequency (CF). The masker was either a CF tone or a tone at a frequency within inhibitory side-bands. For further details, see Materials and methods

PA 4). Using the workstation's stereo output, stimuli could be presented simultaneously by mixing the two channels in a Yamaha A-520 hi-fi amplifier. The stimuli were played through a single mid-range speaker (McFarlow 100MT) mounted on the ceiling of the booth.

Prior to masking experiments, frequency-tuning curves (FTC) were measured for each recording site by analyzing responses to 169 different frequency-level combinations of 250-ms tone bursts (five repetitions; 750-ms interstimulus interval). A FTC and inhibitory side-bands were constructed for each recording site using a statistical criterion. The response threshold was defined as an increase in the mean response rate during presentation of the 250-ms signal by 1.8 SDs above the mean spontaneous rate. The spontaneous rate was determined from the number of impulses in 100-ms intervals preceding the presentation of the tone burst for all frequency-level combinations. The threshold for inhibition was a decrease in the mean response rate during signal presentation by 1.8 SDs below the mean spontaneous rate. From these FTCs, the CF, the threshold at CF, and the threshold at the cluster's most sensitive inhibitory frequency were measured.

The stimulus to study simultaneous masking consisted of a 10-ms pure-tone signal (probe) that was presented simultaneously with an 80-ms pure-tone signal (masker). The phase angle of the probe relative to the masker's phase angle was random. This stimulus was repeated 20 times for all recording sites. The probe and masker had 1-ms and 3-ms rise and fall times, respectively, with a Gaussian rise and fall. The interstimulus interval was 920 ms. The masker frequency was either at the CF of the multi-unit cluster, or a frequency in the center of an inhibitory side-band. The masker level was varied in 5-dB-steps from 10 dB SPL to 70 dB SPL. The probe was always a CF tone presented at a level 20 dB above the recording site's threshold. To study masking effects right at the onset of the masker (the masking sound reached maximum amplitude after 3 ms), the probe was presented 3 ms after masker onset (3-ms probe-delay). In the second configuration, the probe was presented temporally delayed from the masker's onset to explore effects occuring during an ongoing masker (63-ms probe-delay; see Fig. 1).

Data analysis

The spectral tuning properties and the latency of each multi-unit cluster were determined by objective criteria as described above (for details, see Nieder and Klump 1999). The response latency was defined as the time of occurrence of the first of two consecutive 1-ms time bins in the peristimulus time histogram (PSTH) containing a number of impulses that was at least 2 SDs greater than the mean spontaneous rate.

In the simultaneous masking experiment, the combined activity evoked by the probe and the simultaneously played masker was analyzed in a 10-ms time window (according to the probe duration) that was shifted by the cluster's latency. Since the activity in this time window was generally dominated by the impulses evoked by the probe (at least at low masker levels), we will call this activity the "probe-driven response."

An identical analysis window was used to analyze the response that was elicited by the masker alone to 60 stimulus repetitions. Thus, the magnitude of masker-driven and probe-driven responses could be compared. The impulse rates produced by the probe alone were derived from 60 repetitions. A neuronal detection threshold for the probe-tones was reached if the impulse rate elicited by probe and masker together was just significantly different from the response evoked by the masker alone (binomial test, criterion: $P<0.01$).

To characterize the masking effects on the probe-evoked response, a response index (RI) was computed: $RI=(P+M_L)/P$, where $P+M_L$ is the combined activity of a probe and the simultaneously presented masker at a given masker level L and P is the response to the probe alone (at a constant level 20 dB above the threshold). A response index of 1 indicates no difference between the unmasked and masked probe response. A response index that is smaller than 1 indicates that the response to the combination of probe and masker is smaller than the response to the unmasked probe. A response index that is larger than 1 indicates that the response to the combination of probe and masker is stronger than the response to the unmasked probe. A Wilcoxon test was employed for comparison of $P+M_L$ and P (two-tailed P-values).

Histological verification of electrode penetration

After lesioning several recording sites, animals were killed by an overdose of pentobarbital sodium, and, after an initial flush of the circulatory system with a solution containing 0.9% NaCl and 0.5% $NaNO_2$, the brains were fixed by transcardial perfusion of 200 ml paraformaldehyde (6%). The skull and the electrodes were removed and the brain was kept in the fixative for several days. Prior to sectioning, the brain was kept in the same fixative with an additional 10% and 30% sucrose, serving as a cryoprotective. After embedding in egg yolk fixed by glutaraldehyde, frozen sagittal sections (50 µm) were cut and counterstained with cresyl violet. Electrode penetrations were marked by prominent cells (probably glia) surrounding the electrode trunk. The location of the electrode tips was either determined by searching for lesion marks or for the end of the glia-cell tube. The data from 32 recording sites presented in this study were found to lie within the thalamorecipient zone of L2, L2a (Wild et al. 1993).

Results

Simultaneous masking was studied in 32 multi-unit clusters recorded in the input layer L2a of the field L complex, the primary auditory forebrain of birds (Wild et al. 1993; Vates et al. 1996). The spectral tuning properties of all recording sites were studied in detail prior to the masking experiments described here (for details see Nieder and Klump 1999). For each cluster, a FTC and inhibitory side-bands were determined and several parameters were extracted (e.g., CF, threshold at CF, bandwidth of excitatory and inhibitory threshold curves). Since we used the CF of each recording site as the probe-tone frequency and the masker was either at the CF or at an inhibitory side-band frequency, the knowledge of these spectral tuning properties was a prerequisite for the current study. Figure 2 shows an example of the FTC and the inhibitory

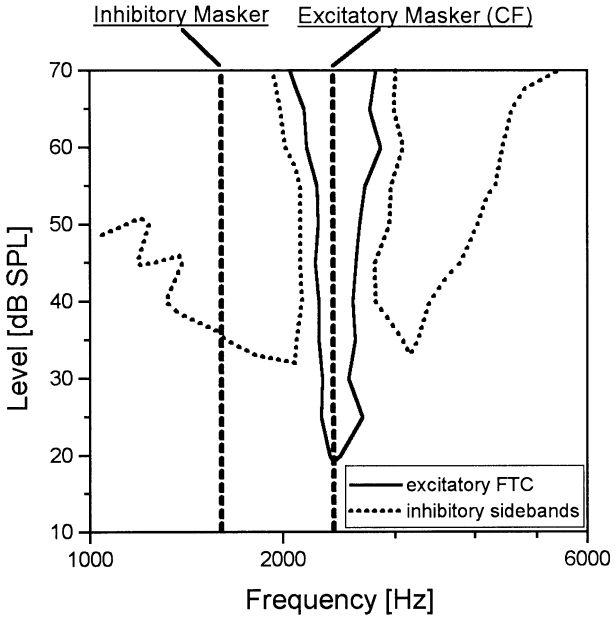

Fig. 2 Spectral relation of excitatory and inhibitory masker-frequencies at a certain recording site. The plot shows the cluster's frequency-tuning curves (*FTC*; *continuous lines*) and the two inhibitory side-bands (*dotted lines*). At this recording site, the CF was 2400 Hz. Therefore, the frequency of the probe and of the excitatory masker was 2400 Hz (indicated by the *right vertical dashed line*). The frequency for the inhibitory masker was 1600 Hz (*left vertical dashed line*), at the spectral center of the low-frequency side-band.

side-bands of a certain recording site. The spectral relation of the excitatory masker (that was always the cluster's CF) and the inhibitory masker are indicated by the two dashed vertical bars. The frequencies of the excitatory and the inhibitory masker were chosen according to each individual recording site's tuning characteristics. For all multiple-unit clusters, the range of CFs was between 1 and 6 kHz; the mean threshold at CF for all cell clusters was 21 dB SPL, with a minimal best theshold of 11 dB SPL.

Responses to the probe alone

The probe-tone was always presented at a level 20 dB above the recording site's threshold. It elicited a high, but not a saturated discharge. Compared with the onset response of an 80-ms CF tone at a level of 70 dB SPL (that is close to the saturated activity), the probe evoked a significantly smaller response in the 10-ms time window ($P< 0.0001$, Wilcoxon, two-tailed, $n=32$). On average, the amount of discharge evoked by the probe alone at a level of 20 dB SPL was 83% of the discharge produced by the 70-dB tone. For both probe delays, the impulse-rate of the response to the probe at a masker level of 10 dB SPL was not different ($P=0.26$; Wilcoxon; two-tailed, $n=32$).

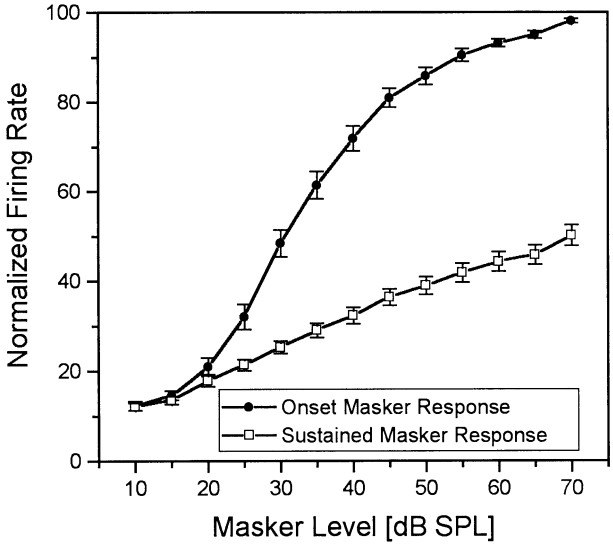

Fig. 3 Normalized mean rate-level functions for responses to the CF masking-tone alone. Functions were derived from measurements in a 10-ms time-window delayed by 3 ms (onset masker response) or 63 ms (sustained masker response) relative to the masker onset, respectively. This time window was identical to the one used to derive the responses to the masked probe in the masking paradigm. *Error bars* ±SE

Simultaneous masking with an excitatory masker

The temporal response characteristics during an ongoing CF-stimulus (i.e., the excitatory masker) were quantitatively determined for all cell clusters (Nieder and Klump 1999). With one exception where the response was found to be purely tonic, all multi-unit clusters recorded in field L2a displayed a temporal pattern called "phasic-tonic" or "primary-like". These responses were characterized by a maximum firing rate near the onset of stimulation ("phasic") and, subsequently, a decreasing impulse rate to a steady ("tonic") level (see for example recording site in Fig. 4). Several phasic-tonic clusters experienced OFF-inhibition.

Figure 3 shows mean rate-level functions produced by the excitatory CF-masker alone during a 10-ms time window at masker onset (3-ms delay) and at a time delay of 63 ms (i.e., the sustained response) shifted by the same latency as used in determining the probe-driven response. The mean rate-level function at masker onset had a sigmoid shape with the steepest slope found for levels between 20 and 45 dB SPL. For the sustained activity, however, the mean rate level function was almost linear. Firing rates during the sustained response at higher CF-masker levels were well below the ones determined for the onset response.

To characterize temporal effects of simultaneous masking, we focus on the probe-evoked response. It was increasingly covered by the masker-evoked response as is shown in Fig. 4 displaying peristimulus time histograms of a typical recording site. At low masker levels (10 and 20 dB SPL), the probe elicited a strong discharge at both probe delays. In the left column it is shown that with in-

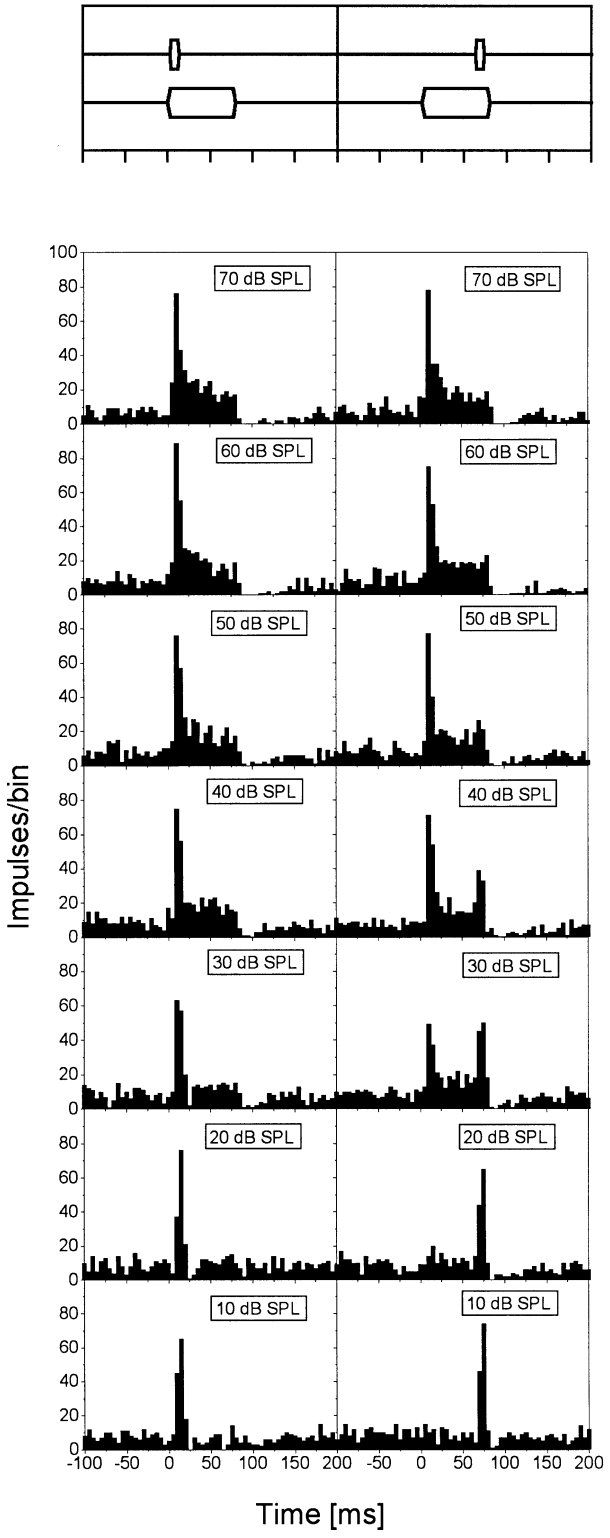

Fig. 4 Peristimulus time histograms (PSTHs) at a multipe-unit cluster for probes masked with a CF masker (same cluster as in Fig. 2). The *upper panels* schematically show the stimulus protocol. The probe was delayed by 3 ms (*left column*) or 63 ms (*right column*). While the probe level was held constant at 20 dB above the threshold, the masker level was increased form 10 dB SPL (*bottom*) to 70 dB SPL (*top*). Although the masker level was raised in 5-dB steps, only decimal masker levels are plotted for clarity. Each PSTH represents 20 stimulus repetitions (5-ms bin width). Note the progressive decline of the discharge within the probe interval at the 63-ms probe delay (*right column*)

creasing masker level the phasic onset discharge produced by the masker increased to the value of the probe-evoked response at the 3-ms delay (for the masker-onset response alone see right column). On the other hand, the probe-elicited activity at the 63-ms delay (right column) exceeded the masker's sustained activity up to higher masker levels than observed for the 3-ms delay condition. In addition, the probe-evoked discharge at the 63-ms delay is increasingly reduced as is obvious for masker levels of between 30 and 50 dB SPL. At the highest masker level of 70 dB SPL, the probe-evoked response could not be distinguished from the response to the masker alone in either delay condition.

To quantify the masking effect, the neuronal detection threshold for the probe was determined (see Materials and methods for definition). The masking effect was significantly stronger at the beginning of the masker. In the 3-ms delay condition, masker levels that were on average 8 dB lower than in the 63-ms probe delay condition prevented probe detection ($P<0.0001$, Wilcoxon, two-tailed; $n=32$). At a 3-ms probe delay, the probe was detectable up to an mean masker-level of 36 dB SPL, whereas the probe-driven response at a 63-ms delay was detectable up to a mean masker level of 44 dB SPL. This detection difference was not correlated with the CF of the recording sites ($r=0.027$, $P=0.89$, Spearman's rank correlation coefficien; $n=32$).

The temporal response pattern of neurons in the input layer of L2a was one of the factors determining the difference in masking between the delay conditions. At a delay of 3 ms, the response elicited by the probe coincided with a strong phasic discharge to the masker onset that had occurred already at low masker levels (see Fig. 4). The number of impulses evoked by the masker onset at a given level was much larger than the sustained activity (Fig. 3). Thus, the probe-driven activity was covered by the onset response to the masker at a lower level. A second effect became obvious when we compared the rate-level functions for the masker alone with the rate-level function obtained with the masker-probe combination. While the neuronal response to the probe plus masker at the 3-ms probe delay remained constant for all masker levels, the discharge elicited at the 63-ms probe delay decreased as the masker level was increased (Fig. 5). This reduction of the probe-driven response was reflected by the number of impulses at the detection threshold. While at the detection threshold the masker-probe combination at the 3-ms delay caused a mean discharge of the multiunit response of 506 impulses/s (single units can reach a discharge rate of up to 160 spikes/s; see Leppelsack 1974), the mean impulse rate for all recording sites was only 308 impulses/s for the 63-ms delay ($P<0.0001$, Wilcoxon, two-tailed; $n=32$). A comparison of the response to the probe alone and the combined probe plus masker response at the detection threshold revealed equal impulse-rates for a delay of 3 ms ($P=0.10$, Wilcoxon, two-tailed; $n=32$). However, significantly reduced impulse-rates at the detection threshold were found at a delay of 63 ms for probe plus masker versus the probe alone ($P<0.0001$, Wilcoxon, two-tailed; $n=32$).

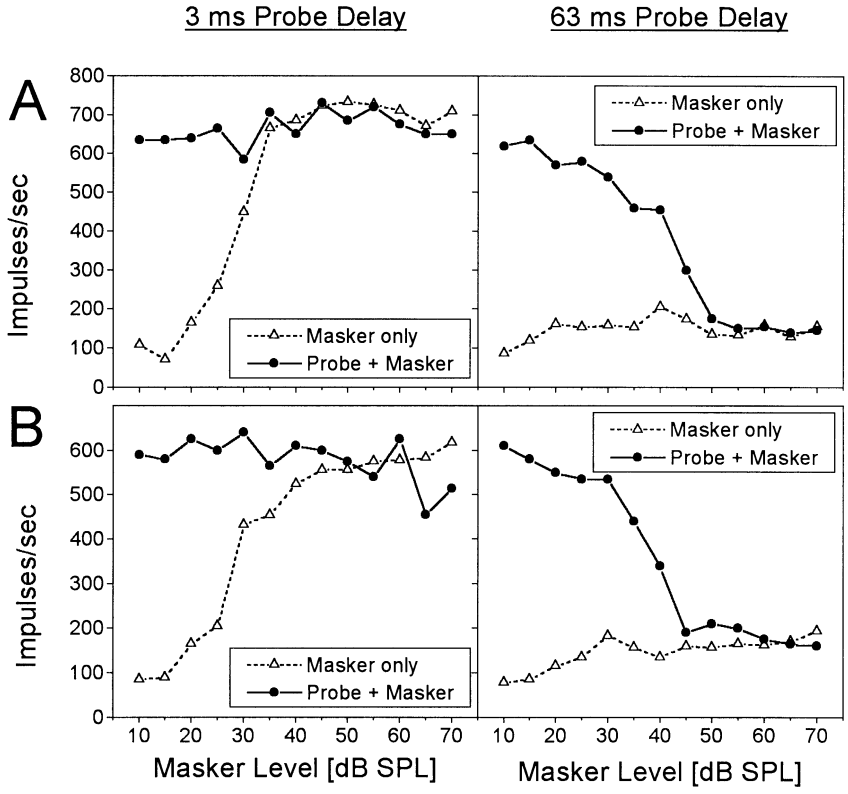

Fig. 5A, B Rate-level functions at two typical recording sites for probe delays of 3 ms (*left column*) and 63 ms (*right column*). Each panel shows the discharge rate produced by the masker alone compared with the combined response of a probe masked with a CF-tone. At probe delays of 63 ms, the probe-driven response declined to the masker's sustained activity level. A reduction of the probe-evoked response at a 3-ms delay was not present with increasing masker levels

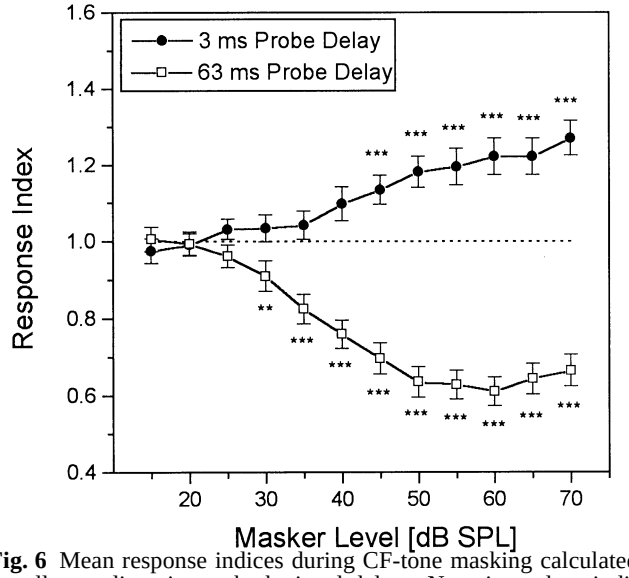

Fig. 6 Mean response indices during CF-tone masking calculated for all recording sites at both signal delays. Negative values indicate a reduction of the probe-evoked response at a 63-ms delay by the masker. For probes delayed by 3 ms, the response experienced an increment of activity at higher masker levels, because the CF-masker elicited an even stronger discharge than the nonsaturated response to the probe alone. *Error bars* ±SE. ***P*<0.01; ****P*<0.001

To quantify the reduction in the probe-driven response, a response index was calculated for all tested masker levels (see Materials and methods). Figure 6 illustrates that the mean probe-driven discharge at a 63-ms delay was significantly reduced at masker levels from 30 to 70 dB SPL (indicated by index values in the graph that are smaller than 1). For probe-delays of 3 ms, however, there was no reduction of the probe-driven rate. At masker levels above 40 dB SPL, the index increased to values above 1. This was due to the masker evoking larger responses than the probe alone and swamping the probe-evoked response at high intensities.

Simultaneous masking with an inhibitory masker

The frequency-tuning curves exhibited prominent inhibitory side-bands at 15 of the investigated recording sites. These inhibitory side-band frequencies may exert a different effect than the masking by tones of an excitatory frequency. At most recording sites, the side-band inhibition was more pronounced at frequencies below the CF. Therefore, the frequency chosen for the inhibitory masker was predominantly lower than the probe frequency. On average, the spectral distance between the frequency of the inhibitory masker and the probe frequency (the CF) was 0.63 octaves. In Fig. 7, the mean response to the inhibitory masker alone at response onset and during the sustained activity was plotted against the masker level. At low levels, the firing rate was more strongly reduced at response onset than during the ongoing response. The reduction at high levels, however, was stronger during the sustained activity than during the onset response. At intermediate masker levels from 35 to 55 dB SPL, the rate reductions for both the onset and the ongoing responses were about equal.

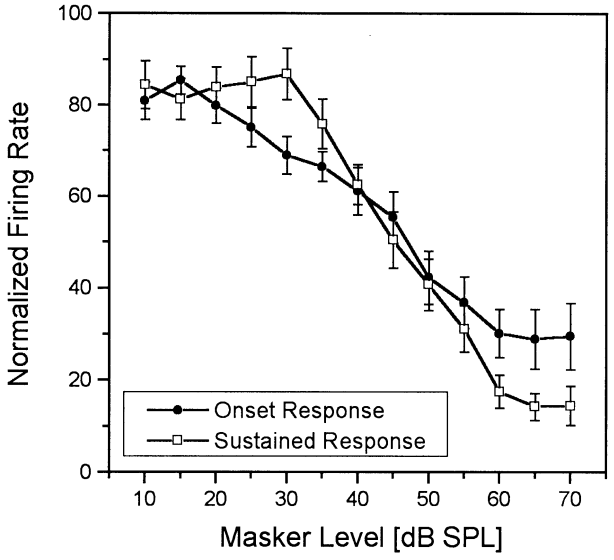

Fig. 7 Normalized mean rate-level functions for responses to a masking tone taken from inhibitory side-bands. The highest discharge at any tone level was set 100%. Again, the time windows for onset and sustained response were identical to the ones used to derive the responses to the probe during masking with a side-band frequency. *Error bars* ±SE

The masking effect of a simultaneously played inhibitory tone on the probe is demonstrated for a typical recording site (Fig. 8). No reduction of the spontaneous activity was evident for maskers with levels ranging from 10 to 20 dB SPL. At a masker level of 40 dB SPL, a reduction of the spontaneous activity could be observed, and at masker levels of 50 dB SPL and above, the probe-driven discharge at both delay conditions experienced a reduction. Correlated with a progressive reduction of the background activity during the on-going masker, the probe-elicited response both at a delay of 3 ms and 63 ms was diminished and finally was totally suppressed at the maximum masker level of 70 dB SPL. On average, probes at both the 3-ms and the 63-ms delay became detectable at about the same masker levels of 54 dB SPL and 55 dB SPL, respectively. No differences were found comparing the detection thresholds for probes at both temporal conditions (P=0.47, Wilcoxon, two-tailed; n=15).

Figure 9 illustrates the decline of the probe-driven rate-level functions at two typical recording sites. In both multi-unit clusters, the discharge caused by the probe at low masker levels was almost completely abolished as the inhibitory masker level was increased from 45 dB SPL to 70 dB SPL. Note that the decline of the activity at a given masker level is almost identical for 3-ms and 63-ms probe delays.

Comparing the masking effect of excitatory and inhibitory maskers

The detection threshold for CF-probes was compared at individual recording sites where both excitatory and in-

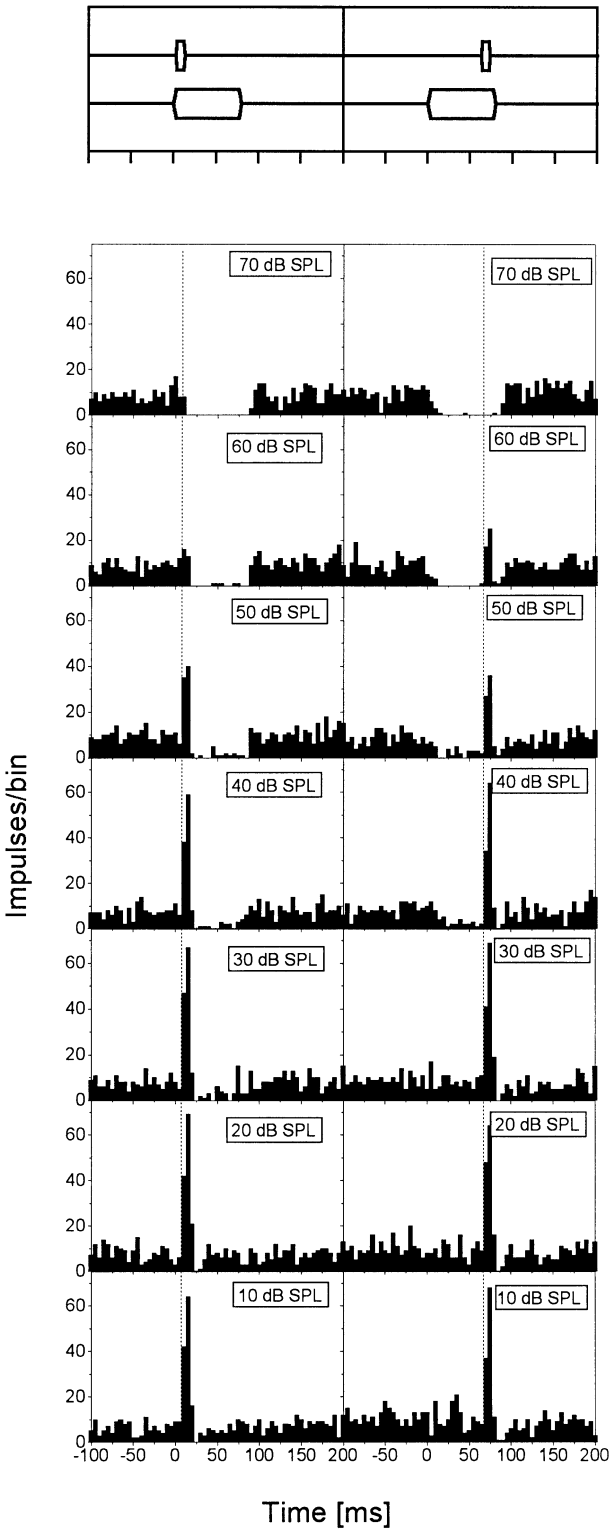

Fig. 8 PSTHs at a typical multiple-unit cluster for probes masked with an inhibitory masker. The *upper panel* represents the temporally correlated stimulus paradigm. As in the experiments with a CF-masker, the probe was delayed by 3 ms (*left column*) or 63 ms (*right column*). Each PSTH represents 20 stimulus repetitions (5 ms bin width). The *dotted vertical line* marks the response onset to the probe. As the masker level was increased form 10 dB SPL (*bottom*) to 70 dB SPL (*top*), the backgound activity became reduced. At a masker level of 70 dB SPL, the probe-evoked discharge was completely suppressed. No differences in the reduction of the probe-driven responses were found for either probe delays

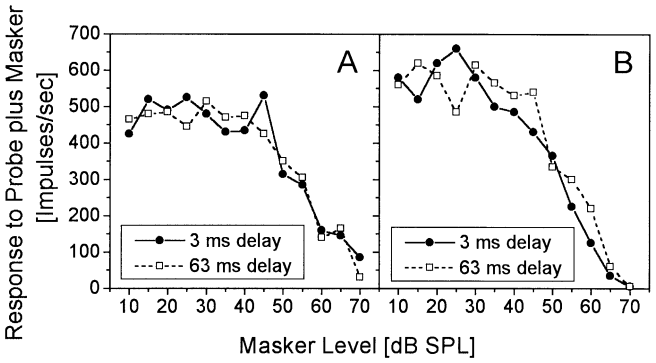

Fig. 9 Rate-level functions derived at two typical recording sites (*A* and *B*) for probes masked with inhibitory tones. Each panel displays the discharge rate produced by probes plus masker delayed by 3 ms and 63 ms, respectively.With increasing masker levels, the probe-driven rate became equally suppressed at masker onset and during the ongoing masker

hibitory frequencies were used as maskers. For probe delays of 3 ms, the level of the masker starting to affect detection of the probe was on average 18 dB higher for inhibitory maskers than for excitatory maskers ($P<0.001$, Wilcoxon, two-tailed; $n=15$). For probe delays of 63 ms, this threshold masker level was on average 11 dB higher for inhibitory than for excitatory maskers ($P<0.005$, Wilcoxon, two-tailed; $n=15$). In summary, frequencies outside the excitatory frequency-tuning curve are capable of masking CF-probe tones. However, the levels of inhibitory maskers have to be considerably higher than those of excitatory CF maskers.

Discussion

Interpretation of multi-unit recordings in the freely moving starling

Multi-unit activity is comprised of action potentials generated by several neurons in the vicinity of the electrode tip. Although we cannot exclude the possibility that neighboring cells in field L possessed different response properties, several studies performed in the auditory forebrain of mammals (Shamma et al. 1993, Gaese and Ostwald 1995; Brosch and Schreiner 1997) and birds (Cohen and Knudsen 1996, 1998) directly comparing multi- and single-unit recordings indicate that temporal information processing of adjacent neurons is rather similar. Multi-unit recordings in a freely moving and nonanesthetized bird have clear benefits. Firstly, the method is less stressful for the animal since no restraint is necessary. Secondly, anesthesia has been shown to considerably affect the response of neurons in the starling's forebrain (Capsius and Leppelsack 1996). A comparison of the spike discharge patterns between alert and anesthetized mammals typically reveals that responses in the auditory cortex of nonanesthetized animals are often more intense and sustained than those in anesthetized animals. (Brugge et al. 1969; Brugge and Merzenich 1973). Also in the auditory forebrain area of the starling, in which the recording sites of the present study were located, a notable modification of the sustained response by anesthesia was found (Capsius and Leppelsack 1996). The sustained response to a masking sound has a prominent influence on the detection of simultaneously presented probes.

Neural mechanisms of masking

The probe-masker interactions reflected by the units' responses in the starling's forebrain differed for excitatory and inhibitory maskers. First we will discuss the results obtained for excitatory maskers. In response to the masker alone, the auditory forebrain neurons in field L2a of the starling typically showed a phasic-tonic discharge pattern. The firing rate was maximal near the onset of the masker and subsequently decreased to a lower sustained level. This mirrors the activity pattern of auditory-nerve fibers (Manley et al. 1985). Such a response pattern is commonly thought to reflect short-term adaptation (Kiang 1965; Smith and Zwislocki 1975). The rapid increase in the impulse rate with increasing masker level during the phasic onset response (i.e., at a probe delay of 3 ms) was sufficient to mask the probe at relatively low masker levels. This effect can be explained solely by the swamping hypothesis (Moore and Glasberg 1982; Pickles 1984), since at the detection threshold the mean impulse rate elicited by the probe plus masker was not different from the impulse rate elicited by the probe alone. The swamping effect cannot be explained by discharge saturation at response onset since the discharge evoked by the stimuli was significantly below saturation. It is the relative amount of impulses elicited by masker plus probe and by the probe alone that is important. Swamping was less effective at a probe delay of 63 ms as the sustained response to the masker at a given level was much lower than the phasic onset response (Fig. 3).

In the auditory periphery, the improved detection of the probe during the tonic part of the response to the masker was explained by an enhanced signal-to-noise ratio resulting from the short-term adaptation in the neural discharge. For example, Coombs and Fay (1989) observed in saccular fibers of the goldfish that the spikes added during probe presentation are about equal, regardless of the state of adaptation of the fiber to the masker (except for saturated responses). The effect of the masker on the response to the probe tended to be independent of the probe's temporal position within the masker (Fay 1991). Studying auditory nerve fibers of the guinea pig, Smith and Zwislocki (1975) found that the change in firing rate produced by adding a probe tone to a masker of the same frequency (i.e., by an increment in the level of a tone) is nearly independent of the time delay between masker and probe onset, and, consequently, of the amount of adaptation. In central auditory neurons of the starling, however, the response to the probe was definitely dependent on its temporal position in the masker. Contrary to

the observation at the masker onset, we found a significantly reduced impulse rate for the probe plus masker versus the probe alone at a probe delay of 63 ms. This may be a feature of central auditory processing that is not observed in the periphery.

For masker frequencies taken from inhibitory side-bands, the probe-driven response was clearly reduced. In central auditory neurons, this response can be explained by the joint action of peripheral suppression and of inhibitory processes in the avian auditory pathway (Manley 1990; Müller and Scheich 1988; Nieder and Klump 1999). In contrast to CF maskers, however, for maskers at inhibitory side-band frequencies the detectability of the probe was not affected by its temporal position.

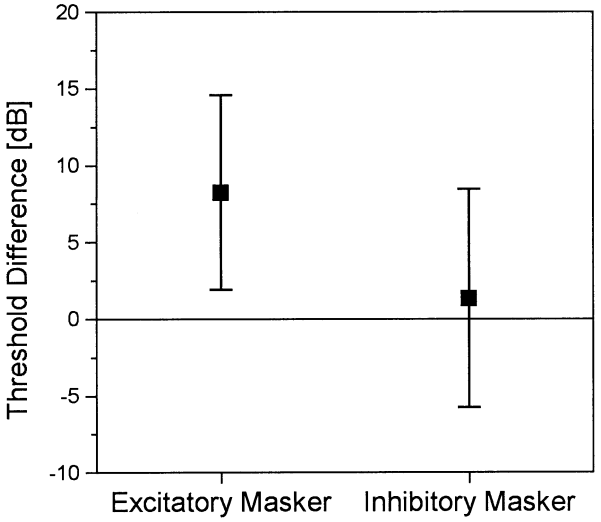

Fig. 10 Detection threshold differences observed for excitatory and inhibitory masker frequencies. A neuronal "overshoot," i.e., a stronger masking effect at masker onset, was found when the frequencies of both masker and probe were the CF. Probes with a 3-ms delay could not be detected anymore with masker levels that were on average 8 dB lower than for probes with a delay of 63 ms. With inhibitory masker frequencies taken from the frequency tuning curves' side-bands, on the other hand, no difference in detection threshold was found. Probes presented with a 3-ms or 63-ms delay relative to masker onset were masked at similar masker levels. *Error bars* ±SD

Evidence for neuronal overshoot

Psychoacoustic experiments on masking revealed a stronger masking of a short probe right at the masker onset compared with the masking some 10 ms after masker onset. This psychophysical effect has been termed "overshoot" (Zwicker 1965a, b; Elliott 1965). Temporal effects of simultaneous masking by a tone that was equal in frequency to the probe tone have been reported by Green (1969), Fastl (1979), and Bacon and Viemeister (1985). In these studies, the amount of overshoot was about 5 dB. For masking paradigms in which broadband noise maskers and brief sinusoidal signals were used, an even greater overshoot of about 10–20 dB was observed (Zwicker 1965a). Masker components above the critical band of the probe tone (i.e., above the frequency channel encoding the probe) were generally more effective in generating an overshoot than masker components below the critical band of the probe tone (Bacon and Viemeister 1985; Schmidt and Zwicker 1991; Wright 1997).

In the current study, a reduction in masking with time after masker-onset could also be observed in the starling's forebrain neurons. This neuronal overshoot was observed when masker and probe were both from the excitatory frequency region of the tuning curve (i.e., from the same excitatory frequency channel that matched psychophysical critical bands in the starling; see Langemann et al. 1995; Nieder and Klump 1999). Its value was, on average, 8 dB (Fig. 10), which matches psychoacoustical within-channel data quite well. In mammalian auditory nerve fibers, Smith and Zwislocki (1971, 1975) calculated that temporal effects in simultaneous masking were between 3 and 5 dB. Data from CAP measurements in the starling indicate an overshoot in the bird's auditory periphery that is in general less than 5 dB (Oeckinghaus and Klump 1990). The overshoot effect is therefore at least about 3–4 dB larger in the starling's forebrain than in the auditory periphery. Whether this additional overshoot is generated in the forebrain or somewhere along the auditory pathway up to field L remains an open question.

The starling's forebrain neurons did not exhibit overshoot if the masker frequency was outside the excitatory frequency-tuning curve, i.e., outside the neuronal critical-band analogue. This result was somewhat surprising, since psychoacoustic studies in humans demonstrated that temporal effects in simultaneous masking are dependent on the spectral relation between masker and probe signal and occur well beyond the limits of the critical band. For example, the magnitude of overshoot increases with increases in the bandwidth of a masker that is centered on the probe frequency (Zwicker 1965a; Bacon and Viemeister 1985). Both processes within one auditory filter and computations across several auditory filters appear to contribute to the temporal masking effects (McFadden and Wright 1992; Wright 1997).

These differences between the psychophysical and neurophysiological evidence for overshoot indicate that the information encoded by the rate response of units in the auditory periphery and of forebrain neurons in field L2a do not provide a sufficient explanation of the psychophysical data. One possibility is that the relevant neural computations occur only in secondary auditory forebrain areas that have not been studied here. However, the question of whether a temporal neural code (de Charms and Merzenich 1996), in addition to a rate code, provides a more comprehensive explanation of the psychophysical data on the time course of across-channel masking also needs to be investigated.

Acknowledgements This work was supported by a grant to G.M.K. from the DFG within the SFB 204 "Gehör". We thank B. J. Frost, D. Geisler, C. Köppl and H. Wagner for critically reading a previous version of the manuscript.

References

Bacon SP, Viemeister NF (1985) The temporal course of simultaneous tone-on-tone masking. J Acoust Soc Am 78:1231–1235

Brosch M, Schreiner CE (1997) Time course of foreward masking tuning curves in cat primary auditory cortex. J Neurophysiol 77:923–943

Brugge JF, Merzenich MM (1973) Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. J Neurophysiol 36:1138–1158

Brugge JF, Dubrovsky NA, Aitkin LM, Anderson DJ (1969) Sensitivity of single neurons in auditory cortex of cat to binaural tonal stimulation: effects of varying interaural time and intensity. J Neurophysiol 32:1005–1024

Capsius B, Leppelsack H-J (1996) Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. Hear Res 96:59–70

Charms RC de, Merzenich MM (1996) Primary cortical representation of sounds by the coordination of action-potential timing. Nature 381:610–613

Cohen YE, Knudsen EI (1996) Representation of frequency in the primary auditory field of the barn owl forebrain. J Neurophysiol 76:3682–3691

Cohen YE, Knudsen EI (1998) The representation of binaural spatial cues in field L of the barn owl forebrain. J Neurophysiol 79:879–890

Coombs S, Fay RR (1989) The temporal evolution of masking and frequency selectivity in the goldfish (*Carassius auratus*). J Acoust Soc Am 86:925–933

Ehret G, Merzenich MM (1985) Auditory midbrain responses parallel spectral integration phenomena. Science 227:1245–1247

Ehret G, Schreiner CE (1997) Frequency resolution and spectral integration (critical band analysis) in single units of the cat primary auditory cortex. J Comp Physiol [A] 181:635–650

Elliott LL (1965) Changes in the simultaneous masked thresholds of brief tones. J Acoust Soc Am 38:738–746

Fastl H (1979) Temporal masking effects. III. Pure tone maskers. Acoustica 43: 282–294

Fay RR (1991) Masking and suppression in auditory nerve fibers of the goldfish, *Carassius auratus*. Hear Res 55:177–187

Fletcher H (1940) Auditory patterns. Rev Mod Phys 12:47–65

Gaese BH, Ostwald J (1995) Temporal coding of amplitude and frequency modulation in the rat auditory cortex. Eur J Neurosci 7:438–450

Green DM (1969) Masking with continuous and pulsed sinusoids. J Acoust Soc Am 46:939–946

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

Jacob R, Krüger J (1991) Manufacture of sharpened microelectrodes from varnished wire. J Neurosci Methods 38:89–93

Kiang NYS (1965) Discharge patterns of single fibers in the cat's auditory nerve. MIT Press, Cambridge, MA

Klump GM, Langemann U (1995) Comodulation masking release in a songbird. Hear Res 87:157–164

Langemann U, Klump GM, Dooling RJ (1995) Critical bands and critical-ratio bandwidth in the European starling. Hear Res 1995:167–176

Leppelsack H-J (1974) Funktionelle Eigenschaften der Hörbahn im Feld L des Neostriatum caudale des Staren. J Comp Physiol 88:271–320

Manley GA (1990) Peripheral hearing mechanisms in reptiles and birds. Springer, Berlin Heidelberg New York

Manley GA, Gleich O, Leppelsack H-J, Oeckinghaus H (1985) Activity patterns of cochlear ganglion neurons in the starling. J Comp Physiol [A] 157:161–181

McFadden D, Wright BA (1992) Temporal decline of masking and comodulation masking release. J Acoust Soc Am 92:144–156

Moore BCJ (1989) An introduction to the psychology of hearing. Academic Press, London

Moore BCJ (1992) Across-channel processes in auditory masking. J Acoust Soc Jpn 13:25–37

Moore BCJ, Glasberg BR (1982) Interpreting the role of suppression in psychophysical tuning curves. J Acoust Soc Am 72: 1374–1379

Müller CM, Scheich H (1988) Contribution of GABAergic inhibition to the response characteristics of auditory units in the avian forebrain. J Neurophysiol 59:1673–1689

Nieder A, Klump GM (1999) Adjustable frequency selectivity of auditory forebrain neurons recorded in a freely moving songbird via radiotelemetry. Hear Res (in press)

Oeckinghaus H, Klump GM (1990) CAP-masking in the European starling (*Sturnus vulgaris*). In: Elsner N, Roth G (eds) Brain – perception – cognition. Thieme, Stuttgart, p 129

Pickles JO (1984) Frequency threshold curves and simultaneous masking functions in single fibres of the guinea pig auditory nerve. Hear Res 14:245–256

Sachs MB (1969) Stimulus-response relation for auditory-noise fibers: two-tone stimuli. J Acoust Soc Am 45:1025–1036

Scharf B (1970) Critical bands. In: Tobias JV (ed) Foundations of modern auditory theory. Academic Press, New York, pp 157–202

Schmidt S, Zwicker E (1991) The effect of masker spectral asymmetry on overshoot in simultaneous masking. J Acoust Soc Am 89:1324–1330

Shamma SA, Fleshman JW, Wiser PR, Versnel H (1993) Organization of response areas in ferret primary auditory cortex. J Neurophysiol 69:367–383

Smith RL, Zwislocki JJ (1971) Responses of some neurons of the cochlear nucleus to tone-intensity increments. J Acoust Soc Am 50:1520–1525

Smith RL, Zwislocki JJ (1975) Short-term adaptation and incremental responses of single auditory-nerve fibers. Biol Cybern 17:169–182

Vates GE, Broome BM, Mello CV, Nottebohm F (1996) Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches (*Taeniopygia guttata*). J Comp Neurol 366:613–642

Wild JM, Karten HJ, Frost BJ (1993) Connections of the auditory forebrain in the pigeon (*Columba livia*). J Comp Neurol 337: 32–62

Wright BA (1997) Detectability of simultaneously masked signals as a function of masker bandwidth and configuration for different signal delays. J Acoust Soc Am 101:420–429

Zwicker E (1965a) Temporal effects in simultaneous masking by white-noise bursts. J Acoust Soc Am 37:653–663

Zwicker E (1965b) Temporal effects in simultaneous masking and loudness. J Acoust Soc Am 38:132–141