Auditory Processing of Acoustic Communication Signals: Sensory Biophysics, Neural Coding, and Discrimination of Conspecific Songs

Andreas V.M. Herz[1], Jan Benda[2], Tim Gollisch[1], Christian K. Machens[3], Roland Schaette[1], Hartmut Schütze[1], and Martin B. Stemmler[1]

[1]Institute for Theoretical Biology, Humboldt-Universität zu Berlin, 10115 Berlin, Germany
[2]Department of Physics, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada
[3]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

I. Introduction

Evolution has led to acoustic communication behaviors of fascinating complexity (see, e.g., Hauser 1996, Bradbury and Vehrenkamp 1998), which are made possible by sophisticated neural systems in both sender and receiver. Remarkably, even small insect auditory systems are capable of astounding computations. Some grasshoppers, for example, reliably detect gaps in conspecific songs as short as 1-2 ms (von Helversen 1972), a performance level similar to that reached by birds and mammals.

These observations raise the question of how a minute insect auditory system processes auditory signals reliably and with high temporal precision. Important insight will come from understanding the auditory periphery. It serves as a strategic bottleneck between the external world and further neural processing stages; every computation and behavioral decision must be based on the primary stimulus representation on the level of auditory receptors.

Understanding the interplay between the dynamics and function of these neurons requires answers to a broad range of questions such as:

- Which physical sound attribute (e.g., sound pressure or energy) actually drives the receptor?
- What are the essential processing steps of auditory transduction and encoding?
- In which way do neural noise sources limit the system's performance?
- Is it possible to "read" the sensory input from the output of a single receptor?
- Are receptor neurons specifically tuned to behaviorally relevant features?

Grasshoppers of the species *Chorthippus biguttulus* provide a suitable model system to study these questions. Their calling and courtship songs are based on broad-band carrier signals with amplitudes that are strongly modulated in time. Although lacking tonal elements, the songs possess an elaborate temporal structure, rhythmically arranged into distinct syllables, that are

separated by short pauses (Fig. 1a). Several differences between individual songs from the same species can be noted: the frequency content of the broad-band carrier, the syllable length, and the precise temporal pattern of amplitude modulations within a syllable (Fig. 1b). This song variability could allow females to choose among different males and thus provide a basis for mate preference and sexual selection (Kriegbaum and von Helversen 1992).

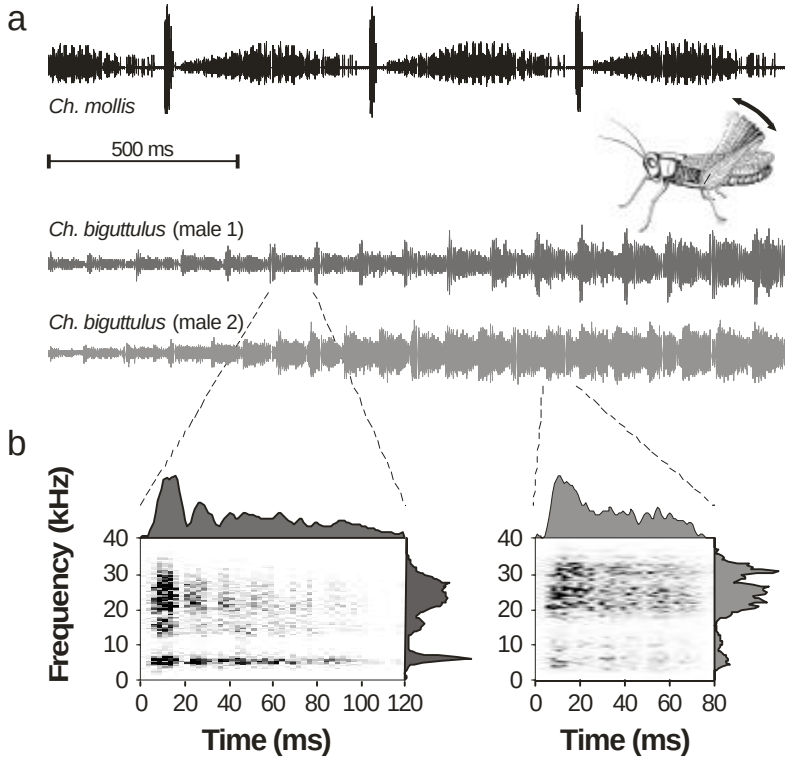

Figure 1: Acoustic communication signals. (**a**) Grasshoppers produce species-specific sound patterns by rasping their hindlegs across their forewings. Songs generated by one *Chorthippus mollis* and two *Chorthippus biguttulus* males are shown. Each song consists of many repetitions of a basic pattern, termed "syllable". (**b**) Within each syllable, the amplitude of the high-frequency broadband carrier is strongly modulated in time as illustrated by the spectrograms.

On the receiver side, the songs are encoded by a total of roughly one hundred auditory receptors (in *Ch. biguttulus*, R. Lakes-Harlan, pers. com.) into trains of action potentials. The cells are located at the two tympana on both sides of the animal; their axons extend through the tympanic nerves to the metathoracic ganglion, where auditory information is processed by local interneurons and then sent to the brain via ascending neurons. The tympanic nerve is easily accessible for intracellular recordings from single axons. This allows one to study the output of individual neurons without damaging the animal's ear. Detailed investigations are best carried

out in locusts (*Locusta migratoria*) whose larger but homologous structures (Stumpner and Ronacher 1991) facilitate long experiments.

In this chapter, we will explore various aspects of the dynamics and signal-processing capabilities of a single type of neuron, the auditory receptor, using a variety of modern and partly novel experimental and theoretical techniques. The methods cover a wide range of disciplines – from biophysics to information theory – and demonstrate that a tight interplay of experiment, data-analysis, and theory can yield valuable new insights. All approaches have one feature in common: no neural parameter or variable, apart from the acoustic input and the final spike output, needs to be measured. The techniques may thus be of use for investigations of other systems that only allow axonal or extracellular recordings.

At this stage, our goal is not centered on producing a single, unified computational model. Rather, we will present a collection of independent studies that aim at elucidating key processes of auditory receptor dynamics – signal transduction (Section II), spike-frequency adaptation (Section III), and spike-train variability (Section IV) – and their consequences for neural coding (Section V) and song discrimination (Section VI). Furthermore, we concentrate on the underlying concepts and main results. Detailed descriptions of the experiments and data-analysis methods can be found in the original articles.

II. Biophysics of auditory signal transduction

Auditory receptors transform an incident sound wave into a train of action potentials. Several sequential steps are required for this signal transduction:

(1) Mechanical coupling. The acoustic stimulus induces vibrations of the ear drum.

(2) Mechanosensory transduction. These vibrations cause the opening of mechanosensory ion channels in the membrane of the receptor neuron, which in grasshoppers and locusts are directly connected to the ear drum via short dendrites.

(3) Electrical integration. The electrical charge of the ions accumulates at the cell membrane.

(4) Spike generation. Action potentials are triggered by voltage-dependent currents and travel down the auditory nerve.

Each of these four steps transforms the signal in a specific way, ranging from nearly linear (the ear-drum response) to strongly non-linear (the generation of action potentials). Spike

generation itself is influenced by (at least) two further processes; **spike-frequency adaptation**, i.e., the gradual reduction of the output activity despite constant stimulation, and **stochastic fluctuations** of the membrane potential due to internal noise sources (such as ion-channel noise). These fluctuations cause jittered spike times as well as "missing" or "additional" spikes from trial to trial.

The entire sequence of signal transduction is completed after about one millisecond, not counting spike-frequency adaptation, which occurs on a longer time scale As the ear drum is delicate and the sensory periphery highly vulnerable, how can one dissect the individual steps in the auditory transduction chain? We are often left with no choice but to make do with the final output, the spike. Nonetheless, much more information can be gathered about the details of transduction that one might surmise at first.

A. Conceptual framework: analysis and comparison of iso-response stimuli

To dissociate and identify the individual steps of auditory signal transmission, we revisit and extend an experimental strategy used for measuring threshold curves in neurobiology (Evans 1975) or applying equivalence criteria in psychophysics (Jameson and Hurvich 1972): Instead of estimating the full input-output functional F describing all four stages (1-4) of the transduction chain sketched above, we focus on F's level surfaces, that is, we vary stimulus parameters such that the investigated neuron stays at a *constant level of final output activity* (Gollisch et al. 2002). For each point on such a "iso-response manifold in stimulus space", the nonlinear process of generating the specified output is identical; thus, by examining the invariances in the input-output relations, we may uncover information about the earlier stages (1-3) of the transduction chain.

In addition, even low-dimensional stimulus subspaces will contain a clear signature of the invariant stimulus manifolds, as shown by a simple example: if the iso-response manifolds of some system form concentric ellipsoids in a high-dimensional input space, any measurement within a two-dimensional planar subspace will reveal ellipses as invariant regions (Fig. 2a). These ellipses can readily be distinguished from the straight lines that result from iso-response manifolds consisting of hyperplanes (Fig. 2b).

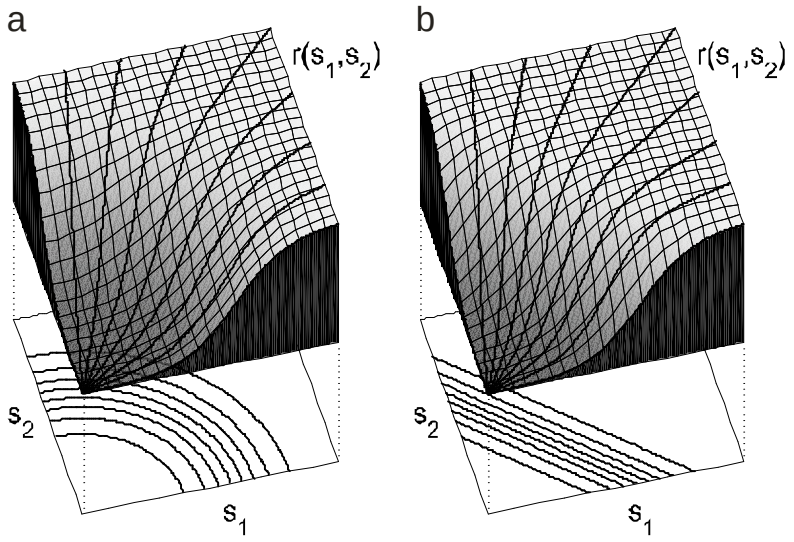

Figure 2: Characterization of input-output relations for two different response scenarios. A hypothetical two-dimensional stimulus space is parameterized by the variables s_1 and s_2. The drawn surfaces represent the resulting response $r(s_1,s_2)$ for two alternative models, which take the quadratic (**a**) and linear sum (**b**) as the argument of a sigmoid nonlinearity. Although the two scenarios are fundamentally different, both produce exactly the same one-dimensional response functions $r(s_1)$ and $r(s_2)$, respectively, as seen by the black areas at the sides of the surface block. Furthermore, any measurement along a radial direction, as is common in experimental practice, will produce similar sigmoid response curves in both cases as seen by the thick black lines running along the surfaces. The iso-response manifolds r=const (here: one-dimensional curves) below the surface plots, however, give a clear signature of the different underlying processes.

Compared to other approaches that involve presenting a whole range of stimuli meant to cover all possible relevant inputs, the approach of only comparing stimuli producing the *same* output is fairly parsimonious. Often enough, experimental time is limited; not needing to take the final nonlinear mapping into account can be a highly valuable benefit.

B. Mathematical framework for sensory processing: LNLN signal cascades

Sensory neurons convert an incoming stimulus into a neural response, such as a firing rate of the occurrence probability of a single action potential. From a mathematical point of view, the operation of a single sensory neuron can be written as

$$r(t) = F[\vec{s}(\cdot)], \tag{1}$$

where the "$\cdot$" in the argument of the stimulus $\vec{s}$ emphasizes that the response r at time t depends not only on the current stimulus but also on its history. Prominent realizations of this general mapping include processing sequences where a linear filter operation L in the time domain is followed by a static non-linearity f,

$$r(t) = f\left[\int_0^\infty L(t)\cdot s(t-t)dt\right]. \tag{2}$$

Such a sequence of a linear filter and a non-linear map is also known as an "LN cascade" (Hunter and Korenberg 1986). If the elementary processes are arranged in the opposite order, an "NL cascade" results,

$$r(t) = \int_0^\infty L(t)\cdot f\left[s(t-t)\right]dt. \tag{3}$$

Transduction processes with more complicated dynamics can be modeled as combinations of elementary cascades. For example, an "LNLN cascade" corresponds to two nested LN cascades,

$$r(t) = f_2\left\{\int_0^\infty L_2(t_2)\cdot f_1\left[\int_0^\infty L_1(t_1)\cdot s(t-t_1-t_2)dt_1\right]dt_2\right\}. \tag{4}$$

C. Firing rates reveal the functional form of the intermediate static nonlinearity

To illustrate how the iso-response method can be applied to find models for the operation of auditory receptors, we first consider sound pressure waves $s(t)$ that consist of superimposed pure tones,

$$s(t) = \sum_{n=1}^{N} A_n \sin\left(2\pi n_n t + \varphi_n\right), \tag{5}$$

where the n_n denote the frequencies, φ_n phase offsets, and the A_n the respective amplitudes. Due to mechanical properties of the ear drum, $s(t)$ is linearly filtered and thereby turned into

$$\tilde{s}(t) = \sum_{n=1}^{N} \frac{A_n}{C_n}\sin(2\pi n_n t + \tilde{\varphi}_n). \tag{6}$$

This means that every tone receives a frequency-dependent gain factor $1/C_n$. In addition, the phase may change from φ_n to $\tilde{\varphi}_n$, but this shift is not relevant for stimulus encoding as the investigated receptors do not phase lock to the sound's carrier (Suga 1960), but rather generate irregular spike trains. Since the resulting discharge rate $r(t)$ stays approximately constant after a brief transient following stimulus onset, the input-output relation (Eq. 1) can be approximated by a static nonlinearity

$$r = f(J) \tag{7}$$

where the "effective stimulus intensity" J describes the cell's spectral integration process.

Turning to an ongoing debate about which stimulus attribute of a sound actually triggers neural output activity, see e.g. Garner (1947), Tougaard (1996) or Heil and Neubauer (2001), the iso-response framework allows us to test three rival hypotheses about the physical nature of J:

Amplitude Hypothesis: J corresponds to the maximum amplitude of $\tilde{s}(t)$. This is the common view of a threshold; a response occurs once the signal reaches a certain value. In the case of few frequency components, J is given by the sum of the scaled amplitudes,

$$J_{AH} = \sum_{n=1}^{N} \frac{A_n}{C_n}. \tag{8}$$

Energy Hypothesis: J corresponds to the temporal mean of the squared signal,

$$J_{EH} = \left\langle \tilde{s}(t)^2 \right\rangle = \frac{1}{2} \sum_{n=1}^{N} \frac{{A_n}^2}{{C_n}^2}. \tag{9}$$

Since the square of the amplitude of a sinusoidal oscillation is proportional to the energy contained in the oscillation, this hypothesis reflects an energy-integration mechanism.

Pressure Hypothesis: J corresponds to the temporal mean of the absolute value of $\tilde{s}(t)$:

$$J_{PH} = \left\langle \left| \tilde{s}(t) \right| \right\rangle. \tag{10}$$

This complies with a pressure-integration mechanism after half-wave rectification.

Based on the iso-response paradigm, we measured firing-rate responses to superpositions of two sine-wave stimuli (Eq. 5) with different relative contributions A_1 and A_2 of the two tones. The stimulus intensities were tuned during the experiment such that always the same firing rate was obtained. This led to different combinations of A_1 and A_2 which can be compared to the predictions of the three hypotheses. As suggested by Fig. 3a, the amplitude as well as the pressure hypotheses can be rejected (Gollisch et al. 2002). The energy hypothesis, on the other hand, provides a good description of the data in the two-tone case. As additional test of the energy hypothesis, we also investigated how iso-firing-rate curves that were obtained separately for different firing rates are related to one another. Fig. 3b shows pairs (A_1, A_2) corresponding to several different firing rates. Pairs corresponding to the same firing rate are accurately fitted by ellipses. To good approximation, all ellipses are scaled versions of one another. This result is in accordance with the energy hypothesis, as the ratio of the ellipses' half-axes should always equal the ratio of the filter constants C_1 and C_2. In addition, the energy model also holds for superpositions of multiple pure tones, and even accurately predicts receptor responses to stationary noise stimuli (data not shown).

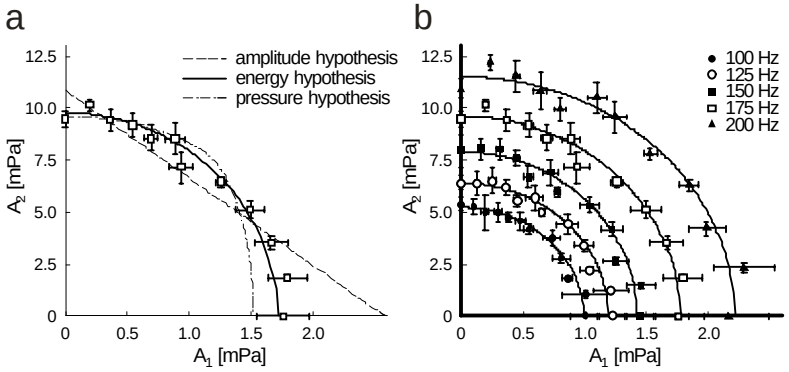

Figure 3: Identification of iso-response manifolds within a firing-rate description. For superpositions of two pure tones, measured pairs of amplitudes corresponding to a discharge rate of 175 Hz (open squares) are shown in (**a**) together with the iso-firing-rate curves for three rival hypotheses about the governing stimulus attribute. The dashed line denotes the fit of the amplitude hypothesis, the solid line the fit of the energy hypothesis, and the dash-dotted line the fit of the pressure hypothesis. While the curves for the amplitude and the pressure hypothesis deviate systematically, the ellipse obtained from the energy hypothesis corresponds well with the data. The different scales on the axes are due to the strong dependence of the sensitivity on the sound frequency and thus reflect the neuron's tuning curve. The points in (**b**) display measured pairs of amplitudes, and the solid lines are corresponding ellipses fitted to the data in accordance with the energy hypothesis. The feedback-adjusted firing rates rise from 100 to 200 Hz in steps of 25 Hz. Note that the fits agree with the data regardless of the firing rate and that ellipses for different firing rates are scaled versions of each other as predicted by the energy hypothesis.

D. Firing probabilities reveal the time scales of mechanical and electrical integration

As the experiments agree with the energy hypothesis (Eq. 9), the auditory transduction dynamics of the investigated receptors can be understood as a sequence of four elementary processes. The first two stages are given by the linear filter operation (Eq. 6) and a static square non-linearity. A second linear filter with a flat kernel accounts for the time average in (Eq. 9), and a final non-linearity describes the firing-rate encoding of the effective sound intensity J_{EH}. Within the general signal-transduction framework, this corresponds to an LNLN-cascade (Eq. 4) with quadratic f_1 and yet undetermined f_2. As the amplitudes of the sound stimuli used so far did not vary in time, temporal details of the transduction process were beyond the experiment's reach.

To uncover the time-resolved dynamics, we extended the iso-response paradigm to the occurrence probability of single spikes by using short time-dependent stimuli that trigger at most one spike (Gollisch and Herz 2004a). A series of two clicks with amplitudes A_1 and A_2 (click duration: 20 μs; inter-click interval < 1 ms) were presented repeatedly while adjusting the stimulus intensity on-line so that the probability of eliciting the single spike remained constant.

Spike probabilities were measured using temporal windows extending to 10 ms after the first stimulus – long enough to make the approach insensitive to the variability in spike timing.

If the inter-click interval is sufficiently large (Fig. 4a), the iso-reponse manifolds are nearly circular (Fig. 4c, filled circles). This is in accordance with the energy hypothesis, which predicts a dependence of the response on the sound energy, $A_1^2 + A_2^2$. For very small inter-click intervals (Fig. 4b), however, the iso-response manifolds are nearly straight lines (Fig. 4c, open squares). This indicates that the sum of both click amplitudes, $A_1 + A_2$, governs the transduction on very short time scales and corresponds to the expectation that the ear drum operates approximately as a linear filter on the sound-pressure wave (Schiolten et al. 1981).

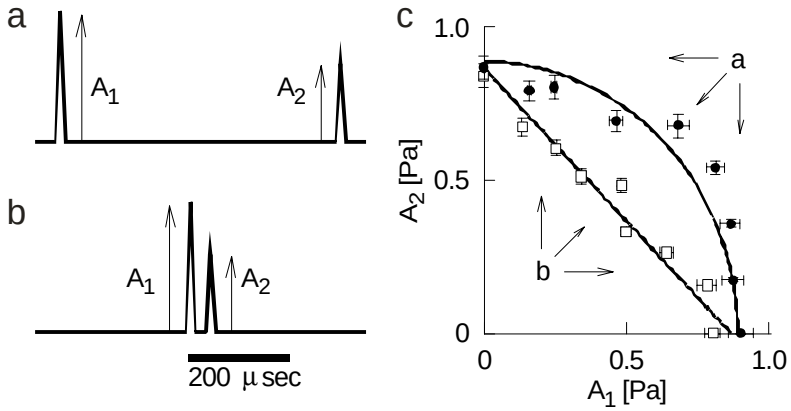

Figure 4: Identification of iso-response manifolds on short time scales. The acoustic stimuli consisted of two short clicks with amplitudes A_1 and A_2 separated by a peak-to-peak interval of (**a**) 550 µs or (**b**) 40 µs. **c**, By adjusting the overall intensity for fixed ratios of A_1 and A_2, stimulus combinations yielding spike probabilities of 70% were obtained (filled circles for stimulus **a**, open squares for stimulus **b**). All error measures display 95% confidence intervals. For the long interval, the data are well fitted by a circle whereas for the short interval, a straight line yields a good fit (solid lines in both cases). This example demonstrates that on different time scales, different stimulus variables can be relevant for the transduction dynamics, here given by the energy A^2 of the sound stimulus (**a**) or simply its amplitude A (**b**).

Together, these findings imply that the amplitude of the sound pressure and its square are the relevant stimulus variables for two sub-processes within the auditory transduction chain. Within an LNLN cascade (Eq. 4), the minimal model compatible with the data, the detailed temporal features of the two processes are reflected by the time courses of L_1 and L_2, respectively. Based on the model structure and the squaring non-linearity connecting L_1 and L_2, exact descriptions of these two functions can be derived from appropriately designed iso-response experiments (Gollisch and Herz 2004a). Measurements show that L_1 behaves like a damped oscillator with a natural oscillation frequency close to the receptor's best frequency,

which lies between a few and forty kilohertz, and a damping time constant in the range of 50 to 500 μs. Thus L_1 most likely reflects the mechanical properties of the eardrum at the receptor's attachment site. Quite differently, L_2 exhibits the exponential characteristics of a leaky integrator and captures the electrical properties of the cell membrane. With values as low as 200 μs, the membrane time constants are very short. These very fast mechanical and electrical processes suggest a tuning of the auditory receptors to the rapid amplitude modulations of behaviorally relevant stimuli such as conspecific communication calls (see also Sections V and VI).

III. Adaptation

Spike-frequency adaptation (Fig. 5a) is a common phenomenon observed in many spiking neurons. It operates on time scales that range from tens of milliseconds to several seconds – i.e. roughly one thousand times longer than the signal-transduction processes considered in the previous section – and serves many functions including gain control, high-pass filtering (Benda et al. 2004, French et al. 2001, Nelson et al. 1997), forward masking (Sobel and Tank 1994, Wang 1998), and synchronizing network activity (Crook et al. 1998, Fuhrmann et al. 2002). Prominent sources of adaptation are the activation of calcium-dependent or slow voltage-dependent potassium currents and the inactivation of fast sodium currents (for an overview, see Benda and Herz, 2003). What all these sources have in common is that they are driven by the output of the neuron, such that spikes provide a "negative feedback" by opening ionic currents to downregulate the frequency of spiking. Adaptation may, however, also contain components that are directly driven by the sensory or synaptic input in a feedforward way. In receptors, e.g., adaptation mechanisms that are independent of the neuron's output act on all steps of the signal-processing chain - the coupling, transduction and encoding of the primary sensory signal. For higher-order neurons, synaptic mechanisms as well as inhibitory inputs contribute to an input-driven adaptation effect. Different sources of spike-frequency adaptation may have different effects on the coding properties of a sensory neuron. A complete understanding of the functional aspects of adaptation therefore requires the localization of its biophyical sources and the identification of the causal relations between sensory input, neural activity, and level of adaptation.

A. Disentangling input-driven and output-driven adaptation in vivo.

To discriminate between input-driven and output-driven sources for adaptation in an auditory receptor one should use an experimental technique that allows one to measure one component independently of the other. This can be done by tuning the intensities for different sound frequencies in such a way that the neuron's steady-state firing rate is the same. In this situation, the level of output-driven adaptation must also be the same. Sudden switches between those sound stimuli then reveal input-driven components of adaptation, as these will approach a new equilibrium value after such a switch. This results in transient deflections of the firing rate, which can be observed in electrophysiological recordings of the spiking activity. The deflections are well-suited to characterize prominent features of the input-driven adaptation component, such as its strength, its time constants, and its correlation with different stimulus and activity parameters (Gollisch and Herz 2004b). Similar insight is to be expected for other neurons exhibiting a mixture of input-driven and output-driven adaptation.

B. Phenomenological model for output-driven adaptation

Within most cellular types of output-driven adaptation spikes directly or indirectly activate ionic currents. Such adaptation currents act subtractively on the time dependent input current *I(t)* since ionic currents are flowing in parallel through the cell membrane. As the dynamics of adaptation are usually slow with respect to the time scale of interspike intervals, we approximate adaptation currents by their temporal average over an interspike interval. To model the firing frequency output of an adapting neuron we describe its spiking dynamics by the onset *f-I* curve, $f_0(I)$, which we obtain by measuring the onset response f_0 of the unadapted neuron (Fig. 5a,b). This approach allows one to derive a universal model for the firing frequency (Benda & Herz, 2003) that is independent of the specific biophysical mechanism of the spike-driven adaptation process. In its simplest form, the model reads

$$\begin{aligned} f(t) &= f_0(I(t) - A(t)) \\ t\,\frac{d}{dt}A(t) &= A_{max}(f(t)) - A(t). \end{aligned} \qquad (11)$$

The adaptation strength A denotes the averaged adaptation current. Depending on the size of A the onset *f-I* curve is shifted to higher input currents (Fig. 5b). The maximum adaptation strength $A_{\max}$ depends on the output firing frequency f and is approached on a time scale governed by the adaptation time constant τ. $A_{\max}$ is given by the difference between the onset *f-I* curve, $f_0(I)$, and

the steady-state *f-I* curve $f_\infty(I)$. The latter is obtained by measuring the steady-state firing frequencies f_∞ for different constant input currents I.

Often, the time course of adaptation is estimated from an exponential fit to the firing-frequency response to current steps (Fig. 5a). The "effective" adaptation time constant t_{eff} obtained from such a measurement differs in general from the adaptation time constant t. Both parameters are approximately related via the slopes $f'_0(I)$ and $f'_\infty(I)$ of the *f-I* curves, or

$$t \approx t_{\text{eff}} \frac{f'_0(I)}{f'_\infty(I)}. \tag{12}$$

It follows that the observed effective time constant (10-75 ms in the auditory receptors of *L. migratoria*) is smaller than the true time constant of the adaptation mechanism.

The adaptation model (11) is completely specified by the onset *f-I* curve, the steady-state *f-I* curve and the adaptation time constant. For a specific neuron, all three quantities can be easily obtained from the firing-frequency responses to step-current inputs (Fig. 5a,b). Once they are determined, the response of that neuron to different stimuli, such as band-limited white noise stimuli or conspecific grasshopper songs, can be predicted (Fig. 5c). As the figure demonstrates the model is in good agreement with the measurements showing that on the level of firing frequencies a single adapation dynamics is sufficient to capture the data. Without the adaptation process the data cannot be explained. In addition the time course of the adaptation strength and thus the position of the current *f-I* curve is obtained. This provides a tool for analyzing how adaptation shifts the neuron's *f-I* curve according to the stimulus.

The approaches presented so far either treat neural responses on the level of mean firing rates or do not describe spike generation at all. The following section presents a stochastic model for the generation of action potential that also captures the observed response variability.

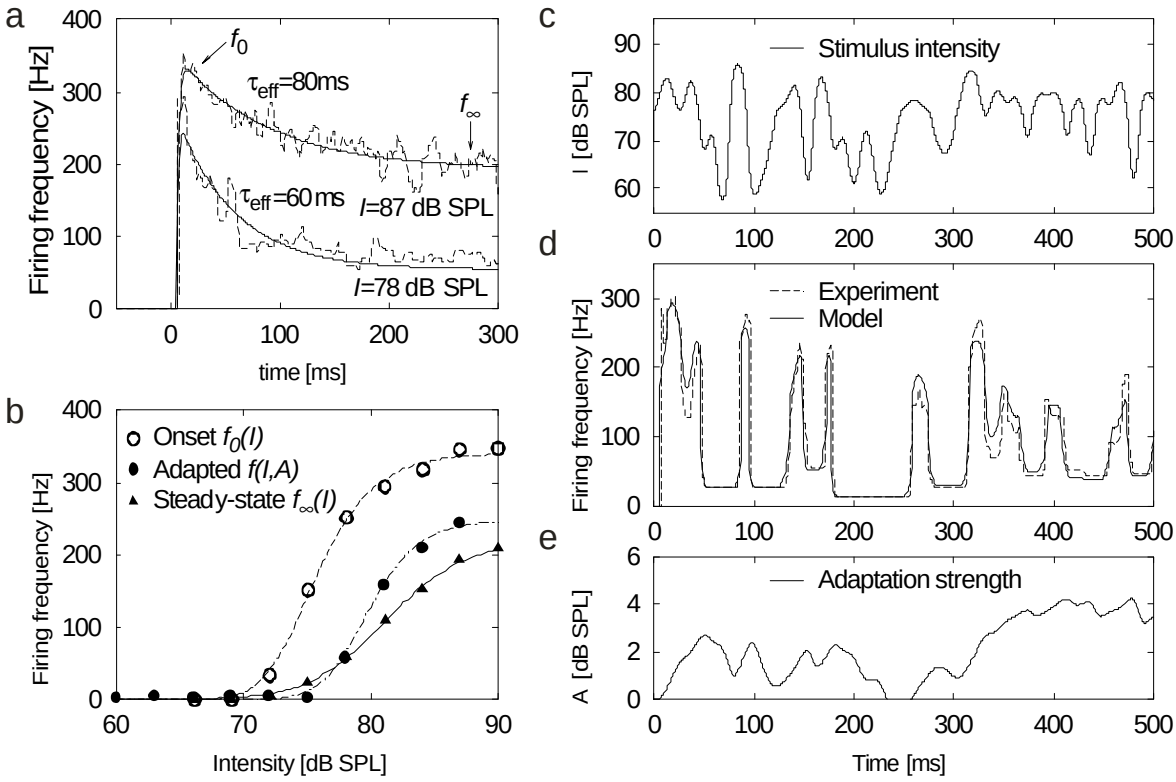

Figure 5: Spike-frequency adaptation. (**a**) Step responses (dashed lines) of a locust auditory receptor to two constant sound stimuli switched on at time zero; sound intensities are indicated by the labels. For simplicity we assume that the intensity measured in dB SPL is proportional to the input current driving the spike generator. (**b**) Onset *f-I* curve (f_0, open circles) and steady-state *f-I* curve ($f_\infty(I)$, filled circles) determined from the measurements shown in a. The triangles display the *f-I* curve for a fixed level of adaptation (*A=const*) corresponding to $I = 78$ dB. (**c**-**e**) Performance of the model. Neural response (**d**, *dashed line*) to the amplitude-modulated stimulus in **c** (carrier frequency: 5 kHz, modulation: broad-band white noise with cut-off frequency 50 Hz). The *f-I* curves from **b** and an adaptation constant $t = 150$ ms resulting from Eq. (12) were used to calculate the firing frequency from the adaptation model Eq. (11) (**d**, *solid line*). The prediction closely follows the experimental data. The firing frequency was computed as the averaged inverse interspike-interval at any time bin. (**e**) Adaptation strength A. Note that a 5 dB shift of the $f_0(I)$-curve may alter the firing rate by more than 100 Hz (see **b**).

IV. Spike generation and trial-to-trial response variability

Even under well-controlled laboratory conditions, repeated presentations of the same stimulus lead to different neural responses. This trial-to-trial variability is caused by stochastic components of the single-cell dynamics, such as ion-channel noise, and results in spike-time jitter and additional or missing spikes. Trial-to-trial variability limits the amount of information

that is transmitted by a sensory neuron about the external stimulus. Reliable accounts of this phenomenon are a prerequisite for the proper interpretation of neural dynamics and coding principles. Models that accurately describe neural variability over a wide range of stimulation and response patterns are thus highly desirable, especially if they can explain the variability in terms of basic neural observables and parameters such as firing rate and refractory period.

A. Mathematical framework for stochastic spike generation: Renewal processes

The simplest mathematical description of spike-time variability treats a neuron as a random point process, i.e. a process that stochastically generates a series of events (spikes) in time. In general, the probability of generating a spike at time t could depend on a variety of factors. Modeling is greatly facilitated if one can assume that only an "effective" stimulus strength $q(t)$ at that very moment and refractoriness caused by the last spike at t_{last} determine the generation of an action potential. For time-independent stimuli, i.e. $q(t) = q$, spike generation becomes a "renewal process" (Cox 1962), so that inter-spike intervals (ISIs) are independent. For time-dependent stimuli, the latter is no longer true and one rather speaks of a "modulated renewal process" (Reich et al. 1998).

The influence of refractoriness on the generation of an action potential depends on the interval Δ that has passed since the last spike was elicited at time $t_{last} = t - \Delta$. Let us denote this memory term, or "recovery function" (Berry and Meister 1998; for a comparison with the "hazard function" see Johnson 1996, or Gerstner and Kistler 2002) by $w(\Delta)$. Note that the memory is independent of the strength of the effective stimulus $q(t)$. Mathematically, spike generation is thus described by a probability per unit time (the "hazard") $r(t|t_{last})$, that is conditional on the last spike occurring at time t_{last}:

$$r(t|t_{last}) = q(t) \cdot w(t - t_{last}) \tag{13}$$

Values of the recovery function w range between zero and unity. Within the absolute refractory period t_a even arbitrarily large stimuli cannot elicit a spike. To capture this property, $w(\Delta)$ vanishes for $0 \le \Delta \le t_a$. It then rises monotonically to account for the relative refractory period during which the neuron relaxes back to its normal level of excitability (Fig. 6a, lower panel).

For a renewal process, it is possible to compute the recovery function directly from the ISI distribution $P_{ISI}(\Delta)$ obtained under constant stimulation ($q(t) = q$), as has been discussed in the literature (Johnson 1996, Berry and Meister 1998, Gerstner and Kistler 2002):

$$w(\Delta) = \frac{1}{q} \cdot \frac{P_{ISI}(\Delta)}{1 - \int_0^\Delta dt\ P_{ISI}(t)} \tag{14}$$

Eq. (14) can be inverted to calculate the ISI distribution from the recovery function for given stimulus strength:

$$P_{ISI}(\Delta) = q \cdot w(\Delta) \cdot \exp\left[-q \cdot \int_0^\Delta dt\ w(t) \right] \tag{15}$$

To use this model framework, one first has to investigate whether the spike trains of a specific neuron are indeed consistent with a renewal process (independent ISIs) for constant stimulation. Second, the effective input q has to be determined as a function of the stimulus intensity – such that the observed and the predicted mean firing rates match (for details, see Schaette et al. 2004).

B. Summary of results (Locusta migratoria)

The variability of receptor spike trains depends strongly on the evoked firing rate. Low firing rates (below 50 Hz) are accompanied by high variability, as reflected in a broad ISI distribution and a Coefficient of Variation (CV) near unity. The Coefficient of Variation is defined as the ratio between the standard deviation of the ISI distribution and the mean ISI. It is unity for Poisson spike trains. For higher firing rates, the ISI distribution becomes sharper (Fig. 6a,b), and the CV decreases down to values of 0.2 (very regular spike trains) for maximum firing rates of around 300 Hz. A similar dependence is observed for dynamic stimuli, but the situation is more complicated as there is also an influence of the temporal structure and modulation depth of the stimulus amplitude modulation.

Extensive data from locust auditory receptors suggest that the renewal assumption is well justified for this system once adaptation effects are taken care off. Each neuron can then be characterized by one unique recovery function $w(D)$. This minimal description provides accurate predictions for ISI distributions caused by input intensities over the entire range of firing-rate responses (see Fig. 6a,b). To cover responses to dynamic stimuli, the renewal process is driven by a time-dependent effective stimulus strength $q(t)$. The model captures the spike-count variability and salient features of the fine temporal structure in response to dynamic stimuli,

although the recovery functions were always calculated from ISI distributions obtained under constant stimulation (Fig. 6c).

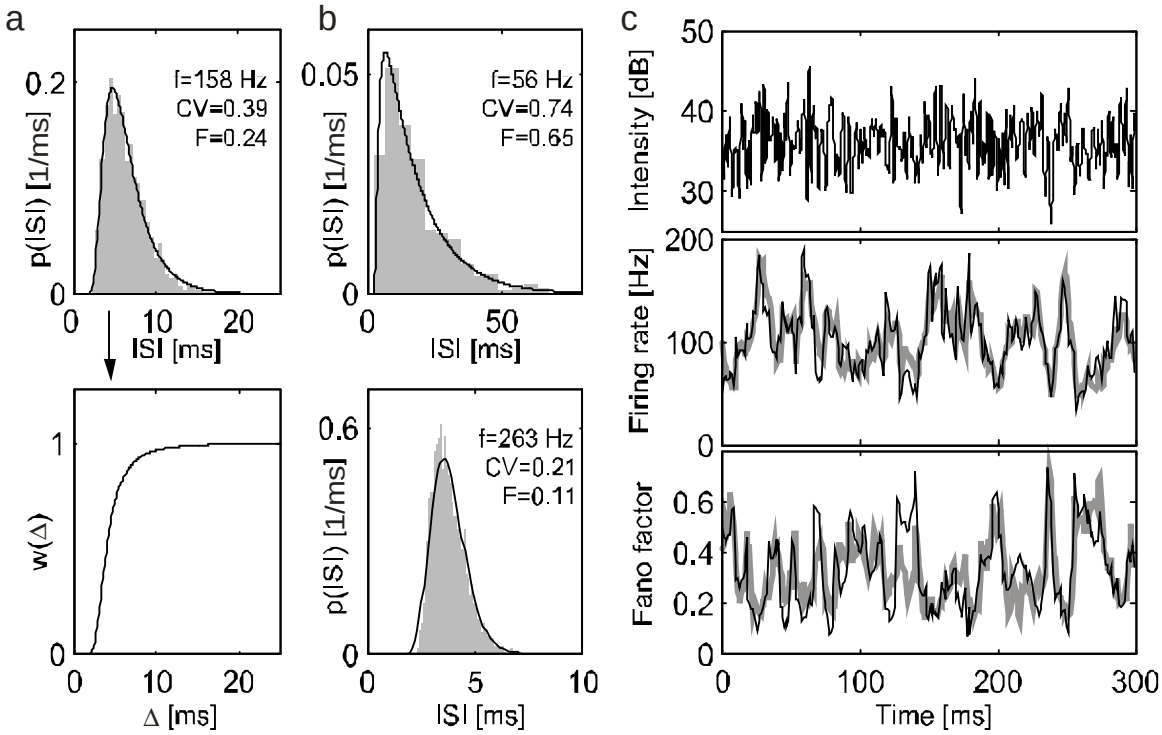

Figure 6: Response variability for constant and time-varying stimuli. (**a**) Upper panel: interspike-interval (ISI) distribution (gray histogram) of a locust auditory receptor resulting from a constant stimulus with intermediate sound intensity (evoked firing rate: 158 Hz). This distribution was used to calculate the recovery function (see Eq. 19) depicted in the lower panel. (**b**) Two examples of ISI distributions from a low (56 Hz, upper panel) and a very high firing rate (253 Hz, lower panel), recorded from the same cell as in (**a**) but at different intensity levels. The black lines are predictions from the recovery function depicted in (**a**) that were obtained by adjusting the stimulus strength q such that experimental and model firing rates matched. In addition, CV values (the ratio between the standard deviation of the ISI distribution and the mean ISI) and Fano factors (the ratio between the spike-count variance and its average, both calculated for a certain counting time, here 10 ms) are shown. With decreasing firing rates, the CV value and the Fano factor increase because refractoriness plays a smaller role at low firing rates. (**c**) Upper panel: amplitude modulation of a Gaussian white-noise sound stimulus (carrier frequency: 4 kHz). Middle panel: time course of the firing rate determined with a 10-ms sliding window. As in the lower panel of this figure, thick gray lines depict experimental data and thin black lines denote the model result. Lower panel: spike-count variability as measured by the Fano factor. As for constant stimuli, spike-train variability is anticorrelated with the firing rate on short time scales; high-activity episodes come with low variability (high precision), and vice versa.

These results demonstrate that key ingredients of stochastic responses are faithfully captured by the renewal model and that a clear mathematical separation between external stimulus and stimulus-independent cell dynamics is possible. Furthermore, spike variability can be modeled using the same stochastic process for constant and strongly time-varying stimuli.

This indicates that there is no principal difference between the dynamics underlying responses to both stimulus classes. Differences in the encoding quality (see Section V) may simply arise from the specific usage of the neuron's dynamic range by the particular stimulus.

Including a detailed adaptation model, such as the one presented in Section III, would enhance the model's predictive power for stimuli varying on multiple time scales, but at the cost of substantially longer recordings needed to calibrate the additional model parameters. The degree of realism could be even further enhanced by incorporating the signal-transduction model of Section II. Together, these elements might yield a biophysically motivated and simple, yet highly accurate description of stimulus encoding.

V. Decoding spike trains

As shown in Sections II to IV, mathematical models of single-cell dynamics provide valuable insight about the response dynamics of auditory receptors. We now go one step further and investigate their signal-processing capabilities using methods from systems analysis and information theory (see also Bialek et al. 1992, Rieke et al. 1997, Borst and Theunissen 1999). These methods provide a quantitative means to estimate the information contained in a spike train that was evoked by a given stimulus. In particular, one can ask whether auditory receptors encode a large range of acoustic stimuli or whether they are specifically tuned to behaviorally relevant features, such as the temporal structure of a grasshopper calling song (Machens et al. 2001).

A. Stimulus reconstruction

In what follows, we focus on linear stimulus reconstruction, i.e., the attempt to recover the original stimulus s(t) from a spike train y(t), as illustrated in Fig. 7. To do so, each spike is replaced by a filter function *h(t)*, resulting in the signal estimate $s_{\text{est}}(t)$,

$$s_{\text{est}}(t) = h_0 + \int_0^T dt\ h_1(t)\, y(t-t), \tag{16}$$

where h_0 is the mean signal level in the absence of spiking. The parameters h_0 and $h(t)$ are determined by minimizing the mean-square error $\left\langle n_{\text{mse}}(t)^2 \right\rangle$ where the angular brackets denote a

time average over the section of the experiment used for parameter estimation and $n_{\mathrm{mse}}(t)$ is the time-dependent reconstruction error $n_{\mathrm{mse}}(t) = s(t) - s_{\mathrm{est}}(t)$.

Let us note that the use of stimulus reconstruction methods does not imply that we think that the auditory system is trying to reconstruct acoustic stimuli from spike trains. Rather, by comparing the reconstruction quality for different stimulus ensembles, we seek to find out which characteristics of acoustic signals are encoded faithfully and which features are discarded.

Auditory receptor neurons of grasshoppers are sensitive to the amplitude modulation (AM) of broad-band sound-pressure waves that exceed a certain threshold. Below this threshold, the cells remain silent. Therefore, the appropriate stimulus $s(t)$ for applying reconstruction techniques (Fig. 7) is *not* the original sound-pressure wave $w(t)$ but rather that part of the AM signal that lies in the sound intensity range covered by the particular receptor. Within the stimulus reconstruction algorithm, therefore, the AM signal should be half-wave rectified at the threshold of each cell and then used for the stimulus reconstruction algorithm. The thresholded AM signal will from now on simply be referred to as the "signal".

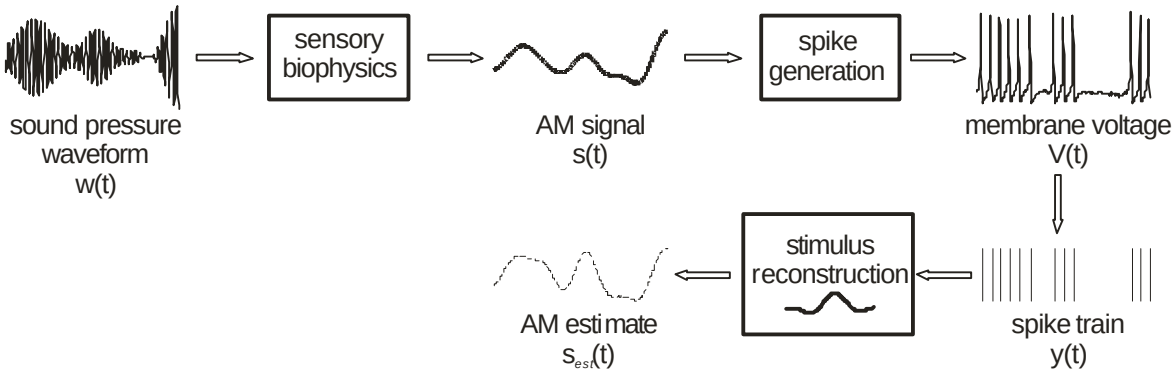

Figure 7: Stimulus preprocessing and reconstruction. The mechanics of the receiver's ear extract the slow amplitude modulation s(t) of a rapidly oscillating sound-pressure wave w(t). Auditory receptor neurons then encode s(t) into the membrane voltage V(t). As a first step of the stimulus reconstruction, the spike train y(t) is extracted from the voltage trace. Within linear reconstruction, each spike is replaced by an optimal filter function to yield $s_{est}(t)$, the estimate of s(t). As shown by this example, stimulus reconstruction does not aim at recovering the original, complete physical stimulus w(t) but instead requires the identification of a representation of the stimulus that is relevant for the animal, in the present case the AM signal s(t).

A nonlinear relationship between the AM signal and the firing rate could require higher-order reconstruction filters for adequate signal reconstruction. Such filters seem to suggest relational codes, i.e., coding schemes that involve higher orders of the spike-train statistics, as in interspike-interval-based codes. Since the firing-rate responses of auditory receptors of

grasshoppers are approximately threshold-linear if amplitude modulations are measured on a logarithmic scale (Römer 1976, Stumpner and Ronacher 1991, Ronacher und Krahe 2000) this potentially misleading interpretation of higher-order kernels can be obviated by transforming $s(t)$ and $s_{\text{est}}(t)$ into the decibel scale. In a more rigorous approach, one could also first measure the *f-I* curve of the neuron under study and use this information to rescale the input signal such that the firing rate depends linearly on the transformed input.

Spike-frequencey adaptation (Section III) can be described, as a rule, by higher-order or time-dependent reconstruction filters, but their estimation requires enourmous amounts of data. In most studies, such complications are circumvented by studying the responses of fully adapted neurons, discarding their responses during the initial phase.

The reconstruction error $n_{\text{mse}}(t)$ can be separated into random and systematic components. Systematic errors occur if one attempts to reconstruct a signal $s(t)$ that is incompatible with the signal the neuron actually encodes. For instance, if only a low-pass-filtered version of the signal is encoded, any attempts to reconstruct higher frequencies have to fail. Systematic errors can be corrected for by introducing a frequency-dependent gain $g(f)$ such that $s_{\text{est}}(f) = g(f)[s(f) + n_{\text{eff}}(f)]$ where $n_{\text{eff}}(f)$ denotes the random errors or "effective noise", as referred to the input (Theunissen et al. 1996, Rieke et al. 1997).

Given the effective noise $n_{\text{eff}}(f)$, the reconstruction success in each frequency band can be measured by the frequency-resolved *signal-to-noise ratio* (SNR),

$$SNR(f) = \frac{S(f)}{N_{\text{eff}}(f)} = \frac{s(f)s^*(f)}{n_{\text{eff}}(f)n_{\text{eff}}^*(f)}, \tag{17}$$

where $S(f)$ and $N_{\text{eff}}(f)$ are the power spectral densities of the signal and the effective noise, respectively. A high *SNR(f)* indicates an accurate reconstruction of the particular frequency component, while a *SNR(f)* of zero implies chance level. The frequency-resolved SNR therefore allows one to assess which stimulus components are best decoded by signal reconstruction. Reconstruction of signals with high bandwidth serves to estimate the cut-off frequency of the system; this cut-off will be unveiled as the frequency where the signal-to-noise ratio approaches zero. To measure the overall reconstruction success, the ratio of the total power of signal and noise will be used, $SNR = \int df\, S(f) / \int df\, N_{eff}(f)$.

B. Information theory

The *mutual information rate* R_{info} quantifies how many bits of information about the signal $s(t)$ are carried by a spike train per second. For example, a value of $R_{\text{info}} = 1\,\text{bit/sec}$ means that the uncertainty about the stimulus can be halved every second by reading the corresponding spike train. Note that the mutual information rate can be large even if the signal is only poorly reconstructed as might occur for stimuli with high bandwidth. If $s(t)$ is a Gaussian random signal, a lower bound on R_{info} (Rieke et al. 1997) is given by

$$R_{\text{info}} \geq \int_0^\infty df \, \log_2[1+\text{SNR}(f)]. \tag{18}$$

Given a time resolution Δt, the efficiency of a neuron at transmitting information can be measured by comparing the estimated mutual information rate $R_{\text{info}}(\Delta t)$ with the information-theoretic limit $R_{\max}(\Delta t)$ which is reached if the spike train is maximally disordered, i.e., Poisson (see Rieke et al. 1995 for details). The *coding efficiency* $e(\Delta t)$ is then defined as

$$e(\Delta t) = \frac{R_{\text{info}}(\Delta t)}{R_{\max}(\Delta t)}, \tag{19}$$

which takes on values between zero and one. While $R_{\max}(\Delta t)$ tends to infinity for $\Delta t \to 0$, this is not the case for $R_{\text{info}}(\Delta t)$, which will instead achieve the value R_{info} given in eq. (18). To yield non-trivial results, the coding efficiency, therefore, has to be evaluated at a finite time resolution that reflects spike-timing variability due to intrinsic noise sources. This time resolution can be estimated by cross-correlation analysis and is about $\Delta t \approx 1$ msec for grasshopper auditory receptor cells.

C. Summary of results

To identify essential features of grasshopper songs and their neural representations, artificial stimuli were designed to vary the most salient statistical properties of grasshopper sounds (see Fig. 1). The communication signals alternate between noise bursts and pauses, resulting in a characteristic double-peak distribution of sound amplitudes (see, e.g., Machens et al. 2001). To study the importance of this structural aspect, two different classes of stimuli were generated.

The first class consisted of random stimuli that have the same amplitude distribution as a typical grasshopper song and thus imitate the gap-infiltrated structure of these songs. Due to their

modulation depth of ~24 dB, these stimuli are called large-modulation-depth (LMD) stimuli. Within the second class, stimuli have a Gaussian amplitude distribution with a modulation depth of 10 dB and are called small-modulation-depth (SMD) stimuli. These random stimuli simulated the combined sound pattern of a group of five to ten grasshoppers singing simultaneously such that the song pauses of individual songs are filled by the other songs.

Since the shortest behaviorally relevant time scales of the AM signals are around one to two milliseconds (von Helversen 1972), frequency components of at least 250 to 500 Hz are required in the random stimuli. To analyze the neural representation at these short time scales, LMD and SMD stimuli were designed with piece-wise flat spectral characteristics and cut-off frequencies of up to 800 Hz. Additionally, to test whether the specific mix of frequency components found in natural songs might be of importance, one of the LMD stimuli exhibited a song-like spectrum (SLS). In all experiments, the amplitude distribution for each stimulus was kept constant by fixing the integrated AM signal power. A larger bandwidth, therefore, corresponds to a lower power spectral density.

Reconstructions from the recorded spike trains of single locust receptor neurons demonstrate that even single cells are capable of encoding amplitude modulations with signal-to-noise ratios of up to 10:1. In this regime, the original stimulus can be faithfully “read” (Bialek et al. 1991) from the spike train (Fig. 8). Most importantly, the success of the linear reconstruction method suggests that the studied auditory receptors encode acoustic stimuli as a strongly time-varying firing rate; we did not find any hints for coding schemes that involve higher orders of the spike-train statistics. The data also show that sounds with large modulation depth are encoded with higher signal-to-noise ratios, information rates (up to 180 bits/s) and coding efficiencies (up to 0.4) than stimuli with small modulation depth. While LMD stimuli exhibit greater raw amplitude variations, the encoding of SMD stimuli is still poorer even when the AM signal power above threshold is identical in the two classes of stimuli. Matching the spectral properties of the songs as in the SLS stimulus, on the other hand, does not increase signal-to-noise ratios.

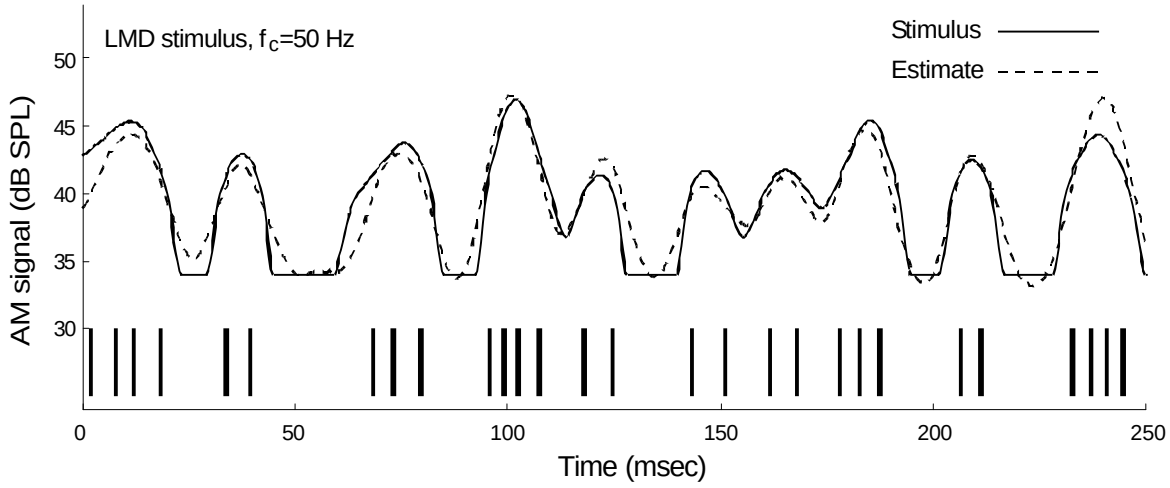

Figure 8: "Reading" a neural spike train. Shown is the linear reconstruction of an LMD signal with 50 Hz cut-off frequency from the responses of a single locust auditory receptor. The stimulus was thresholded at 34 dB, the threshold of this particular neuron. The close correspondence between stimulus and linear stimulus estimate demonstrates that the sound signal can be faithfully inferred from the time-resolved firing rate; higher-order correlations of the spike train are not needed to reconstruct the stimulus. However, both the timing of individual spikes as well as their local rate contribute to the high reconstruction quality.

Spikes are triggered with high reliability and temporal precision when the sound intensity rapidly passes the firing threshold — as occurs at the beginning of a syllable of the grasshopper calling song (see Fig. 1). This phenomenon emphasizes the paramount importance of gaps and pauses for the recognition of acoustic stimuli, as the precision in spike timing leads to a faithful representation of the supra-threshold sound pattern. Grasshoppers seem to utilize this effect in the design of their songs which consist of repeated patterns of sound and (relative) quiet.

Highest rates for the information transfer of single cells are observed for stimuli with large modulation depth and a cut-off frequency of 200 Hz. This finding should be compared with behavioral studies in which various artificial auditory stimuli were presented that were generated by filtering the Fourier components of model songs with regular or irregular syllable composition (von Helversen and von Helversen, 1998). These studies demonstrate that depending on the original syllable structure, Fourier components between 150 and 300 Hz are required by *Ch. biguttulus* females to reliably detect gap signals. Together, these two results suggest that the response properties of single receptor neurons are optimized for features of the acoustic environment that are of prime importance for behavioral decisions.

VI. Song discrimination

The experimental results summarized in the previous section clearly indicate that single auditory receptors are well suited to transmit information about conspecific communication signals. This raises the question whether they are also capable to discriminate between individual songs from the same species. This would provide a basis for mate preference and sexual selection on the single-cell level.

A. Spike-train metrics and discrimination matrices

Successful discrimination requires that spike trains elicited by repeated presentations of the same song be more similar to each other than spike trains elicited by different songs. To compute the similarity between two spike trains, each spike was first replaced by an alpha function that mimics the time course of an excitatory postsynaptic potential (EPSP) in a hypothetical downstream neuron of the grasshopper's auditory system. The distance between the two spike trains was then defined as the mean square distance between their EPSP convolved traces (van Rossum 2001). By varying the width τ of the EPSP function effects of the temporal resolution can be studied: if τ is large, only differences in the average spike rate contribute to the distance measure; if τ is small, even small differences in spike timing matter.

To discriminate the songs based on their evoked responses, one spike train was arbitrarily chosen as a template for each of the eight presented song. The remaining spike trains were classified by assigning each spike train to the closest of the eight templates. Averages were then computed by permuting all possible template choices, yielding classification matrices (Fig. 9a). The diagonal elements of the matrices correspond to correctly classified, the off-diagonal elements to misclassified spike trains. Replacing the alpha function by exponential or Gaussian kernels of the same width led to similar results, as did the use of cost-based metrics (Victor and Purpura, 1997). Using more sophisticated supervised and unsupervised clustering algorithms to assign spike trains to the songs that elicited them provided, at best, a marginal improvement in overall discrimination performance.

B. Summary of results

Too broad or too narrow an EPSP filter will impair the discrimination of songs, leading to many off-diagonal entries in the corresponding matrices. However, for an EPSP width of $\tau \approx 5$ ms,

most spike trains are correctly assigned. This time scale also corresponds to the shortest unique song features that auditory receptors accurately encode. Signal variations on a scale of a few milliseconds are thus important for song evaluation; downstream neurons with much longer EPSP's will not be suited to process the information relevant for discrimination.

To analyze how the discrimination capability unfolds in time, we varied the time window T in which spikes were used to discriminate, extending T from 0 to 1000 msec. Calculating the percentage of correctly classified spike trains for every (τ,T)-pair yields the contour plot shown in Fig. 9b. The eight tested songs can be discriminated with high reliability (> 80%) after a mere four or five syllables (T = 400 - 500 ms).

The question arises as to which song cues are relevant for evaluating potential mates. For instance, the songs differ in their syllable length (60 - 120 ms) and broadband carrier frequency content (Fig. 1b). However, such variations are unlikely to play a role in the females' assessment of males: syllable duration is greatly affected by body temperature and can vary more than two-fold when the male moves from sun to shadow; similarly, the frequency content depends strongly on the distance of the listening female grasshopper to the singer, with higher frequencies decaying more quickly with distance. Differences in spectral shape result in different firing rates, thereby facilitating discrimination. In fact, close inspection of Figure 9b reveals that using a moving average of the neuron's firing rate alone (filter width τ = 1000 ms), almost 40% of the spike trains can be assigned to the correct song.

To eliminate any spurious discrimination cues that simply reflect differences in syllable length or broadband frequency content, we complicated the discrimination task and generated a set of rescaled songs that each had a common syllable length of 100 ms and the same carrier spectrum. The overall stimulus power was adjusted such that the last second of each song had the same mean sound intensity. The detailed structure of amplitude modulations within a syllable, which differed from song to song, remained as a discrimination cue.

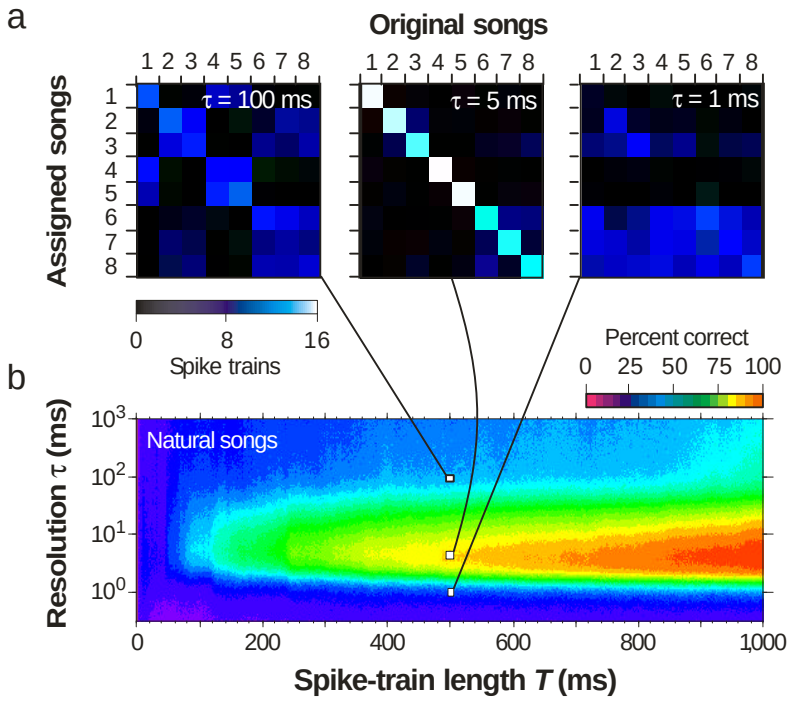

Figure 9: Song discrimination. (**a**) Discrimination matrices computed at T = 500 ms for τ = 1 ms, τ = 5 ms, and τ = 100 ms. T denotes the spike-train length used for the data analysis and τ represents the time scale on which time-varying firing rates are resolved. Rows and columns of the matrices denote the consecutively numbered songs. The matrix entries N(a,b) give the number of spike trains that were originally elicited by song # b (horizontal axis) and in turn classified as belonging to song # a (vertical axis). A spike train is correctly classified if a = b. Perfect discrimination is achieved if all spike trains are assigned onto the diagonal elements. (**b**) Discrimination performance as a function of the response duration T and resolution τ . Within the red area, at least 95% of the spike trains are classified correctly; the half-width of the EPSP-like functions used to filter the spike trains before discriminating the songs are then within the range of τ = 3-10 ms.

Surprisingly, the ability to differentiate these artificial songs does not decrease substantially (data not shown, but see Machens et al. 2003). A few hundred milliseconds, or 40-50 spikes, suffice to distinguish more than 80% of the stimuli; indeed, two (instead of eight) songs can often be discriminated perfectly after a few tens of milliseconds, or four to five spikes. This demonstrates that the remarkable temporal precision of the receptor neurons allows one to recover even slight differences between the rescaled songs from the neural response.

With a population of 40-50 receptors per ear, even more information is available about the fine details of the amplitude modulation pattern. To discriminate conspecific from heterospecific signals, however, such a high precision is superfluous. Instead the acquired information could be used to evaluate male singers. In this context, it should be noted that a female grasshopper does not need to recognize an individual male; it is sufficient to respond to those males whose signals indicate a good genetic constitution.

Grasshoppers produce their songs by a rhythmic movement of the hindlegs against the forewings. During a syllable, each hindleg moves up and down three to four times. Each of these strokes lasts for about 7-12 ms and is succeeded by a short gap (~ 2 ms) at the movement reversal point. In a healthy male these brief gaps are camouflaged by systematic phase shifts between the movement of both legs, leading to a fairly smooth total amplitude modulation pattern within a syllable (von Helversen and von Helversen 1997). However, should the sound produced by one of the hindlegs be weaker, the modulation pattern of the other side will dominate. Males with developmental faults often damage a hindleg or cripple a forewing when moulting. Their songs exhibit 2-ms gaps within each syllable. Females refuse to mate with such impaired singers (Kriegbaum 1989). Poor genetic quality may also result in an inadequate co-ordination between the two hemisegmental central pattern generators that are responsible for the stridulation movement, creating a deviant amplitude modulation pattern (Ronacher 1989). It may thus be speculated that it would be advantageous for a female grasshopper to carefully "inspect" the amplitude modulations of the song of a potential mate. Our data demonstrate that the information required to do so is faithfully encoded in the auditory periphery.

VII. Conclusions and Outlook

As we have tried to demonstrate in this chapter, several fundamental questions about the dynamics and signal processing capabilities of auditory receptors can be quantitively answered by closely combining experiment, modelling, and theory. For example, we have been able to show that sound energy, but not sound pressure, governs the auditory transduction process. This finding sets tight constraints for any future biophysical transduction model. The time scales of the mechanical and electrical integration have been revealed with high accuracy from recording neural output activity more than one millimeter away from the relevant processes, thus leaving the vulnerable mechanical structures of the ear unimpaired. The measured electrical time constants are well below one millisecond and reflect the high demand for temporal resolution in the auditory periphery. Traditional correlation techniques are limited by spike-time variability and thus cannot capture the dynamics on time scales as short as in our example.

Together with the low internal noise level, short integration times also explain the large information rates of the receptor neurons and their astounding discrimination capability. It thus appears that various microscopic biophysical parameters are tuned to optimally encode

conspecific communication signals. Similarly, stimulus reconstructions from spike trains evoked by grasshopper songs suggest that even the time scale of adaptation may be tuned such that it balances the rising overall song intensity (see Fig. 1) to provide a fairly invariant sound representation for down-stream neurons (Machens et al 2001).

These findings indicate that the dynamics and coding strategies of grasshopper auditory receptors are well adjusted to important behavioral tasks. Whether the auditory receptors have evolved to process grasshopper songs, or selection pressure has forced the song patterns to match the properties of receptors, or both, is a currently unanswered question. One may therefore ask whether super-optimal stimuli exist that receptor spike trains can encode even better than conspecific calling songs. To investigate this question, we have started to systematically search for such stimuli. Instead of exploring the huge space of natural and naturalistic stimuli, we use information from the responses of a given receptor neuron to identify those regions in stimulus space that are encoded best. An information-theoretic foundation (Machens 2002), reliable online analysis, and automatic feedback to the stimulus generation are central aspects of this approach.

From a conceptual point of view, the underlying strategy is similar to the iso-response method (Section II) whose key ingredient is a systematic exploration of stimuli causing the *same* output. Investigating such regions implies a radical change of the traditional perspective regarding neural input-output relations. Instead of asking what output is produced by a given input, one seeks to identify input ensembles that are associated with a fixed output. Fast online techniques and feedback-driven stimulation are again crucial ingredients of this approach, too. With growing computer power already integrated into the modern lab bench, both types of investigations will soon become suitable for various other sensory systems.

These methodological observations suggest that future advances in sensory neurobiology will strongly benefit from close interactions between theoreticians and experimentalists. In fact, the requirements of an approach grounded in modern data-analysis techniques need to be considered long before the first experiment is carried out, and necessitates a tight integration of theoreticians into the experimental design process. It is our strong belief that many of the currently unsolved questions regarding the dynamics and coding strategies of auditory and other sensory systems will be solved through such interactions.

Computer-aided searches for invariant stimulus regions may also help us to better understand the astounding computations that follow the first steps of transducing and encoding the external stimulus discussed in the present contribution. Female grasshoppers, for example, seem to use the ratio between syllable length and pause length for recognizing male songs as "conspecific" (von Helversen and von Helversen 1997). They even tolerate a global time-warp of more than 100% if the syllable-to-pause ratio remains the same. This implies that small insect neural systems are capable of performing non-trivial division operations in the temporal domain. They also solve similarly complex computations when localizing a sound source. Discovering the physical principals and biological mechanisms underlying these behaviors remains a fascinating challenge.

Acknowledgements: Astrid Franz, Olga Kolesnikova, Rüdiger Krahe, Petra Prinz, and in particular Bernd Ronacher were strongly involved in various projects summarized in this review. We would like to thank them for their specific contributions as well as for many fruitful discussions. Figs. 4, 8, and 9 are reproduced with permission; copyright 2001, 2002 Society for Neuroscience.

References

Benda J, Herz AVM (2003) A universal model for spike-frequency adaptation. *Neural Comp.* **15**, 2523-2564.

Benda J, Longtin A, Maler L (2004) Rapid spike-frequency adaptation enhances the detection of transient communication signals. Preprint.

Berry MJ, Meister M (1998) Refractoriness and neural precision. *J Neurosci* **18**, 2200-2211.

Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. *Science* **252**:1854-57.

Borst A, Theunissen FE (1999) Information theory and neural coding. *Nat Neurosci* **2**, 947-957.

Bradbury JW, Vehrenkamp SL (1998) Principles of animal communication. Sunderland, MA: Sinauer.

Cox DR (1962) *Renewal Theory*. Methuen, London.

Crook SM, Ermentrout GB, Bower JM (1998) Spike frequency adaptation affects the synchronization properties of networks of cortical oscillators. *Neural Comput.* **10**, 837-854.

Evans EF, in *Handbook of Sensory Physiology, vol. 5.2, Auditory Systems*, eds. W.D. Keidel and W.D. Neff, 1–108, Springer, Berlin (1975).

French AS, Höger U, Sekizawa S-I, Torkkeli P H (2001) Frequency response functions and information capacities of paired spider mechanoreceptor neurons. *Biol. Cybern.* **85**, 293–300.

Fuhrmann G, Markram H, Tsodyks M. (2002) Spike frequency adaptation and neocortical rhythms. *J Neurophysiol.* **88**, 761-770.

Garner WR (1947) The effect of frequency spectrum on temporal integration of energy in the ear. *J Acoust Soc Am* **19**, 808–814.

Gerstner W, Kistler WM (2002) *Spiking Neuro Models.* Cambridge University Press, Cambridge.

Gollisch T, Herz AVM (2004a) Identification of sensory transduction chains *in vivo*. Preprint.

Gollisch T, Herz AVM (2004b) Input-driven and output-driven adaptation can be disentangled *in vivo*. Preprint.

Gollisch T, Schütze H, Benda J, Herz AVM (2002) Energy integration describes sound-intensity coding in an insect auditory system. *J.Neurosci.* **22**, 10434-10448.

Hauser MD (1996) The evolution of communication. Cambridge, MA: MIT.

Heil P, Neubauer H (2001) Temporal integration of sound pressure determines thresholds of auditory-nerve fibers. *J Neurosci* **21**:7404-7415.

Hunter IW, Korenberg MJ. (1986) The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol Cybern.* **55**(2-3):135-144.

Jameson D, Hurvich LM (1972) Eds., *Handbook of Sensory Physiology*, vol. 7.4, Visuall Psychophysics, Springer, Berlin.

Johnson (1996) Point process models of single-neuron discharge. *J Comp Neurosci* **3**, 275-299.

Kriegbaum H, von Helversen O (1992) Influence of male songs on female mating behaviour in the grasshopper Chorthippus biguttulus. *Ethology* **91**:248-254.

Machens CK (2002) Adaptive sampling by information maximization. *Phys. Rev. Lett.*, **88**:228104.

Machens CK, Schütze H, Franz A, Kolesnikova O, Stemmler MB, Ronacher B, Herz AVM (2003) Single auditory neurons rapidly discriminate conspecific communication signals. *Nature Neurosci.* **6**(4):341-342

Machens CK, Stemmler MB, Prinz P, Krahe R, Ronacher B, Herz AVM (2001) Representation of acoustic communication signals by insect auditory receptor neurons. *J Neurosci* **21**:3215–3227.

Marmarelis PN, Marmarelis VZ (1978) *Analysis of physiological systems: The white noise approach*. Plenum Press, New York.

Nelson ME, Xu Z, Payne JR (1997) Characterization and modeling of p-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish. *J. Comp. Physiol. A* **181**, 532– 544.

Reich DS, Victor JC, Knight BW (1998) The power ration and the interval map: spiking models and extracellular recordings. *J Neurosci* **18**, 10090-10104.

Rieke F, Bodnar DA, Bialek W (1995) Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc R Soc Lond B Biol Sci* **262**, 259 –265.

Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W (1997) Spikes—exploring the neural code. Cambridge, MA: MIT.

Römer H (1976) Die Informationsverarbeitung tympanaler Rezeptorelemente von Locusta migratoria (Acrididae, Orthoptera) *J Comp Physiol A* **109**, 101-122.

Ronacher (1989) Stridulation of acridid grasshoppers after hemisection of thoracic ganglia: evidence for hemiganglionic oscillators. *J. Comp. Physiol. A* **164**, 723-736.

Ronacher B, Krahe R (2000) Temporal integration vs. parallel processing: coping with the variability of neuronal messages in directional hearing of insects. *Eur J Neurosci* **12**, 2147-2156.

Schaette R, Gollisch T, Herz AVM (2004) Spike-train variability of auditory neurons *in vivo*: Dynamic responses follow predictions from constant stimuli. Preprint.

Schiolten P, Larsen ON, Michelsen A (1981) Mechanical time resolution in some insect ears. *J Comp Physiol* **143**, 289–295.

Sobel EC, Tank DW (1994) In vivo Ca^{2+} dynamics in a cricket auditory neuron: an example of chemical computation. *Science* **263**, 823–826.

Stumpner A, Ronacher B (1991) Auditory interneurones in the metathoracic ganglion of the grasshopper Chorthippus biguttulus. I. Morphological and physiological characterization. *J Exp Biol* **158**, 391– 410.

Suga N (1960). Peripheral mechanisms of hearing in locusts. *Jpn J Physiol* **10**, 533-546.

Theunissen FE, Roddey J, Stufflebeam S, Clague H, Miller J (1996) Information theoretic analysis of dynamical encoding by four identified primary sensory interneurons in the cricket cercal system. *J Neurophysiol* **75**, 1345–1364.

Tougaard J (1996) Energy detection and temporal integration in the noctuid A1 auditory receptor. *J Comp Physiol* [A] **178**, 669–677.

van Rossum M. (2001) A novel spike distance. *Neural Comp.* **13**, 751-763.

Victor JD, Purpura KP (1997) Metric-space analysis of spike trains: theory, algorithms, and application. *Netw. Comput. Neural Syst.* **8**, 127-164.

von Helversen D (1972) Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke Chorthippus biguttulus. *J Comp Physiol* **81**:381– 422.

von Helversen O, von Helversen D (1997) Recognition of sex in the acoustic communication of the grasshopper Chorthippus biguttulus. *J Comp Physiol A* **180**, 373–386.

von Helversen O, von Helversen D (1994), in *Neural Basis of Behavioural Adaptations*, eds. K. Schildberger and N. Elsner, 253–284, Gustav Fischer, Stuttgart (1994).