

Spectro-Temporal Receptive Fields of Auditory Neurons in the Grassfrog

II. Analysis of the Stimulus-Event Relation for Tonal Stimuli

A. M. H. J. Aertsen, P. I. M. Johannesma, and D. J. Hermes

Department of Medical Physics and Biophysics, University of Nijmegen, Nijmegen, The Netherlands

Abstract. The nature of the stimulus-response relation for single auditory neurons is reflected in the properties of the Pre-Event Stimulus Ensemble: the ensemble of stimuli, preceding the occurrence of an action potential (neural event). This paper describes methods to analyse the spectro-temporal properties of this ensemble. These methods are based on the *analytic signal* representation of acoustic signals and functionals derived from it: the *instantaneous amplitude* and *instantaneous frequency* and the *dynamic power spectrum*. The procedures have been applied to a number of extra-cellular single unit recordings from the grassfrog, *Rana temporaria* L., recorded during presentation of an ensemble of tonal stimuli. The outcome of this analysis describes the *spectro-temporal receptive field* of the neuron under the present stimulus conditions. The procedure, based on the dynamic power spectrum is applicable to an *arbitrary* stimulus ensemble, thus allowing a comparison of the spectro-temporal receptive fields for different types of stimuli.

I. Introduction

The investigation of stimulus-response relations of auditory neurons, in general, is performed on the base of a low-dimensional type of analysis. One aspect of the neural activity (e.g. firing rate) is related to one or more parameters (e.g. frequency or intensity) of the stimulus. Experiments with tone burst stimuli play a dominant role. Well known examples of this approach are the *tuning curve* and the *PST-histogram* (e.g. Kiang et al., 1965), which characterize the *spectral* resp. the *temporal* behaviour of the neuron. The question whether the response of a neuron to different types of stimuli (e.g. noise, natural sounds) can be understood on the base of these characteristics has been answered positively for peripheral neurons (Evans, 1975). For

more central neurons the picture evidently is more complex (Worden and Galambos, 1972; Bullock, 1977; Smolders et al., 1979). Frequency selectivity on itself, in general, proves to be an inadequate characteristic of neural function. The same holds for simple temporal characteristics.

A commonly used way to characterize sound, especially more complex sound, is the *spectrogram* or *sonogram* (e.g. Potter et al., 1947; Singleton and Poulter, 1967): the intensity density of the signal is given as a function of both *frequency* and *time*. Inspired by this observation procedures were developed to analyse the *spectro-temporal selectivity* of auditory neurons for a variety of acoustic stimuli. The present paper is concerned with the methodical aspects of this approach. The spectro-temporal intensity density of the particular stimuli which precede the occurrence of action potentials is evaluated against the intensity density of the overall stimulus ensemble. This analysis, basically a crosscorrelation-type of approach, leads to the "*spectro-temporal receptive field*" of an auditory neuron. The approach is based on the spectro-temporal characterization of sound, described in the accompanying paper (Part I, Aertsen and Johannesma, 1980).

Spectro-temporal receptive fields upon presentation of an ensemble of tonal stimuli were determined for single neurons in the auditory midbrain area (torus semicircularis) of the grassfrog (*Rana temporaria* L.). Findings in several anuran species (e.g. Capranica, 1965; Gerhardt, 1974, 1978; Van Gelder et al., 1978) indicate that both spectral and temporal aspects of species-specific vocalizations are behaviourally significant. These findings suggest that the combined spectro-temporal approach may provide information about the neural basis of behaviour in the natural acoustic environment. Two types of spectro-temporal analysis have been used: a *parametric* description (*instantaneous frequency* and *amplitude*) and a de-

scription by means of the *dynamic power spectrum*. The first one is especially suited for narrowband stimuli, the second one is more general and can be used for an arbitrary stimulus ensemble. The latter one will also be used in the investigation of spectro-temporal receptive fields upon stimulation with natural stimuli. This will be the subject of a forthcoming paper (Part III).

2. Methods

Preparation and Recording

Extracellular single unit recordings were made in the torus semicircularis of adult grassfrogs from Ireland. Before the preparation the frog was anaesthetized in a 0.1% solution of MS-222 (ethyl-*m*-aminobenzoate methanesulphonate). After the disappearance of the hindleg reflex upon pinching between the toes the concentration was lowered to 0.05%. After removal of a portion of the skin on the head, a hole was drilled into the region of the skull overlying the tectum opticum. The dura mater was removed. The wound was covered with paraffin oil to prevent it from drying.

During the experiments the frog, wrapped up in a thin layer of moist cotton gauze, was fixated into a small holder and was moistened continuously. The purpose of the latter was to prevent the frog from drying and to maintain the skin respiration. The moistening fluid was a 0.025–0.05% solution of MS-222, this in order to maintain anaesthesia at a level at which corneal reflex returned but no other movements occurred.

Experiments were carried out at a temperature of 14–16°C. A microelectrode was advanced in the brain by a remotely controlled hydraulic microdrive (Trent Wells). Two types of microelectrodes were used: teflon coated, Insl-X insulated tungsten electrodes (typical impedance 1–3 MΩ at 1000 Hz) and glass micropipettes, filled with woods metal-indium alloy, with platinum tip (Dowben and Rose, 1953) (typical impedance 300–800 kΩ at 1000 Hz). Recordings from both types of electrodes did not show strikingly different results. Single unit recording times ranged from 5 min up to 2 h.

Stimulus Presentation and Data Registration

Experiments were performed with the frog in an electrically shielded acoustic chamber (IAC 1202A). The stimulus ensemble consisted of sequences of short tonal elements, the Gamma-tones (γ -tones). The characteristics of this tonal ensemble and the technical procedures of generation are described in detail in the companion paper (Part I, Aertsen and Johannesma, 1980). A segment of a sequence of γ -tones is shown in Fig. 1.

The stimulus ensemble was generated by a Programmable Stimulus Generator (PSG), controlled by a PDP 11/10. The signal was amplified (Luxman 507X) and the sound was delivered by two electrodynamic microphones (Sennheiser MD 211N), which were coupled to the frog's tympanic membranes by means of closed cavities, filled with damping material to reduce resonances. The sound pressure level was measured using half inch condenser microphones (Bruel and Kjaer 4134), connected to the acoustic couplers. The frequency response of the sound system, as measured in situ, was flat within 10 dB for frequencies between 100 and 3000 Hz, with considerable loss for frequencies exceeding 3000 Hz (some 24 dB/oct). Identical stimuli were presented to both ears in all cases. The average sound intensity was in the range of 60–80 dB SPL.

Neural activity was amplified (Grass P16), action potentials were transformed into unitary pulses by means of a level discriminator and fed to a Data Acquisition System (DAS), built around a PDP 11/45. The timing of both the PSG and the DAS was controlled

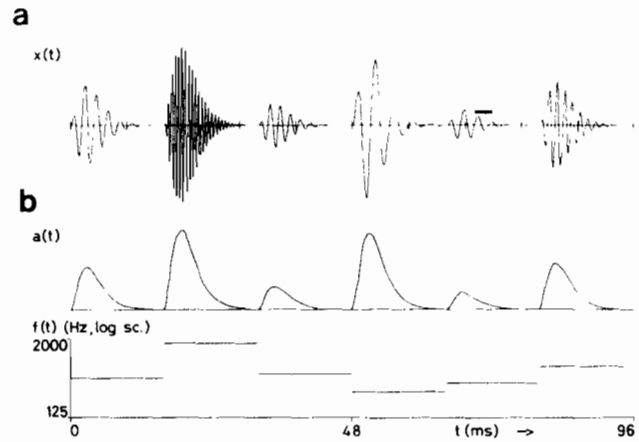


Fig. 1a and b. Segment from a sequence of γ -tones. **a** Signal $x(t)$ as a function of time. **b** Upper trace: instantaneous amplitude $a(t)$. Lower trace: instantaneous frequency $f(t) = \frac{1}{2\pi} \omega(t)$

by the same external clock, providing a time resolution of neural events, relative to the stimulus, of 10 μ s.

Spectro-temporal analysis of the stimulus-event relation was performed by means of software (AFT-cube, Average IFT; Sect. 3.2) and hardware (Dynamic Spectrum Analyser; Aertsen and Johannesma, 1980). For the purpose of the latter type of analysis the DAS could be operated in an "inverse" mode, providing unitary pulses in real time from filed spike trains, synchronized with the stimulus from the PSG.

3. Spectro-Temporal Analysis of the Stimulus-Event Relation

3.1. The Pre-Event Stimulus Ensemble (PESE)

Extracellular single unit recordings from auditory neurons during presentation of an acoustic *Stimulus Ensemble* (SE) result in sequences of action potentials. Because of the all or none character of the action potentials the single unit activity can be modelled as a series of neural events, a stochastic point process:

$$z(t) = \sum_{n=1}^N \delta(t - t_n) \quad (0 \leq t \leq T) \quad (3.1)$$

with $\delta(t)$ = Dirac delta function; $\{t_n\}$ = moments of occurrence of action potentials, recorded during presentation of the Stimulus Ensemble; T = duration of the SE; N = number of action potentials.

The question which stimulus properties determine the neural events can be investigated by means of an analysis of the properties of the stimulus-event (*s-e*) relation. The nature of the *s-e* relation is reflected in the characteristics of the *Pre-Event Stimulus Ensemble* (PESE), defined as the ensemble of stimuli which precede a neural event (Johannesma, 1972;

Johannesma et al., 1973; Grashuis, 1974):

$$PESE = \{x_n(\tau) = x(t_n - \tau); n = 1, N; \tau \geq 0\}. \quad (3.2)$$

The concept of the PESE is illustrated in Fig. 2. By definition the PESE forms a subset of the SE. A comparison of the characteristics of the PESE as compared to those of the SE provides a description of the function of the unit in the neural representation of acoustic stimuli.

3.2. Parametric Analysis of the PESE :
AFT-cube and Average IFT

The concepts of the parametric signal representation, described in the accompanying paper (Part I), and the Pre-Event Stimulus Ensemble (Sect. 3.1) will now be combined for the analysis of the single neuron response to a γ -tone ensemble. Apart from a phase constant each stimulus preceding a neural event is fully described by its instantaneous amplitude $a(\tau)$ and instantaneous frequency $f(\tau)$. Therefore each element from the PESE can be geometrically depicted by a trajectory in a 2-dimensional (a, f) -space (i.e. a parameter space) or, equivalently, by a curve in a 3-dimensional (a, f, τ) -space. The latter representation will be used in the following.

The PESE corresponds to the set of pre-event curves in (a, f, τ) -space. The functional properties of the neuron, generating the events, are reflected in this set of PESE-curves as compared to the a priori set of SE-curves. A graphical representation of the parametric description of the PESE (and likewise of the SE) is

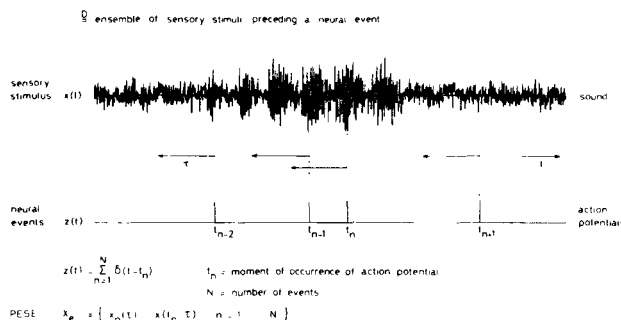


Fig. 2. Definition of the Pre-Event Stimulus Ensemble (PESE)

given by the AFT-cube, shown in Fig. 3. Sampled replicas of PESE-curves are mapped into the AFT-cube, the axes of which denote, respectively, instantaneous amplitude $a(\tau)$ (vertical axis), instantaneous frequency $f(\tau)$ (horizontal axis, log scale), and time τ before the occurrence of action potentials. Values of $a(\tau)$ have been normalized with respect to the maximum amplitude occurring in the γ -tone ensemble. The AFT-cube is made up of a number of (f, τ) -cross sections taken at equidistant values of a . Each cross section contains a number of dots which represent the projections of portions of sampled (a, f, τ) -curves within a small range Δa around the selected value of a . Throughout this paper Δa was chosen 0.1, the sample interval of (PE)SE-curves was 1.01 ms.

The AFT-cube shown in Fig. 3a represents the PESE for the torus semicircularis unit 26-6. This result should be confronted with the characteristics of the SE.

UNIT 26-6

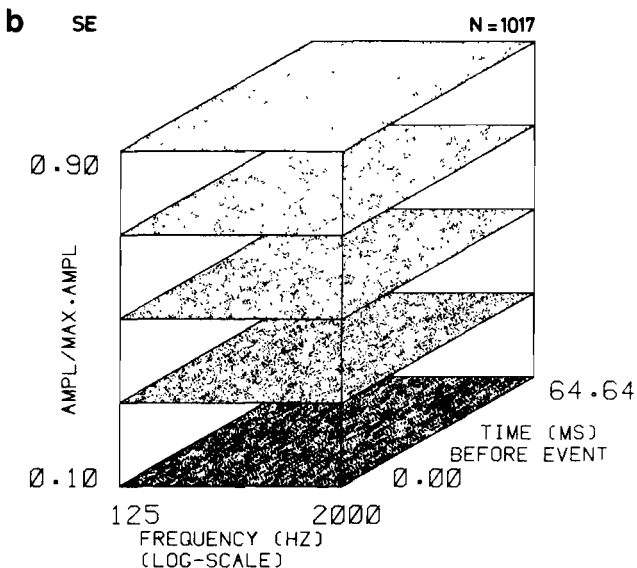
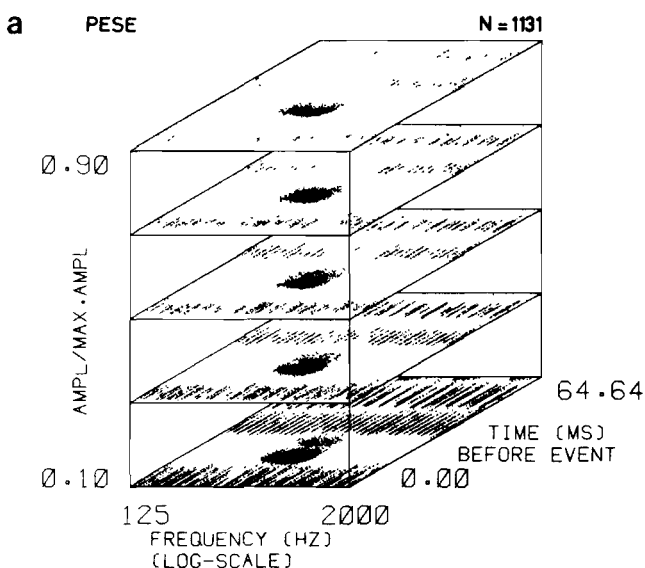


Fig. 3a and b. AFT-cubes: parametric representation of tonal elements from the Pre-Event Stimulus Ensemble of the torus semicircularis unit 26-6 (a) and the Stimulus Ensemble (b). The parameters of interest are: instantaneous amplitude a , instantaneous frequency f and time before the event τ . The cube's axes correspond to these parameters. The number of elements N from the PESE resp. the SE is indicated in the figures. Further explanation in text

UNIT 26-6

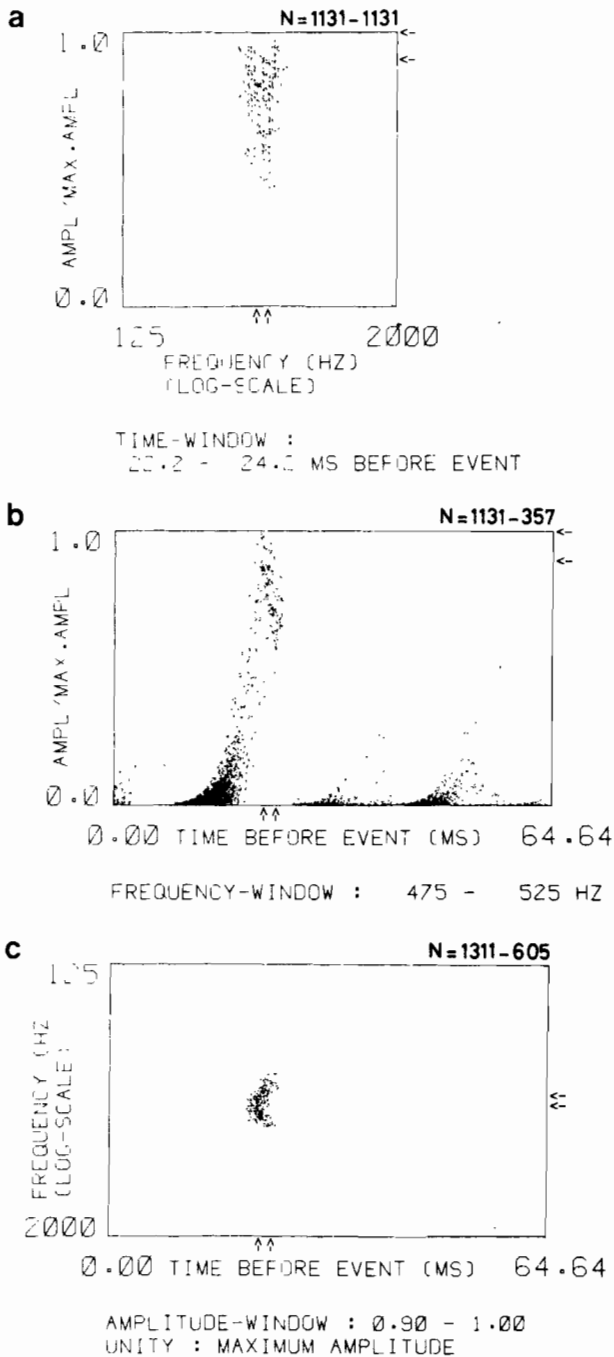


Fig. 4a-c. Cross sections in parameter space of the Pre-Event Stimulus Ensemble of the torus semicircularis unit 26-6 (cf. Fig. 3a). **a** Instantaneous amplitude vs. instantaneous frequency: (a, f) . **b** Instantaneous amplitude vs. time before neural event: (a, τ) . **c** Instantaneous frequency vs. time before neural event: (f, τ) . For each of the cross sections the numerical values of the parameter window and the number of elements from the PESE having values within this window are indicated. In all three cases the complete PESE was subjected to the analysis. The location in parameter space of the various windows (as indicated by arrows) causes the difference between the numbers of elements within the various cross sections. Further explanation in text

These are depicted in the AFT-cube in Fig. 3b. In this case a random sequence of events which bears no specific relation to the γ -tone ensemble supplied the triggering moments. This random sequence of events was made up of the moments of positive going level crossings of a Gaussian wideband noise. From the results in Fig. 3 we observe for unit 26-6 a relative abundance of dots in a specific region of the AFT-cube. More detailed information about the characteristics of this specific region can be obtained from different types of cross sections in (a, f, τ) -space. Some examples for unit 26-6 are shown in Fig. 4. An (a, f) -cross section (Fig. 4a) gives the amplitude-frequency selectivity of the unit at a certain time τ before the spike; dimensionally it may be compared to a "tuning curve". An (a, τ) -cross section (Fig. 4b) gives the temporal amplitude sensitivity at a certain frequency, whereas an (f, τ) -cross section (Fig. 4c) shows the spectro-temporal selectivity at a certain amplitude. The periodic structure of the (a, τ) -cross section (Fig. 4b), also appearing in the AFT-cube (Fig. 3a), is a combined effect of the periodic structure of the stimulus ensemble (every 16 ms one γ -tone) and the neuron's firing characteristics. From Fig. 4b we note that the PESE did not contain any low amplitude tones with the preferred frequency at around 23 ms before the event, which points to an amplitude threshold mechanism. Furthermore we note from Fig. 4b the relative low frequency of occurrence of higher amplitude tones at around 39 ms before the event as compared to the more uniform density at around 7 and 55 ms before the event, which reflects a post-activation suppression phenomenon. From Fig. 4c we note the frequency dependence of the latency. The small regularities along the frequency axis (Figs. 3a and 4c) do not reflect a true neuron property, but are due to the interaction of statistical properties of the γ -tone ensemble and the neural response properties (cf. Discussion and Appendix).

For tonal stimuli both the SE and the PESE lead to a collection of points in (a, f, τ) -space, visualized by the AFT-cube. If the stimulus ensemble is rich enough the number of points per unit volume in (a, f, τ) -space $r(a, f, \tau)$ forms a reliable estimate of the density $q(a, f, \tau)$ of stimuli in parameter space. The chosen resolution of a , f , and τ (i.e. the dimensions of the unit volume in parameter space) and the number of events N , used to compile the AFT-cube determines the precision Δr of the estimate. After proper normalization the density $q(a, f, \tau)$ may be interpreted as a probability density function: the probability per unit extension in parameter space to find particular (a, f, τ) -combinations in the stimulus ensemble.

Similarly the AFT-cube of the PESE leads to the histogram $r_e(a, f, \tau)$ as an estimate of the pre-event

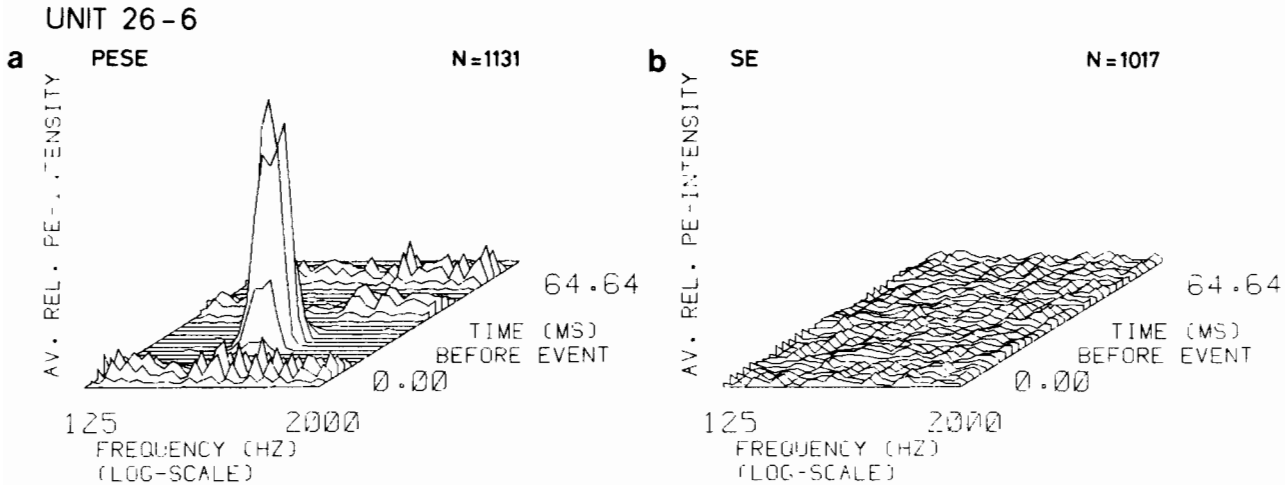


Fig. 5a and b. Average IFT: average intensity I as a function of frequency f and time before the event τ of tonal elements from the Pre-Event Stimulus Ensemble of unit 26-6 (a) and the Stimulus Ensemble (b) (cf, Figs. 3 and 4). The number of elements N from the PESE resp. the SE considered in the averaging procedure is indicated in the figures. The intensity scales in a and b are in terms of arbitrary, identical units. Further explanation in text

probability density in parameter space $\rho_e(a, f, \tau)$. The reliability of this estimate is determined by the a priori precision Δr . The probabilistic interpretation of results from the parametric analysis of the SE and PESE is inspired by the ensemble theory as used in statistical physics (e.g. Tolman, 1938), which is rendered feasible by the stochastic nature of the mechanism underlying neural firing.

On the base of the densities $\rho(a, f, \tau)$ and $\rho_e(a, f, \tau)$ the expected value of any quantity $q(a, f, \tau)$ over the SE respectively the PESE can be calculated. In the study of the auditory system the intensity of a signal as a function of both frequency and time is of interest. The expected value of the spectro-temporal intensity of the sound in the PESE is given by

$$E(I_e(f, \tau)) = \int da a^2 \rho_e(a, f, \tau). \quad (3.3)$$

This function can be estimated by the corresponding average over the PESE

$$\langle I_e(f, \tau) \rangle = \frac{\sum_a a^2 r_e(a, f, \tau)}{\sum_a r_e(a, f, \tau)}. \quad (3.4)$$

The integration in (3.3) and the summation in (3.4) extend over all amplitude values present in the PESE. Similar expressions hold for the a priori expected value over the SE.

The average spectro-temporal intensity of sound in the PESE and the SE for an ensemble of γ -tones is shown for unit 26-6 in Fig. 5.

Ordinate values in Fig. 5a and b refer to identical, arbitrary units. The result of this type of analysis will be referred to as *Average IFT*. The Average IFT of

unit 26-6 is shown in Fig. 6a in another representation, the grey density coding, where intensity is represented by means of a number of discrete grey-levels.

Regarding Figs. 5 and 6a the following observations can be made:

- I. Figure 5b shows that for the SE $\langle I(f, \tau) \rangle$ is a constant value I_0 within statistical variations.
- II. Figure 5a and 6a indicate that for the PESE on the base of values of $\langle I_e(f, \tau) \rangle$ four spectro-temporal domains can be discerned:
 1. The average intensity is relatively high in the domain $(f, \tau) \in (300-750 \text{ Hz}, 15-30 \text{ ms})$.
 2. The average intensity is relatively low (but not zero) in the domain $(f, \tau) \in (300-750 \text{ Hz}, 30-50 \text{ ms})$.
 3. The average intensity is relatively low (down to zero) in the domains $(f, \tau) \in (125-300 \text{ Hz}, 15-30 \text{ ms})$ and $(f, \tau) \in (750-2000 \text{ Hz}, 15-30 \text{ ms})$.
 4. The average intensity is, apart from the regularities already noted, similar to the a priori expected value for all other (f, τ) -combinations.

These observations are summarized schematically in Fig. 6b.

The experimental observations lead to the following conclusions regarding unit 26-6 (see Fig. 6c):

1. There is an increased probability of neural firing upon adequate stimulation with tones: *activation A* with a simple frequency selectivity and timing preference around $(f, \tau) = (520 \text{ Hz}, 22 \text{ ms})$.
2. Once the neuron has fired to an adequate tonal stimulus, there is a decreased probability of firing to the same type of stimulus: *suppression S*, which may be due to refractoriness, adaptation and/or post-excitatory inhibition. The suppression may last up to

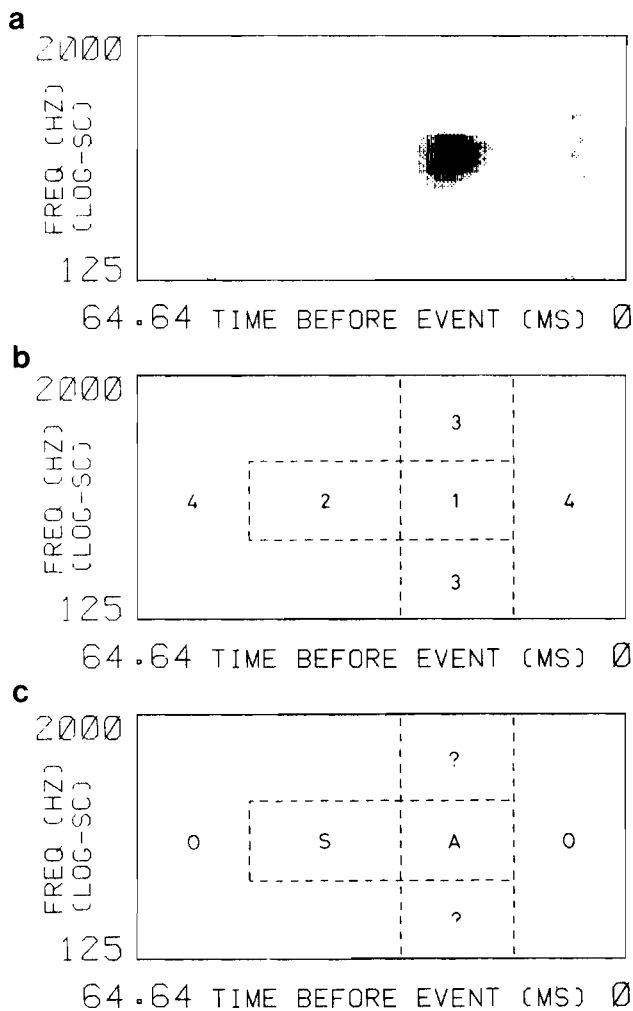


Fig. 6a-c. Average IFT of tonal elements from the Pre-Event Stimulus Ensemble of unit 26-6. **a** Grey density representation of result as shown in Fig. 5a. Note that here the frequency- and time-axes have been interchanged with respect to Fig. 5a. **b** Schematic description of Figs. 5a and 6a. The figures in Fig. 6b refer to the text. **c** Schematic interpretation of Figs. 5a and 6a: spectro-temporal receptive field, consisting of activation (A) and suppression (S). Further explanation in text

20 ms. Due to the definition of τ as time *before* the neural event this suppression domain is *preceding* the activation domain in Fig. 6c.

3. There cannot be made precise conclusions regarding the spectro-temporal domain 3 (possible lateral effects), due to the absence of spontaneous activity, the specific spectro-temporal selectivity of the neuron and the single tone character of the SE. At any rate, in this domain a possible activation is not dominant over a possible suppression.

4. The neuron shows a "memory" which, including the latency of some 15 ms, is in the order of 50 ms.

Summarizing, the average spectro-temporal intensity of the PESE, the particular selection from the SE

as it was made by the neuron, deviates significantly from what could be expected a priori in a specific domain in frequency and time: the *spectro-temporal receptive field* of the auditory neuron as determined by an *ensemble of γ -tones* on the base of a *parametric analysis*. The receptive field for unit 26-6, made up of an activation domain and a suppression domain, is shown schematically in Fig. 6c.

It should be noted that neither the (a, f) -space nor the (a, f, τ) -space fulfils the requirements of a vector space (Lang, 1973). As a consequence simple operations on the trajectories in parameter space may result in trajectories corresponding to signals which do not belong to the original ensemble. As a simple example the addition or averaging of two tonal stimuli does not normally result in a tonal stimulus. Likewise the Average IFT of the PESE based on the use of tonal stimuli represents sound but not necessarily a tonal sound. In general the Average IFT (like the AFT-cube) represents the degree of optimality of sounds for the neuron (optimal in the sense of probability of spike generation). The sounds have been optimized with respect to the frequency and amplitude of the separate components, the optimum phase relations of the components still have to be determined.

3.3. Average Dynamic Power Spectrum of the PESE

A meaningful interpretation of the AFT-cube and Average IFT, as introduced in the preceding section, is restricted to a stimulus ensemble of narrowband sounds, where instantaneous amplitude and frequency, apart from being well defined, are intuitively understandable. Although formal problems not necessarily have to arise, this approach is less fruitful for wideband sounds like noise, clicks, tonal complexes and most natural sounds.

A more general spectro-temporal description of signals is given by the *Dynamic Power Spectrum* (e.g. Flanagan, 1965) or *spectrogram* (e.g. Koenig et al., 1946; Potter et al., 1947). The dynamic power spectrum $P(f, t)$ of a signal gives the intensity as a function of time for the different frequency components of the signal. It can be measured by a *Dynamic Spectrum Analyser* (DSA). The mathematical background and principles of operation of this instrument are described in the accompanying paper (Part I, Aertsen and Johannesma, 1980). The DSA determines the dynamic power spectrum of a signal in the form of a set of functions of time, representing the instantaneous intensity of the signal in a set of neighbouring $1/3$ octave bands.

In general the dynamic spectrum $P(f, t)$, which is a nonlinear functional defined on an arbitrary signal, describes the spectro-temporal structure of that signal.

UNIT 26-6

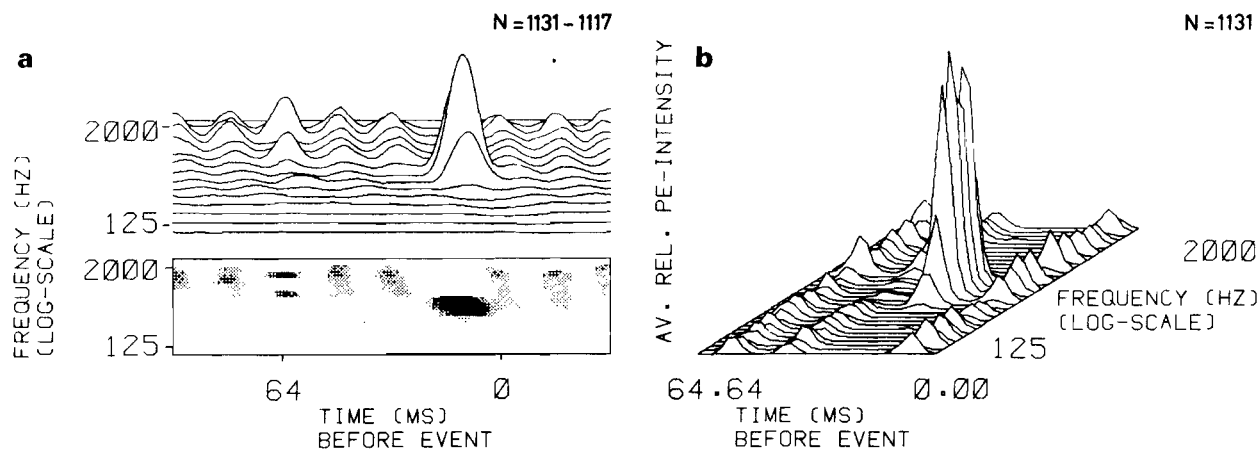


Fig. 7a and b. Spectro-temporal receptive field of torus semicircularis unit 26-6, determined by an ensemble of γ -tones. **a** Different representations of the Average Pre-Event Dynamic Power Spectrum, determined by the DSA. The number of elements N_1 presented to the DSA and the number of elements N_2 actually considered in averaging ($N_2 < N_1$) are indicated in the figure. **b** Average IFT (cf. Figs. 5a and 6a). Note that the frequency- and time-axes have been interchanged with respect to Fig. 5a

In this approach the dynamic spectrum is considered to be an important characteristic of an acoustic stimulus. This leads to the hypothesis that the dynamic spectrum may differ distinctly for signals belonging to the PESE as compared to signals from the SE. A possible way to investigate this hypothesis is the computation of the *Average Dynamic Power Spectrum* of the PESE:

$$\langle P_e(f, \tau) \rangle = \frac{1}{N} \sum_{n=1}^N P_n(f, \tau), \quad (3.5)$$

where $P_n(f, \tau)$ is the dynamic spectrum of the n -th element of the PESE. This procedure of spike triggered averaging of pre-event dynamic spectra is implemented in the DSA (cf. Part I). The experimental results for the PESE of unit 26-6, using again the γ -tone stimulus ensemble, is given in Fig. 7a. For comparison the results of the Average IFT for this neuron, using the same γ -tones, is shown in Fig. 7b. The latter result is identical to the one shown in Fig. 5a; to facilitate a comparison the frequency- and time-axes have been interchanged.

Due to limitations of the DSA ("dead time" upon triggering, in this case about 32.3 ms; cf. Part I) not all of the elements of the PESE have been considered in averaging. The numbers involved are indicated in Fig. 7a. The average dynamic power spectrum of the SE (not shown here), like the Average IFT, gives an approximately constant value for all (f, τ) -combinations.

From this result we conclude that the Average Dynamic Power Spectrum of the PESE for unit 26-6

indeed differs distinctly from the a priori expected result i.e. the Average Dynamic Power Spectrum of the SE. For narrowband signals, remaining within a given 1/3 octave band the dynamic power spectrum reduces to the parametric IFT-representation. Therefore it is not surprising that the spectro-temporal receptive field for γ -tones as determined by the DSA (Fig. 7a), apart from obvious differences in resolution, agrees well with the receptive field determined on the basis of a parametric description (Figs. 6a and 7b). The distortion, caused by the frequency dependent delay in the DSA (cf. Part I), however, should be taken into account in the estimation of time constants like latency.

In the frame of the dynamic power spectrum analysis we have defined and constructed a spectro-temporal analysis of the PESE for a neuron which, contrary to the parametric approach, can be applied for an *arbitrary* choice of the stimulus ensemble. In particular this procedure enables to determine the spectro-temporal receptive field of a neuron in a (pseudo) natural acoustic environment (Aertsen and Johannesma, 1979). This leads to the question whether the spectro-temporal receptive field is dependent on the type of stimulus ensemble and, if so, whether normalizing procedures can be developed, to account for this. In view of the dimensions of the analysis, i.e. frequency and time, a natural first choice would be to correct the Average Dynamic Power Spectrum of the PESE for the overall spectro-temporal structure of the SE. The spectro-temporal receptive fields of auditory neurons for a natural stimulus ensemble will be investigated in more detail in the forthcoming paper (Part III).

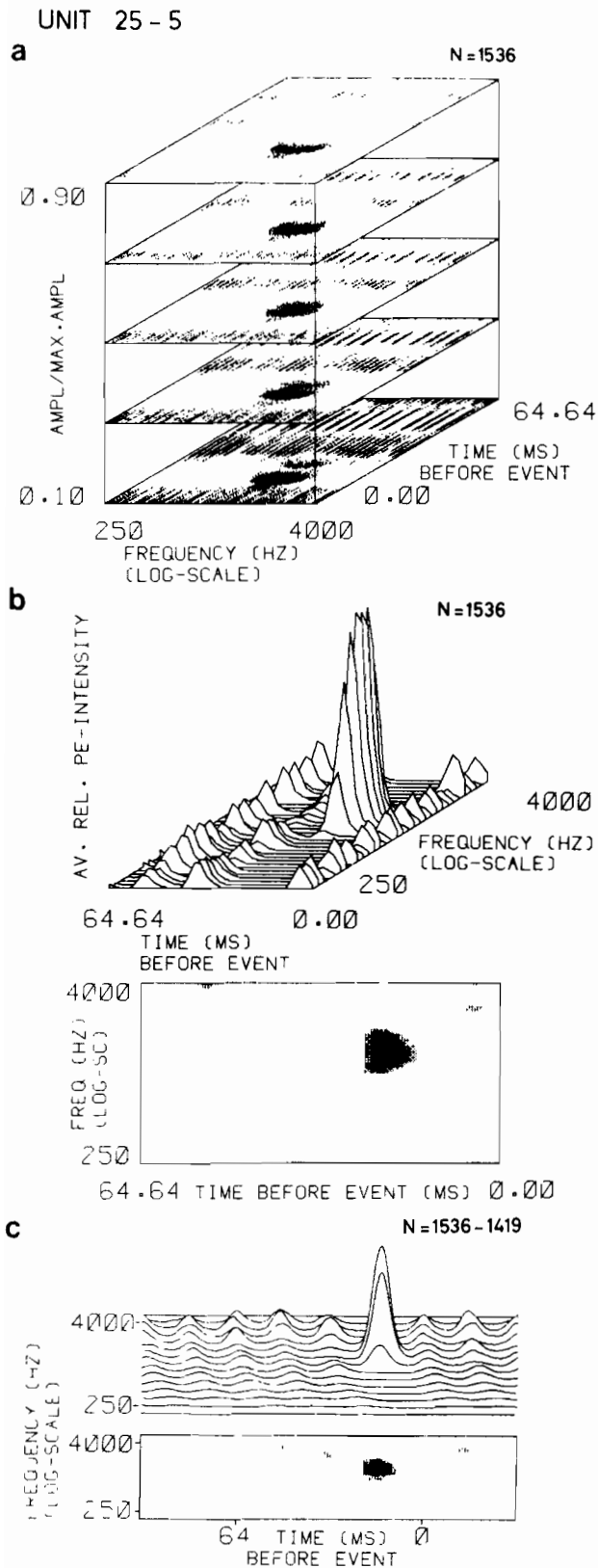


Fig. 8a-c. Spectro-temporal analysis of torus semicircularis unit 25-5, using an ensemble of γ -tones. **a** AFT-cube. **b** Average IFT in different representations. **c** Average Pre-Event Dynamic Power Spectrum in different representations. The numbers of elements from the PESE, taken into account in the various procedures, are indicated in the figures. Further explanation in text

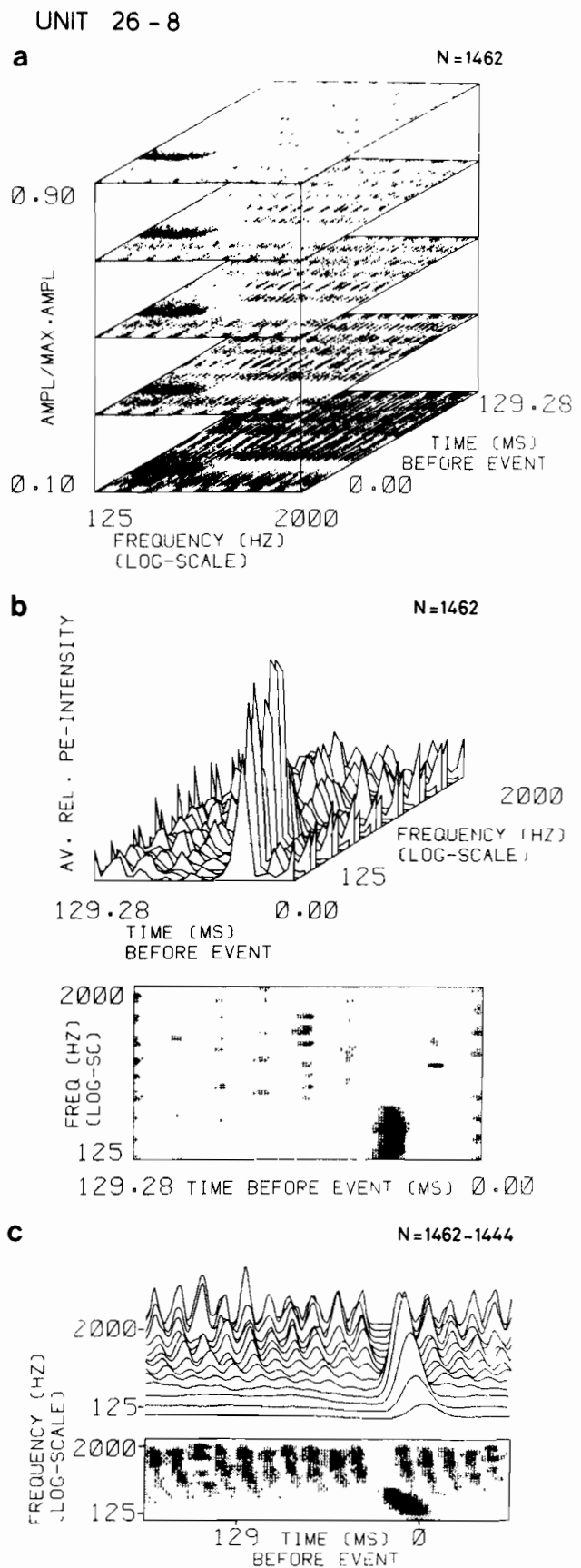


Fig. 9. Spectro-temporal analysis of torus semicircularis unit 26-8, using an ensemble of γ -tones. Details as in Fig. 8

4. Results

The analysis as described in this paper was applied to 29 torus semicircularis recordings, obtained from 11 grassfrogs. This leads to a variety of results, four examples of which are shown in Figs. 8–11.

In all examples shown we present the AFT-cube (a), the Average IFT in two representations (b) and the average dynamic power spectrum, likewise in two representations (c). Note the different frequency and time scales in the various cases. The values of the average measures (b and c) have been scaled to the individual extrema. The “dead-time” in the DSA-analysis (cf. Part I) in these examples amounted to about 32.3 ms (Fig. 8c), resp. 64.6 ms (Figs. 9c, 10c, and 11c). None of the neurons did show any spontaneous activity. The units 25-5 (Fig. 8) and 26-8 (Fig. 9) represent neurons with relatively high firing rate under stimulus conditions: on the average 2.96 resp. 2.82 spikes/s, whereas the units 38-1 (Fig. 10) and 26-4 (Fig. 11) showed a much lower activity: on the average 0.05 resp. 0.12 spikes/s. In the latter two cases the activity decreased considerably during the first few seconds of presentation of the γ -tone ensemble.

In all 29 cases it was possible to distinguish an activation type of spectro-temporal receptive field. In a number of cases the receptive field also showed a more (e.g. Figs. 8 and 9) or less (e.g. Fig. 10) clear suppression domain preceding the activation domain. This shows up most clearly in the grey density representations.

The limited number of recordings does not allow to draw any inferences regarding a possible classification of torus semicircularis neurons into subpopulations.

A closer inspection of the examples shows the following: *Unit 25-5* (Fig. 8) is responsive to tones with frequency within a band of about 1 octave around 1350 Hz. The latency, defined as the interval between the neural event and the moment in time of the maximum value of the Average IFT, is about 20 ms. This particular definition, as arbitrary as any other one, was adopted because it can still be used when the result contains more noise (e.g. Fig. 11). From Fig. 8a we note that the latency in firing to individual tones is frequency dependent, with shortest latency for a frequency of about 1350 Hz and gradually increasing with increasing or decreasing frequency. The quite sharply defined activation type receptive field is preceded by a suppression domain with similar frequency characteristics, which does not extend in time to more than 50 ms before the action potential (Fig. 8b and c).

Unit 26-8 (Fig. 9) shows a wide frequency selective band for activation at low frequencies (125–300 Hz) without a pronounced best frequency. The lower frequency bound of the γ -tone ensemble (125 Hz) obviously is still within this unit's receptive field. The latency, which in this case is hardly frequency dependent (Fig. 9a), is about 33 ms. The sharply defined

activation domain is preceded by a suppression domain at the same low frequencies which comprise the activation domain, lasting up to about 100 ms before the event (Fig. 9b and c). Note that the suppression domain becomes narrower in frequency with increasing time before spike. Especially in this case (Fig. 9c) the temporal distortion, caused by the frequency-dependent delay in the DSA (cf. Part I) is very obvious.

Unit 38-1 (Fig. 10) shows a less sharply defined activation domain with best frequency around 1900 Hz, bandwidth 0.5–1 octave. The latency equals about 55–60 ms. The activation domain is preceded by a suppression domain with the same frequency characteristics, extending in time to as much as about 125 ms before the event (Fig. 10b and c). In this case we could not exclude the possibility that the recording also contained some spikes from one or more other neurons.

Unit 26-4 (Fig. 11) shows a vague activation type receptive field for frequencies around 400 Hz, with a latency of about 40–45 ms. This neuron does not show any sign of a suppression domain in the receptive field. The receptive field, although vague, is reproducible under repeated stimulation.

The location and shape of the activation domain is rather varied as is clear from the examples. In general it was not possible to predict the existence of a suppression domain from the information about the activation domain. For instance a unit was found which looked very similar to unit 26-8 (Fig. 9), except that it did not show any sign of suppression preceding the activation domain.

As expected for narrowband stimuli (Sect. 3.3) the Average Pre-Event Dynamic Power Spectra (c) agree with the results of the Average IFT (b).

As already noted the regular fine structure along the frequency axis (e.g. Fig. 9) is not solely attributable to the properties of the neuron. This effect is due to the interaction of the statistical characteristics of the γ -tone sequence and the neuron's receptive field properties (cf. Discussion and Appendix). For the same reason additional evidence had to ensure that the suppression domain in the receptive field truly represents a neural property. This evidence could be obtained from the receptive field properties deduced from stimulation with a γ -tone sequence with frequency values shifted over one octave, and statistical properties shifted correspondingly. In all cases shown and in most of the remaining cases this additional evidence ensured the neural origin of the suppression domain.

5. Discussion

In the present paper we have described methods to investigate the nature of the stimulus-event relation for single auditory neurons. The characteristics of the *Pre-Event Stimulus Ensemble* (PESE) as compared to those

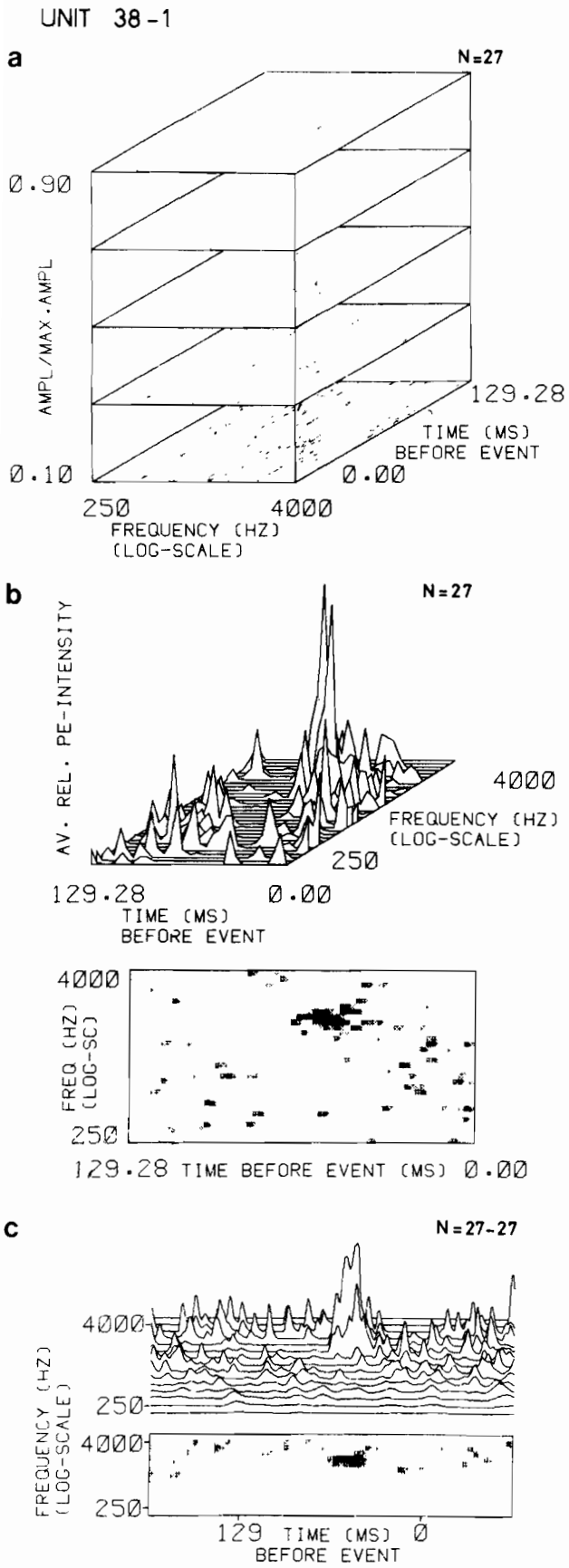


Fig. 10a-c. Spectro-temporal analysis of torus semicircularis unit 38-1, using an ensemble of γ -tones. Details as in Fig. 8

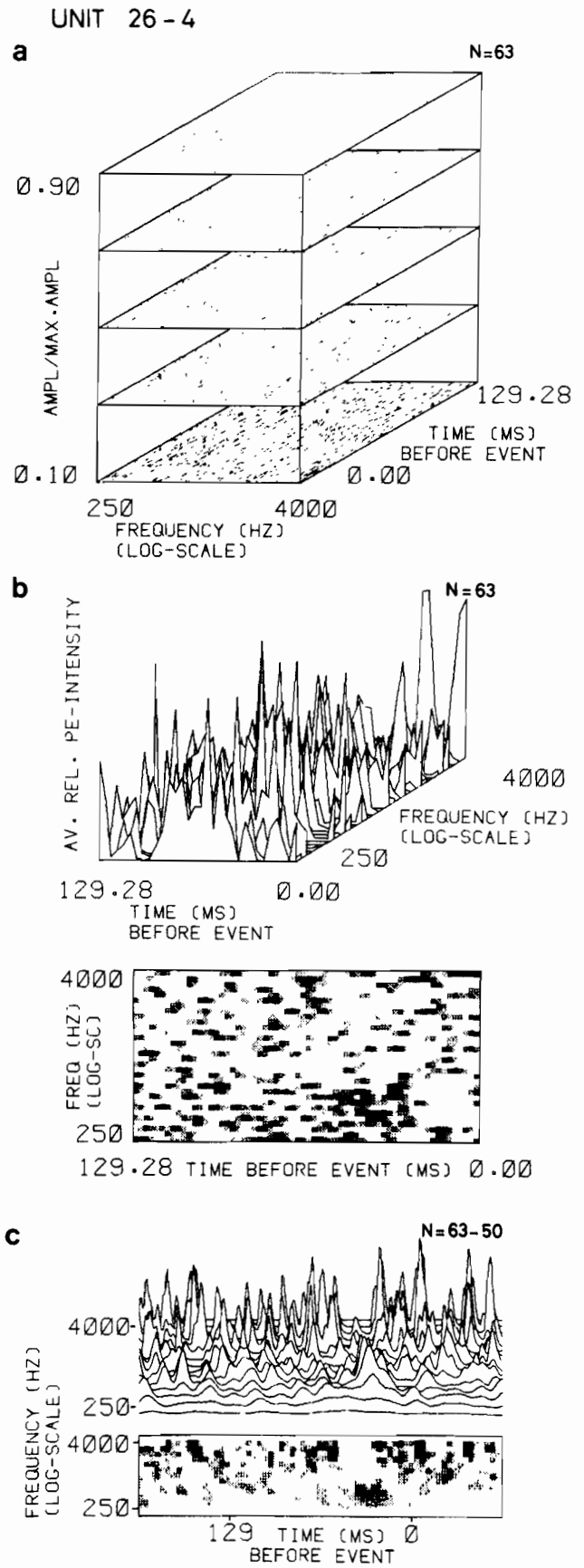


Fig. 11a-c. Spectro-temporal analysis of torus semicircularis unit 26-4, using an ensemble of γ -tones. Details as in Fig. 8

of the total *Stimulus Ensemble* (SE) indicate which stimulus properties determine the occurrence of action potentials. In this investigation we have focused on a spectro-temporal analysis of the PESE. Two approaches for this analysis have been described. The first one is based on a *parametric* description of sound in terms of amplitude and frequency as functions of time. A parametric representation of the PESE has been given: the AFT-cube. Quadratic averaging over different amplitude values gives the average intensity of stimuli in the PESE as a function of frequency and time before the neural event: *Average IFT*. A useful application of the parametric approach is restricted to narrowband stimuli. An approach, applicable to an arbitrary ensemble of stimuli is based on the description of sound by the *Dynamic Power Spectrum*. In this approach the introduction of averaging procedures leads to the *Average Pre-Event Dynamic Power Spectrum*. In order to evaluate these methods, they have been applied to the event sequences recorded from single neurons in the torus semicircularis of the grassfrog, stimulated with an ensemble of narrowband stimuli: the γ -tones.

Results and Restrictions

As described in the methods section (Sect. 2) recordings were made from anaesthetized animals (MS-222). This may have affected the firing characteristics of neurons in the torus semicircularis. It is known for some anuran species that already at the level of the auditory nerve barbiturate anaesthesia has a depressant effect on neural activity (Capranica, 1976). Neurons in our recordings did not show any spontaneous activity, whereas most of the cells recorded from the torus semicircularis of curarized green treefrogs (*Hyla cinerea*) were spontaneously active, albeit in general at very low rates (Feng, 1975; quoted by Capranica, 1976). Results of experiments, currently in progress, using a lower level of MS-222 anaesthesia indicate a richer and more complex firing behaviour of neurons in the torus semicircularis.

The stimulus ensemble used in the present study consisted of single γ -tones. Apart from the obvious limitations imposed by the use of one particular envelope waveform, the single-tone character combined with the absence of spontaneous activity does not provide information about possible lateral suppression. Moreover the composite spectral pattern of vocalizations (cf. Fig. 1 in Part I) is not represented in the γ -tones. Ethological investigations have demonstrated the behavioural significance of these spectral features for some anuran species (e.g. Capranica, 1965; Gerhardt, 1974). These findings, combined with the demonstrated non-linearities in the peripheral auditory system of anurans (Capranica and Moffat, 1979), point to the importance of extending the single-tone ensemble to a multi-tone ensemble.

In order to avoid systematic effects of monotonous parameter variation the γ -tones have been presented as a pseudo-random sequence. The statistical properties of this sequence, however, are not ideal in the sense that the different order serial correlations show particular regularities (cf. Appendix). For instance the probability of having a tone with frequency f_2 , given a tone with frequency f_1 immediately preceding it, is not uniform, the same holds for higher order combinations. This is reflected in the spectro-temporal receptive field of more complex neurons, determined by presentation of such a sequence of tones. The spectro-temporal selectivity of a neuron, composed of an activation and a suppression domain, separated in time, effectively results in a neural analysis of serial correlations in the γ -tone sequence. This affects the composition of the PESE for that neuron. The regular fine structure along the frequency axis in the AFT-cube and Average IFT as well as the dips in the activation areas of e.g. units 26-6 (Figs. 3 and 5a) and 26-8 (Fig. 9) are caused by this interaction of stimulus properties and receptive field properties. These findings once more point to the important influence of statistical properties of a stimulus ensemble, based on pseudo-random algorithms (e.g. Swerup, 1978; Eckhorn and Pöpel, 1979). For the neurons as mentioned here a monotonous parameter variation would have caused even more drastic effects. A general solution to this type of problems is not available, particular solutions will be dictated by a specific statement of the question of interest and experimental restrictions as for instance limited recording time.

Recordings from the torus semicircularis did not always result in (quasi-)stationary sequences of neural events. In a number of cases (e.g. the units 26-4 and 38-1 in Sect. 4) the firing rate after a few seconds of stimulation decreased to less than 5% of the initial firing rate. This "long-term adaptation" appeared to be reproducible, provided that a long enough silent period had elapsed. Adaptation is not measured by application of our analysis procedures to the total recording of neural activity. Application of the analysis to segments of these recordings suggests an increased spectro-temporal selectivity in the adapted stage as compared to the initial stage. Apparently for a number of neurons, using this tonal stimulus ensemble, the pre-event stimulus ensemble is inhomogeneous. Obviously a further investigation into this adaptation phenomenon demands more specific experiments and analysis. The results presented in this paper bear to the integral recording in all cases. The neurons, which show this "long-term adaptation", have spectro-temporal receptive fields which are, under the present stimulus conditions, more vaguely defined than for the "non-adapting" neurons. Moreover they generally show long latencies (~ 40 – 70 ms) as compared to the "non-adapting" neurons (~ 10 – 35 ms). These findings point

to a possible sub-population of more complex neurons in the torus semicircularis with long integration time and/or of higher order.

Since in all experiments stimuli were presented binaurally it is not possible to draw any conclusion regarding binaural interactions (Feng, 1975; quoted by Capranica, 1976).

Frequency and Time

The goal of the analytic procedures presented here is a *spectro-temporal* representation of the stimulus-event relation instead of *separate spectral and temporal* descriptions. This choice was influenced mainly by two arguments. The first argument concerns the neuronal characteristics to be encountered. In general there are no a priori arguments to expect that the spectro-temporal selectivity of auditory neurons is separable in two independent dimensions of frequency and time. This implies that analytic procedures which aim at a general applicability have to refrain from an a priori separation of the relevant variables. Separability should be an outcome of the analysis rather than an assumption. For instance the cases of "cf/fm-sensitive" neurons (e.g. Suga et al., 1978) and neurons with composite intensity, frequency and time selectivity (e.g. Miller et al., 1974) are examples where an a priori separation may give rise to misleading conclusions. The second argument concerns the stimulus ensemble. In order to find out to what extent the stimulus-event relation is invariant with respect to the type of stimuli presented (e.g. tones, noise, vocalizations), it is necessary that in all cases the *s-e* relation is analysed in terms of the same variables by means of preferably identical or at least comparable methods. Complex sounds such as vocalizations and noise possess an intricate spectro-temporal structure (see for example the numerous sonograms in Potter et al., 1947). This implies that the use of a spectro-temporal representation of simple tonal stimuli facilitates a comparison to results for more complex stimuli. The arguments given here lead to analytic procedures which are applicable to a wide class of stimulus ensembles and a wide class of neural characteristics within the spectro-temporal domain. A complementary property of these general procedures is that they provide more global information on the stimulus-event relation. Once these global properties have been found more specific hypotheses can be tested using more specific stimuli and more specific analytic procedures.

An a posteriori inspection of results for a number of neurons in the torus using the present analytic procedures does not reveal a very specific spectro-temporal coupling in the receptive fields, apart from a frequency dependence of the latency in some cases. Recent experiments in our laboratory, however, indicate that, at least for low frequency units, the picture is more complex.

The AFT-cube provides a parametric representation of the Pre-Event Stimulus Ensemble. This type of result is closely related to a dot display analysis of frequency and/or intensity scans (Nelson et al., 1966; Van Gisbergen et al., 1975). From the AFT-cube and its different cross sections it is possible to extract neural characteristics like best frequency, latency, amplitude sensitivity, bandwidth etc. Several of these parameters may turn out to be a function of one or more of the three variables in the AFT-cube (e.g. latency as a function of frequency, bandwidth as a function of time).

The AFT-cube in this context provides the most complete representation of the *s-e* relation, while both the Average IFT and the Average Pre-Event Dynamic Power Spectrum provide an *average* characteristic of the PESE, thereby necessarily discarding information present in the AFT-cube. These two average measures describe the average distribution of intensity over frequency and time of stimuli which preceded a neural event. This interpretation of these measures is expressed in the term "*spectro-temporal receptive field*". In this context the concept of "*receptive field*" of a neuron should be interpreted in an abstract sense as the collection of characteristics of stimuli which affect the neural firing behaviour (e.g. Sejnowski, 1976). This definition is a generalization and a projection into the *stimulus domain* of the original definitions, which were expressed in terms of the receptor surface (Adrian et al., 1931; Hartline, 1940; quoted by Grüsser and Grüsser-Cornehls, 1973). In the auditory literature the concept has been used with reference to the *spatial* properties of adequate acoustic stimuli (Wickelgren, 1971; Knudsen and Konishi, 1978). In the present study the dimensions of interest are abstract parameters of the stimulus, namely frequency and time, which is expressed in the term *spectro-temporal* receptive field. The receptive field may consist of different subdivisions: *activation* and/or *suppression*, their position in parameter space is given by the (f, τ) -coordinates of the extrema of the average intensity density (or, alternatively, of the mean). The "*receptive field size*" can be defined by the product $2\pi\sigma_f\sigma_t$ of the intensity density over the relevant area(s), which, from the point of view of the animal, is analogous to the "uncertainty product" as defined for signals (Gabor, 1946). For the examples shown the receptive field subdivisions were fairly large (5–25) as compared to the "*size*" of the γ -tones (0.65) (cf. Part I) and, likewise, the main frequency components of the vocalization elements. We did not find double activation areas, neither was their frequency extent large enough to encompass different frequency bands in the vocalizations. It is striking to note the similarity of time constants in the composite receptive fields (activation and suppression) and the temporal composition of the vocalizations (sound elements, separated by silent periods, cf. Fig. 1 in Part I).

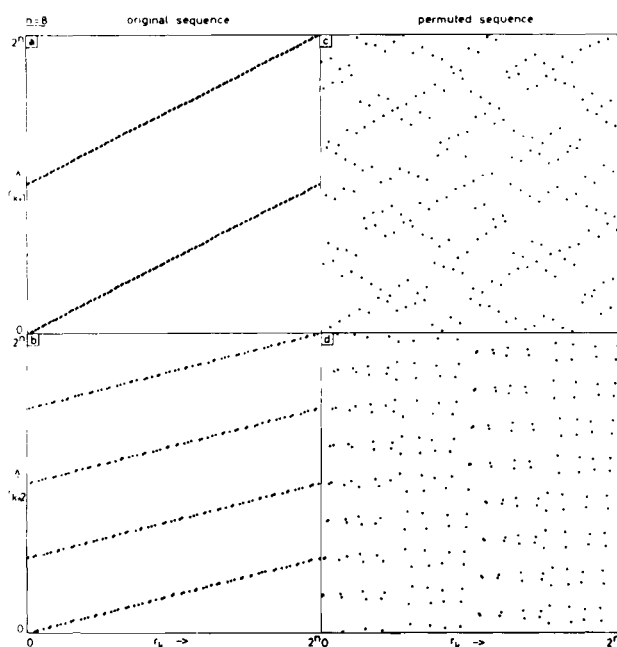


Fig. 12a-d. Scatter diagrams of serial correlation of numbers r in pseudo-random sequences, based on the original "HP-algorithm" (a and b) and the "permutation-algorithm" (c and d). a and c refer to serial correlation with lag 1, b and d refer to lag 2. The number of bits n in the shift register in this case is 8, which results in a sequence of 255 numbers. Further explanation in text

Both the Average IFT and the Average Dynamic Spectrum are cross correlation functions of the neural activity and functionals of the stimulus. In this sense they are related to a more general type of analysis of nonlinear systems, the Wiener approach (e.g. Lee and Schetzen, 1965). At this level of analysis the correspondence is of an intuitive type. The Average IFT and the Average Dynamic Power Spectrum can be formally connected by introduction of a general Complex Spectro-Temporal Intensity Density function of signals (CoSTID), defined on the analytic signal (Rihaczek, 1968; Johannesma and Aertsen, 1979). This underlying function can be tied up with the second order Wiener kernel.

The Average IFT technique can provide a detailed description of the spectro-temporal receptive field, the resolution is determined by the spectro-temporal resolution of the stimulus ensemble. The Average Dynamic Power Spectrum gives a more global measure of the receptive field, the resolution is determined by the filter characteristics of the spectrum analyser. The Average IFT is applicable for an ensemble of narrowband stimuli, the dynamic spectrum analysis can be applied to an arbitrary stimulus ensemble.

In the present investigation we have studied the responses from single torus semicircularis neurons to γ -tones. The waveform of the γ -tones was chosen in an attempt to mimic some of the characteristics of the male

grassfrog's vocalizations. It is tempting to interpret the describing parameters of the γ -waveform as "bioacoustic dimensions" of the vocal repertoire of the male grassfrog. Variation of the parameters – in this case A and ω – in neurophysiological experiments may then be interpreted as an investigation into the possible "focal zones" at the single neuron level (Scheich, 1977). The receptive field properties of torus semicircular neurons for a wide ensemble of natural stimuli (vocalizations, environmental sounds), using the techniques as described in this paper, will be the subject of the forthcoming paper (Part III).

Appendix

The Pseudo-Random Sequences

The pseudo-random modulation of the amplitude factor A and the carrier frequency ω_γ in the sequence of γ -tones (cf. Part I) was generated by a maximum length series algorithm as implemented in the HP3722A noise generator. The contents of the n -bits shift register with feedbacks, used in the closed loop condition for generating pseudo-random binary sequences, was interpreted as a sequence of n -bits integer numbers. This results in a sequence of $2^n - 1$ different values, being a permutation of the natural numbers in the range $(0, 2^n)$. The amplitude sequence consisted of 127 values ($n=7$), the frequency sequence contained 255 values ($n=8$). These values can be converted to values in the desired range for A resp. ω_γ , with additional exponential scaling in the case of ω_γ . Thereby the amplitude factor and the carrier frequency can be considered as random variables of a lattice type (Papoulis, 1965) with a uniform probability on the lattice points.

In order to reduce the strong serial correlations between succeeding values, the sequence was permuted to a new one by taking as succeeding elements those from the original sequence with order numbers $1, n+1, 2n+1$, etc., with cyclic continuation of the original sequence. Thereby the contents of the shift register has been totally "refreshed" for each new element of the sequence. This procedure reduces the serial correlation in the sequence to a large extent, as is demonstrated for the case $n=8$ by the scatter diagrams of the serial correlations with lags 1 and 2 in Fig. 12.

It is evident, however, that the restrictions of short pseudo-random sequences cannot be fully overcome. Some degree of structure in the sequence remains. Note for instance the non-uniform distribution of points in the scatter diagram with lag 1 (Fig. 12c), with rather large holes, also along the diagonal and the regular fine structure in the scatter diagram with lag 2 (Fig. 12d). These observations have to be taken into account in the interpretation of results from experiments using these pseudo-random sequences.

Acknowledgements. This investigation forms part of the research program of the Workgroup Neurophysics at the Department of Medical Physics and Biophysics, University of Nijmegen (The Netherlands) and was supported by the Netherlands Organization for Advancement of Pure Research (Z.W.O.). The authors wish to thank Jos Eggermont and Ton Vendrik for helpful discussions on this manuscript. The expertise of Koos Braks, Jan Bruijns, Wim van Deelen, and Jos Olders was a great help in the experiments and data analysis.

References

- Adrian, E.D., Cattell, M., Hoagland, H.: Sensory discharges in single cutaneous nerve fibers. *J. Physiol. (London)* **72**, 377–391 (1931)
- Aertsen, A.M.H.J., Johannesma, P.I.M.: Spectro-temporal analysis of auditory neurons in the grassfrog. In: *Hearing mechanisms*

- and speech, pp. 87–93. Creutzfeld, O., Scheich, H., Schreiner, Chr. (eds.). *Exp. Brain Res.* [Suppl. II] (1979)
- Aertsen, A.M.H.J., Johannesma, P.I.M.: Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli. *Biol. Cybernetics* **38**, 223–234 (1980)
- Bullock, T.H. (ed.): Recognition of complex acoustic signals (*Life Sciences Res. Rep.*, Vol. 5). Berlin, Heidelberg, New York: Springer 1977
- Capranica, R.R.: The evoked vocal response of the bullfrog: a study of communication by sound. Cambridge, MA: MIT Press 1965
- Capranica, R.R.: Morphology and physiology of the auditory system. In: *Frog neurobiology, a handbook*, pp. 551–575. Llinas, R., Precht, W. (eds.). Berlin, Heidelberg, New York: Springer 1976
- Capranica, R.R., Moffat, A.J.M.: Nonlinear properties of the peripheral auditory system. In: *Comparative studies of hearing in vertebrates*. Berlin, Heidelberg, New York: Springer 1979
- Dowben, R.M., Rose, J.E.: A metal-filled microelectrode. *Science* **118**, 22–24 (1953)
- Eckhorn, R., Pöpel, B.: Generation of Gaussian noise with improved quasi-white properties. *Biol. Cybernetics* **32**, 243–248 (1979)
- Evans, E.F.: Cochlear nerve and cochlear nucleus. In: *Handbook of sensory physiology*, Vol. V/2: Auditory system, pp. 1–108. Keidel, W.D., Neff, W.D. (eds.). Berlin, Heidelberg, New York: Springer 1975
- Feng, A.S.: Sound localization in anurans: an electrophysiological and behavioral study. Ph. D. Thesis, Cornell Univ., Ithaca, New York 1975
- Flanagan, J.L.: *Speech analysis, synthesis, and perception*. Berlin, Heidelberg, New York: Springer 1965
- Gabor, D.: Theory of communication. *J. IEE (London) Part III*, **93**, 429–457 (1946)
- Gelder, J.J. van, Evers, P.M.G., Maagnus, G.J.M.: Calling and associated behaviour of the common frog, *Rana temporaria*, during breeding activity. *J. Anim. Ecol.* **47**, 667–676 (1978)
- Gerhardt, H.C.: The significance of some spectral features in mating call recognition in the green treefrog (*Hyla cinerea*). *J. Exp. Biol.* **61**, 229–241 (1974)
- Gerhardt, H.C.: Mating call recognition in the green treefrog (*Hyla cinerea*): the significance of some fine-temporal properties. *J. Exp. Biol.* **74**, 59–73 (1978)
- Gisbergen, J.A.M. van, Grashuis, J.L., Johannesma, P.I.M., Vendrik, A.J.H.: Spectral and temporal characteristics of activation and suppression of units in the cochlear nuclei of the anaesthetized cat. *Exp. Brain Res.* **23**, 367–386 (1975)
- Grashuis, J.L.: The pre-event stimulus ensemble: an analysis of the stimulus-response relation for complex stimuli applied to auditory neurons. Ph. D. Thesis, Nijmegen 1974
- Grüsser, O.-J., Grüsser-Cornehls, U.: Neuronal mechanisms of visual movement perception and some psychophysical and behavioral correlations. In: *Handbook of sensory physiology*, Vol. VII/3A: Central processing of visual information, pp. 333–429. Jung, R. (ed.). Berlin, Heidelberg, New York: Springer 1973
- Hartline, H.K.: The receptive fields of the optic nerve fibers. *Am. J. Physiol.* **130**, 690–699 (1940)
- Johannesma, P.I.M.: The pre-response stimulus ensemble of neurons in the cochlear nucleus. In: *Proc. of the IPO Symp. on Hearing Theory*, pp. 58–69. Cardozo, B.L. (ed.). Eindhoven 1972
- Johannesma, P.I.M., Gisbergen, J.A.M. van, Grashuis, J.L., Vendrik, A.J.H.: Forward and backward analysis of stimulus response relations of single cells in the cochlear nucleus. In: *Proc. IV International Biophysics Congress (IUPAB)*, Vol. 3, pp. 508–523. Moscow: USSR Academy of Sciences 1973
- Johannesma, P.I.M., Aertsen, A.M.H.J.: Neural image of sound in the grassfrog. In: *Hearing mechanisms and speech*, pp. 79–86. Creutzfeld, O., Scheich, H., Schreiner, Chr. (eds.). *Exp. Brain Res.* [Suppl. II] (1979)
- Kiang, N.Y.-S., Watanabe, T., Thomas, E.C., Clark, L.F.: Discharge patterns of single fibers in the cat's auditory nerve. Cambridge, MA: MIT Press 1965
- Knudsen, E.I., Konishi, M.: Center-surround organization of auditory receptive fields in the owl. *Science* **202**, 778–780 (1978)
- Koenig, W., Dunn, H.K., Lacy, L.Y.: The sound spectrograph. *J. Acoust. Soc. Am.* **18**, 19–49 (1946)
- Lang, S.: *Linear algebra*. Reading, MA: Addison-Wesley 1973
- Lee, Y.W., Schetzen, M.: Measurement of the Wiener kernels of a nonlinear system by crosscorrelation. *Int. J. Control* **2**, 237–254 (1965)
- Miller, J.M., Beaton, R.D., O'Connor, T., Pflingst, B.E.: Response pattern complexity of auditory cells in the cortex of unanesthetized monkeys. *Brain Res.* **69**, 101–113 (1974)
- Nelson, P.G., Erulkar, S.D., Bryan, J.S.: Responses of units of the inferior colliculus to time-varying acoustic stimuli. *J. Neurophysiol.* **29**, 834–860 (1966)
- Papoulis, A.: *Probability, random variables and stochastic processes*. Tokyo: McGraw-Hill Kogakusha 1965
- Potter, R.K., Kopp, G.A., Green, H.C.: *Visible speech*. New York: Van Nostrand 1947
- Rihaczek, A.W.: Signal energy distribution in time and frequency. *IEEE Trans. Inf. Theory* **IT-14**, 369–374 (1968)
- Scheich, H.: Central processing of complex sounds and feature analysis. In: *Recognition of complex acoustic signals*, pp. 161–182. Bullock, T.H. (ed.). (*Life Sciences Res. Rep.*, Vol. 5.) Berlin, Heidelberg, New York: Springer 1977
- Sejnowski, T.J.: On global properties of neuronal interaction. *Biol. Cybernetics* **22**, 85–95 (1976)
- Singleton, R.C., Poulter, T.C.: Spectral analysis of the call of the male killer whale. *IEEE Trans. Audio Electroacoust.* **AU-15**, 104–113 (1967)
- Smolders, J.W.T., Aertsen, A.M.H.J., Johannesma, P.I.M.: Neural representation of the acoustic biotope: a comparison of the response of auditory neurons to tonal and natural stimuli in the cat. *Biol. Cybernetics* **35**, 11–20 (1979)
- Suga, N., O'Neill, W.E., Watanabe, T.: Cortical neurons sensitive to combinations of information bearing elements of biosonar signals in the mustache bat. *Science* **200**, 778–781 (1978)
- Swerup, C.: On the choice of noise for the analysis of the peripheral auditory system. *Biol. Cybernetics* **29**, 97–104 (1978)
- Tolman, R.C.: *The principles of statistical mechanics*. London: Oxford University Press 1938
- Wickelgren, B.C.: Superior colliculus: some receptive field properties of bimodally responsive cells. *Science* **173**, 69–71 (1971)
- Worden, F.G., Galambos, R. (eds.): *Auditory processing of biologically significant sounds*. (*Neurosci. Res. Prog. Bull.*, Vol. 10.) Brookline, MA: Neurosciences Research Program 1972

Received: May 11, 1980

A. M. H. J. Aertsen
Department of Medical Physics and Biophysics
University of Nijmegen
Geert Grooteplein Noord 21
NL-6525 EZ Nijmegen
The Netherlands