

Spectro-Temporal Receptive Fields of Auditory Neurons in the Grassfrog

III. Analysis of the Stimulus-Event Relation for Natural Stimuli

A. M. H. J. Aertsen, J. H. J. Olders, and P. I. M. Johannesma

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

Abstract. The stimulus-event relation of single units in the auditory midbrain area, the torus semicircularis, of the anaesthetized grassfrog (*Rana temporaria* L.) during stimulation with a wide ensemble of *natural stimuli*, was analysed using first and second order statistical analysis techniques. The *average stimulus* preceding the occurrence of action potentials, in general, did not prove to give very informative results. The second order procedure consisted in the determination of the *average dynamic power spectrum* of the *pre-event stimuli*, following procedures as described elsewhere (Aertsen and Johannesma, 1980; Aertsen et al., 1980). The outcome of this analysis was filtered with the overall power spectrum of the complete stimulus ensemble in order to correct for its non-uniform spectral composition. The "stimulus-filtered" average pre-event dynamic spectrum gives a first indication of the "*spectro-temporal receptive field*" of a neuron under natural stimulus conditions. Results for a limited number of recordings are presented and, globally, compared to the outcome of an analogous analysis of experiments with tonal stimuli.

1. Introduction

The nature of the stimulus-response or stimulus-event relation of single auditory neurons under natural stimulus conditions has mostly been investigated on a rather qualitative level (e.g. Worden and Galambos, 1972; Bullock, 1977; Smolders et al., 1979). Post-stimulus time histograms (PSTH) are determined for (species-specific) vocalizations and their properties are compared to the spectral tuning and/or temporal characteristics like latency, determined with technical stimuli, e.g. tones. This approach has proven to be successful in peripheral parts of the auditory system, however, it has met considerable problems in reconciling the results in more central parts of the brain.

The present paper is the third one in a series, dealing with the spectro-temporal analysis of the response properties of auditory neurons under various stimulus conditions. This analysis basically follows a *statistical* approach, inspired by procedures which have been used for noise stimuli (De Boer and Kuyper, 1968; Johannesma, 1972, 1980; De Boer and DeJongh, 1978). In the first paper (Part I: Aertsen and Johannesma, 1980) methods were described to give a spectro-temporal characterization of sounds, both narrowband (e.g. tones) and wideband (e.g. natural sounds). The application of these methods to the analysis of the stimulus-event relation using a *tonal* stimulus ensemble was the subject of the second paper (Part II: Aertsen et al., 1980). The present paper (Part III) is concerned with a first attempt to study the single neuron response to *natural* sounds, using the procedures as developed in the Papers I and II. The stimulus ensemble consisted of a wide selection of sounds as they occur in the natural environment, i.e. the biotope, of the animal under investigation, the grassfrog (*Rana temporaria* L.). The stimulus sequence had a duration of approximately 4 min and contained such various sounds as species-specific vocalizations, vocalizations from other frogs and toads, sounds from other animals (prey and predators) and sound from abiotic origin (e.g. wind and rain). The composition and characteristics of this natural stimulus ensemble, the *Acoustic Biotope*, have been described in detail in Part I.

The properties of the stimulus-event relation have been investigated by analysing the ensemble of stimuli, preceding the occurrence of action potentials, the *Pre-Event Stimulus Ensemble* (PESE; Johannesma, 1972). A comparison of the characteristics of the complete *Stimulus Ensemble* (SE) and this subensemble, the PESE, provides information about the selecting mechanism of the neuron and leads to a functional description of its role in the neural representation of the

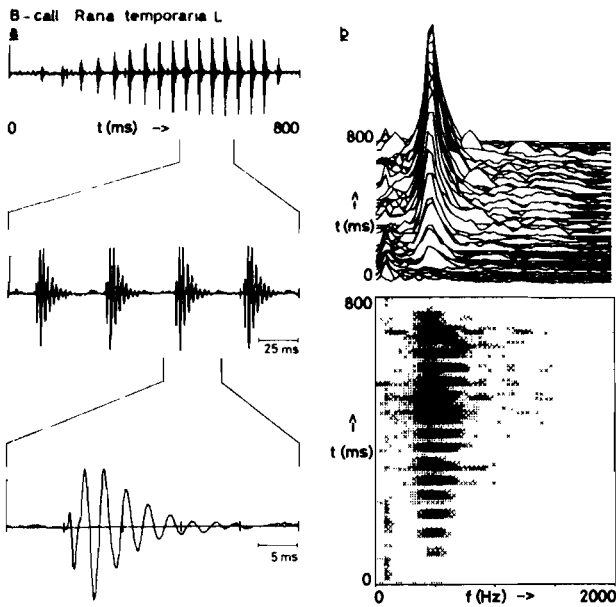


Fig. 1a and b. The "B-call" of the male grassfrog (*Rana temporaria* L.). **a** Time course of vocalization on different time scales, indicated by horizontal bars. Vertical bars refer to identical, arbitrary amplitude values. **b** Sonogram (dynamic amplitude spectrum) of vocalization in different representations. Frequency resolution 39.1 Hz, time resolution 12.8 ms

sensory environment. The properties of the PESE can be investigated by an analysis of the ensemble average of various relevant stimulus functionals. In the present investigation this cross correlation approach to the study of the stimulus-event relation has been applied for first and second order signal characteristics. The *first order* analysis leads to the average pre-event stimulus, in fact the "revcor"-procedure, first introduced by De Boer and Kuyper (1968). The *second order* properties of the PESE have been studied by analysing the spectro-temporal structure of pre-event stimuli, leading to the average pre-event dynamic power spectrum (Part II, Aertsen et al., 1980). The outcome of the second order analysis has to be *normalized* for the spectro-temporal properties of the stimulus ensemble in order to arrive at a correct representation of the neural spectro-temporal selectivity, the *spectro-temporal receptive field* (cf. Part II). As a first, incomplete, attempt the average pre-event dynamic spectra have been "corrected" for the overall spectral composition of the acoustic biotope: "stimulus filtering".

In the present paper the procedures will be described, furthermore the concepts will be demonstrated by application to a number of extra-cellular single unit recordings from the grassfrog's auditory midbrain area, the torus semicircularis, recorded during presentation of the natural stimulus ensemble. The results of the spectro-temporal analysis will be, globally,

compared to the findings from tonal experiments, analysed in the very same way (cf. Part II). Such a comparison should provide an indication to what extent the receptive field is a *stimulus-invariant* neural characteristic.

A short, preliminary report on the approach as followed in this paper and the preceding ones (Parts I and II) has been published elsewhere (Aertsen and Johannesma, 1979).

2. Methods

Preparation and Recording

Extracellular single unit recordings were made in the torus semicircularis of adult grassfrogs from Ireland. The methods used for the preparation of the animal and the recording of single unit activity have been described in detail in Part II. During the experiments the animal was maintained at a level of anaesthesia (MS-222) at which corneal reflex returned but no other movements occurred

Stimulus Presentation and Data Registration

Experiments were performed with the frog in an electrically shielded acoustic chamber (IAC 1202A). The stimulus ensemble consisted of a selection of various sounds as they occur in the natural environment of the animal under investigation, i.e. the grassfrog. The stimulus sequence had a duration of approximately 4 min and contained, amongst others, vocalizations of various species of frogs and toads, vocalizations of predators and prey as well as sounds of abiotic origin (e.g. wind and rain). The characteristics of this natural stimulus ensemble, the *Acoustic Biotope*, and the technical procedures of sound recording and composition of the stimulus tape have been described in detail in Part I. An example of one of the stimulus segments, a "B-call" of the male grassfrog (typology according to Van Gelder et al., 1978) is given in Fig. 1.

The stimulus tape was played back to the animal from an AM-recorder (Otari MX 5050 QXH, tape speed 38 cm/s). After low-pass filtering (2 cascaded -24 dB/oct Bessel filters, each -3 dB at 5 kHz) the signal was amplified (Luxman 507X) and presented to the animal by two electro-dynamic microphones (Sennheiser MD 211 N), coupled to the frog's tympanic membranes by means of closed cavities. 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, measured in the acoustic couplers by half inch condenser microphones (Bruel and Kjaer 4143) was in the range of 60-80 dB SPL.

Neural activity was amplified (Grass P16) and recorded on the stimulus delivering tape recorder. It was transformed into unitary trigger pulses (neural events) by means of a level discriminator. The analysis of the stimulus-event relation was performed by means of a hardware signal averager (Biomac 1000, Datalab) and a real-time dynamic spectrum analyser (DSA; Part I). The results of both types of analysis could be read out and transferred to the laboratory's PDP 11/45 for further elaboration. Furthermore in a number of cases the neural events, together with a synchronizing clock signal (20 kHz) from the stimulus recorder, were fed to a Data Acquisition System (DAS), built around a PDP 11/45. These data were used for evaluation of the cross coincidence histograms (Sect. 3.1).

In these experiments we also recorded the neural responses to an ensemble of *tonal* stimuli. Procedures for the analysis of the

stimulus-event relation as well as the results of those experiments have been described in Part II.

3. The Stimulus-Event Relation

Extra-cellular recording of single unit activity during presentation of an ensemble of stimuli, e.g. a long sequence of natural sounds, results in a sequence of action potentials which can be modelled as a series of neural events:

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

with $\delta(t)$ = Dirac delta function, $\{t_n\}$ = moments of occurrence of action potentials, N = number of action potentials during presentation of stimulus ensemble (duration T).

In general, repetition of the experiment will not lead to an identical neural activity $z(t)$. Therefore $z(t)$ has to be considered as a particular realization of a stochastic point process.

A functional description of the neuron with respect to its role in the neural representation of sensory stimuli can be given by an analysis of the *stimulus-event relation*. This analysis can be separated logically into 2 components:

1. Does a stimulus-event relation *exist*, i.e. are the statistics of the stochastic point process governed in some way by the stimulus ensemble or is it an autonomous process.
2. If a stimulus-event relation is shown to exist, what is the *nature* of it, i.e. which stimulus properties affect the point process and in which way can their effect be described.

3.1. Existence: Event-Event Correlation

The existence of the *s-e* relation was investigated by presenting the natural stimulus ensemble twice. The resulting two sequences of neural events $z_1(t)$ and $z_2(t)$ were tested for similarity by evaluation of the Cross Coincidence Function, based on the crosscorrelation function $R_{z_1 z_2}(\sigma)$ of the event sequences $z_1(t)$ and $z_2(t)$: *event-event correlation* (Aertsen et al., 1979)

$$R_{z_1 z_2}(\sigma) = \frac{1}{T} \int_0^T dt z_1(t) z_2(t + \sigma). \quad (2)$$

A relatively high number of coinciding events near $\sigma=0$ indicates that the event sequences are synchronized more than would be expected for independent realizations of an autonomous point process. This synchronization is attributed to stimulus effects: the more distinct the cross coincidence function, the more

strict the coupling between the stimulus ensemble and the neural activity. An example of this analysis for two torus semicircularis neurons is shown in Fig. 2. Each column corresponds to one neuron. The dot displays represent the neural activity during consecutive presentations of the stimulus ensemble, $z_1(t)$ and $z_2(t)$. The form of the cross coincidence function is represented by $N_{1,2}(\sigma)$: the number of coinciding events within a time bin Δ for varying time shift σ . The arrow in the graph of $N_{1,2}(\sigma)$ represents the expectation value $E(N_{1,2}) = N_1 N_2 \frac{\Delta}{T}$ for uncorrelated stationary sequences, it may serve as a reference. Note the highly different values of the binwidth Δ and the time scale σ for the two neurons.

Both units show a distinct peak around $\sigma=0$, which indicates that indeed the neural activity is influenced reproducibly by presentation of the stimulus ensemble. The width of the peak, measured at half the maximum value, however, differs significantly: about 4 ms for unit 26-6 (Fig. 2, left) versus as much as 250 ms for unit 28-3 (Fig. 2, right). Furthermore we note the regular fine structure in the coincidence histogram for unit 26-6: smaller peaks at time intervals of about 35 ms. This is connected to the rather stereotyped response of this neuron (cf. the dot displays in Fig. 2) to the regular, pulsatile grassfrog vocalizations in the stimulus ensemble (cf. Fig. 1). This fine structure is altogether absent in the firing pattern of unit 28-3; the form of $N_{1,2}(\sigma)$ does not change much when analysed with a smaller binwidth Δ . These findings indicate that, whereas the overall activity N of the neurons differs not more than a factor of about 2.5, there is a dramatic difference in the temporal precision of the coupling between stimulus ensemble and neural activity. This difference is expected to be reflected in the nature of the *s-e* relation, which will be investigated in the following sections. On the other hand, the form of the coincidence histogram provides a test for the adequateness of a quantitative description of the *s-e* relation. The two examples shown in Fig. 2 more or less portray the extremes which were encountered in a relatively small population of torus semicircularis neurons which were investigated. A more rigorous statistical approach to the question discussed here has recently been described by Pedersen (1980).

3.2. Nature: Stimulus-Event Correlation

The analysis of the neural representation of natural stimuli is nearly always performed on the base of the Peri- (or Post-) Stimulus Time Histogram (PSTH) (e.g. Kiang and Moxon, 1974; Worden and Galambos, 1974; Bullock, 1977; Smolders et al., 1979): a natural

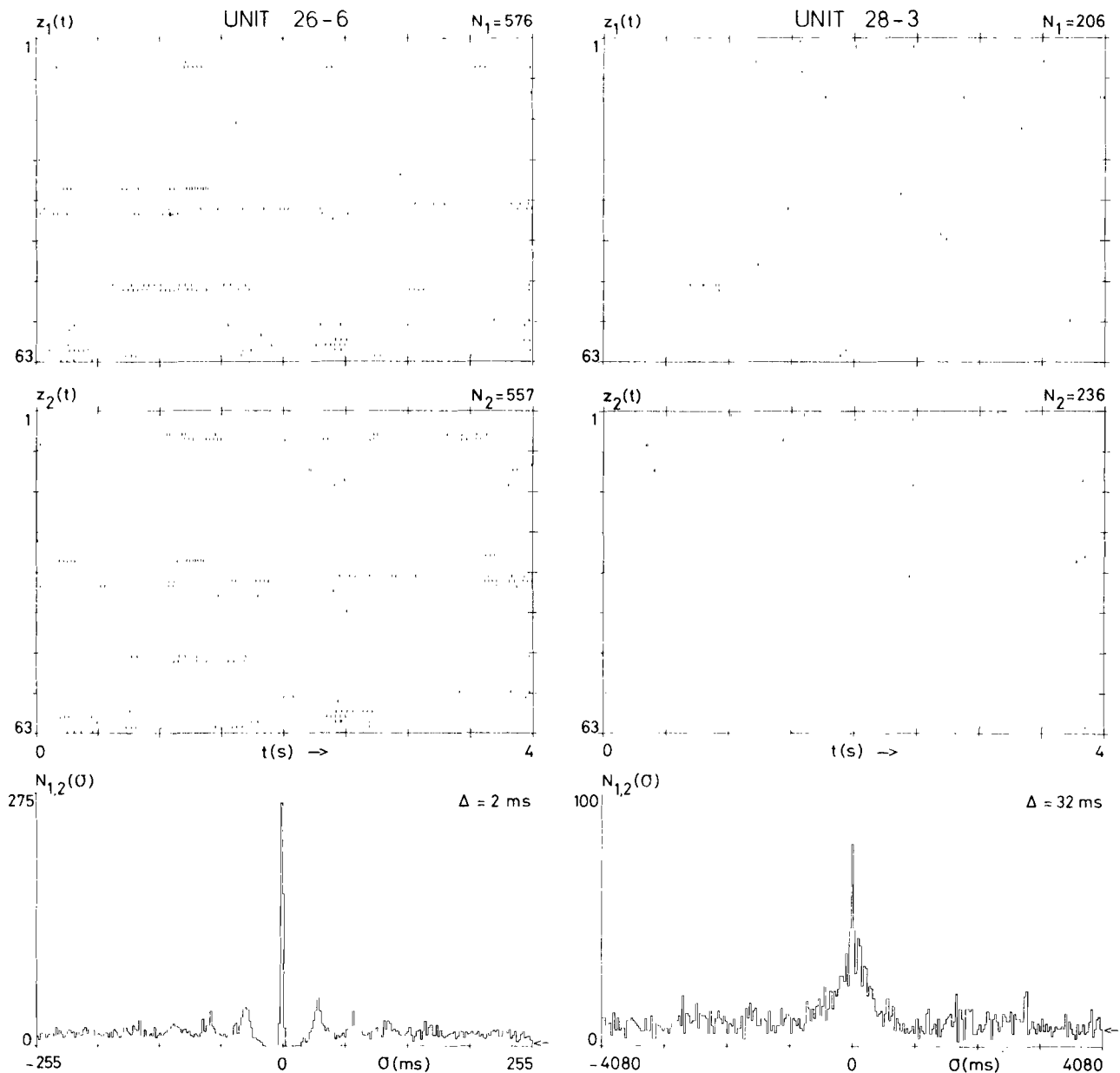


Fig. 2. Existence of stimulus-event relation. Dot displays and cross coincidence histograms of neural activity recorded from two neurons in the torus semicircularis of the grassfrog during subsequent presentations of the acoustic biotope. Left column: unit 26-6, right column: unit 28-3

stimulus, necessarily of short duration (~ 1 s), is presented repetitively and the associated neural activity is summated in the form of a histogram to yield an estimate of the *event density* $n(t)$ under stimulus conditions. Conspicuous features in the PSTH, i.e. significant deviations from an a priori expected level, are related to the spectro-temporal properties of the stimulus. A necessary prerequisite for this approach is full *controllability* of the stimulus.

This requirement, however, excludes the possibility to analyse the neural activity recorded from normally

behaving animals in a natural (acoustic) environment. Therefore we will follow another approach in which the more modest requirement of *observability* of the stimulus has to be fulfilled. This approach, which makes use of correlation functions of stimulus and neural activity, is based on the concept of the *Pre-Event Stimulus Ensemble* (PESE) (Johannesma, 1972, 1980; Grashuis, 1974; Aertsen et al., 1980). The PESE is defined as the ensemble of stimuli $x_n(\tau) = x(t_n - \tau)$ which precede a neural event $\delta(t - t_n)$ and, by definition, forms a subset of the *Stimulus Ensemble* (SE).

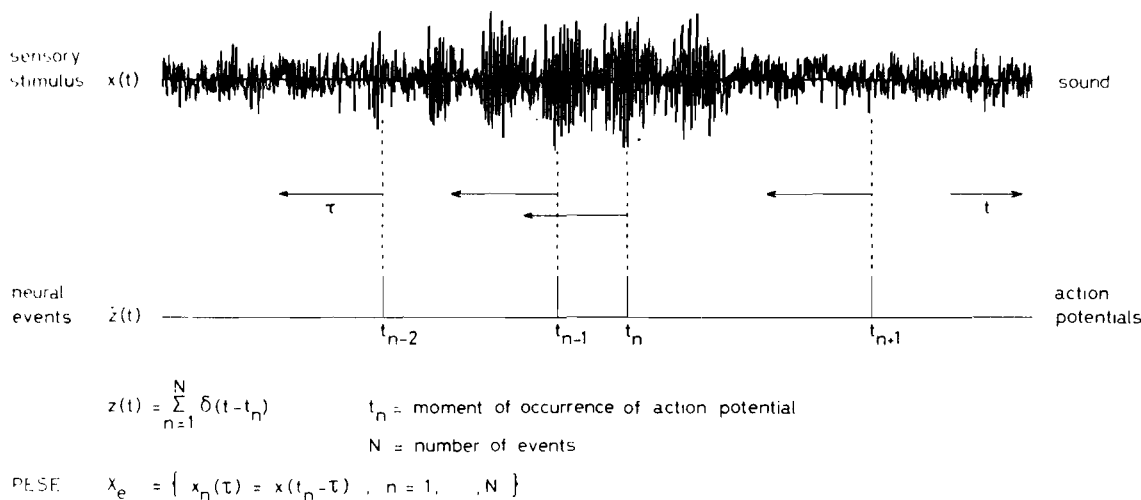


Fig. 3. Definition of the Pre-Event Stimulus Ensemble (PESE)

The concept of the PESE is illustrated in Fig. 3. A comparison of the characteristics of the PESE of a neuron as compared to those of the SE provides a description of the function of that neuron in the neural representation of acoustic stimuli. This approach is rendered feasible by the following observations (Johannesma, 1972):

1. The information on the stimulus-event relation, contained in the PSTH $n(t)$ is also present in the different order cross correlation functions of stimulus $x(t)$ and neural event sequence $z(t)$, or, equivalently, in the different order moment functions of the PESE.

2. The definition of the PESE as well as the determination of its characteristics by means of the moment functions do not require the stimulus to be repetitive. In fact, complex non-repetitive stimuli may be more adequate because of the potential richness of the SE, and, consequently, the PESE. This enables the use of a complex stimulus ensemble like the acoustic biotope for natural stimuli and Gaussian noise in the case of technical stimuli.

3. Since the stimulus does not have to be repetitive a correlation-type of approach does not require full controllability of the stimulus, the observability of the stimulus is sufficient.

These observations indicate that if the experimental situation does not allow the stimulus to be fully controllable, the formation and analysis of the PESE, also for natural stimuli may be an attractive alternative. In the present investigation we have applied this approach and studied the first and second order characteristics of the $s-e$ relation for a wide natural

stimulus ensemble: the acoustic biotope (Aertsen and Johannesma, 1980) by means of *stimulus-event correlation*.

4. The Average Stimulus Preceding a Neural Event

A natural first choice to study the PESE is to determine its average value. The first order moment function or average value $r_e(\tau)$ of the PESE is given by

$$r_e(\tau) = \langle x_e(\tau) \rangle = \frac{1}{N} \sum_{n=1}^N x_n(\tau) \quad (3)$$

with τ = time *before* the neural event ($\tau > 0$), $x_n(\tau)$ = n -th element from the PESE, N = number of action potentials recorded during presentation of the SE (duration T).

This type of analysis was first introduced into sensory neurophysiology as the "revcor" procedure (De Boer and Kuypers, 1968) and used to study the stimulus-event relation for auditory nerve fibres (e.g. De Boer and De Jongh, 1978) and cochlear nucleus neurons (e.g. Grashuis, 1974) for Gaussian white noise. It has been applied in visual electrophysiology as well (e.g. Marmarelis and Marmarelis, 1978). It has been shown that $r_e(\tau)$, apart from a scaling factor, equals the first order crosscorrelation function of stimulus $x(t)$ and event density $n(t)$ (Johannesma, 1972).

In the present investigation the average stimulus preceding an event has been determined by means of a hardware signal averager with an additional delay line to account for the necessary shift ("*triggered correlation*", De Boer and Kuypers, 1968). The outcome of this procedure for the torus semicircularis unit 26-6 is

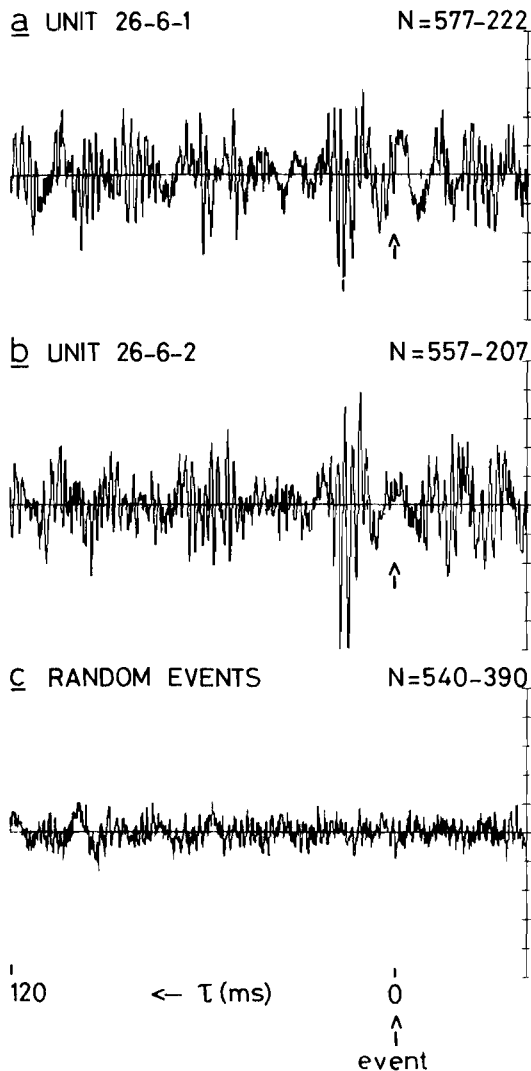


Fig. 4a-c. Average pre-event stimulus $r_e(\tau)$ of the torus semicircularis unit 26-6 during first **a** and second **b** presentation of the acoustic biotope, vs. the a priori expected result for this stimulus ensemble **c**, determined by using a random sequence of events. The vertical scale in all three figures refers to arbitrary, identical units. The numbers N_1 and N_2 in the figures indicate the number of events presented to the signal averager (N_1), respectively the number of events actually considered in averaging (N_2)

shown in Fig. 4. In order to test the reproducibility of this result, the acoustic biotope was presented twice and the resulting averages of both recordings are given separately (Fig. 4a and b).

These results should be compared to the a priori expected outcome for the acoustic biotope. The latter is shown in Fig. 4c, using the same amplitude scale as in Fig. 4a and b, to allow for a proper comparison. This result was obtained using a random sequence of events which bears no relation to the SE and which consisted of the moments of positive going high-level crossings of Gaussian wideband noise. Events occur-

ring during the processing sweep of the signal averager are ignored. Due to this "dead time" only N_2 of the total number of elements N_1 in the PESE ($N_2 \leq N_1$) are taken into account in averaging. The numbers N_1 , N_2 are indicated in Fig. 4. A comparison of the results in Fig. 4 shows that the average pre-event stimulus of unit 26-6 (Fig. 4a and b) contains oscillations and amplitude variations, which are not present to that extent in the result for random events (Fig. 4c). The most prominent component is an oscillatory waveform with a duration of about 10-15 ms and its maximum amplitude some 15 ms prior to the neural event. The dominant frequency of the waveform is in the order of 500 Hz. This waveform is also present at other values of τ (including $\tau < 0$), now with smaller amplitude and less distinct, as though smeared out. This repetition should be attributed to the quasi-periodic structure of the grassfrog vocalizations in the natural stimulus ensemble (cf. Fig. 1). The gross properties of the average stimulus preceding a neural event in this case appear to be quite reproducible upon repeated stimulation, the reproducibility of the detailed waveform, however, is not very impressive. The latter is not very surprising since phase-lock in this frequency region is hardly expected to be found at this level of the frog's auditory system [cf. the time-locked response to trains of clicks which at the auditory nerve level of the bullfrog already disappeared for a repetition rate of 200 Hz (Frishkopf and Goldstein, 1963)]. Already the coincidence histogram for this unit (cf. Fig. 2, left) indicated a temporal precision of neural firing of as much as 4 ms.

A tentative conclusion from the results in Figs. 2 and 4 might be that unit 26-6 is mainly sensitive for stimulus frequencies around 500 Hz, whereas temporally it might be coupled to a slowly varying parameter of the stimulus, like the envelope or intensity, with a latency of 15-20 ms. This interpretation is preliminary, however, for several reasons. First, the missing of triggers (more than half in this case) may lead to erroneous results (e.g. Wilson and Evans, 1975). Secondly, the stimulus ensemble is certainly not like a Gaussian white noise. It is highly structured in time and frequency and the 500 Hz component from the species-specific vocalizations is abundantly present and moreover in a very specific temporal organization (cf. Fig. 1). Third, higher frequencies for which the neuron might be sensitive may have been lost altogether in the averaging procedure as described, due to a poor quality of phase-lock.

For the moment the relevance of first order averaging of the PESE for a natural stimulus ensemble, while using the experimental techniques as described in Sect. 2 (e.g. anaesthesia, simple hardware averaging) awaits further evidence. A possible way to overcome

some of the problems indicated here is described in the following section.

5. Spectro-Temporal Averaging of the Pre-Event Stimulus Ensemble

The next obvious step, after a first order analysis of the PESE, would be to proceed with a second order analysis, i.e. to study the second order moment function of the PESE $r_e(\sigma, \tau)$, defined by

$$r_e(\sigma, \tau) = \langle x_e(\sigma)x_e(\tau) \rangle = \frac{1}{N} \sum_{n=1}^N x_n(\sigma)x_n(\tau). \quad (4)$$

This function is also referred to as the *time-dependent autocorrelation function* of the PESE. Analogously to the first order average it proves to be proportional to the second order crosscorrelation function of stimulus and neural activity. In the case of a Gaussian white noise stimulus this approach leads to the second order Wiener kernel of the nonlinear system describing the transformation of the stimulus $x(t)$ to the neural activity $z(t)$ (Lee and Schetzen, 1965; Johannesma, 1972). This approach, however, is not intuitively fitted to an understanding of auditory information processing. We prefer an alternative approach, based on a second order representation of sound as a function of *both frequency and time*: the Dynamic Power Spectrum $P(\omega, \tau)$, also referred to as short-time power spectrum or spectrogram (e.g. Potter et al., 1947; Flanagan, 1965). The dynamic power spectrum $P(\omega, \tau)$ and the time-dependent autocorrelation $r(\sigma, \tau)$ can be formally related by an appropriate Fourier transform (Fano, 1950; Schroeder and Atal, 1962). On the other hand this approach is a generalization of the parametric description of sound in terms of instantaneous frequency and amplitude, which is applicable in the case of narrowband signals (Aertsen and Johannesma, 1980).

We have investigated the spectro-temporal structure of the PESE by measuring the ensemble average of dynamic power spectra, associated with the individual pre-event stimuli: the *Average Pre-Event Dynamic Power Spectrum* $p_e(\omega, \tau)$, given by

$$p_e(\omega, \tau) = \langle P_e(\omega, \tau) \rangle = \frac{1}{N} \sum_{n=1}^N P_n(\omega, \tau), \quad (5)$$

where $P_n(\omega, \tau)$ indicates the dynamic power spectrum of the n -th pre-event stimulus $x_n(\tau)$. The dynamic spectrum analysis has been implemented in hardware in order to enable the analysis to be real time: the *Dynamic Spectrum Analyser* (DSA). Basically it consists of a set of $\frac{1}{3}$ octave filters followed by envelope detectors, the result is the intensity as a function of time for the different frequency components of the

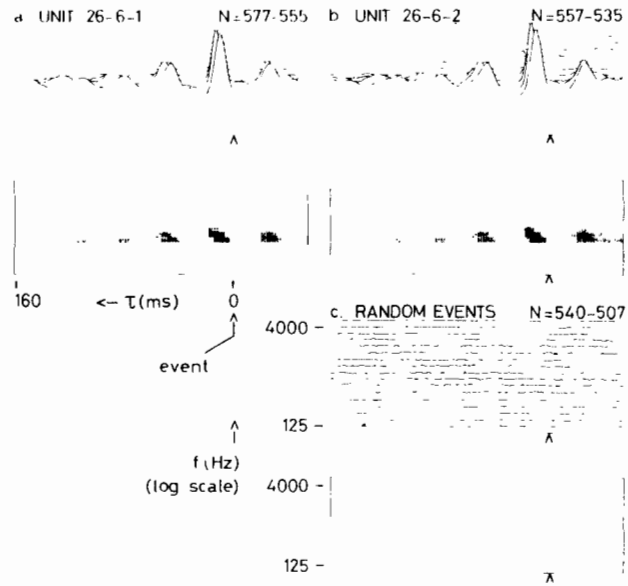


Fig. 5a-c. Average pre-event dynamic power spectrum $p_e(f, \tau)$ of the torus semicircularis unit 26-6 during first **a** and second **b** presentation of the acoustic biotope, vs. the a priori expected result for this stimulus ensemble **c**. Results are presented in two representations: 3-D (upper graph) and grey-coding (lower graph). The intensity scale in all three figures refers to arbitrary, identical units. The numbers N_1 and N_2 in the figures indicate the number of events presented to the DSA (N_1), respectively the number of events actually considered in averaging (N_2).

signal. The DSA also offers the facility of spike-triggered averaging of individual spectrograms. A description of the mathematical background and principles of operation is given in Part I (Aertsen and Johannesma, 1980). It should be noted here that the phase characteristics of the bandpass filters in the DSA give rise to a frequency dependent delay in the dynamic spectrum (cf. Fig. 6 in Part I). Obviously this delay should be taken into account when estimating temporal characteristics, like latency. Numerical values throughout the present paper have been corrected for this effect.

The average dynamic power spectrum of pre-event stimuli of unit 26-6 is shown in Fig. 5a in two ways: in a 3-D form and a grey-scale representation. The degree of reproducibility of this result is indicated by Fig. 5b, which was obtained upon repeated stimulation with the natural stimulus ensemble. The a priori expected outcome for this stimulus ensemble is shown in Fig. 5c, which was determined using a random sequence of events. All three averages were scaled to the same, arbitrary intensity scale in order to allow a proper comparison. Like in the first order analysis (Sect. 4) not all events were taken into account in the averaging procedure, due to the "dead time" in the analysis equipment, which equals 29.5 ms in all cases

OVERALL STIMULUS SPECTRUM

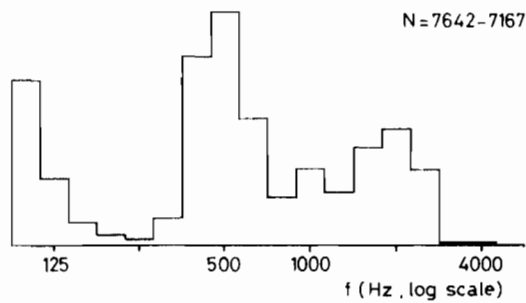


Fig. 6. Average spectral intensity density $J(f)$ of the acoustic biotope, determined with the DSA (resolution: $\frac{1}{3}$ octaves), using a random sequence of events and additional averaging over time τ .

considered in the present paper. The actual numbers N_1, N_2 ($N_1 \geq N_2$) are indicated in the figures. It is clear that, due to the better instrument specifications far fewer events are lost in this type of analysis than in the first order analysis.

From the results in Fig. 5 we learn that the average pre-event dynamic power spectrum for unit 26-6, using a natural stimulus ensemble, obtains conspicuously high values in the frequency range of about 350–600 Hz, some 20 ms prior to the neural event, as compared to the a priori expected dynamic spectrum. The duration of this elevation is about 15 ms and it is preceded as well as followed by remarkably low values in the dynamic spectrum, over an interval of about the same size. For the very same reason as indicated with relation to Fig. 4 the same elevation appears again for other values of τ , with gradually decreasing amplitude, moreover becoming less distinct. This result, again, points to a rather simple activation type frequency selectivity for unit 26-6, with a best frequency around 400–500 Hz and a latency in the order of 20 ms.

The outcome of this analysis, however, and as a consequence the interpretation of its results, is influenced strongly by the spectro-temporal structure of the stimulus ensemble, as was noted already in the case of first order averaging. The average pre-event dynamic power spectrum should be “normalized” with respect to the spectro-temporal characteristics of the stimulus ensemble in order to arrive at a measure of the spectro-temporal selectivity of the neuron: its “Spectro-Temporal Receptive Field” (Aertsen et al., 1980). In the case of a non-white Gaussian stimulus ensemble this normalization can be obtained simply by a spectral weighting in the bifrequency domain (“post-whitening”, e.g. Lee and Schetzen, 1965). In the case of a highly structured stimulus ensemble like the acoustic biotope an appropriate normalization procedure has not yet been formally derived. For the moment we will follow a more intuitive approach, which obviously

cannot lay any claim to formal justification. The primary goal is to compensate for the non-uniform spectral intensity density of the acoustic biotope. To this end an estimate was determined of the overall spectral composition of the stimulus ensemble. In order to enable its use in connection to the results as in Fig. 5, this estimate was determined using the DSA with the settings as in Fig. 5. In order to suppress the variance in the spectral estimate due to the temporal structure of the stimulus ensemble, averaging was performed with a random sequence of as many as about 7500 events. In order to avoid strong dead time effects in the analysis this sequence had a time span of 15 times that of the biotope, with an event rate comparable to that in Fig. 5c, i.e. about 500 events per complete stimulus presentation. The resulting dynamic power spectrum did still contain some temporal variations, this was eliminated by averaging over time τ . The resulting overall stimulus intensity spectrum $J(\omega)$, with a frequency resolution of $\frac{1}{3}$ octaves, shown in Fig. 6, is obviously not “white”. The most dominant peak (around 500 Hz) is caused mainly by the grassfrog’s vocal repertoire. The low frequency peak (below some 250 Hz) mainly reflects background noise, like wind and water, whereas the high frequency contribution (above some 1250 Hz) mainly originates from vocalizations of other animals. In order to compensate for this non-uniform spectral intensity density of the stimulus ensemble, the average pre-event dynamic power spectrum $p_e(\omega, \tau)$ was divided by the overall stimulus spectrum $J(\omega)$ to yield the “stimulus-filtered” average pre-event dynamic spectrum $q_e(\omega, \tau)$:

$$q_e(\omega, \tau) = \frac{p_e(\omega, \tau)}{J(\omega)} \quad (6)$$

The outcome of this procedure, applied to the results in Fig. 5 is shown in Fig. 7. The small peaks in the high frequency region (upper 3 filters) of the spectrograms merely reflect numerical instabilities due to the division by relatively small numbers (cf. Fig. 6). The “whitening”-effect of the “stimulus-filtering” can be noted from a comparison of the “random”-results in Figs. 5c and 7c. The main effect on the “neural”-results of unit 26-6 (Figs. 7a, b vs. 5a, b) is a broadening of the frequency range of the domain of high intensity values, without altering the peak frequency value very much. This implies that, also after a correction has been made for the relative abundance of stimulus intensity in the spectral region around 500 Hz, the average pre-event dynamic spectrum of unit 26-6 maintains the relatively high intensity values in that region. This clearly points to a genuine activation type frequency selectivity for frequencies around 500 Hz. The tuning of unit 26-6 furthermore appears to be simple and quite broad.

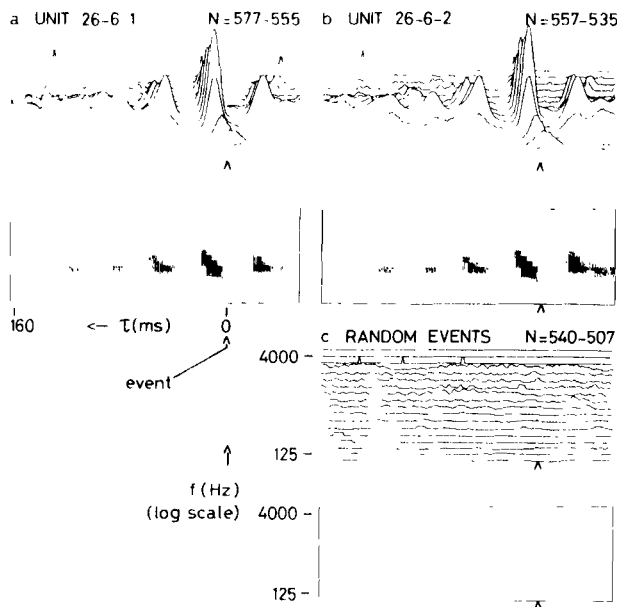


Fig. 7a-c. Average pre-event dynamic spectrum divided by stimulus spectrum. Spectro-temporal analysis of the stimulus-event relation of the torus semicircularis unit 26-6, corrected for the spectral composition of the acoustic biotope (“stimulus-filtering”); average pre-event dynamic power spectrum (Fig. 5), divided by the overall stimulus spectrum (Fig. 6)

Obviously the “stimulus-filtering” procedure has *not* been able to account for stimulus artefacts in the *temporal* structure of the average dynamic spectrum. This problem remains to be solved, it escapes from the intuitive type of reasoning which was applied here. It even remains to be seen whether a formally derived normalization procedure allows for an *uncoupling* of spectral and temporal normalization. We expect that a proper normalization will rather turn out to be a *combined* spectro-temporal normalization, because of the specific spectro-temporal coupling present in the acoustic biotope. Still, we maintain that some type of “stimulus-filtering” as described in (6) should be contained in an appropriate normalization, which is why the results in Fig. 7 should be regarded as a first step towards a naturally determined spectro-temporal receptive field. An interpretation of the results in Fig. 7 would be: activation in a frequency region around 500 Hz with a latency of some 20 ms. Noteworthy is furthermore the distinct temporal boundary of the activation domain, which covers an interval of approximately 15 ms and which is separated from the other elevation domains by “silent” intervals of about the same size. When connected with the stereotyped pulsatile structure of the male grassfrog vocalizations in the stimulus ensemble (note the time scale in Fig. 1) this indicates a rather strict temporal coupling of the neural firing process to the acoustic stimulus. This

finding is in accordance with the results of the coincidence analysis for this unit (Fig. 2). Unit 26-6 has also been investigated with a sequence of short tone pips. This revealed a composite spectro-temporal receptive field for tonal stimuli, containing both an *activation* and a *suppression* domain (cf. Figs. 6 and 7 in Part II). The characteristics of the “tonal” and “natural” activation domains match quite well. The suppression for tonal stimuli covers the same frequency range as the activation domain, whereas in time the suppression domain precedes the activation, i.e. it covers larger values of τ in the spectrogram. This points to a “*post-activation-suppression*” phenomenon: once the neuron has fired to an adequate tonal stimulus, the probability of firing to the same type of stimulus immediately following it (in this case: within some 20 ms) decreases considerably. It is tempting to connect this suppression phenomenon, revealed with tonal stimuli, with the “silent” intervals in Fig. 7 and the temporal structure of the grassfrog vocalizations (short sound elements, separated by silent intervals of about the same size).

Summarizing, the approach described here to analyse the average dynamic spectrum of pre-event stimuli may well be a possible way to investigate the spectro-temporal structure of the single neuron sensitivity domain for acoustic stimulation, i.e. its spectro-temporal receptive field, under natural stimulus conditions. Results of this analysis, however, have to be normalized with respect to the characteristics of the stimulus ensemble used. The “stimulus filtering” as described may be regarded as an intuitive first step towards this goal.

6. Results

The analysis as described in this paper was applied to 24 single unit recordings in the torus semicircularis from 11 grassfrogs. In all cases but one the natural stimulus ensemble was presented twice, in all cases a tonal analysis was performed as well.

For all neurons the *existence* of a stimulus-event relation could be established, whether simply by mere visual inspection of the dot displays or by use of the cross coincidence histogram. Unit 28-3 (Fig. 2, right) was the least clear in this respect.

All recordings were subjected to the first and second order analysis of the PESE as described in this paper. The results for 8 units will be presented as an illustrative sample. The average pre-event stimulus is shown in Fig. 8a-h, together with the result for a random sequence of events (Fig. 8i). The average dynamic power spectrum of pre-event stimuli of these units together with the result for random events is shown in Figs. 9 (3-D representation) and 10 (grey vague (Fig. 9f) or even very vague (Fig. 9g and h). The

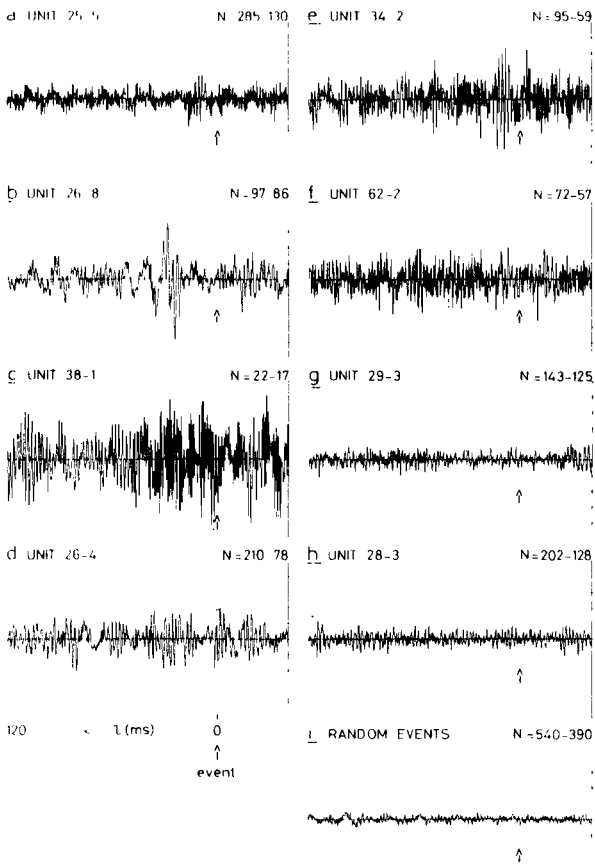


Fig. 8a-i. Average pre-event stimulus $r_e(t)$ of eight torus semi-circularis units a-h, determined with the acoustic biotope, vs. the a priori expected result for this stimulus ensemble i. The vertical scale in all figures refers to arbitrary, identical units. Further details as in Fig. 4

coding). Finally, Figs. 11 (3-D) and 12 (grey) contain the “stimulus filtered” average pre-event dynamic power spectrum for the 8 neurons and the random event sequence. In each figure the results were scaled using the same amplitude respectively intensity scale, to allow for a comparison. The numbers of events, incorporated in averaging (N_1, N_2) are indicated in the figures.

The analysis of the average pre-event stimuli (Fig. 8) in general does not lead to very informative results. The reproducibility upon repeated stimulation appears to be rather low, i.e. in general definitely less reproducible than in the case of unit 26-6 (Fig. 4). Consequently the interpretation of these results remains difficult for the time being, be it merely for the reasons already mentioned in Sect. 4.

The average pre-event dynamic power spectrum (Figs. 9 and 10) gives rise to a variety of results. In all cases one or more (f, τ)-domains can be distinguished with notably higher intensity values. These domains vary from being quite distinct (e.g. Fig. 9b, d, and e) to

elevations cover various frequency regions, also temporally the results differ to quite a degree. The reproducibility of the results upon repeated presentation of the stimulus ensemble, in general, is rather high, be it that the variance increases with decreasing number of neural events. For an interpretation in terms of spectro-temporal sensitivity of the neurons these original averages do not satisfy for reasons explained in Sect. 5. A better impression can be gained from an observation of the “stimulus filtered” average pre-event dynamic power spectra (Figs. 11 and 12), weighed for the overall spectral composition of the stimulus ensemble. When compared to the original averages, the effect of “stimulus filtering” varies from only minor (e.g. Fig. 9a and e) to quite drastic (e.g. Fig. 9b and c). Quite clear activation domains appear for high frequencies (e.g. Fig. 9a, c, and e) as well as low frequencies (e.g. Fig. 9b and d), together with units with only a very vague spectral preference (Fig. 9f-h). Latency varies from short (e.g. Fig. 9a, b, and e: about 20 ms) to longer (e.g. Fig. 9c and f: about 60 ms), very long (Fig. 9g: about 100 ms) or even no clear latency (Fig. 9h). In this context it is interesting to note that unit 28-3 with the most vague activation domain (Fig. 9h) corresponds to the very broad coincidence histogram shown in Fig. 2 (right).

Although the “stimulus-filtered” average pre-event dynamic power spectra do not really represent the spectro-temporal receptive fields for natural stimuli, it still is worthwhile to compare the present findings with the main characteristics of the receptive fields of these units, as they were determined using an ensemble of *tonal* stimuli (Part II: Aertsen et al., 1980). Findings from the latter experiments are summarized in Table 1, more detailed results for the units in Fig. 7 and in Figs. 11 and 12a-d are given in Part II. Most of the information in Table 1 is self-explanatory, the last column deserves some explanation. The firing rate of neurons in the torus to a long (~ 8 min) sequence of short (16 ms) tone pips, immediately following each other, appeared to behave in, grossly spoken, two different ways: either it remained constant during the whole sequence (“stationary” units: +) or, after a few seconds of stimulation, it decreased strongly to stay at a very low level or, after some time (~ 1 min) recovered slightly but never retained the initial level (“long-term adapting” units: -). The characteristics of the tonal receptive fields appeared to be coupled to this phenomenon (cf. Table 1 and Part II). The comparison of these results will receive further attention in the Discussion.

7. Discussion

In the present paper procedures were described to analyse the stimulus-event relation of auditory neu-

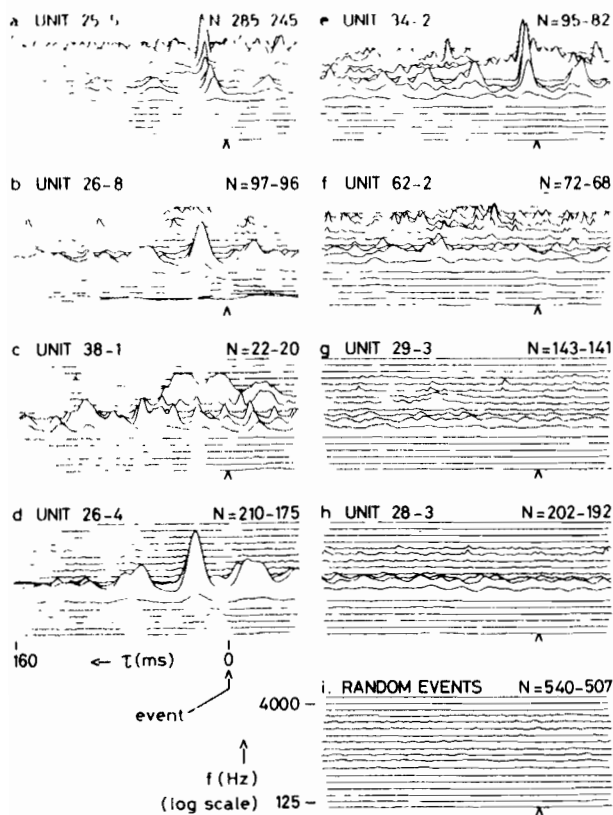


Fig. 9a-i. Average pre-event dynamic power spectrum $p_s(f, \tau)$ of eight torus semicircularis units **a-h**, determined with the acoustic biotope, vs. the a priori expected result for this stimulus ensemble **i**. The intensity scale in all figures refers to arbitrary, identical units. Further details as in Fig. 5

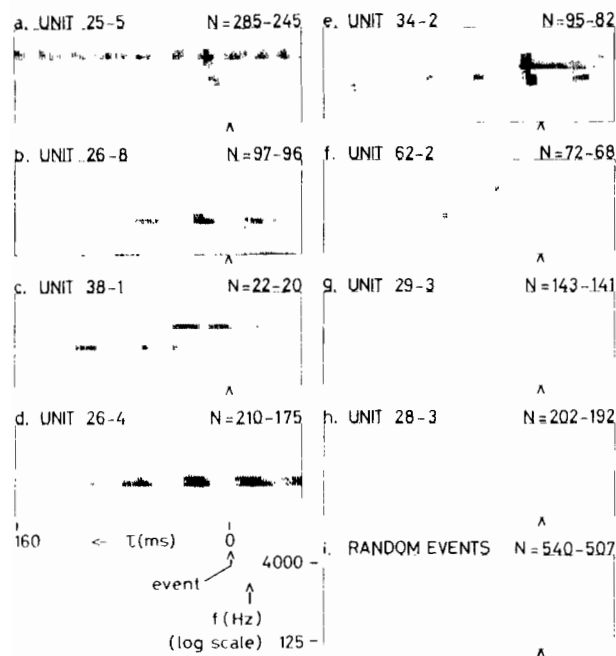


Fig. 10. Grey-scale representation of the results in Fig. 9

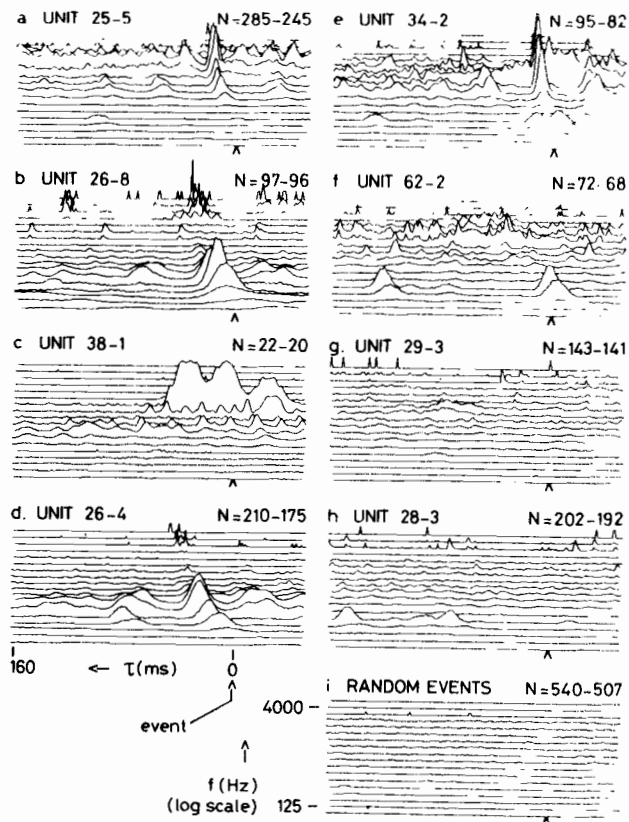


Fig. 11. Spectro-temporal analysis of the stimulus-event relation of eight torus semicircularis units, corrected for the spectral composition of the acoustic biotope ("stimulus-filtering"): average pre-event dynamic power spectrum (Fig. 9) divided by the overall stimulus spectrum (Fig. 6)

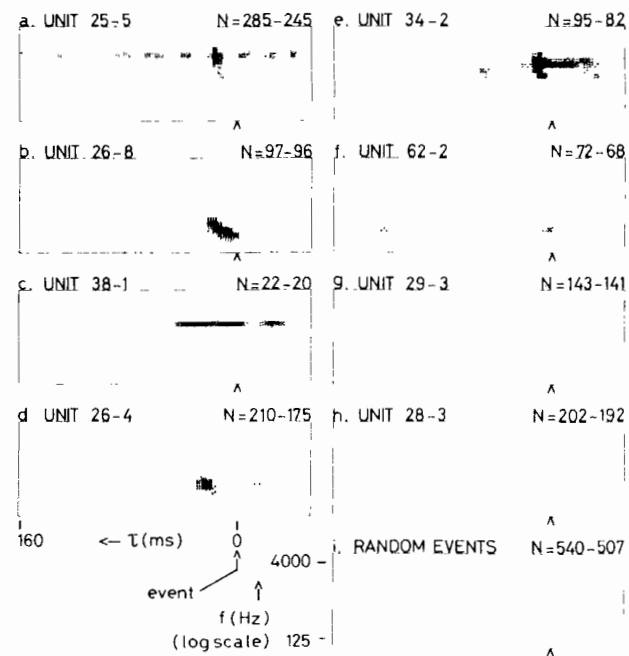


Fig. 12. Grey-scale representation of the results in Fig. 11

Table 1. Characteristics of the stimulus-event relation of nine torus semi-circularis neurons, determined with an ensemble of tonal stimuli (Aertsen et al., 1980). Numerical values are given in round numbers. Further explanation in text

Unit	Receptive field type (activation A and/or suppression S)	Best frequency (Hz)	Spectral width (octaves)	Latency (ms)	Temporal width (ms)		Response type [stationary (+) or adapting (-)]
					Activation	Suppression	
26-6	A, S	520	1	20	15	20	+
25-5	A, S	1350	1	20	15	20	+
26-8	A, S	≤ 300	≥ 1.5	30	15	60	+
38-1	A, S	1900	0.5-1	60	30	60	-
26-4	A	400	1	40	30		-
34-2	A	1140	0.5	15	15		+
62-2	A	> 2000	≥ 0.5	70	30		-
29-3	A	200-300	1-1.5	120	50		-
28-3							-

rons, using an ensemble of natural stimuli. Basically these methods consist of the determination of different types of average measures, defined on the ensemble of stimuli, which precede the occurrence of action potentials. In order to illustrate the concepts, these procedures were applied to a limited number of extracellular single unit recordings from the torus semi-circularis of the grassfrog.

Various points concerning the experimental conditions have already been discussed elsewhere (Aertsen et al., 1980), among which the possible effects of anaesthesia (MS-222) and the binaural presentation of stimuli.

The approach as described is based on the concept of the Pre-Event Stimulus Ensemble (PESE), the special selection from the stimulus ensemble, made by the neuron by means of the neural events it fires. The properties of this selection are thought to characterize the neuron's behaviour in the representation of acoustic stimuli. A characterization of this sub-ensemble by means of various average measures results in a global description of the stimulus-event relation, a property which is common to any statistical approach. In fact the various averages represent as many *projections* of the pre-event stimulus ensemble, each projection emphasizing one or more special aspects of the nature of the stimulus-event relation. Formally the determination of these average measures is identical to the calculation of the cross correlation functions of the neural activity and various functionals of the stimulus. These cross correlation functions play a central role in the Wiener-like approach to the theory of non-linear systems (Lee and Schetzen, 1965; Marmarelis and Marmarelis, 1978). Accordingly the procedures as outlined in the present paper and preceding ones (Parts I and II) may be interpreted as elements of a *probabilistic formulation* of the non-linear analysis of neural systems (Johannesma, 1980).

First Order Averaging of the PESE

Results in the present study of first order averaging of the PESE in general were not very informative. Occasionally the average pre-event stimulus showed a distinct waveform (e.g. Figs. 4, 8b, and 8c). The reproducibility of the results was low. In a number of cases the waveform provided an indication of a possible activation-type frequency selective mechanism. The few units which did produce a distinct waveform in the average pre-event stimulus do not necessarily have to show phase-lock (Lavine, 1971) to the stimulus waveform. The results of the dynamic spectrum analysis (Figs. 5, 9, and 10) indicate that in a number of cases the neural firing behaviour is coupled temporally to the intensity of certain frequency components. This phenomenon, in combination with the stereotyped waveforms of various vocalizations in the natural stimulus ensemble may well give rise to a "virtual" phase-lock in the response. The general correspondence in magnitude between the average pre-event stimuli (Figs. 4 and 8) and the average pre-event dynamic spectra (Figs. 5, 9, and 10) seems to support this possibility. The presence of genuine phase-lock to the stimulus waveform can be investigated by additional stimulation with a sign reversed copy of the stimulus ensemble and a detailed inspection of the form of the cross coincidence histograms of the resulting recordings of neural activity around $\sigma=0$. It is interesting to note that this same extension of the stimulus ensemble has been suggested in another context: the possibility to improve the statistical properties of pseudo-random noise sequences (Swerup, 1978).

Summarizing: given the limited amount of experimental evidence the results on the whole so far, together with the problems already mentioned in Sect. 4, suggest that for a natural stimulus ensemble the procedure of first order averaging of the PESE of neurons at this level of the auditory system of the

grassfrog, under the present experimental conditions, leads to results which are not simply interpretable given our present-day understanding of the auditory nervous system.

Spectro-Temporal Averaging of the PESE

The second order analysis of the PESE was performed with respect to the variables of *frequency* and *time*. These dimensions of sound have traditionally played an important role in the investigation of the auditory system, furthermore they have been shown to be of behavioural significance in several anuran species (e.g. Capranica, 1965; Gerhardt, 1974, 1978; Van Gelder, 1978). Moreover the analysis aims at a *combined spectro-temporal* representation of the stimulus-event relation instead of *separate spectral* and *temporal* descriptions. The arguments for this choice have been discussed in Part II. Recent findings indicate that indeed the spectro-temporal receptive fields of units in the grassfrog's auditory midbrain, determined under various stimulus conditions, in general cannot be factorized into separate spectral and temporal characteristics (in preparation).

A representation of sound in terms of frequency and time, which is applicable to any type of signal, is provided by the dynamic power spectrum as implemented in the DSA (cf. Part I). The DSA, like any sonograph, however, shows several shortcomings. The spectral resolution, and the temporal resolution coupled to it, have to be selected a priori. These characteristics of the frequency analysing mechanism of the DSA appear in the spectro-temporal signal description. Phase relations between different frequency components are lost, due to the quadratic operation, inherent to any sonograph. This implies that it is impossible to make an a posteriori trade-off between spectral and temporal resolution. The particular selection of filters, used in the present investigation ($\frac{1}{3}$ octave bands), is an arbitrary one; the method of "constant percentage bandwidth" analysis, however, is quite common in auditory research. The combined spectro-temporal resolution of the filters in the DSA, the "uncertainty product" (Gabor, 1946), equals 0.78, which is only slightly above the theoretical limit of 0.5. This is evidently much smaller than the "size" of the activation domains which were found for the torus semicircular units. This suggests that, in view of a combined spectro-temporal analysis, the specifications of the DSA, under the present conditions, are quite satisfactory.

Quite another technical limitation of the DSA is formed by the "dead-time" in averaging (here: 29.5 ms; Part I), which results in a missing of neural triggers. Thanks to the relatively low firing rates of neurons in

the torus semicircularis this did not impose very severe limitations (cf. the numbers N_1 and N_2 , indicated in the Figs. 5, 9, and 10). In cases of higher event rates, however, one should be aware of possible systematic effects. A solution to this problem is purely a technical matter, a simple solution might be a repeated analysis of the same data combined with the marking of events, already used in averaging.

Several of the spectro-temporal shortcomings of the dynamic power spectrum analysis are the consequence of its nature: an instrument-mediated projection of an underlying spectro-temporal intensity density of signals. It can be shown that any DSA can be expressed as a linear functional of an underlying *Complex Spectro-Temporal Intensity Density* (CoSTID) (Rihaczek, 1968; Johannesma and Aertsen, 1979). This CoSTID as such has no physical meaning but various important signal characteristics can easily be derived from it. A spectro-temporal analysis of the PESE by means of the CoSTID provides a close formal link to the more conventional type of second order cross correlation analysis (e.g. Lee and Schetzen, 1965). This more formal approach to the spectro-temporal description of sound may very well prove to be indispensable in the derivation of adequate stimulus-normalization procedures, necessary to attain a correct representation of the neural spectro-temporal receptive field, which takes into account the a priori knowledge of the characteristics of the stimulus ensemble (Aertsen and Johannesma, in preparation). Preliminary results of simulation studies, following that approach and using as a stimulus ensemble various mixtures of natural stimuli and Gaussian wideband noise, appear to be promising. The "stimulus-filtering" as described in the present paper also appears to play an important role in that more formal approach, be it that other aspects (i.e. the autocorrelation structure) of the stimulus ensemble have to be taken into account as well. At the present level of analysis it can be stated that the "stimulus-filtered" average pre-event dynamic spectrum provides a useful first step towards the determination of the neural spectro-temporal receptive field.

Stimulus-Invariance of Receptive Field

A comparison of the present findings for an ensemble of natural stimuli to the outcome of an analysis using tonal stimuli provides an indication to what extent the spectro-temporal receptive field of a neuron still depends on the stimulus ensemble used, after the necessary corrections for stimulus properties have been made. The question of interest is, whether the receptive field is possibly completely or partially *stimulus-invariant*. This evidently is a necessary prerequisite if

one hopes to reconcile the results for different types of stimuli. In general there is, however, no obvious a priori reason why this stimulus-invariance should hold for such various stimuli as tones, natural sounds or noise. Consequently, the stimulus-invariance should be *investigated*, rather than be *presupposed*. This is the main argument for an analysis of the stimulus-event relation for various stimulus ensembles, mutually as remote as noise, tones and natural sounds, where in each case identical, or at least comparable, procedures are followed. At the same time this approach implies that in all cases the stimulus ensemble should be wide enough to encompass all relevant aspects of the particular projection of the *s-e* relation one is interested in. In the present case this implies that the natural stimulus ensemble should be constructed as wide as possible in the sense that all sorts of spectro-temporal intensity patterns occur. An ensemble, restricted to species-specific vocalizations will evidently be less adequate in this respect, other natural sounds have to be included as well, possibly together with synthetical variations, to form an acoustic (pseudo-)biotope.

A comparison of the present findings with natural stimuli and the results of the same type of analysis of tonal experiments (cf. Table 1 and Part II) leads to the following observations. The torus semicircularis units which showed a "stationary" response to a sequence of short tone pips (denoted by "+" in Table 1), had a tonal receptive field which, in general, was well defined, relatively small-sized, had short latency and could consist of an activation domain with or without a preceding suppression domain. The units which showed a "nonstationary" response to tones ("long-term adaptation", denoted by "-" in Table 1) had receptive fields which were vaguer, larger (especially in the temporal dimension), had longer latency, and contained an activation domain. Sometimes a preceding suppression domain could be distinguished, but not always was the receptive field clear enough to decide on this. Occasionally "nonstationary" units were found e.g. unit 28-3 for which hardly any sign of a tonal receptive field could be found at all, although the cross coincidence analysis clearly pointed to the existence of a stimulus-event relation. It appears that the results of natural stimulation of the "stationary" units, on the whole, agree quite well with the tonal findings for these units. For the "adapting" neurons the acoustic biotope, in general, appeared to be relatively more effective in eliciting spikes than the tonal sequences, probably due to its more complex character (e.g. non-stationarity, spectral composition). Their naturally determined spectro-temporal selectivity was less distinct than for the stationary units, but in general contained as much as or even more structure than their tonally determined receptive fields did. Especially the temporal

spread was very large. Unit 28-3 (Figs. 11a and 12a), for instance, only showed a clear frequency selectivity (activation around 250 Hz) after the time-window of the analysis was enlarged to several hundreds of ms. This result is in accordance with the large temporal uncertainty in the cross coincidence histogram for this unit (Fig. 2, right). Even at these large values of the time-window the tonal results did not give rise to any spectro-temporal selectivity.

Summarizing: for the "stationary" units the results of tonal and natural analysis agree quite well, the "non-stationary" units are evidently more complex and, at the present level of analysis, a proper comparison of the spectro-temporal selectivity under different stimulus conditions cannot really be made in a satisfactory way. For these latter units this comparison awaits the development of more adequate stimulus-normalization procedures. Only then the outcome truly reflects stimulus dependence of the receptive fields.

A possible way to evaluate the adequateness of the spectro-temporal selectivity of a neuron, remaining within the *same* stimulus ensemble, might be found by making use of the cross coincidence function (Sect. 3.1). The receptive field of a neuron can be translated, by making use of a probabilistic formulation, into a predictor of neural activity, given a certain spectro-temporal intensity pattern of the stimulus (Johannesma, 1980). This predictor might be used to generate a theoretically expected event density upon stimulation with the acoustic biotope. The cross correlation function of this expected event density with an actual recording of that neuron may be compared to the experimental cross coincidence histogram of that same neuron under identical stimulus conditions. This comparison may provide an indication to what degree the receptive field analysis forms an adequate description of the stimulus-event relation for that particular stimulus ensemble, or whether important characteristics may have been overlooked or may be hidden in higher order properties of the PESE.

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 and Koos Braks, Jan Bruijns, Wim van Deelen, and Dik Hermes for expert assistance and/or advice in the experiments and data analysis.

References

- Aertsen, A.M.H.J., Smolders, J.W.T., Johannesma, P.I.M. Neural representation of the acoustic biotope: on the existence of stimulus-event relations for sensory neurons. *Biol. Cybern.* **32**, 175-185 (1979)

- 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. Cybern. **38**, 223-234 (1980)
- Aertsen, A.M.H.J., Johannesma, P.I.M., Hermes, D.J.: Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli. Biol. Cybern. **38**, 235-248 (1980)
- Boer, E. de, Kuyper, P.: Triggered correlation. IEEE Trans. Biomed. Eng. BME-**15**, 169-179 (1968)
- Boer, E. de, Jongh, H.R. de: On cochlear encoding: potentialities and limitations of the reverse correlation technique. J. Acoust. Soc. Am. **63**, 115-135 (1978)
- 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: M.I.T. Press 1965
- Fano, R.M.: Short-time autocorrelation functions and power spectra. J. Acoust. Soc. Am. **22**, 546-550 (1950)
- Flanagan, J.L.: Speech analysis, synthesis, and perception. Berlin, Heidelberg, New York: Springer 1965
- Frishkopf, I.S., Goldstein, M.H., Jr.: Responses to acoustic stimuli from single units in the eighth nerve of the bullfrog. J. Acoust. Soc. Am. **40**, 469-472 (1963)
- Gabor, D.: Theory of communication. J.I.E.E. (London), Part III **93**, 429-457 (1946)
- Gelder, J.J. van, Evers, P.M.G., Magnus, 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 cell 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)
- Grashuis, J.I.: The pre-event stimulus ensemble: an analysis of the stimulus-response relation for complex stimuli applied to auditory neurons. Ph. D. Thesis, Nijmegen 1974
- 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., 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)
- Johannesma, P.I.M.: Functional identification of auditory neurons based on stimulus-event correlation. In: Psychophysical, physiological, and behavioural studies in hearing, pp. 77-84. Brink, G. van den, Bilsen, F.A. (eds.). Delft: Delft University Press 1980
- Kiang, N.Y.-S., Moxon, E.C.: Tails of tuning curves of auditory nerve fibers. J. Acoust. Soc. Am. **55**, 620-630 (1974)
- Lavine, R.A.: Phase-locking in response of single neurons in cochlear nuclear complex of the cat to low-frequency tonal stimuli. J. Neurophysiol. **34**, 467-483 (1971)
- Lee, Y.W., Schetzen, M.: Measurement of the Wiener kernels of a nonlinear system by cross-correlation. Int. J. Control **2**, 237-254 (1965)
- Marmarelis, P.Z., Marmarelis, V.Z.: Analysis of physiological systems. The white-noise approach. New York, London: Plenum Press 1978
- Pedersen, J.G.: On the existence of stimulus-event relations for sensory neurons: a statistical method. Biol. Cybern. **38**, 201-212 (1980)
- Potter, R.K., Kopp, G.A., Green, H.C.: Visible speech. New York: Van Nostrand 1947
- Richaczek, A.W.: Signal energy distribution in time and frequency. IEEE Trans. Inf. Theory IT-**14**, 369-374 (1968)
- Schroeder, M.R., Atal, B.S.: Generalized short-time power spectra and autocorrelation functions. J. Acoust. Soc. Am. **34**, 1679-1683 (1962)
- 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. Cybern. **35**, 11-20 (1979)
- Swerup, C.: On the choice of noise for the analysis of the peripheral auditory system. Biol. Cybern. **29**, 97-104 (1978)
- Wilson, J.P., Evans, E.F.: Systematic error in some methods of reverse correlation. J. Acoust. Soc. Am. **57**, 215-216 (1975)
- 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: September 1, 1980

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