

# Spectro-temporal response fields in the inferior colliculus of awake monkey

43.64.QH

Versnel, Huib; Zwiers, Marcel; Van Opstal, John

Department of Biophysics  
University of Nijmegen  
Geert Grooteplein 21  
6525 EZ Nijmegen  
The Netherlands  
phone: +31-24-3614689  
fax: +31-24-3541435  
email: [huibv@mbfys.kun.nl](mailto:huibv@mbfys.kun.nl)

## ABSTRACT

The spectro-temporal response field (STRF) characterises the excitatory and inhibitory response properties of an auditory neuron, and it can be used to linearly predict responses to a wide range of sound stimuli. Using broadband sounds with spectro-temporal modulations ('ripples'), we determined STRFs from single-neuron recordings in the inferior colliculus of awake rhesus monkeys. A wide variety of ripple responses was found. For most neurons, responses to natural sounds could be well predicted on the basis of the STRF. These results suggest that spectro-temporal features are extracted at the level of the IC.

## INTRODUCTION

Natural sounds as speech and music are generally broadband and characterized by complex dynamic spectral features. An important theme in auditory neuroscience concerns the neural processing of such sounds: how and where along the auditory pathway are features such as pitch, loudness and timbre represented? Auditory neurons have been classified on the basis of responses to simple stimuli such as tones and amplitude modulated (AM) broadband noise. Responses to tones provide spectral information: e.g. the neuron's frequency response yields its best frequency (BF) and tuning bandwidth. Responses to AM noise provide temporal information: e.g. the modulation transfer function (Langner and Schreiner, 1988). However, these stimuli are very different from natural sounds, and the response features thus obtained are not sufficient to understand the mechanisms that underlie the processing of complex sounds. More appropriate stimuli should preferably integrate the spectral and temporal properties of sounds, yet allowing for a straightforward parameterization of these properties. One class of such stimuli is the so-called 'ripple': sinusoidal envelope modulations in the spectro-temporal domain, which are elementary 2D Fourier components (Kowalski et al. 1996a, Klein et al. 2000, Depireux et al 2001). These stimuli are fully characterized by two parameters, one temporal and one spectral, and their spectro-temporal characteristics resemble those of natural sounds. From the neural responses to different ripples the spectro-temporal response field (STRF) can be derived. On the basis of the STRF and the spectrogram of an arbitrary stimulus, the response to that stimulus can be predicted by applying linear systems theory. A comparison of this linear prediction and the actual response yields information on the linearity of the neural processing.

Most electrophysiological studies on the auditory system, including those using spectro-temporal stimuli, have been performed in anesthetized animals. It is conceivable that

anaesthesia affects the neural response characteristics. Therefore, we studied auditory processing of spectro-temporal sounds in awake rhesus monkeys. We performed extracellular single-unit recordings in the inferior colliculus (IC) to ripples, tones, and animal vocalizations. The distribution of STRFs was examined and the spectral dimension of STRFs was compared to the tuning curves obtained with tones. Finally, we examined whether neural responses to natural vocalizations could be predicted on the basis of STRFs.

## METHODS

### Recordings

Single-neuron responses were recorded with glass-coated tungsten microelectrodes (impedance 0.5-1.5 M $\Omega$ ) in the IC of two adult male rhesus monkeys (*macaca mulatta*; weight  $\pm$ 7.5 kg). The chronic recording chamber (16 mm diameter) was placed over a trepane hole in the skull, centered above the midline, and 2 mm posterior of the interaural line (Goossens and Van Opstal, 2000, for details). Monkeys were rewarded with a drop of water for fixating a visual target at a straight ahead location for 2 s during each trial. In monkey A, recordings were done in each hemisphere; in monkey B, in the right hemisphere only. Recording sites were evenly selected across an area of about 20 mm<sup>2</sup>. In a typical penetration, 1-3 units were recorded at any depth where acoustically evoked responses were encountered (no restriction to the central nucleus of IC). Prior to entering the IC the electrode passed through the superior colliculus. The BF of auditory neurons was found to increase in an orderly manner in dorsal-ventral direction, indicative for the IC (Langner and Schreiner, 1988).

### Sound stimuli

Stimuli were presented in the free field at the frontal central position at a distance of 0.9 m from the monkey. Pure tones were presented over a wide range of frequencies (223 – 20159 Hz; 1/6 octave steps) at three sound levels (40 - 50 - 60 dB SPL), with a duration of 200 ms. Ripple stimuli consisted of a broadband complex of 126 components equally distributed (20/octave) from 0.25 to 20 kHz. The envelopes were sinusoidally modulated in the spectro-temporal domain, such as shown in Fig. 1. In a typical experimental run, we presented all combinations of the ripple density  $\Omega$  (-2:0.4:2 cyc/oct), and the ripple velocity  $W$  (8:8:40 Hz), once. Note that  $\Omega < 0$  corresponds to an upward direction of the spectral envelope (Fig. 1, *top*),  $\Omega > 0$  to a downward direction (Fig. 1, *middle*), and that  $\Omega = 0$  is equivalent to a pure amplitude modulation (Fig. 1, *bottom*). The modulation depth was 100 %, the sound level was 56 dB SPL, and the duration was 2500 ms.

We presented 6 vocalizations, 3 macaque calls and 3 bird-song segments, which were obtained from other labs ([www.wjh.harvard.edu/~mnkylab/media/rhesuscalls](http://www.wjh.harvard.edu/~mnkylab/media/rhesuscalls) and [www.mathlab.sunysb.edu/~tony/birds](http://www.mathlab.sunysb.edu/~tony/birds)) and resampled to 50 kHz. Artefacts and background noises were mostly removed. The calls were presented at 3 sound levels (40-50-60 dB SPL), durations varied from 300 to 1600 ms, and the number of repetitions was 5.

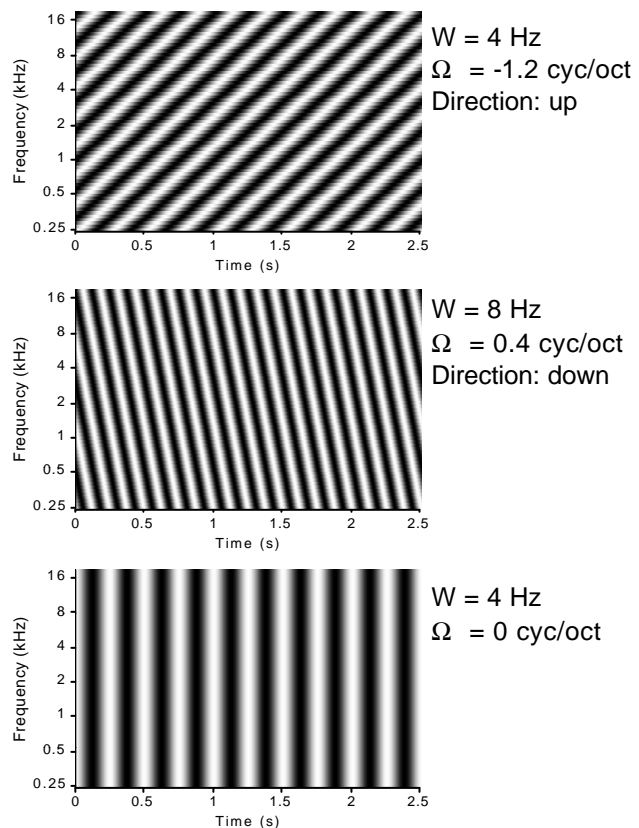


Fig. 1. Spectrograms of three ripple stimuli.

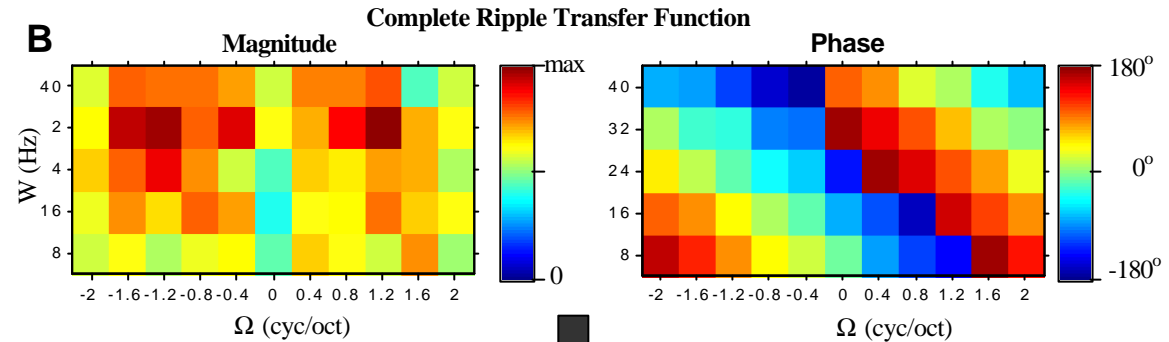
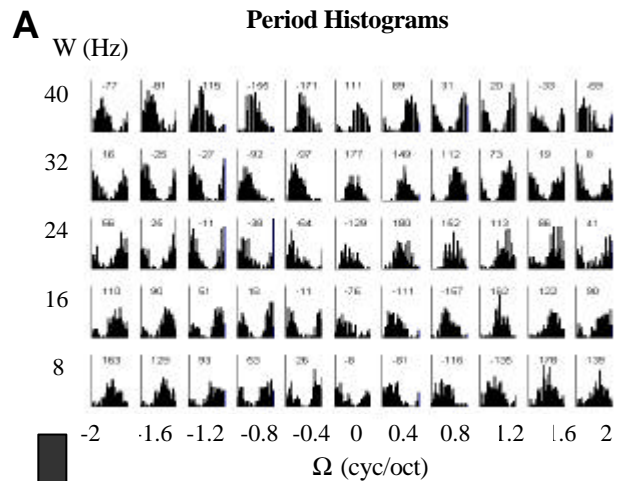
### Ripple-data analysis

A response window of 2250 ms was used, from 250 ms after stimulus onset (thus avoiding transient onset responses) to stimulus offset. Period histograms (Fig. 2A) were constructed based on the period of the temporal modulation of the stimulus (inverse of  $W$ ). A Fourier analysis was performed on the period histograms, and a magnitude  $A(W, \Omega)$  and phase  $\phi(W, \Omega)$  were derived (Fig. 2B), yielding a complete spectro-temporal transfer function  $T(W, \Omega) = A(W, \Omega) * \exp(i\phi(W, \Omega))$ .

Inverse Fourier transformation of  $T(W, \Omega)$  produces the STRF (Fig. 2C):

$$\text{STRF}(t, x) = F^{-1}[T(w, \Omega)] \quad (1)$$

with  $x$  the tonotopic distance in octaves. The spectral section of the STRF reflects



frequency tuning, the temporal section reflects the impulse response. The temporal response pattern of a cell to a sound stimulus,  $R(t)$ , can be predicted by time-convolution and spectral integration with the stimulus spectrogram  $S(x, t)$ :

$$R(t) = \sum_x \{ \text{STRF}(x, t) \otimes_t S(x, t) \} \quad (2)$$

Three response parameters were derived from the transfer function  $T$ : the cell's best ripple velocity, its best ripple density, and its direction preference  $D$ :

$$D = (R_{\text{up}} - R_{\text{down}}) / (R_{\text{up}} + R_{\text{down}}) \quad (3)$$

A best frequency ( $\text{BF}_{\text{strf}}$ ) and latency were derived from the STRF.

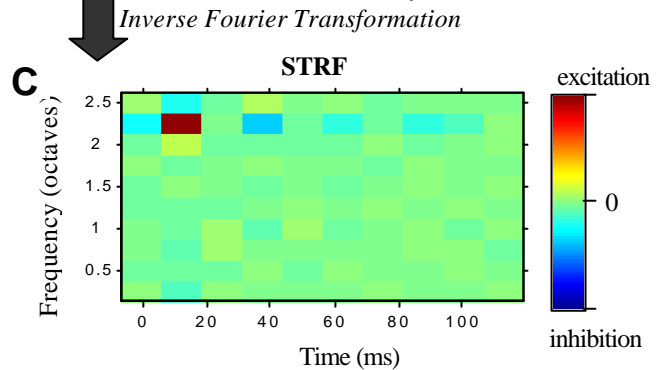
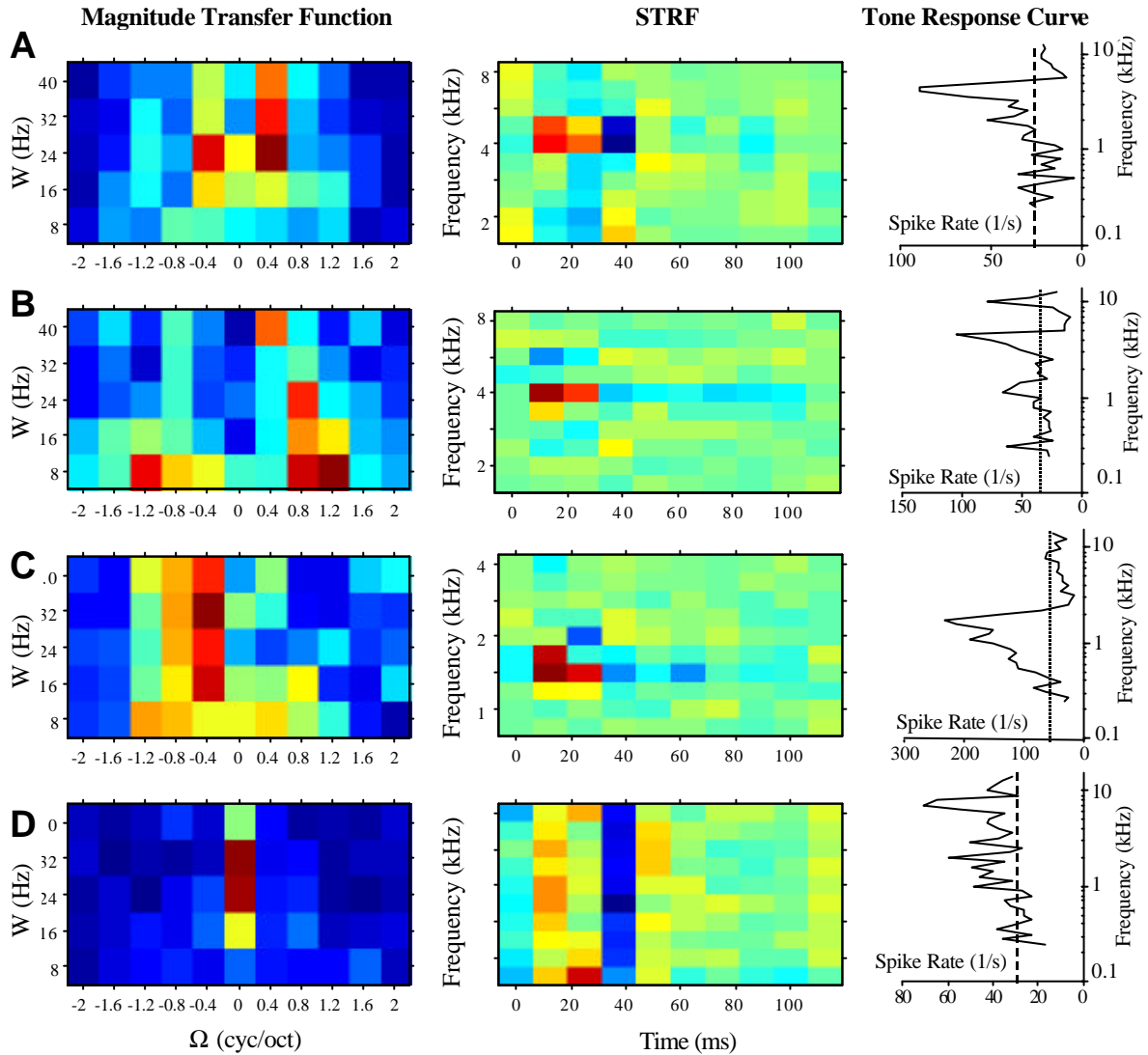


Fig. 2. Analysis of ripple responses.

## RESULTS

### Ripple responses

Most neurons in IC followed the phase of the ripple modulation (e.g., see the period histograms in Fig. 2A). We found that these neurons were bandpass tuned to certain ripple velocities and densities (magnitude TFs in Figs 2B, 3). The ripple preference clearly varied between cells (compare Fig. 3A to 3B). Often the transfer function was symmetric: the upward-direction responses were similar to downward-direction responses, including the best ripple density and best ripple velocity (Figs. 2B, 3A,B). However, we also found cells with a clear preference for one direction (Fig. 3C), 12/89 cells showed such a preference (Fig. 4). A substantial proportion



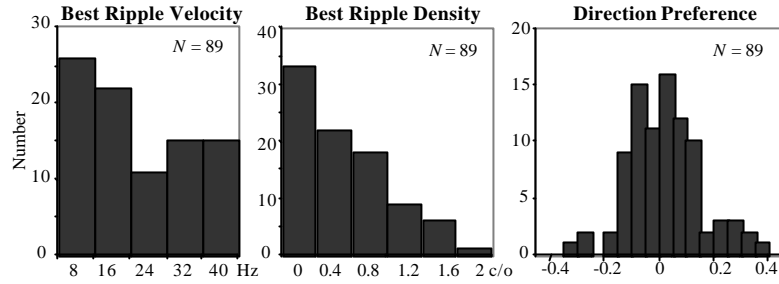
**Fig. 3.** The magnitude transfer function, STRF and tone response curve (frequency tuning) for four different IC neurons. A. Example of high best ripple velocity (24 Hz) and low ripple density (0.4 cyc/oct).  $D=0.01$ .  $BF_{STRF}=4.0$  kHz. B. Example of low best velocity (8 Hz) and high best density (1.2 cyc/oct).  $D=-0.09$ .  $BF_{STRF}=4.0$  kHz. C. Example of asymmetric TF, direction preference is upward.  $D=0.26$ .  $W_b=32$  Hz,  $W_b=0.4$  cyc/oct.  $BF_{STRF}=1.4$  kHz. D. Example of only zero- $W$  response.  $W_b=32$  Hz,  $D=-0.07$ . In all examples the spectral dimension of the STRF is comparable to the frequency tuning curve. BFs are similar, and even inhibitory regions agree, e.g. inhibition above BF in B. Note scale difference between STRF and tone response curves.

of cells (33/89) preferred a ripple density of zero, i.e., a pure amplitude modulation (Figs. 3D,4). When these cells responded to ripples with  $\Omega \neq 0$ , they did not follow the phase of the envelope.

The STRFs generally showed excitation for a narrow range of frequencies with a latency of 10-20 ms so that a  $BF_{strf}$  could easily be defined (Fig. 3A-C). Inhibition occurred about 20 ms after excitation, and, in several cases the STRF showed inhibition to other frequencies, e.g., in Fig. 3B above  $BF_{strf}$ . The frequency tuning given by the STRF was roughly similar to the frequency tuning obtained from responses to pure tones. In particular, BFs agreed for the majority of cells (58/66): the differences between their ripple and tone estimates of BF were less than 0.5 octave. Often, agreement was also found for the inhibitory sidebands, as in Fig. 3B. Broad frequency tuning coupled to zero ripple density preference was typically observed in cells with a broad tuning to tones (Fig. 3D). Note that temporal response properties given by the STRF are not reflected by the tone response curves.

Figure 4 shows that IC neurons were tuned to a wide range of ripple densities and ripple velocities. Further, symmetric responses ( $D=0$ ) as well as asymmetric ( $D \neq 0$ ) were found and,

Fig. 4. Distributions of ripple response parameters over 89 IC neurons. A few of the cells with high best velocities (32 or 40 Hz) which were additionally tested in a larger velocity range (8-80 Hz; fixed density) sometimes had their best response above 40 Hz.



within the latter group, some cells preferred upward ( $D>0$ ) and others the downward direction ( $D<0$ ).

#### Predictions of vocalization responses

In 19 cells for which an STRF had been obtained, responses to the six different vocalizations were also recorded. Following Eqn. 2, we predicted these responses using a time bin of 12.5 ms. Figure 5 shows two examples of predicted and measured responses. Almost all positive peaks in the response patterns were reproduced by the predictions, in particular in case A. Low-response segments corresponded to small deflections or inhibitory troughs of the predicted trace. The discrepancy between the predicted troughs and recorded near-zero response is a direct result of the straightforward linear prediction which does not account for half-wave rectification as exhibited by neural firings. The STRF in Fig 5B is interesting in that it reveals onset inhibition followed by excitation. Consequently, the vocalization response showed a large excitatory peak after stimulus offset, which was also predicted (albeit too early). Note that while most energy of the bird call was below  $BF_{strf}$ , the prediction was fair.

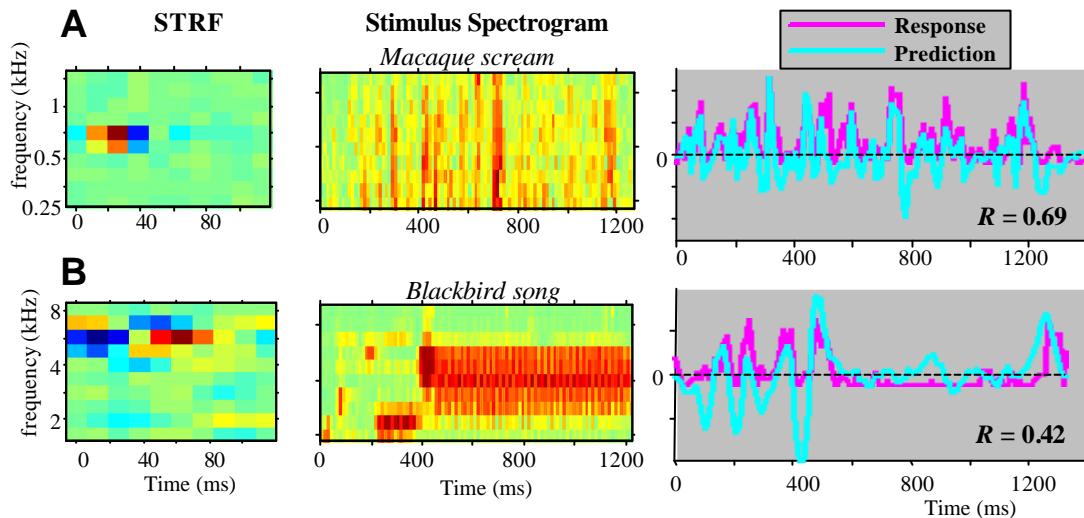
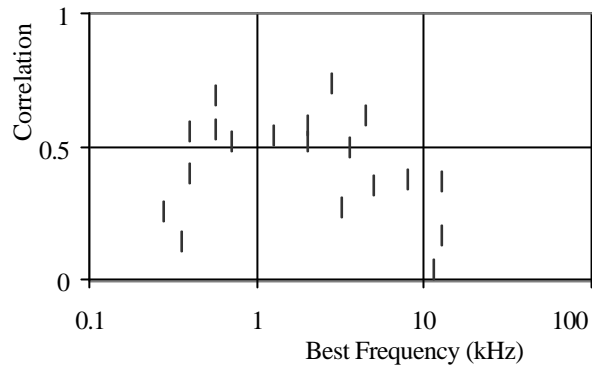


Fig. 5. Prediction of responses to vocalizations on the basis of a cell's STRF and the stimulus' spectrogram (Eqn 2), for two different IC neurons. The spectrogram is shown over the frequency range of the STRF. The spontaneous spike rate (recorded before stimulus onset) has been subtracted from the response. A. Prediction with high correlation to response. B. Prediction with medium-high correlation to response.

In general, for a given neuron, the best prediction is found for the vocalization which spectrum has the largest overlap with the STRF. Figure 6 shows for each cell the best correlation between prediction and response (out of six vocalizations). In 10 cells, correlations were higher than in the example of Fig. 5B ( $R>0.4$ ). In 4 cells, predictions were poor ( $R<0.25$ ). Two of those cells had BFs above 10 kHz and were narrowly tuned, consequently, these cells did not respond to the vocalizations. The two other cells had BFs below 400 Hz, and since the ripples lacked energy below 250 Hz, their STRF might not have been estimated well. The 5 cells with mediocre correlations (0.25-0.4) had all nonmonotonic rate-level curves (tested with broadband noise), and in four of those cells, the correlations were higher ( $>0.4$ ) for lower levels (-10 or -20

Fig. 6. Correlations of predictions with recorded responses to vocalizations for 19 IC neurons. For each neuron, the largest of 6 correlations found for 6 different vocalizations (at a high sound level) is shown.



dB). Taken together, the linear prediction of IC cell responses was remarkably good for the majority of cells. Cases for which the prediction failed could be well explained.

## DISCUSSION

This report presents two important findings on spectro-temporal representation of sounds in the monkey IC. First, many IC neurons show a distinct preference for specific spectro-temporal features, and these preferences vary substantially in both the temporal and spectral dimension across the IC. Second, responses to complex broadband sounds with distinct spectro-temporal features like the animal vocalizations used in our study can be linearly predicted on the basis of responses to ripples, which are the elementary spectro-temporal stimuli. We conclude from these two points that a spectro-temporal Fourier representation of the stimulus spectrum (“ripple analysis”), is already present at the level of the IC. The ripple analysis has previously been demonstrated in the auditory cortex (Kowalski et al. 1996b; Depireux et al. 1998). Other studies (e.g., Langner and Schreiner, 1988) have shown that IC neurons are tuned to temporal modulations over a wide range of modulation frequencies (‘velocities’ in this report), leading to the hypothesis of IC neurons as temporal filters. Our study suggests an extension of this notion to IC neurons as spectro-temporal filters. The auditory cortex might subsequently refine the analysis. The range of temporal, spectral and direction scales might be modified from IC to auditory cortex to enhance the animal’s sensitivity to the acoustic range that dominates its environment and is most relevant to its survival.

## REFERENCES

- Depireux, D.A., Simon, J.Z. and Shamma, S.A. (1998) Measuring the dynamics of neural responses in primary auditory cortex. *Comments in Theoretical Biology* 5: 89-118.
- Depireux, D.A., Simon, J.Z., Klein, D.J. and Shamma, S.A. (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J. Neurophysiol.* 85: 1220-1234.
- Goossens, H.H.L.M. and Van Opstal, A.J. (2000) Blink-perturbed saccades in monkey. II. Superior colliculus activity. *J. Neurophysiol.* 83: 3430-3452.
- Klein, D., Depireux, D., Simon, J. and Shamma, S.A. (2000) Robust spectro-temporal reverse correlation for the auditory system: optimizing stimulus design. *J. Comput Neurosci* 9: 85-111.
- Kowalski, N., Depireux, D.A. and Shamma, S.A. (1996a) Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. *J. Neurophysiol.* 76: 3503-3523.
- Kowalski, N., Depireux, D.A. and Shamma, S.A. (1996b) Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. *J. Neurophysiol.* 76: 3524-3534.
- Langner, G. and Schreiner, C.E. (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60: 1799-1822.

**Acknowledgements** - Supported by the Netherlands Organization for Scientific Research, the Nijmegen Institute for Cognition and Information, and the University of Nijmegen