

Release from masking in fluctuating background noise in a songbird's auditory forebrain

Georg M. Klump^{CA} and Andreas Nieder¹

Institut für Zoologie, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

¹Present address: Andreas Nieder, Center for Learning and Memory, Department of Brain and Cognitive Sciences, E25-249, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

^{CA}Corresponding Author

Received 15 March 2001; accepted 4 April 2001

Fluctuations in the ubiquitous masking background noise can be exploited by the vertebrate auditory system to considerably improve signal detection. Here we demonstrate neuronal masking release in amplitude-modulated background noise on the level of the European starling's auditory forebrain, an area that is the analogue of the mammalian primary auditory cortex. Tone-evoked responses in the presence of modulated and unmodulated maskers were recorded in unrestrained birds via

radiotelemetry. Based on a rate code, the average amount of neuronal masking release was similar to that observed in a psychoacoustic study on the starling with stimuli confined to a single auditory filter. The results suggest that the neurons exploited predominantly temporal features of the acoustic background to improve signal detection. *NeuroReport* 12:1825–1829 © 2001 Lippincott Williams & Wilkins.

Key words: Auditory scene analysis; Bird; CMR; Hearing; Release from masking; Signal detection field

INTRODUCTION

Signal detection in the natural environment is typically compromised by ubiquitous background noise (for review see [1]). Background noise is often temporally structured, as has been demonstrated by an analysis of recordings from the Cornell sound library by Nelken *et al.* [2] and is shown here by an example representing a morning chorus of singing birds (Fig. 1a). Such a temporal structure, in which positively correlated envelope fluctuations (i.e. a coherent modulation, termed comodulation) can be found over a wide range of frequencies, results from the intermittent sound production of the sources and from the action of air turbulences upon the signal on its path of transmission [3]. It has been suggested that coherent changes in the acoustic features of signals may aid the human auditory system in the analysis of complex auditory scenes (for review see [4]). Given the adaptive value of auditory scene analysis, it is not surprising that other vertebrates, such as the European starlings (Aves), also possess this ability [5,6].

The ability of an improved segregation of a signal from an acoustic background should also affect masked thresholds. A paradigm in which an improved detection of a signal in a temporally structured masker has been attributed to mechanisms that can also affect auditory grouping has first been applied in human psychophysical experiments by Hall *et al.* [7] (see also [8], and for review see [9]). Hall and colleagues found that contrary to the expectation formed on the basis of the concept of auditory frequency

filters (i.e. critical bands [10]), thresholds for tones may be considerably reduced in a comodulated masker compared with an unmodulated masker of equal sound energy. This effect has been termed comodulation masking release, and humans can experience a reduction in hearing threshold of about 12 dB in slowly amplitude-modulated maskers (similar to the ones used in this study) compared to unmodulated maskers of the same bandwidth. This psychophysical experiment demonstrating comodulation masking release has also been done in European starlings that exhibit a psychophysical release from masking of up to 18 dB on average [11] in broadband background noise with a comodulated envelope. Some of the cues that may enable the auditory system to improve signal detection in comodulated background noise are illustrated in Fig. 1b.

Despite the relevance of comodulation masking release and other processes related to auditory grouping and the analysis of auditory scenes for understanding real-life communication mechanisms, so far only few studies explored their neural bases. Studies of the neural mechanisms underlying comodulation masking release have been conducted in the primary auditory cortex of the anesthetized cat [2,12], in the cochlear nucleus of the anesthetized guinea pig [13], and in the auditory nerve of the anesthetized chinchilla [14]. The results have been compared to human psychophysical data. Here we present a study of the neural basis of comodulation masking release in a songbird, the European starling, that compares neural

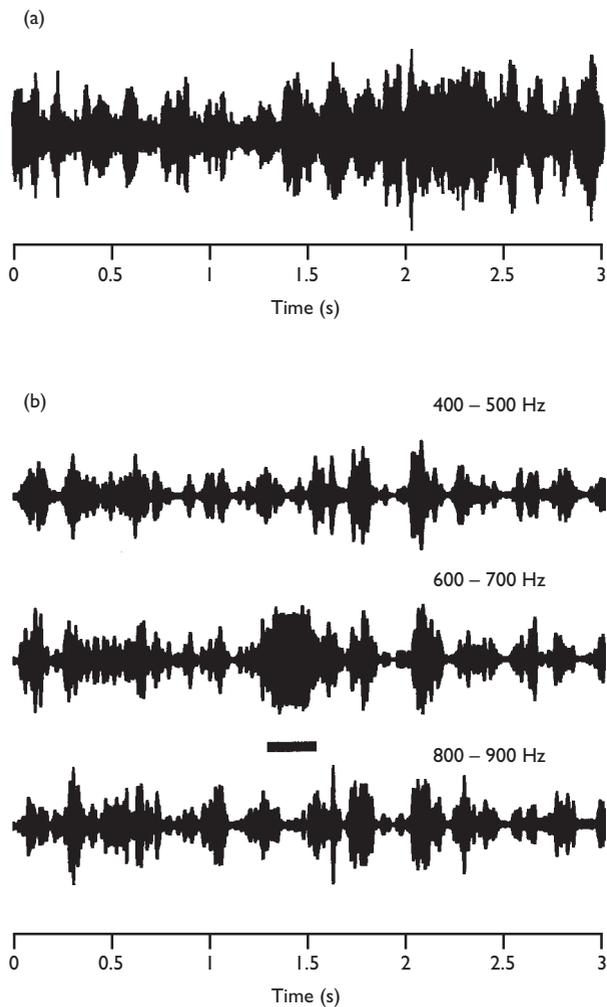


Fig. 1. (a) An oscillogram of 3 s recording of a dawn chorus of birds in a deciduous European wood in spring. (b) Illustration of the effect of comodulation on signal detection. The graph shows amplitude fluctuations in three exemplary 100 Hz frequency bands cut out of an 800 Hz wide band of comodulated noise by digital filtering. Comodulation creates a distinct temporal structure of the envelope of the masking background noise. The envelope fluctuations are correlated between different frequency bands that can be analyzed in separate frequency filters of the auditory system (see parallel traces). When a 200 ms tone is added to one of the frequency bands of the background noise (600–700 Hz, indicated by the position of the bar underneath the amplitude vs time trace in the middle), a constant high amplitude is found in this frequency band and the envelope fluctuations are reduced. This steady increase in signal energy within a single frequency channel may provide one type of cues that the neurons that can exploit for improved signal detection in comodulated noise. In addition, during the presence of the tone the correlations of the envelopes between the frequency band with the added signal and the other frequency bands in the noise are reduced (i.e. the envelopes become more dissimilar). Thus, comparisons across frequency channels may provide another type of cue serving the neurons to improve signal detection in comodulated noise. A more in-depth discussion of the different cues that may lead to a release from masking by comodulation can be found in the review by Moore [9].

responses in an awake, freely-moving animal with its psychophysical performance using the classical paradigm for demonstrating comodulation masking release that was pioneered by the work of Hall *et al.* [7].

MATERIALS AND METHODS

Experiments were performed on nine wild-caught adult European starlings (*Sturnus vulgaris*) of both sexes. Polyimid-coated CrNi-wires (diameter 17 μm) were sharpened at the tip and chronically implanted in the birds' auditory forebrain in clusters of up to 14 microelectrodes. Surgery was performed under full anesthesia (0.8–3% halothane, moisturized oxygen as the carrier) and the birds were placed in a stereotactic apparatus to enable the electrodes to be guided through a slit in the dura to the input layer of the field L complex in the forebrain that is analogous to the primary auditory cortex (for review see [15]). Localization of recording sites was confirmed by standard histology [16]. The electrodes were fixed to the skull using dental acrylic, and an additional socket was glued to the electrode cluster for carrying a transmitter [16,17]. Surgical wounds were treated with a local anesthetic (lidocain). In all birds, the surgical wounds healed rapidly and the electrodes stayed in place up to several months. The procedures of animal experimentation were 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. Multiple-unit activity was recorded via radiotelemetry using a small FM radiotransmitter (Frederick Haer type 40-71-1, 5.3 g including battery) with a high-impedance input stage. During the recording sessions the individual was sitting in a cage inside a sound-proof booth and was totally unrestrained. The transmitted signal was received by a commercial FM tuner with the antenna placed in the booth, band-pass filtered (500–5000 Hz), amplified and stored at a sampling rate of 32 kHz on the disk of a SiliconGraphics Indy workstation. After rejection of recording artifacts, spikes were detected using a software window discriminator.

Pure-tone stimuli and noise maskers were synthesized digitally by the SGI workstation using its 16-bit digital-to-analogue converter at a sampling rate of 32 kHz. All noise-maskers were generated by digital FIR-filtering of Gaussian white noise. To generate 'unmodulated' maskers of a certain bandwidth, the Gaussian noise was only band-pass filtered (cut-off ≥ 1400 dB/octave). To generate 'comodulated' maskers, the Gaussian white noise was multiplied with a 12.5 Hz low-passed noise (the modulator) prior to band-pass filtering. In the experiments, the masking noise had a bandwidth of 50, 200, 800 or 3200 Hz. A detailed description of stimulus generation can be found in Klump and Langemann [11]. The sound level of the comodulated noise was increased by an amount that ensured that unmodulated and comodulated noise of the same bandwidth had the same long-term acoustic energy.

Tone bursts and noise-maskers could be presented simultaneously by mixing the two output channels of the workstation in a hi-fi amplifier (Yamaha A 520). The stimuli were played through a midrange speaker (McFarlow 100MT) mounted at the ceiling of the booth. The set-up in the sound-proof booth was calibrated using a General Radio 1982 precision sound-level meter with the microphone placed 80 cm below the speaker, i.e. at which the bird's head would be in the experiment.

For each multiple unit, first a frequency tuning curve (FTC) was determined using statistical criteria derived from signal detection theory and the unit's characteristic

frequency (CF, the frequency that elicited a response at a minimal level) and best threshold (the sound-pressure level at the CF) were measured (for details see [16]). The unit's noise threshold was determined from peristimulus time histograms (PSTHs) that were constructed from responses to a pulsed 100 Hz wide band of noise (400 ms duration, the center frequency was always the unit's CF) presented at 13 different spectrum-levels in 5 dB increments. Data from 20 artifact-free responses were averaged.

Analogous to the psychoacoustical experiments in the same species [11], we studied the masking of a test tone (CF, 200 ms duration, 10 ms linear ramp) by a continuously presented noise. Maskers had a spectrum level of, on average, 16.4 dB SPL (i.e. were presented at a level that was on average 20 dB above the neurons' threshold determined with a noise stimulus of 100 Hz bandwidths at each recording site). The test tone was presented in 5 dB increments from 10 dB SPL to 70 dB SPL. At each level, tone bursts were repeated until 20 artifact-free recordings were obtained. The response threshold indicating detection of the test tone in noise by the neurons was defined as the mean spike rate observed while the masker alone was presented plus the standard deviation multiplied by a factor of 1.8 (this corresponds to a d' of 1.8 in signal detection theory terms as the threshold criterion). Similar to the psychoacoustic study in the starling, the neuronal release from masking was determined as the difference between the threshold for detection of the test-tone when masked with unmodulated and with a comodulated noise of the same bandwidth. Positive values for the release from masking indicate lower detection thresholds for tones in comodulated noise.

A Wilcoxon matched-pairs signed-ranks test was employed for comparisons on pairs of measurements from the same recording site. Multiple comparisons of repeat measurements were done with a Friedman test. Data from independent samples of clusters were compared with a Mann-Whitney U-test. All p values are two-tailed.

RESULTS

Multi-unit recordings from a total of 26 recording sites in the forebrain provided the data base for the analysis. When stimulated with tones, all clusters showed phasic-tonic (i.e. primary-like) response characteristics followed by an off response or by spontaneous activity in 19 and seven clusters, respectively. Average (\pm s.d.) spontaneous activity was 73 ± 22 impulses/s. The multi-unit responses were tuned to CFs ranging from 1.2 kHz to 6.0 kHz covering the starling's range of sensitive hearing. The unit's best thresholds ranged from 11.4 to 39.6 dB SPL (21.7 ± 7.1 dB SPL). The majority (69%) of the clusters had inhibitory sidebands flanking the excitatory tuning curves [16]. Units were sharply tuned with an average Q10dB of 5.37 ± 2.36 and an average Q40dB of 3.16 ± 0.94 . The slopes of the excitatory tuning curves were very steep; median high-frequency and low-frequency slopes were 163 and -172 dB/octave, respectively. The bandwidth of the excitatory frequency tuning curve 10 dB above the neurons' threshold was 743 Hz on average.

The responses that were recorded from the multi-unit clusters when stimulated with unmodulated band-pass noise varied significantly with the bandwidth of the noise

(range 50–3200 Hz; $p < 0.00005$, Friedman test, see Fig. 2a). Note that the spectrum level of the noise was held constant resulting in an increase of 6 dB for each step of increasing noise bandwidth. The response strength first increased from 101 impulses/s at a noise bandwidth of 50 Hz to 127 impulses/s at a noise bandwidth of 800 Hz and then decreased to 117 impulses/s with a further increase in the noise bandwidth to 3200 Hz (all pair-wise comparisons except the one between noise bandwidth 200 Hz and 3200 Hz revealed significant differences; $p < 0.05$, Wilcoxon test). A similar response pattern was found when the units were stimulated with comodulated noise of the same range of bandwidths and an intensity that was equal to the intensity of the unmodulated noise of the same bandwidth. As in the case of the unmodulated noise, the response strength varied significantly with the bandwidth of the noise (range 50–3200 Hz; $p < 0.0001$, Friedman test, see Fig. 2a). The response strength first increased from 97 impulses/s at a noise bandwidth of 50 Hz to 123 impulses/s at a noise bandwidth of 800 Hz and then decreased to 111 impulses/s with a further increase in the noise bandwidth to 3200 Hz (all pair-wise comparisons except the one between noise bandwidth 200 Hz and 3200 Hz revealed significant differences; $p < 0.05$, Wilcoxon test). With the exception of the responses at a noise bandwidth of 50 Hz ($p < 0.005$), there were no significant differences in the units' responses when stimulated with the unmodulated or the comodulated noise.

Although the response strength to the noise alone did not differ for noise bandwidths ranging from 200 Hz to 3200 Hz, the detection thresholds for the CF tones differed significantly between the unmodulated and the comodulated condition (all $p < 0.01$, Wilcoxon test, see Fig. 2b). Compared with the units' tone-derived best threshold in quiet, the units' average tone threshold in the presence of the noise maskers was elevated between 16.9 and 27.3 dB. The signal to noise ratio (i.e., the level of the tone relative to the spectrum level of the noise masker) for the masker of bandwidth 3200 Hz was 29.2 dB. On average, masked thresholds for detection of the tone were decreased by 3.3 dB, 5.2 dB, 4.7 dB and 6.5 dB in the comodulated condition versus the unmodulated condition for maskers of a bandwidth of 50 Hz, 200 Hz, 800 Hz and 3200 Hz, respectively. For masker bandwidths of ≥ 200 Hz between 23% and 31% of the recording sites showed a masking release of ≥ 10 dB and between one and three of the recording sites showed a masking release that was close to the value of the average behavioral masking release [11] of 18 dB or even better. The relative number of clusters showing a masking release was relatively independent of bandwidth of the noise ranging from 69% for maskers of 800 Hz bandwidth to 81% for maskers of 3200 Hz bandwidth. There was no significant variation in the neurons' masking release when stimulated with noise of differing bandwidth ($p = 0.60$, Friedman test), i.e. the masking release did not increase with increasing bandwidth (Fig. 2c). There was no difference in masking release between the condition in which the bandwidth of the noise was restricted to the excitatory tuning curve of the cluster versus the largest bandwidth of 3200 Hz. Recording sites with substantial inhibitory sidebands did not show more CMR for the 3200 Hz maskers than those lacking inhibitory sidebands (U-test, $p > 0.1$).

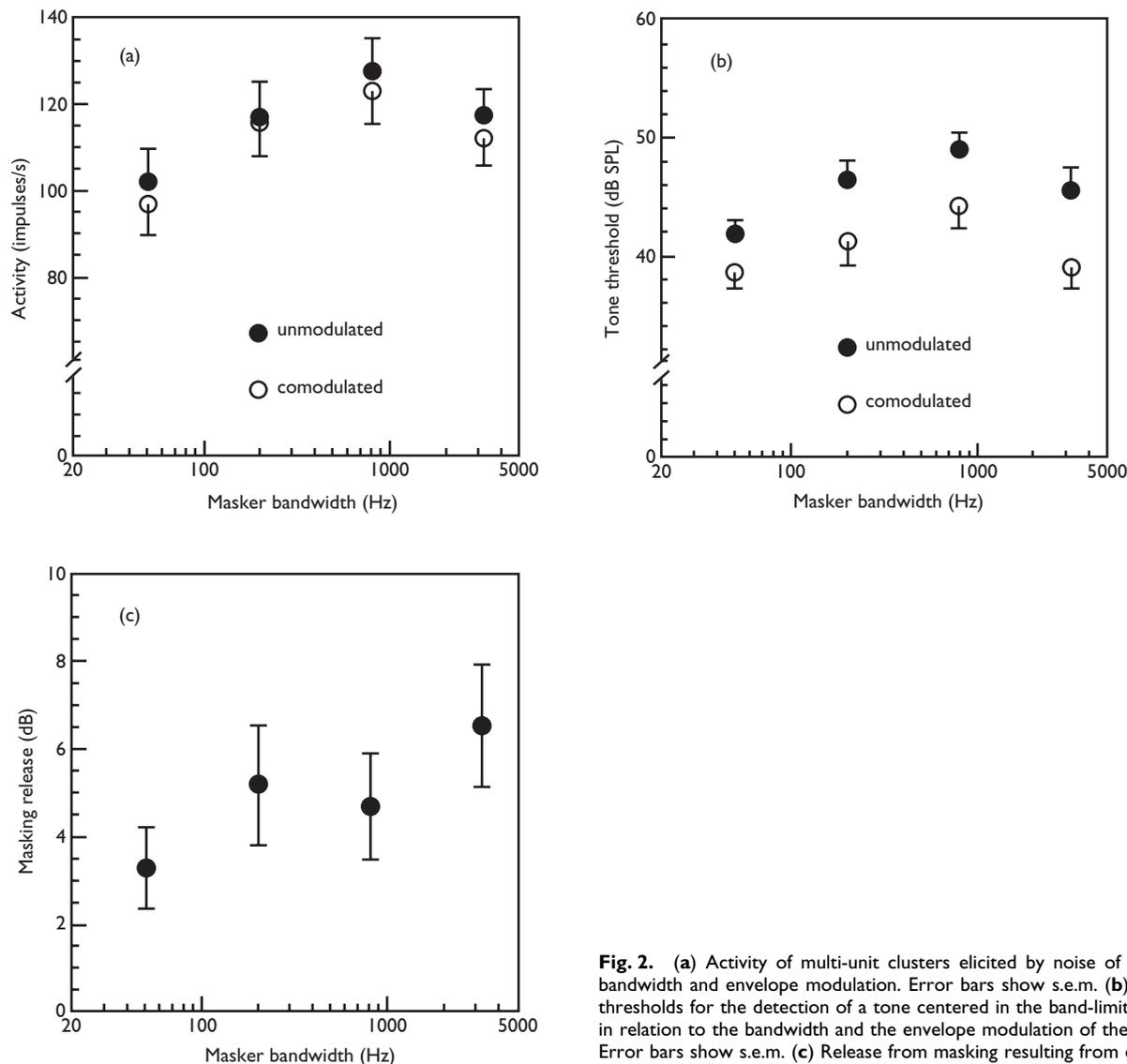


Fig. 2. (a) Activity of multi-unit clusters elicited by noise of different bandwidth and envelope modulation. Error bars show s.e.m. (b) Masked thresholds for the detection of a tone centered in the band-limited noise in relation to the bandwidth and the envelope modulation of the masker. Error bars show s.e.m. (c) Release from masking resulting from envelope modulation in relation to the masker bandwidth. Error bars show s.e.m.

DISCUSSION

Unmodulated masking noise stimulated the units and affected neuronal signal-detection thresholds for tones as predicted from the clusters' excitatory tuning curves and a previous study of simultaneous tone-on-tone masking in the starling [16,17]. Since the noise maskers had a constant spectrum level, an increase of the masker bandwidth resulted in a larger overall intensity of the masker providing excitation and should thus have led to an increase of the neural response up to a noise bandwidth corresponding to the width of the clusters' excitatory tuning curve. The observed increase in the clusters' impulse rate up to a noise bandwidth of 800 Hz matched quite well the bandwidths of the units' excitatory tuning curves of 743 Hz measured 10 dB above their threshold. A further increase in the bandwidth of the stimulating noise resulted in a reduced response. This can be expected given that most of the clusters displayed inhibitory sidebands together with their excitatory tuning curves [17]. The signal-to-noise ratio

of 29 dB, which was needed by the neurons for detecting a tone in an unmodulated wide-band masker, was as would be predicted from the Q10-dB bandwidth of the tuning curves of 743 Hz that was determined with tones. If we assume that the sound energy of the tone at the neuronal detection threshold is equal to the sound energy of the noise passed through the neuronal frequency filter (this assumption is commonly made in psychophysical studies of masking using wide-band noise [10,18]), we can compute the neurons' equivalent rectangular filter bandwidth from the signal to noise ratio at threshold. The sample of neurons of this study had a noise-derived equivalent rectangular filter bandwidth of 787 Hz, indicating that the neuronal filter bandwidth derived in a masking paradigm is similar to the Q10-dB bandwidth of the tone-derived excitatory tuning curve. The neuronal critical masking ratio of 29 dB was slightly larger than the value of the psychophysical critical masking ratio that varied between 21 and 27 dB for frequencies from 1 to 6.3 kHz, respectively.

In general, comodulated bands of noise resulted in the same amount of neuronal excitation as unmodulated noise of similar bandwidth and overall long-term sound energy. This would be predicted if the neurons acted as integrators with a long integration time (which has been estimated from psychophysical data to have a value of about 0.5 s [19]). We cannot explain why a significant reduction of the response occurred in cases in which the noise of a bandwidth of 50 Hz was comodulated. Despite the similarity of the response elicited by the unmodulated and the comodulated masker alone, the detection threshold for the tone was significantly lower in the comodulated masker. Contrary to results from a behavioral study in the starling [11] in which an average release from masking by comodulation of 18 dB was found for a noise of a bandwidth of 1600 Hz, the average neuronal release from masking was only 5.6 dB (data from noise bandwidth 800 Hz and 3200 Hz averaged). It has been argued, however, that psychophysical detection thresholds may often be represented by the responses of the most sensitive neurons [20]. If we follow this argument, we can interpret our results as showing a good match between neural and behavioral data.

The average neuronal release from masking was much more similar to the starling's psychophysical release from masking of 5.4 dB that was observed for maskers of 200 Hz, i.e. those confined to a single auditory frequency filter [21]. It is interesting to note in this context that we did not observe an increase in the neuronal masking release for maskers having a larger bandwidth than the bandwidth of the units' excitatory tuning curves. Thus, inhibitory interactions between different frequency filters in the starling's primary auditory forebrain do not appear to play a role in this masking paradigm. This is emphasized by the observation that the masking release was not reduced for clusters that had no clear inhibitory sidebands bordering the excitatory response.

Our neurophysiological data from the starling's auditory forebrain are different in at least some respect from that obtained from other animals. So far, the only other neuronal data are from anesthetized mammals. They provide different explanations for comodulation masking release. The data from the cochlear nucleus of the guinea pig suggest, for example, that inhibitory processes may be involved in generating the release from masking [13] which in the forebrain of the starling, i.e. in the area that is the analogue of the mammalian primary auditory cortex, do not appear to play a role. In the starling, it was the increased rate of neuronal discharge that indicated the occurrence of the tone signal in the masking background and that increase occurred at lower levels of the tone in comodulated maskers. In the auditory cortex of the cat, however, neurons respond to a tone that is added to an amplitude modulated masker by reducing the neuronal discharge and changing their temporal response pattern [2,12].

CONCLUSION

So far, we cannot describe a universal neuronal mechanism

for explaining patterns of comodulation masking release that are observed to be at least qualitatively similar in psychophysical experiments with some birds and mammals. Neurons in the starling's brain react differently in a comodulation masking release paradigm, both with respect to temporal aspects of their discharge and with respect to across-frequency inhibition, than the neurons in other species that have been investigated. It needs to be shown by studying more species whether this may reflect different coding strategies in the brain of birds and mammals. Recent experiments [22,23] analyzing aspects of comodulation masking release using brief probe-tone signals and sinusoidally amplitude-modulated maskers also indicate some differences between starling and human auditory processing. More studies of psychophysics and physiology of neuronal processing mechanisms in awake preparations are needed to provide a better understanding of the basic mechanisms underlying such demanding auditory tasks necessary to analyze natural acoustic scenes.

REFERENCES

1. Klump GM. Bird communication in the noisy world. In: Kroodsma DE and Miller EH, eds. *Ecology and Evolution of Acoustic Communication in Birds*. Ithaca, NY: Cornell University Press; 1996, pp. 321–338.
2. Nelken I, Rotman Y and Bar Yosef O. *Nature* **397**, 154–157 (1999).
3. Richards DG and Wiley RH. *Am Nat* **115**, 381–399 (1980).
4. Bregman AS. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press; 1990, p. 773.
5. MacDougall-Shackleton SA, Hulse SH, Gentner TQ *et al.* *J Acoust Soc Am* **103**, 3581–3587 (1998).
6. Wisniewski AB and Hulse SH. *J Comp Psychol* **111**, 337–350 (1997).
7. Hall JW, Haggard MP and Fernandes MA. *J Acoust Soc Am* **76**, 50–56 (1984).
8. Schooneveldt GP and Moore BCJ. *J Acoust Soc Am* **85**, 273–281 (1989).
9. Moore BCJ. *J Acoust Soc Jpn (E)* **13**, 25–37 (1992).
10. Scharf B. Critical bands. In: Tobias J, ed. *Foundations of Modern Auditory Theory*. New York: Academic Press; 1970, p. 157–202.
11. Klump GM and Langemann U. *Hear Res* **87**, 157–164 (1995).
12. Nelken I, Jacobson G, Ahdut L *et al.* Neural correlates of co-modulation masking release in auditory cortex of cats. In: Houtsma AJM, Kohlrausch A, Pries VF and Schoonhoven R, eds. *Physiological and Psychophysical Bases of Auditory Function*. Shaker Publishing; 2001, pp. 282–289.
13. Winter IM, Pressnitzer D and Meddis R. *Abstr Assoc Res Otolaryngol* **23**, 181–182 (2000).
14. Mott JB, McDonald LP and Sinex DG. *J Acoust Soc Am* **88**, 2682–2691 (1990).
15. Carr CE and Code RA. The central auditory system of reptiles and birds. In: Dooling RJ, Popper AN and Fay RR, eds. *Comparative Hearing: Birds and Reptiles*. New York: Springer; 2000, pp. 197–248.
16. Nieder A and Klump GM. *Hear Res* **127**, 41–54 (1999).
17. Nieder A and Klump GM. *Exp Brain Res* **124**, 311–320 (1999).
18. Langemann U, Klump GM and Dooling RJ. *Hear Res* **84**, 167–176 (1995).
19. Buus S, Klump GM, Gleich O *et al.* *J Acoust Soc Am* **98**, 112–124 (1995).
20. Parker AJ and Newsome WT. *Annu Rev Neurosci* **21**, 227–277 (1998).
21. Klump GM, Langemann U and Nieder A. Mechanisms that improve signal detection in noise: a study of co-modulation masking release in a songbird. In: Palmer AR, Rees A, Summerfield AQ and Meddis R, eds. *Psychophysical and Physiological Advances in Hearing*. London: Whurr Publishers; 1998, pp. 270–276.
22. Nieder A and Klump GM. *Eur J Neurosci* **13**, 1033–1044 (2001).
23. Langemann U and Klump GM. *Eur J Neurosci* **13**, 1025–1032 (2001).

Acknowledgements: This study was funded by grants from the DFG, SFB 204 and FOR 306.