The Journal of Neuroscience, May 1, 2001, 21(9):3215–3227

Representation of Acoustic Communication Signals by Insect Auditory Receptor Neurons

Christian K. Machens,[1,2] Martin B. Stemmler,[1,2] Petra Prinz,[2] Rüdiger Krahe,[2] Bernhard Ronacher,[1,2] and Andreas V. M. Herz[1,2]

[1]*Innovationskolleg Theoretische Biologie,* [2]*Institut für Biologie, Humboldt-Universität zu Berlin, 10099 Berlin, Germany*

Despite their simple auditory systems, some insect species recognize certain temporal aspects of acoustic stimuli with an acuity equal to that of vertebrates; however, the underlying neural mechanisms and coding schemes are only partially understood. In this study, we analyze the response characteristics of the peripheral auditory system of grasshoppers with special emphasis on the representation of species-specific communication signals. We use both natural calling songs and artificial random stimuli designed to focus on two low-order statistical properties of the songs: their typical time scales and the distribution of their modulation amplitudes.

Based on stimulus reconstruction techniques and quantified within an information-theoretic framework, our data show that artificial stimuli with typical time scales of >40 msec can be read from single spike trains with high accuracy. Faster stimulus variations can be reconstructed only for behaviorally relevant amplitude distributions. The highest rates of information transmission (180 bits/sec) and the highest coding efficiencies (40%) are obtained for stimuli that capture both the time scales and amplitude distributions of natural songs.

Use of multiple spike trains significantly improves the reconstruction of stimuli that vary on time scales <40 msec or feature amplitude distributions as occur when several grasshopper songs overlap. Signal-to-noise ratios obtained from the reconstructions of natural songs do not exceed those obtained from artificial stimuli with the same low-order statistical properties. We conclude that auditory receptor neurons are optimized to extract both the time scales and the amplitude distribution of natural songs. They are not optimized, however, to extract higher-order statistical properties of the song-specific rhythmic patterns.

Key words: auditory receptor; neural coding; acoustic communication; natural stimuli; stimulus reconstruction; insect

Evolutionary processes have shaped acoustic communication behaviors of remarkable complexity (Hauser, 1996; Bradbury and Vehrenkamp, 1998). These behaviors are made possible by sophisticated neural systems in both sender and receiver. In human beings, for example, highly specialized cortical areas process auditory stimuli, extract language information, and generate fine-tuned motor signals required for proper speech production (Levelt, 1993; Ehret and Romand, 1997).

Auditory systems of insects have a much simpler architecture, and with up to a few hundred neurons, they are orders of magnitude smaller than those of most vertebrates. Nevertheless, these systems are capable of astounding computations. Some grasshoppers, for instance, detect gaps in conspecific songs as short as 1–2 msec (von Helversen, 1972), a performance level similar to that reached by birds and mammals.

These observations trigger the general question of how a small insect auditory system could possibly be organized to process acoustic signals reliably and with high temporal precision. Important clues will come from understanding the auditory periphery. Do receptor neurons encode a large range of acoustic stimuli or are they specifically tuned to behaviorally relevant features, such as the temporal structure of a grasshopper calling song? Is the information carried by the spike train of a single auditory receptor sufficient to identify a given stimulus, or are several neurons required to do so?

To analyze these questions, we focus on acridid grasshoppers of the insect order *Orthoptera*. Their calling, courtship, and rivalry 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 separated by short pauses.

On the receiver side, such songs are encoded by roughly 100 auditory receptors into discrete trains of action potentials. The receptor cells are located within the two tympana on both sides of the animal; their axons extend through the tympanal nerves to the metathoracic ganglion, where auditory information is processed by local interneurons and then sent to the brain via ascending neurons.

The characteristics of their songs and the simplicity of their auditory system make grasshoppers an ideal candidate for addressing questions of auditory signal processing. Already early on, system-identification methods (Marmarelis and Marmarelis, 1978) were applied to this system in an effort to understand the (nonlinear) encoding of sound within a firing-rate picture (Sippel and Breckow, 1983).

Modern stimulus reconstruction methods (Bialek et al., 1991; Rieke et al., 1997) allow us to advance these approaches and study single-trial responses instead of sample averages. Specifically, we

Received July 28, 2000; revised Feb. 1, 2001; accepted Feb. 1, 2001.

This work has been supported by the Deutsche Forschungsgemeinschaft, the Alexander von Humboldt Foundation, and the Human Frontier Science Program. We are grateful to Jan Benda, Astrid Franz, Fabrizio Gabbiani, Matthias Hennig, and Hartmut Schütze for their insightful input during various stages of this work, and to Dagmar and Otto von Helversen for providing us with natural grasshopper songs.

Correspondence should be addressed to Andreas V. M. Herz, Innovationskolleg Theoretische Biologie, Invalidenstrasse 43, 10115 Berlin, Germany. E-mail: a.herz@biologie.hu-berlin.de.

use both natural calling songs and artificial stimuli that are designed to vary the most salient features of the songs, and we quantify our experimental findings within an information-theoretic framework.

This study is thus part of a larger ongoing enterprise to analyze and compare the tuning properties of auditory systems under naturalistic stimulation and extends previous studies in cats (Attias and Schreiner, 1998), birds (Theunissen et al., 2000), and frogs (Rieke et al., 1995) to insects. Our results support the view (Suga, 1989) that despite the large evolutionary distance between various auditory systems, important aspects of their information-processing strategies follow common design principles.

MATERIALS AND METHODS

Stimulus design

Acridid grasshoppers generate chirping sound patterns by rasping their hindlegs across their forewings. The songs are characterized by a broadband carrier signal with frequencies in the range of 5–40 kHz and amplitudes that are modulated in a species- and task-specific temporal pattern (Elsner, 1974; Meyer and Elsner, 1996). This amplitude-modulation signal (AM signal) is used for song recognition (von Helversen and von Helversen, 1997). Accordingly, artificial stimuli were designed to capture the most salient statistical properties of the AM signals: their typical time scales and their amplitude distributions.

Time scales of natural AM signals. As representative examples of grasshopper stimuli, calling songs of *Chorthippus biguttulus* males recorded at 30°C were used, as kindly provided by D. and O. von Helversen (University of Erlangen, Germany). Each phrase of these songs lasts for 2–4 sec and consists of many repetitions of a basic pattern, termed "syllable," separated by short pauses (von Helversen and von Helversen, 1997), as illustrated in Figure 1, *A* and *B*. Depending on the individual animal and the ambient temperature, each combination of syllable and pause spans between 60 and 140 msec. Loss or injury of a hindleg or forewing results in short yet pronounced gaps within a syllable (see Fig. 1*C*). Gaps in a male song model as short as 1–2 msec reduce the frequency of a behavioral response on the part of the female almost down to zero (von Helversen, 1972; von Helversen and von Helversen, 1997). The time scales relevant to auditory recognition thus span three orders of magnitude, from 1 msec to several seconds.

The overall rhythmic structure of a song is evident in the power spectral density of the AM signal (see Fig. 1*B*, *C*, *right panels*). Gaps within a syllable result in more prominent higher-frequency spectral components (Fig. 1*C*, *arrow*).

Distribution of natural AM signals. To restrict attention to carrier frequencies that match the sensitivity range of low-frequency receptors, the calling songs of *Ch. biguttulus* males were first low-pass-filtered, keeping only frequencies below 10 kHz. An estimate of the AM signal was then calculated by taking the Hilbert transform (Haykin, 1994) of the song waveform. A typical AM signal, corresponding to three syllables of the calling song in Figure 1*A*, is shown on the *left* of Figure 1*B*.

The distribution of modulation amplitudes was calculated by choosing a 1-sec-long segment of the example song in Figure 1*A* for which mean and variance of the AM signal computed over 40 msec windows vary least. The measured amplitude distribution displays a double-peak structure with one peak at low amplitudes for the pauses between distinct syllables and one peak at higher amplitudes for the syllable segments (see Fig. 1*B*, *center*). The distribution is, therefore, highly non-Gaussian. Note that the low-amplitude peak is centered away from zero because the pauses do not consist of silence but of relative quiet.

Amplitude modulations will be quantified by their modulation depth, defined as the range covered by the central 95% of the amplitude distribution. For the natural song shown in Figure 1*B*, this definition implies a modulation depth of 24 dB.

Artificial stimuli with large modulation depth. The double-peak distribution of modulation amplitudes is a signature of all grasshopper species. One set of artificial stimuli was designed to capture this characteristic feature. To investigate the importance of spectral properties of the AM signal, one subtype of stimulus was chosen to exhibit the power spectral density prescribed by the natural song shown in Figure 1*B*. Having both the large modulation depth (LMD) of a grasshopper song and a song-like spectrum (SLS), this artificial stimulus (Fig. 1*D*) comes closest to the properties of the natural song.

The remaining stimuli contain a uniform or "white" mix of all modulation frequencies, from zero up to a cut-off frequency of either 25, 50, 100, 200, 400, or 800 Hz (Fig. 1*E*). According to the Nyquist criterion (Press et al., 1992), these AM signals require sampling frequencies f_{sampling} ranging from 50 to 1600 Hz, i.e., sampling intervals from 20 down to 0.625 msec.

All LMD stimuli were created in two steps. In a first step, Fourier components with the specified spectral amplitudes but random phases were chosen, thereby generating Gaussian random-amplitude modulations.

In a second step, these Gaussian AM signals were used to generate AM signals that have the same amplitude distribution as a typical calling song of a male *Ch. biguttulus*, while maintaining the desired spectral characteristics of the AM signal. To do so, the measured amplitude values of a calling song were sorted into increasing order; the same number of random, Gaussian variables was also sorted. Finally, the Gaussian set of AM values was mapped in a one-to-one fashion onto the set derived from the calling song. Without corrective measures, this procedure could generate nontrivial higher-order correlations among the phases and distort the spectrum (Li and Hammond, 1975); our correction scheme involved an iterative procedure, alternately shaping the spectrum and then mapping the cumulative distribution of the artificial variables onto that of the target distribution shown in the *center* of Figure 1, *B*, *D*, and *E*. In the investigated cases, however, the distortion of the spectrum was negligible, so that the simple one-to-one mapping nonlinearity sufficed for the transformation. This also allowed us to use the reverse transformation back into Gaussian stimuli for the purpose of calculating information-theoretic quantities.

By construction, these artificial stimuli have the same broad distribution of modulation amplitudes as the male grasshopper song of Figure 1, *A* and *B*. However, they are stationary, random AM signals and therefore lack the regular syllable structure of natural songs (see Fig. 1*B*).

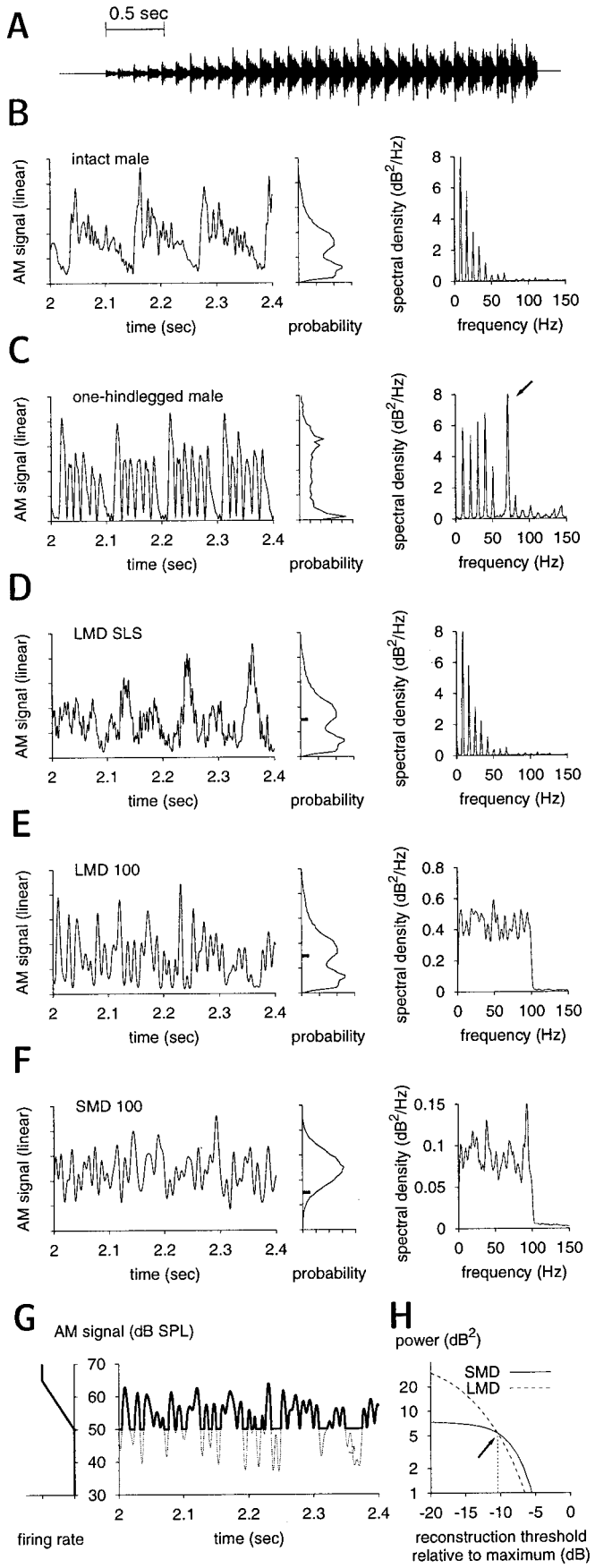
Artificial stimuli with small modulation depth. Artificial stimuli of another, second set mimic a situation where pauses and gaps of individual songs are blurred by other sounds as might occur when many grasshoppers sing simultaneously. Because the acoustic waves sum linearly and different individuals do not synchronize their song patterns, the amplitude modulations of the summed sound pressure waves will have a nearly Gaussian distribution. Assuming that 5–10 grasshoppers sing at the same time, the modulation depth of their summed sound waves was estimated using recorded songs. This yielded values between 8 and 12 dB, i.e., much less than the modulation depth of an individual song. On the basis of these observations, Gaussian-distributed AM signals with a modulation depth of 10 dB were generated (Fig. 1*F*). To restrict the amplitude values within a finite range, the tails of the Gaussian distributions were cut off at 3.5 standard deviations. Because the peak sound pressure level was always twice the average sound pressure level, the peak sound intensity was always 6 dB above the average sound intensity. These stimuli will be referred to as small-modulation-depth (SMD) stimuli.

For all types of artificial stimuli, the final AM signal was used to modulate the amplitude of a 5 kHz sine tone, the carrier frequency preferred by low-frequency receptors of acridid grasshoppers (Michelsen, 1971; Römer, 1976; Meyer and Elsner, 1996). Stimuli were digitized to a resolution of 20 kHz, and each stimulus lasted for 10 sec.

Electrophysiology

All experiments were performed on adult, male and female *Locusta migratoria*, because these are available throughout the year, and the physiological properties of their auditory receptors closely match those of *Ch. biguttulus* (Ronacher and Krahe, 2000). Legs, wings, head, gut, and pronotum were removed to immobilize the animals and to facilitate access to the metathoracic ganglion and tympanal nerve. Preparations were fixed with wax onto a Peltier element, and their temperature was kept constant at 30°C, as controlled by a sensor inserted into the abdomen. All experiments were performed in a Faraday cage lined with sound-attenuating foam to reduce echoes. The preparation was placed between two speakers (D28-2; Dynaudio, Skanderborg, Denmark) that were oriented toward the animal's ears at a distance of 35 cm.

The digitized stimuli were played back using Turbolab (Stemmer Software GmbH, Puchheim, Germany). To allow exact control of sound intensity, the signal was sent to the speakers via an attenuator (Heinecke, Seewiesen, Germany) and an amplifier (Diora WS 502 C, Conrad, Hirschau, Germany). The responses of single low-frequency receptor neurons to auditory stimuli were recorded intracellularly in the tympanal nerve with glass microelectrodes (GC-100F, Clark Electromedical Instruments, Reading, UK) filled with 1 M KCl (20–60 MΩ).



Receptor thresholds were characterized using a 5 kHz sine tone at different intensities and ranged from 30 to 65 dB sound pressure level (SPL). In one set of experiments the peak stimulus intensity was chosen to be 10 dB above the receptor's threshold. This scheme guaranteed that all stimuli types used have the same AM signal power above the receptor's threshold (compare also Fig. 1*H* and section on stimulus preprocessing and calibration).

Stimuli and responses were recorded with a DAT recorder (PC 204 A, Sony, Tokyo, Japan) at a resolution of 20 kHz and analyzed off-line. The experimental protocol complied with German law governing animal care.

Stimulus reconstruction

Information theory and systems analysis provide quantitative measures for the signal-processing capabilities of a neural system. In particular, the information contained in the spike trains of sensory afferents about an external stimulus can be estimated using stimulus reconstruction methods (Bialek et al., 1991; Rieke et al., 1997). Their application is strongly facilitated by using a priori knowledge about what qualitative aspects of the stimulus could potentially be encoded by the neuron.

Note that the use of stimulus reconstruction methods does not imply that we assume that auditory receptor neurons try to map acoustic stimuli in a one-to-one fashion on their spike trains. On the contrary, 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.

Stimulus preprocessing and calibration. Auditory receptor neurons of grasshoppers are sensitive to amplitude modulations of broad-band sound-pressure waves that exceed a certain response threshold. Below this threshold, the cells remain silent. Therefore, the appropriate preprocessed stimulus $s(t)$ for applying reconstruction techniques (see Fig. 2) is not the original sound-pressure wave $w(t)$ but 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

←

Figure 1. Stimulus design and preprocessing. *A*, Sound-pressure wave of the calling song from a *Ch. biguttulus* male, characterized by a pronounced amplitude modulation of a broad-band carrier in the 5–40 kHz range. The song itself is composed of many repetitions of a basic pattern, termed syllable, plus the adjacent pause. Depending on the individual animal and the ambient temperature, a syllable plus pause lasts for 60–140 msec and is repeated up to 40 times. *B*, *Left*, The AM signal of a short song section (three syllables) obtained from the sound-pressure wave shown in *A*. *Middle*, Distribution of modulation amplitudes. *Right*, Power spectral density of the AM signal. The first peak, at ~8 Hz, corresponds to the mean duration of a syllable (~125 msec). *C*, *Left*, Section of a calling song from a *Ch. biguttulus* male that has lost one hindleg. Short, yet pronounced gaps of 2–3 msec appear within the four syllables. Their regular occurrence causes a large spectral peak at ~70 Hz (*right*, *arrow*). *D*, Design of artificial stimuli with the same amplitude distribution and the same spectrum as the natural song in *B*. As for all artificial stimuli, the AM signal is used to modulate a 5 kHz carrier sine wave. *E*, Design of artificial stimuli with the same amplitude distribution as in *B* and a spectrum that is "white," i.e., flat up to a certain cut-off frequency, here 100 Hz (*right*). Deviations from the ideal, flat spectrum result from the finite signal length. Because of their large modulation depth (24 dB), defined as the range covered by 95% of the amplitude distribution, such stimuli are called LMD stimuli. *F*, Design of artificial stimuli with a Gaussian amplitude distribution, used as a model of the sound of several grasshoppers singing simultaneously. These stimuli have a modulation depth of 10 dB and are referred to as small-modulation-depth (SMD) stimuli. *G*, Transformation of the AM signal. Within a finite range of sound intensities above their response thresholds, receptors discharge approximately in proportion to the logarithm of the signal amplitude; therefore, transforming the AM signal logarithmically yields a piecewise linear curve of firing rate versus sound intensity, as shown schematically on the *left*. The rising part of this curve has a typical range of 10–20 dB. The resulting preprocessed LMD stimulus is depicted as *thick line* (*center*) and exhibits a short pause whenever the original amplitude modulation meanders subthreshold. *H*, Calibration of SMD and LMD stimuli. Shown is the suprathreshold power of the respective AM signal. To allow for a fair comparison of SMD and LMD stimuli, both must have the same suprathreshold AM signal power in the experiments. This point is indicated by the *vertical line* in *H* and by the *short horizontal bars* in *D–F*.

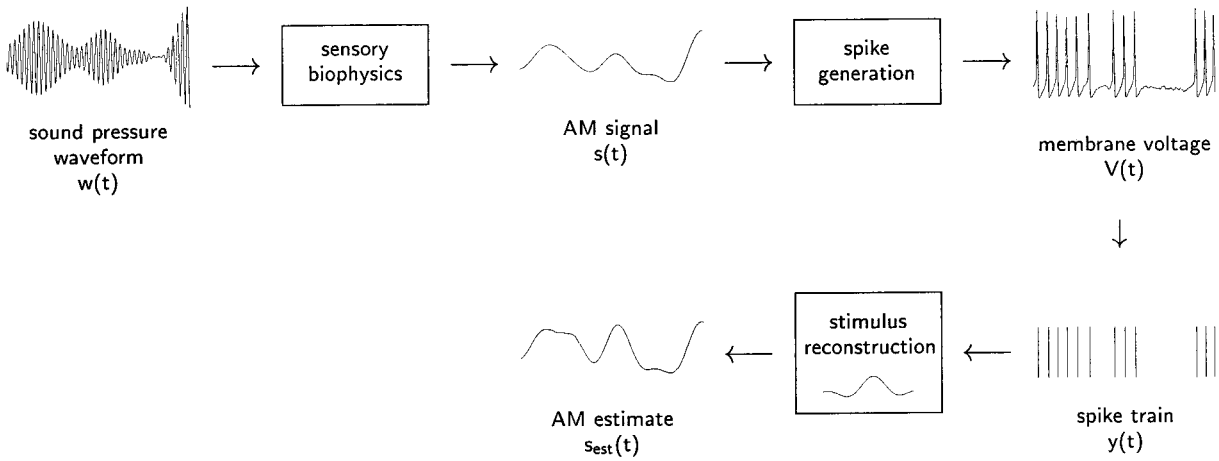


Figure 2. Stimulus preprocessing and reconstruction. The mechanics of the receiver's ear extract the slow amplitude modulation $s(t)$ of a rapidly oscillating sound-pressure wave $w(t)$. Auditory receptor neurons then encode $s(t)$ into the membrane voltage $V(t)$. As a first step of the stimulus reconstruction, the spike train $y(t)$ is extracted from the voltage trace. Within linear reconstruction, each spike is then 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 $s(t)$.

AM signal was first half-wave rectified at the threshold of each cell (see Fig. 1*G*) and then used for the stimulus reconstruction algorithm. From now on, the thresholded AM signal will be referred to simply as "signal."

The total power of the half-wave-rectified signal depends on the applied threshold (see Fig. 1*H*). Both SMD and LMD stimuli have the same total power in the amplitude modulations, if the peak signal intensity is chosen to be 10 dB above the receptor's threshold. The threshold applied in this case is also indicated by *small black bars* in the *center panels* of Fig. 1, *D–F*. Given such a calibration of the threshold, the stimulus ensembles retain differences in two primary traits: containing pauses (LMD) or lacking them (SMD), and having a periodic (SLS) or white spectrum (all others).

Linear reconstruction and filter calculation. To reconstruct the signal $s(t)$ each spike in the recorded spike train $y(t)$, a series of Dirac impulse functions, is replaced by a linear reconstruction filter $h_1(\tau)$, resulting in the signal estimate $s_{est}(t)$:

$$s_{est}(t) = h_0 + \int_0^T d\tau\, h_1(\tau) y(t-\tau), \tag{1}$$

where h_0 is the mean signal level in the absence of spiking. The parameters h_0 and $h_1(\tau)$ are determined by minimizing the mean-square error $\langle n_{mse}(t)^2 \rangle$ where the angular brackets $\langle \cdots \rangle$ denote a time average over the section of the experiment used for parameter estimation, and $n_{mse}(t)$ is the time-dependent reconstruction error $n_{mse}(t) = s(t) - s_{est}(t)$.

To analyze the activity of a population of N receptor neurons, different reconstruction filters $h_{1,i}(\tau)$, $i = 1, \ldots, N$ are allowed for each spike train $y_i(t)$, i.e.:

$$s_{est}(t) = h_0 + \frac{1}{N} \sum_{i=1}^{N} \int_0^T d\tau\, h_{1,i}(\tau) y_i(t-\tau). \tag{2}$$

As before, h_0 and $h_{1,i}(t)$ are obtained by minimizing the mean-square reconstruction error.

To restrict the number of parameters to be estimated, each reconstruction filter was expanded into an orthonormal series of Hermite functions (Arfken, 1985) of order up to 20. The mean-square-error minimization then results in a linear system of equations for the expansion parameters, which can be solved numerically (Rieke et al., 1997, appendix A.8.2; Press et al., 1992).

By design, reconstruction filters perform above average on the data section used to estimate the filter parameters. To avoid this bias, filters were always estimated on 9 of 10 segments of a recording and then evaluated on the remaining segment. Sampling errors were reduced by taking averages over repeated permutations of this procedure.

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 (Theunissen and Miller, 1995). Because 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 and Krahe, 2000), this potentially misleading interpretation of higher-order kernels was obviated by transforming $s(t)$ and $s_{est}(t)$ into the decibel scale.

In reconstructions from multiple spike trains, nonlinearities arise if the intensity ranges of the different neurons do not fully overlap. This problem was avoided in reconstructions from multiple traces by including only neurons that had approximately the same threshold.

Firing-rate adaptation can be described in principle by higher-order or time-dependent reconstruction filters, but their estimation requires enormous amounts of data. In the studies involving artificial stimuli, these complications were circumvented by discarding the first second of the 10 sec neural response patterns.

Signal-to-noise ratio. The reconstruction error $n_{mse}(t)$ 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 $s_{est}(f) = g(f)[s(f) + n_{eff}(f)]$, where $n_{eff}(f)$ denotes the random errors or "effective noise," as referred to the input (Theunissen et al., 1996; Rieke et al., 1997).

Given the effective noise $n_{eff}(f)$, the success of a stimulus reconstruction for each frequency can be measured by the signal-to-noise ratio (SNR):

$$\mathrm{SNR}(f) = \frac{S(f)}{N_{eff}(f)} = \frac{s(f)s^*(f)}{n_{eff}(f)n^*_{eff}(f)}, \tag{3}$$

where $S(f)$ and $N_{eff}(f)$ are the power spectral densities of the signal and the effective noise, respectively. In the linear filter case, the gain $g(f)$ is related to the signal-to-noise ratio by $g(f) = \mathrm{SNR}(f)/(1 + \mathrm{SNR}(f))$. Using this relation, one can also calculate the signal-to-noise ratio based on the power spectral densities of estimate and reconstruction error, $\mathrm{SNR}(f) = S_{est}(f)/N_{mse}(f)$.

A high SNR indicates an accurate reconstruction, whereas an SNR of zero implies chance level. The SNR allows one to assess which frequency 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.

A measure for the overall success of a reconstruction can be defined by using the total power of signal and noise:

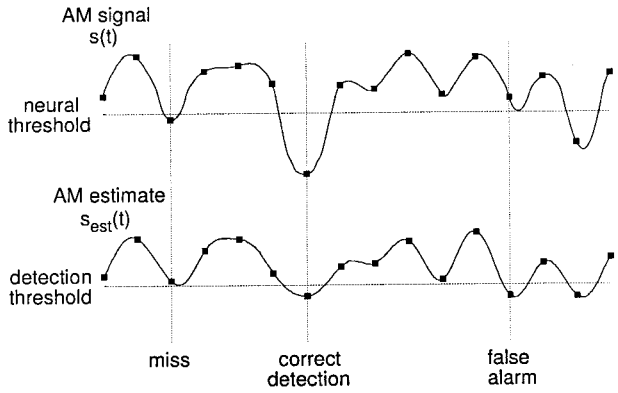


Figure 3. Gap detection. After sampling a stimulus at twice its cut-off frequency (*filled squares*), gaps were defined to be all those parts of the stimulus that fell below threshold for exactly one sampling point (*top trace*). A gap was classified as correctly detected if the reconstructed stimulus (*bottom trace*) was smaller than a given detection threshold; otherwise the gap was classified as missed. A false alarm occurs when the reconstructed stimulus falls below threshold but the signal contains no gap.

$$\mathrm{SNR} = \frac{\int_0^\infty df\, S(f)}{\int_0^\infty df\, N_{\mathrm{eff}}(f)}. \tag{4}$$

Information transfer. 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 $R_{\mathrm{info}} = 1$ bit/sec means that under ideal circumstances, the uncertainty about the stimulus can be halved every second by reading the corresponding spike train. Note that themutual 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:

$$R_{\mathrm{info}} \geq \int_0^\infty df \log_2[1 + \mathrm{SNR}(f)]. \tag{5}$$

LMD and SMD signals were generated from Gaussian distributions by nonlinear but invertible transformations. Such transformations conserve the information carried by the signal. After performing the corresponding inverse transformation on both the AM signal and its reconstruction, the lower bound on R_{info} is therefore still given by Equation 5. Accordingly, the signal-to-noise ratio in Equation 5 was calculated using the original Gaussian AM signal and the inversely transformed reconstruction. Furthermore, the artificial signals were reconstructed without applying a threshold to the preprocessed signal, because the reconstruction errors thus obtained followed a Gaussian distribution more closely, making it more likely that Equation 5 is a tight lower bound. Because the receptor neurons do not encode details of the subthreshold signal (see also Fig. 4*D*), the information rates thus computed must be conveyed almost exclusively by the suprathreshold signal.

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

$$\epsilon(\Delta t) = \frac{R_{\mathrm{info}}(\Delta t)}{R_{\mathrm{max}}(\Delta t)}, \tag{6}$$

and takes values between 0 and 1. Although $R_{\mathrm{max}}(\Delta t)$ tends to infinity for $\Delta t \to 0$, this is not the case for $R_{\mathrm{info}}(\Delta t)$, which will instead achieve the value R_{info} given in Equation 5. To yield nontrivial results, the coding efficiency, therefore, has to be evaluated at a finite time resolution that reflects spike-timing variability caused by intrinsic noise sources. This time resolution was estimated by cross-correlation analysis. The full width at half maximum of the cross-correlation peak was calculated for each two spike trains recorded from the same stimulus. The smallest width obtained, $\Delta t \approx 1$ msec, was then taken to be the approximate time resolution of the system. Values of Δt ranging from 0.5 to 2 msec yielded comparable results and underline the robustness of the method.

Redundancy. Let $R_{\mathrm{info}}(1)$, $R_{\mathrm{info}}(2)$, and $R_{\mathrm{info}}(1, 2)$ denote the mutual information rates obtained from reconstructions based on two individual spike trains and from the corresponding population reconstruction, respectively, with all quantities calculated at a time resolution of $\Delta t = 0.1$ msec. A measure of redundancy ρ of these two cells was defined as:

$$\rho = \frac{R_{\mathrm{info}}(1) + R_{\mathrm{info}}(2) - R_{\mathrm{info}}(1, 2)}{\min\{R_{\mathrm{info}}(1), R_{\mathrm{info}}(2)\}}, \tag{7}$$

where $\min\{x, y\}$ denotes the smaller of the two variables x and y. A value of $\rho = 1$ implies complete redundancy and $\rho = 0$ corresponds to complete independence. Negative values of ρ occur if the two cells are synergistic.

Because identical spike trains carry the same information, their redundancy is 1. The reverse, however, is not true: a redundancy of 1 does not imply that the spike trains were identical, because even two different spike trains might carry identical information. Therefore, redundancy is not simply a measure of spike-train variability, but a measure of the information-theoretic consequences of that variability.

Gap detection. A gap is a brief, silent interruption of an acoustic stimulus. Sampled at twice the cut-off frequency of the AM signal, a stimulus was defined to exhibit a gap whenever the AM signal remained below the neuron's threshold for exactly one sampling point. Therefore, the average length of a gap is simply given by the inverse of the corresponding sampling frequency. Note that in this definition a gap is not a completely silent part of the stimulus, but rather a part that appears to be silent as perceived by the investigated neuron. When stimuli are presented with a peak amplitude of 10–20 dB above threshold, LMD stimuli contain a significant number of gaps, whereas SMD stimuli comprise at most a few gaps during the whole course of the stimulus.

A gap was called detected if the reconstructed stimulus at that instant was smaller than a detection threshold (see Fig. 3, *correct detection*). Varying the detection threshold balances the tradeoff between the two types of error that might occur: a miss when the stimulus exhibits a gap that was not detected, or a false alarm when the stimulus exhibits no gap, but the reconstruction falsely indicates a gap (see Fig. 3, *miss* and *false alarm*).

The tradeoff between miss and false alarm can be quantified by the receiver operating characteristics (Poor, 1994), in which the probability of correct detection is plotted against the probability of false alarm, both being parametrized by the detection threshold. This measure differs from the previous measures in that it focuses solely on the reliability of gap detection and does not take into account how accurately the stimulus is encoded between two gaps.

RESULTS

To identify essential features of acoustic communication signals and their neural representations in grasshoppers, auditory receptor neurons were stimulated with natural and artificial sounds. Recordings were performed in the migratory locust *L. migratoria*, a well established model system (Stumpner and Ronacher, 1991; Ronacher and Krahe, 2000). The artificial stimuli were designed to vary the most salient statistical properties of grasshopper sounds, which consist of a broad-band carrier. Its amplitude modulations (the AM signal), illustrated in Figure 1*A–C*, carry the behaviorally relevant information. Typically, the songs alternate between noise bursts and pauses, leading to a characteristic double-peak distribution of sound amplitudes (Fig. 1*B*, *center*). To investigate the importance of this structural aspect, two different classes of stimuli were generated.

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

modulation depth of ~24 dB (Fig. 1*B*, *D*, *E*), these stimuli are called large-modulation-depth (LMD) stimuli.

Within the second class, stimuli have a Gaussian amplitude distribution (Fig. 1*E*) with a modulation depth of 10 dB and are called small-modulation-depth (SMD) stimuli. These random stimuli simulate the combined sound pattern of a group of 5–10 grasshoppers singing simultaneously, such that the song pauses of individual songs are filled by the other songs. Additionally, the Gaussian distribution facilitates the comparison with previous stimulus-reconstruction studies in other sensory systems (Bialek et al., 1991; Rieke et al., 1995; Theunissen et al., 1996; Wessel et al., 1996).

Because the shortest behaviorally relevant time scales of the AM signals are ~1–2 msec (von Helversen, 1972; von Helversen and von Helversen, 1998), frequency components of at least 250–500 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.

This set of artificial AM signals with well defined amplitude distributions and spectral characteristics allowed us to test whether auditory receptors can encode arbitrary stimulus features down to the millisecond time scale and whether the pauses and gaps in individual songs are of any importance for the encoding procedure. Moreover, estimating the amount of information that receptors transmit becomes a straightforward task (see Materials and Methods).

Low-frequency auditory receptors of acridid grasshoppers respond best to amplitude-modulated sounds with carrier frequencies in the 4–8 kHz range (Römer, 1976; Stumpner and Ronacher, 1991). Above their threshold and below saturation, the firing rate of the receptors increases approximately linearly with sound pressure level if the latter is measured on a logarithmic scale (Römer, 1976; Stumpner and Ronacher, 1991). All attempts to reconstruct the original sound pressure wave failed (data not shown). A representation of the signal appropriate for stimulus reconstruction rather consists of a thresholded AM signal in decibels. To reconstruct the preprocessed AM signal from a spike train, each spike is replaced by a filter (Fig. 2), resulting in a smooth time-varying function. After minimization of the mean-square distance between this function and the original AM signal, the (optimal) linear reconstruction of the signal is obtained (see Materials and Methods).

In all, responses were recorded from $n = 27$ receptor neurons, and each individual stimulus class was presented to up to 10 neurons. Stimuli were presented with peak intensities ranging from 3 to 21 dB above the threshold of individual receptors, entailing firing rates from 40 to 160 Hz.

Single auditory receptors are capable of encoding sound-amplitude modulations with high signal-to-noise ratios

As illustrated in Figure 4, amplitude modulations of sound pressure waves can be reconstructed from the spike trains of individual auditory receptors. Figure 4*A* displays the amplitude modulation of a 5 kHz tone, the resulting spike train from an auditory receptor, and the estimated amplitude modulation as reconstructed from this spike train. The stimulus is an LMD stimulus with a spectrum of amplitude modulations that is flat up to a cut-off frequency f_c of 50 Hz. The original AM signal was thresholded at 34 dB, corresponding to the experimentally determined threshold of the investigated cell. The signal estimate closely follows the thresholded signal, with deviations of at most a few decibels. As quantified in Figure 4*B*, the distribution of reconstruction errors has a standard deviation of ~1.4 dB.

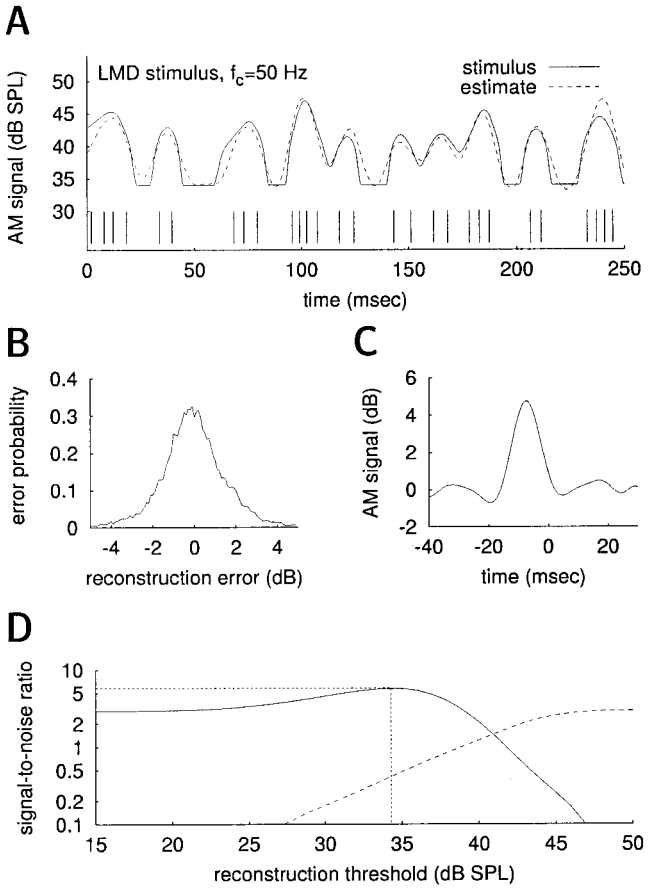


Figure 4. Reconstruction of an LMD signal with 50 Hz cut-off frequency from the responses of a single receptor. *A*, Preprocessed stimulus, spike train, and stimulus estimate. The stimulus was thresholded at 34 dB, the threshold of this particular neuron. *B*, Distribution of reconstruction errors. The distribution has a standard deviation of only 1.4 dB, which indicates that the preprocessed stimulus was reconstructed quite accurately from the spike train, depicted below the stimulus in *A*. *C*, Linear reconstruction filter. The estimated signal is obtained by a convolution of the spike train with this filter, which amounts to replacing each spike by the filter function. *D*, Dependence of the reconstruction quality on the stimulus representation. The AM signal was split into a suprathreshold and a subthreshold signal, each of which was independently reconstructed from the spike train. The calculated signal-to-noise ratios of the suprathreshold signal (*solid line*) reach their maximum when the threshold used for splitting matches the threshold of the specific neuron. At this threshold, reconstruction of the subthreshold AM signal fails almost completely, as shown by a signal-to-noise ratio of <0.5:1 (*dashed line*).

The linear reconstruction filter is shown in Figure 4*C* and represents the contribution of a single spike, situated at time zero, to the reconstruction of the AM signal. The central peak of the filter is shifted by 7–8 msec from zero, reflecting the intrinsic delay between the stimulus presentation and the response of the auditory receptor. Results obtained for cut-off frequencies from

25 to 400 Hz and intermediate firing rates indicate that the full width at half maximum, w, of the optimal reconstruction filter is roughly given by $w \approx (2f_c)^{-1}$. The qualitative shape of the filter remains the same as that seen in Figure 4*C* (data not shown).

The signal-to-noise ratio (SNR) quantifies the reconstruction success by the ratio of AM signal variance to the variance of the effective noise as referred to the input (see Materials and Methods). In the example illustrated in Figure 4, the signal-to-noise ratio was 6:1, demonstrating that even a single auditory receptor can accurately represent the time course of the AM signal. Without proper thresholding of the AM signal, the SNR would decrease significantly, to values of ~3:1, as depicted in Figure 4*D*. By contrast, applying an upper threshold and trying to reconstruct only the subthreshold portion of the AM signal leads to considerably worse results (Fig. 4*D*, *dashed line*). At the threshold indicated by the *vertical line*, the SNR in this case is <0.5:1. Hence, little if any information can be recovered from the subthreshold part of the stimulus. These results underline the importance of carefully adjusting stimulus-reconstruction techniques to the specific properties of the neural system under study.

Stimuli with gaps are transmitted with higher information rates and coding efficiencies than stimuli without gaps

To allow for a direct comparison of stimuli with large and small modulation depth, a set of experiments was performed in which the AM signal power above the receptor threshold was made to be equal for both classes of stimuli. For this purpose, the AM signals were thresholded, and the remaining variance in the AM signals was computed. As shown in Figure 1*H*, the above-threshold AM signal power is the same for LMD and SMD stimuli if the peak stimulus intensity is adjusted to be 10 dB above the receptor threshold.

To compare the reconstructions obtained for different stimuli, we applied three measures: (1) signal-to-noise ratios, to determine how accurately a given reconstruction follows an AM signal (Fig. 5*A*); (2) information rates, to assess how much information is conveyed by each spike train about the stimulus (Fig. 5*B*); and (3) coding efficiencies, to measure how efficient receptor neurons use their resources to transmit that information (Fig. 5*C*) (see also Materials and Methods).

With most of the original power in the LMD stimuli well below receptor threshold, there exists no a priori reason to expect any difference in the signal reconstruction success for the two calibrated signals. Nonetheless, for all three measures, LMD stimuli clearly outperform SMD stimuli. No matter for which cut-off frequency, LMD stimuli can always be reconstructed more accurately than SMD stimuli. In addition, spike trains always convey more information about LMD than SMD stimuli, and receptor neurons use their resources more efficiently when presented with LMD stimuli. Hence, receptor neurons are much better suited to convey information about stimuli that feature a natural amplitude distribution and, therefore, gaps.

On the other hand, matching the spectral properties of the songs, such as the one depicted in the *right panels* of Figure 1, *B* and *D*, has almost no effect. Although the sharp spectral peaks of the songs confer strong rhythmicity to the natural grasshopper calls, rhythmicity neither impairs nor aids the quality of the reconstructions. Given that the predominant portion of the natural spectrum falls between 0 and 50 Hz, the signal-to-noise ratios, information rates, and coding efficiencies of the spectrally matched stimulus (LMD SLS) should be compared with that with

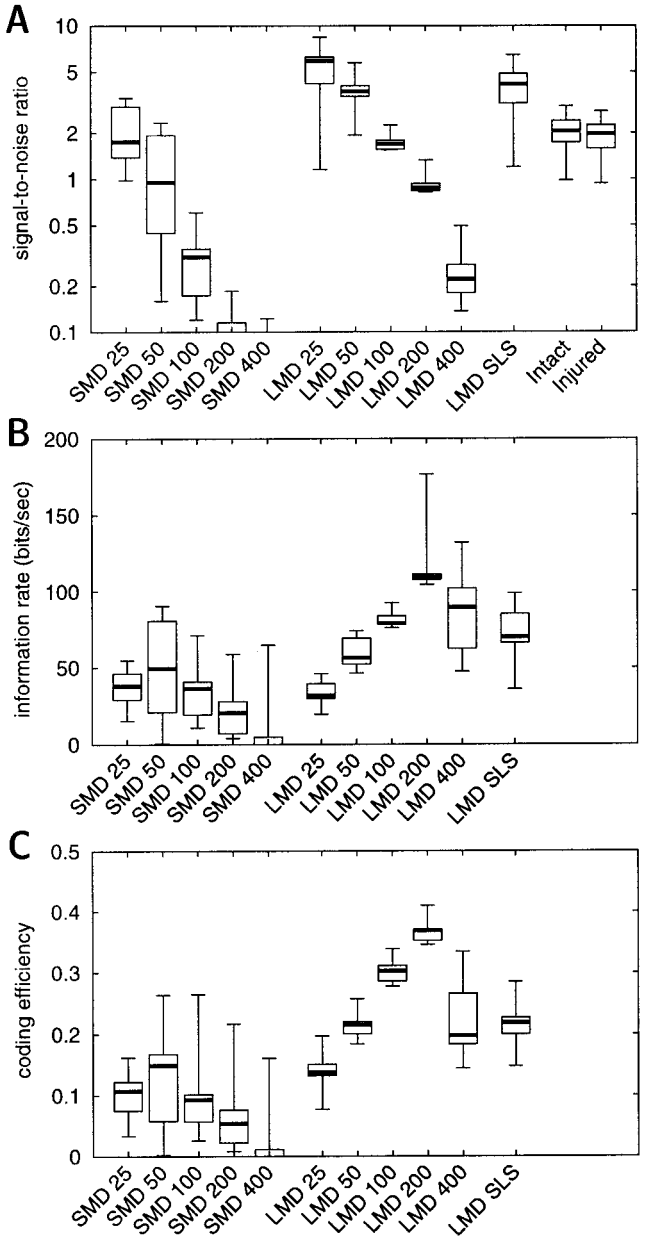


Figure 5. Summary of all experiments with equal suprathreshold AM signal power. Altogether 27 cells from 14 animals were analyzed, with mean firing rates ranging from 40 to 100 Hz. The sound intensity relative to the receptor thresholds was chosen such that the suprathreshold power of the AM signals of LMD and SMD stimuli was equal. *A*, Signal-to-noise ratio for natural and artificial stimuli. Because the signal variance was kept constant, the signal-to-noise ratio, measuring the reconstruction quality, falls off with increasing cut-off frequency f_c for LMD as well as SMD stimuli. The signal-to-noise ratio for stimuli with a song-like spectrum (SLS) is comparable to that of LMD stimuli with $f_c = 50$ Hz. *B*, Mutual information rate R_{info}, quantifying the information carried by the spike train about the AM signal. *C*, Coding efficiency ϵ, measuring the fraction of the maximum possible information transfer. Because the probability distribution of natural songs is unknown, neither R_{info} nor ϵ can be calculated for these stimuli. LMD stimuli with a cut-off frequency of 200 Hz result in the largest values for R_{info} and ϵ, suggesting that single receptor neurons are optimized for stimuli with such statistics. *Thick lines* indicate the median, *boxes* indicate the quartiles, and *bars* indicate the maximum and minimum observed values.

a 50 Hz cut-off frequency (LMD 50). Similar values for all three quantities indicate that filling in the valleys between the peaks in the spectrum to create the "white" LMD stimulus has almost no consequences.

The highest information rates (up to 180 bits/sec) and coding efficiencies (up to 40%) are reached for the LMD stimulus with a cut-off frequency of 200 Hz. This stimulus has a natural amplitude distribution and varies randomly on time scales down to 2.5 msec. Among all stimuli tested, this stimulus best exploits the operating regime of a receptor neuron.

For all stimuli, signal-to-noise ratios decrease with increasing cut-off frequency. For a linear system, this is to be expected because the power density of the AM signal decreases with increasing cut-off frequency, a consequence of keeping the total power of the AM signals constant.

High-frequency amplitude modulations of random stimuli cannot be reconstructed from single receptors

To analyze the frequency–response properties of the system under investigation, signal-to-noise ratios were resolved in the frequency domain (Fig. 6). Our data show that for stimuli with high cut-off frequencies, slow temporal variations are represented better than fast variations. The decrease of the measured SNRs with frequency suggests that the auditory receptors act as band-pass filters for the AM signal, instead of being tuned to any particular modulation frequency. High signal-to-noise ratios can be achieved even for artificial broad-band stimuli by matching the amplitude distribution of the natural grasshopper calling songs, as in the LMD stimuli.

Signal-to-noise ratios decrease for high cut-off frequencies, as seen in Figure 6*D–F*. Faster variations can be reconstructed only for LMD stimuli, in which the AM signal repeatedly falls below threshold. Even then, no significant amount of information about the stimulus is retrievable for frequencies >400 Hz (Fig. 6*F*). On the basis of the spike train of a locust auditory receptor, decoding arbitrary signal features at time scales of 1–2 msec is nearly impossible. This drop-off at higher frequencies does not depend on the mean firing rate of the receptor (data not shown).

Reading from multiple spike trains improves the reconstruction of high-frequency amplitude modulations

Distributed representations based on the activity patterns of a population of many auditory receptors admit an improved resolution of stimulus features in both time and intensity. Additionally, the population serves to increase the overall range of sound intensities covered, because the 40–60 low-frequency receptor neurons on each body side of the animal have thresholds spreading over >40 decibels SPL (Römer, 1976; Jacobs et al., 1999; Ronacher and Krahe, 2000). The question then poses itself: can the population encode arbitrary stimulus features of 1–2 msec duration, the minimal time needed to detect the gap in the song of a "one-hindlegged" male?

No evidence has been found to date for any physiological coupling between auditory receptors in acridid species. On the assumption that the responses of different receptors are independent and the receptor properties do not change during the experiment, sequential recordings from different cells may be pooled to estimate the information carried by a population of auditory receptors.

To quantify the information that can be gained by pooling responses, we computed the information redundancy of two cells as a function of the stimulus type presented. Both cells had the same threshold and encoded the same stimulus range. As shown in Figure 7*A* (*left panel*), repeated LMD stimuli with low cut-off frequencies (25, 50, and 100 Hz) result in highly redundant spike trains. For these stimuli, then, pooling responses yields almost no additional information. By contrast, SMD stimuli with higher cut-off frequencies (100 Hz and more) elicit spike trains that carry largely independent information.

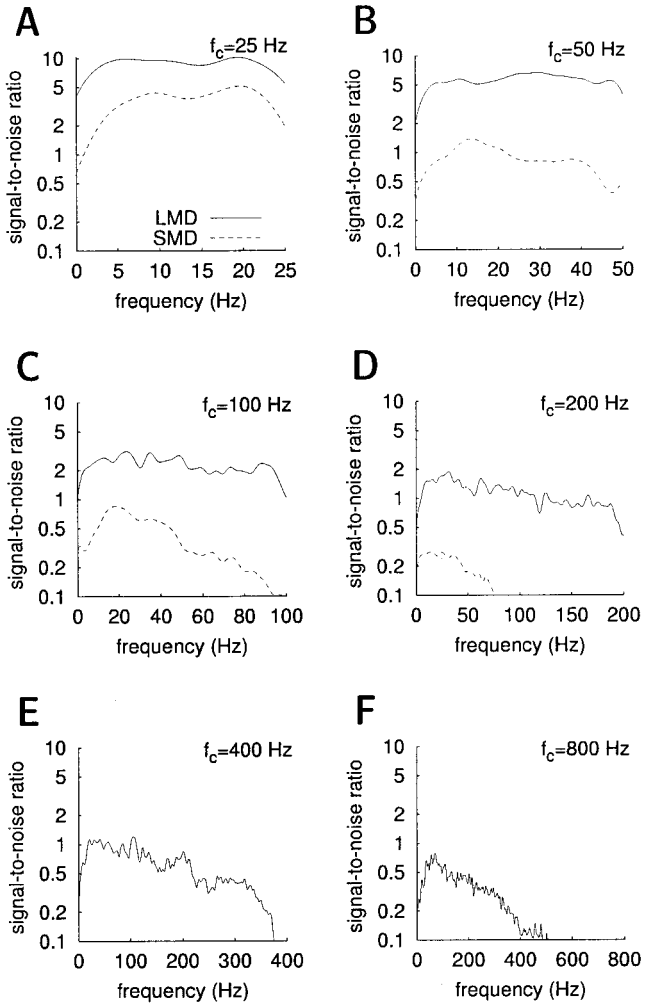


Figure 6. Signal-to-noise ratio for the artificial stimuli, shown in the frequency domain. The cut-off frequency f_c of the AM signal was varied while the integrated AM signal power was kept constant. Again, sound intensities were chosen such that above the receptor thresholds, LMD and SMD stimuli had the same AM signal power (4.8 dB^2). Firing rates were 40–60 Hz for the LMD stimuli (*solid lines*) and 60–80 Hz for the SMD stimuli (*dashed lines*). Although LMD stimuli are encoded by fewer spikes, their longer subthreshold periods and steeper onsets result in more accurate reconstructions, as shown by the higher signal-to-noise ratios. The fine structure of the curves is not statistically significant because the signal-to-noise values have relative errors of 17.5%.

The information redundancy calculated over repeated stimulations of the same cell (Fig. 7*A*, *right panel*) is almost the same as that from different cells (Fig. 7*A*, *left panel*). This indicates that the encoding procedure of cells sensitive to the same stimulus range is almost identical. It thus matters little whether spike trains from the same or different neurons are being compared, as long as the neurons have the same sound-intensity threshold.

The information gained by additional spike trains must come from the variability between these spike trains, because identical

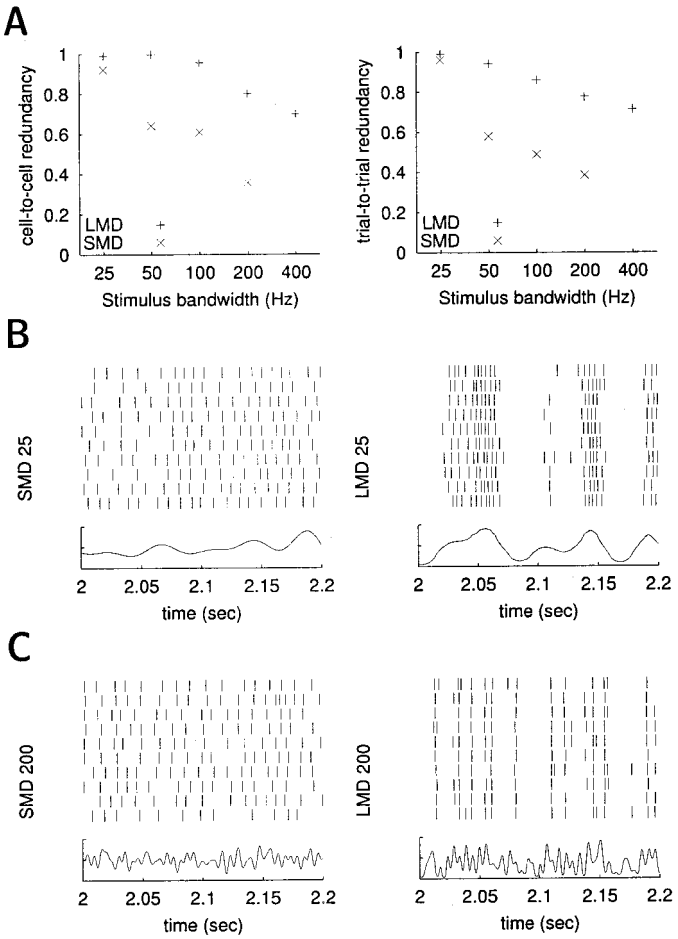


Figure 7. Redundancy and reliability of neural responses. *A*, *Left*, Redundancy of spike trains from two different cells as a function of stimulus type and cut-off frequency. The peak stimulus intensity was 60 dB, and both cells had a threshold of 50 dB. Neurons convey identical information about a stimulus if the redundancy equals 1, whereas if the redundancy is 0, they convey independent information. *Right*, The redundancy of spike trains from one cell responding to repeated presentations of the same stimulus. LMD stimuli (+) cause larger redundancy than SMD stimuli (×), and in both cases, cell-to-cell redundancy is comparable to trial-to-trial redundancy. *B*, *C*, Spike raster plots for one of the cells in response to the stimuli denoted on the abscissa. In terms of spike timing precision, the LMD 200 stimulus (*C*, *right*) stands out clearly, but other stimuli result in higher trial-to-trial redundancies (*A*, *right*).

spike trains are necessarily completely redundant. The trial-to-trial variability for some of the tested stimuli can be inferred from Figure 7, *B* and *C*. Spike-train variability, however, does not always correlate with nonredundant information. For instance, although spike responses elicited from an LMD stimulus with f_c = 200 Hz (Fig. 7*C*, *right*) are more reliable than those elicited from an SMD stimulus with f_c = 25 Hz (Fig. 7*B*, *left panel*), the latter has a higher redundancy.

If two receptor neurons respond to distinct, non-overlapping stimulus intensity ranges—in other words, if each receptor "hears" a different part of the stimulus—the information redundancy in the spike trains decreases correspondingly (data not shown).

Given the reduced redundancy of information for stimuli with higher bandwidth and/or small modulation depth, pooling neuronal responses should aid most in uncovering information about short-time and small-intensity features of the AM signal. This is indeed true. For instance, Figure 8*A*, *left panel*, shows that the information rate for the SMD stimulus with a cut-off frequency of 100 Hz increases much faster with the number of spike trains than the information rate for the LMD stimulus, which begins to saturate already when four or five spike trains are used for the reconstruction.

Signal-to-noise ratios based on a pool of eight spike trains (Fig. 8*B*) are significantly improved as compared with SNRs in the reconstruction from a single spike train, especially at high frequencies (Fig. 6*D–F*). Pooling spike trains is similarly efficient and important for stimuli with a larger bandwidth. A pool of eight spike trains evoked by the SMD stimulus with a cut-off frequency of 200 Hz provides more than four times as much information as a single spike train (Fig. 8*A*). Even more dramatic results could be expected for SMD stimuli with higher cut-off frequency, but the measured values for R_{info} are too small and too variable to give statistically significant results after cross-validation. Finally, for cut-off frequencies f_c = 200, 400, or 800 Hz, the information rate for LMD stimuli increases by roughly 100 bits/sec when pooling eight spike trains (Fig. 8*A*), as compared with reading from a single spike train.

Signal-to-noise ratios decay quickly for very high modulation frequencies but are still ≳1:1 at 300 Hz, corresponding to a time resolution that samples stimulus features every 1.5–2 msec (Fig. 8*B*) (f_c = 400 Hz). Stimulus features are thus faithfully represented at all behaviorally relevant time scales. This increase in the coding performance is possible because cell-to-cell variations in the response patterns of individual receptors may be exploited on the population level.

Reading from multiple spike trains improves the detection of gaps

Although high signal-to-noise ratios indicate that the reconstruction accurately captures the time course of the AM signal, it is useful to focus on brief gaps in the AM signal, because these are of specific behavioral relevance. The AM signal was defined to exhibit a gap during all times when it was in the subthreshold range for exactly one time step, where the time step is defined as the inverse of the sampling frequency (twice the cut-off frequency) of the signal. A gap was considered to be detected if the reconstructed AM signal fell below a detection threshold during the same time instant (see Materials and Methods and Fig. 3).

Figure 8*C* shows the receiver operating characteristics (ROCs) for different LMD stimuli (SMD stimuli will only rarely fall below threshold and were excluded therefore from this analysis). Two conclusions can be drawn from this figure. (1) Using eight spike trains instead of a single one significantly enhances the detection of gaps. (2) In all cases, the ROC curve is asymmetric regarding the balance between correct detection and false alarm; even for detection thresholds that lead to few false alarms, almost all gaps can be detected.

The spurious recognition of gaps can be attributed to the fact that stimulus parts that come very close to the threshold of the neuron from above without crossing it might easily be mistaken for gaps. The nature of the stimulus as a whole, however, may help to assess whether coming close to the threshold of a particular neuron truly constitutes a gap. Later processing stages, taking into account larger sections of the stimulus or multiple spike trains, can therefore refine the decision of what is a gap and what is not.

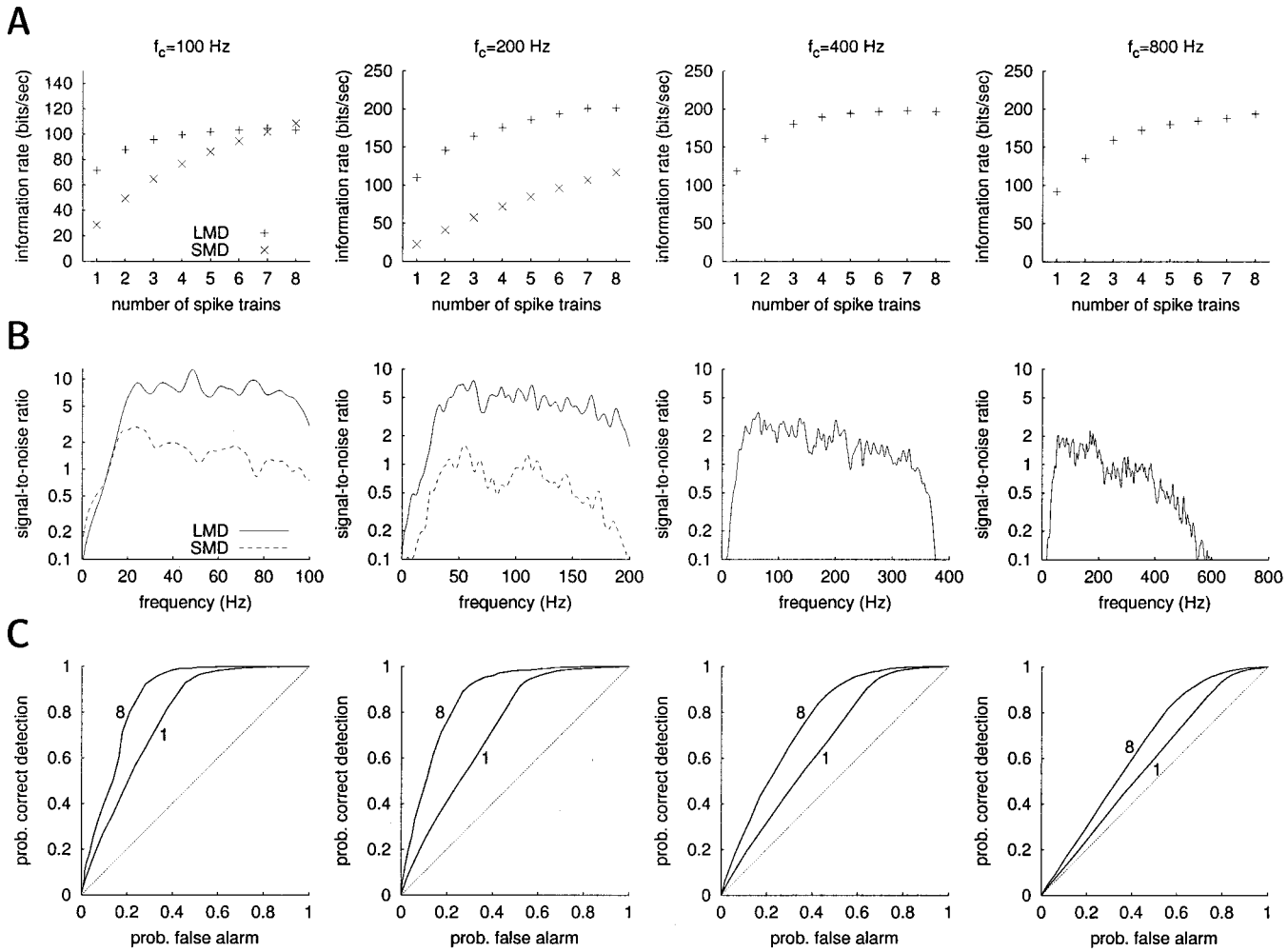


Figure 8. Reconstruction from multiple spike trains and gap detection. *A*, *Left*, Mutual information as a function of the number of spike trains used in the reconstruction for stimuli with cut-off frequency f_c. Depicted are mean values from a pool of eight spike trains of two cells having the same threshold (50 dB). The peak stimulus intensity was 60 dB. As the information redundancy shown in Fig. 7 already suggests, information rates improve particularly strongly for SMD stimuli and stimuli with high cut-off frequency. *B*, Signal-to-noise ratio when all eight spike trains are combined. SMD stimuli with cut-off frequencies of 400 and 800 Hz are not shown because they led to insignificant information transfer and SNR values near unity. Signal-to-noise ratios obtained from the full pool of spike trains are significantly larger than those calculated for single cells but decrease rapidly at high frequencies. Signal components >400 Hz are thus still poorly represented in the stimulus reconstruction. *C*, Receiver operating characteristics for gap detection as a function of the number of spike trains (1 vs 8) used for the reconstruction. Shown are the probability of detecting a gap (correct detection) versus the probability of falsely predicting a gap (false alarm). The *dashed* identity *curve* marks chance level. Only LMD stimuli are shown, because SMD stimuli contain almost no gaps. Using all spike trains significantly increases the gap detection performance.

Natural songs can be reconstructed with high quality from multiple spike trains

On the basis of the results obtained for artificial stimuli, a population of receptor neurons seems to be capable of representing stimulus features on time scales down to 1.5–2 msec. Therefore, we predicted that also the natural song of an injured, i.e., one-hindlegged *Ch. biguttulus* male can be reconstructed from a small group of spike trains. This is indeed the case, as shown in Figure 9. Here, the AM signal of the song is depicted together with pooled reconstructions from four spike trains of two receptors that had roughly the same intensity thresholds. If the sound pressure level of the song is just above threshold, only the onset of each syllable is encoded. With increasing sound intensity the full syllable structure is recovered, including the gaps within each syllable. Interestingly, at the highest sound intensity, 21 dB above threshold, the maximum amplitude of each reconstructed syllable is fairly invariant from syllable to syllable (Fig. 9*B*, *bottom row*), despite the slowly rising overall intensity of the recorded auditory input (Fig. 9*A*). This phenomenon coincides with the adaptation of the firing rate as the sound intensity increases.

Comparing Figure 9 with Figure 5, these adaptation effects also explain why the signal-to-noise ratios from the natural *Ch. biguttulus* songs do not reach those obtained from LMD stimuli with a natural spectrum (LMD SLS) or cut-off frequencies of 25 or 50 Hz (LMD 25 and LMD 50), although the songs contain their major spectral power in this frequency range. In reconstructions using artificial stimuli, the first second of the 10 sec response pattern was discarded to avoid adaptation effects. When reconstructions of the respective stimuli are based on the first 2–4 sec, as in the reconstructions of natural songs, the obtained signal-to-noise ratios decrease and reach values similar to those obtained from the songs (data not shown).

Retrieval of the AM signal from the entire set of auditory receptors characterized by staggered intensity thresholds must

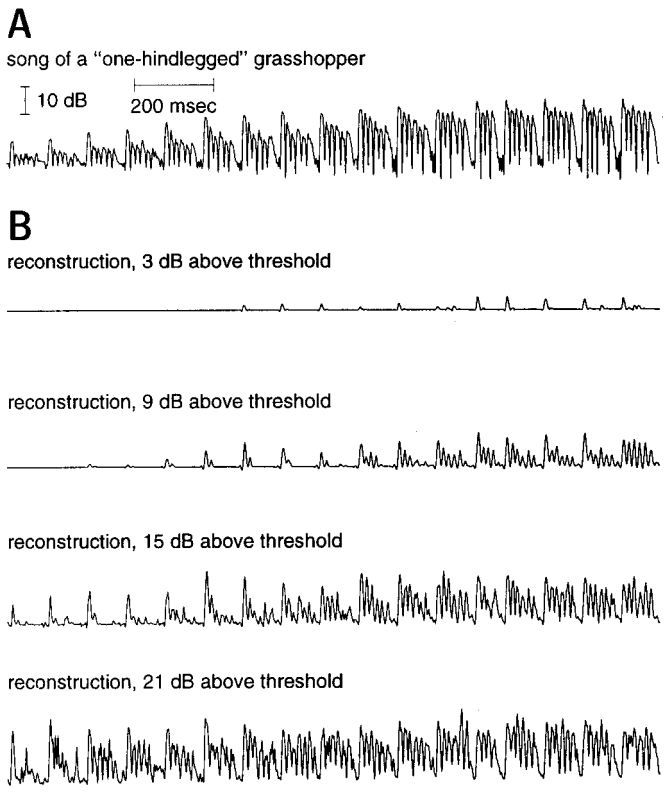


Figure 9. Reconstruction of the calling song of a "one-hindlegged" grasshopper. *A*, AM signal. This signal has not yet been thresholded and is displayed on a decibel scale. *B*, Reconstruction of the signal, thresholded at 55 dB, for different stimulus intensities (peaks at 58, 64, 70, and 76 dB). Four spike trains from two cells with thresholds of 55 dB were pooled. At sound-pressure levels just exceeding the firing threshold, only the onset of each syllable is encoded in the spike train. With increasing sound intensity, more and more details of the song appear in the reconstruction, until at 21 dB above threshold even the short gaps of 2 msec length are almost perfectly preserved. Note that in the last reconstruction, the maximum amplitude of each reconstructed syllable remains approximately the same throughout the entire song. This demonstrates that adaptation effects balance the rising overall intensity of the song. Downstream processing stages therefore receive a fairly invariant representation of each syllable onset.

emphasize the start of each syllable, because many spikes in different trains coincide at the syllable upstroke (Adam, 1977; Ronacher and Römer, 1985). Stimulus reconstructions from the full receptor population therefore will inevitably display systematic deviations from the original AM signal structure, emphasizing certain features, in particular rapid increases of the sound intensity after a pause or gap, and downplaying others.

DISCUSSION

Mate finding of various insects, such as cicadas, crickets, and grasshoppers, largely relies on acoustic communication. This requires the reliable detection and recognition of conspecific acoustic signals that might be corrupted by various noise sources. Often, signal recognition is not based exclusively on the acceptance of the correct species-specific sound pattern, but also involves active rejection of signals that contain wrong or suspicious components. For example, female *Ch. biguttulus* grasshoppers do not respond to calling songs if these contain short gaps that are not part of the song of intact males (von Helversen, 1972; von Helversen and von Helversen, 1998).

In the present study, stimulus reconstructions were performed to analyze the representation of such behaviorally relevant signals. Reconstructions from the spike trains of single receptor neurons demonstrate that even single cells are capable of encoding amplitude modulations with high signal-to-noise ratios (Fig. 4). In addition, our data show that sounds with large modulation depth are encoded with much higher signal-to-noise ratios, information rates, and coding efficiencies than stimuli with small modulation depth. Although LMD stimuli show greater raw amplitude variations, in general the encoding of SMD stimuli is still poorer even when the AM signal power above threshold is identical in the two classes of stimuli (Figs. 5, 6). Matching the spectral properties of the songs as in the SLS stimulus, on the other hand, does not increase signal-to-noise ratios.

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 (Adam, 1977; Ronacher and Römer, 1985). This phenomenon emphasizes the paramount importance of gaps and pauses for the recognition of acoustic stimuli, because the precision in spike timing leads to a faithful representation of the suprathreshold sound pattern. Grasshoppers seem to exploit 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 (Fig. 5). 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.

Against background noise generated by competing grasshoppers, other sound sources, and multiple sound reverberations in the habitat, the effective modulation depth of an individual song, as evaluated by a female, decreases rapidly with the distance of the male singer. With increasing distance, therefore, the song no longer resembles an LMD stimulus and is expected to become more and more similar to an SMD stimulus in its modulation depth. This implies that the precise shape of the amplitude modulations of the song can no longer be reconstructed faithfully because SMD stimuli lead to much lower signal-to-noise ratios than LMD stimuli (Fig. 5*A*). Heard at a distance in the field, there is thus no large difference between the reconstructed song of an intact and an injured male grasshopper. We thus predict that females only discriminate against one-hindlegged males if they are nearby. In fact, *Ch. biguttulus* females avoid mating with such males (Kriegbaum and von Helversen, 1992). At close distances, the high signal-to-noise ratios for LMD stimuli (Fig. 5) should also allow the detection of much finer details in a song, which might provide the female with additional information about the male's fitness.

Distributed codes involving many receptor neurons help to represent the acoustic environment in greater detail, especially improving the resolution for stimuli that cannot be reconstructed well from single spike trains (Fig. 8). Combining receptor neurons that cover the same sound intensity range helps in two important ways. (1) The bandwidth of modulation frequencies that can be

faithfully encoded is increased, which leads to a greater detectability of very short gaps (down to 1.5–2 msec). In fact, the resolution achieved by the receptor neurons matches the resolution limits that have been found in behavioral experiments. (2) The representation of stimuli with a small modulation depth is enhanced. Information that is gained about such stimuli might help grasshoppers to detect acoustic communication signals in a noisy environment.

Strong correlations in the response patterns across neurons limit the information gained by considering multiple spike trains; the net information rate saturates at five to eight combined spike trains. This number should be compared with the number of receptor cells that have a linear firing rate characteristic in a specified intensity range. Given the threshold distribution measured by Römer (1976), a maximum of five receptors covering the same intensity range appears to be a realistic estimate. Interestingly, similar numbers for information saturation have been found for peripheral neurons in other sensory systems (Warland et al., 1997; Stanley et al., 1999).

Because the grasshopper calling songs are matched to the properties of single auditory receptor neurons, it is natural to ask whether the receptors are optimally adapted to the environment at the population level. The distribution of receptor thresholds is species dependent; hence, the thresholds could reflect the adaptation of a given grasshopper species to a specific environment. Although a thorough evaluation awaits further study, one fact is already clear: the population of spike trains is less efficient in the Shannon-information sense than a single spike train. The representation of an LMD stimulus recoverable from a single spike train captures up to 40% of the theoretical maximum information about the stimulus that the spike train could possibly convey. A set of eight receptor spike trains yields only up to 8% of the corresponding theoretical maximum, because receptor neurons encoding the same intensity range do not spike independently of each other.

The high cell-to-cell redundancy in the response to stimuli with large modulation depth suggests that synchronous response patterns are caused solely by stimulus locking and do not carry additional information based on the relative timing of different spike trains. Synchronicity may nevertheless serve as an important signal for downstream neurons and improve their capability to reliably detect rapid stimulus onsets. In addition, a high cell-to-cell redundancy endows the system with fault tolerance with respect to failure of individual receptors.

With appropriate preprocessing of the stimulus, the linear stimulus reconstruction from the response of a single neuron can be interpreted as an estimate of the firing rate of that neuron. The high values obtained for signal-to-noise ratios, information rates, and coding efficiencies (Fig. 5) thus indicate that a firing-rate code suffices to recover large amounts of information. Our results do not suggest the use of a relational code, i.e., a code that involves higher-order moments of the spike train, as in interspike-interval-based coding schemes.

In closing, let us note that the application of systems analysis methods based on white-noise signals (Marmarelis and Marmarelis, 1978) to the auditory system has a long history; see Eggermont (1993) for an overview. When studying auditory receptor neurons of *L. migratoria*, Sippel and Breckow (1983) pointed out that the response properties depend strongly on the test stimuli used, which in their case were sinusoidal and Gaussian white-noise signals. Motivated by findings of this type, various recent studies have specifically addressed the role of natural-scene statistics for the processing of acoustic stimuli. For example, Rieke et al. (1995) showed that auditory afferents in bullfrogs transmit the most information about artificial stimuli with a spectrum that matches that of conspecific calls. For auditory receptor neurons of grasshoppers, it is not the spectrum but the amplitude distribution of the song that is important. The importance of amplitude distributions has also been demonstrated by Attias and Schreiner (1998) in the inferior colliculus of cats showing that highest information rates are reached for stimuli with an amplitude distribution that resembles that of natural sounds.

A close match of the response properties of a sensory system with the statistics of the most relevant natural stimuli has long been identified as a functionally important evolutionary design principle. The relevance of this principle for auditory systems has mainly been discussed with respect to the frequency domain (Suga, 1989). The studies of Rieke et al. (1995), Attias and Schreiner (1998), and Theunissen et al. (2000) suggest that also in the temporal domain, auditory coding has evolved to optimally exploit the regularities of behaviorally significant signals, most importantly conspecific communication signals. Our results on the neural processing of grasshopper calling songs demonstrate that the same principle holds even in an insect system. Together, these studies imply that any quantitative analysis of computations performed by auditory systems requires a proper understanding of the statistics of natural stimuli.

REFERENCES

Adam L-J (1977) The oscillating summed action potential of an insect's auditory nerve (*Locusta migratoria*, Acrididae). II. Underlying spike pattern and causes of spike synchronization. Biol Cybern 28:109–119.

Arfken G (1985) Mathematical methods for physicists. San Diego: Academic.

Attias H, Schreiner CE (1998) Coding of naturalistic stimuli by auditory midbrain neurons. In: Advances in neural information processing systems 10 (Jordan MI, Kearns MJ, Solla SA, ed), pp 103–109. Cambridge, MA: MIT.

Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. Science 252:1854–1857.

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

Eggermont JJ (1993) Wiener and Volterra analysis applied to the auditory system. Hear Res 66:177–201.

Ehret G, Romand R (1997) The central auditory system. New York: Oxford UP.

Elsner N (1974) Neuroethology of sound production in gomphocerine grasshoppers (Orthoptera: Acrididae). J Comp Physiol 88:67–102.

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

Haykin S (1994) Communication systems. New York: Wiley.

Jacobs K, Otte B, Lakes-Harlan R (1999) Tympanal receptor cells of *Schistocerca gregaria*: correlation of soma positions and dendrite attachment sites, central projections and physiologies. J Exp Zool 283:270–285.

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

Levelt WJM (1993) Speaking: from intention to articulation. Cambridge, MA: MIT.

Li ST, Hammond JJ (1975) Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients. IEEE Trans Syst Man Cybern 4:557–561.

Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems. New York: Plenum.

Meyer J, Elsner N (1996) How well are frequency sensitivities of grasshopper ears tuned to species-specific song spectra? J Exp Biol 199:1631–1642.

Michelsen A (1971) The physiology of the locust ear. Z Vergl Physiol 71:49–128.

Poor HV (1994) An introduction to signal detection and estimation. New York: Springer.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in C. New York: Cambridge UP.

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 109:101–122.
Ronacher B, Römer H (1985) Spike synchronization of tympanic receptor fibres in a grasshopper (*Chorthippus biguttulus* L., Acrididae). J Comp Physiol [A] 157:631–642.
Ronacher B, Krahe R (2000) Temporal integration versus parallel processing: coping with the variability of neuronal messages in directional hearing of insects. Eur J Neurosci 12:2147–2156.
Sippel M, Breckow J (1983) Non-linear analysis of the transmission of signals in the auditory system of the migratory locust *Locusta migratoria*. Biol Cybern 46:197–205.
Stanley GB, Li FF, Dan Y (1999) Reconstruction of natural scenes from ensemble response in the lateral geniculate nucleus. J Neurosci 19:8036–8042.
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 (1989) Principles of auditory information-processing derived from neuroethology. J Exp Biol 146:277–286.
Theunissen FE, Miller JP (1995) Temporal encoding in nervous system: a rigorous definition. J Comput Neurosci 2:149–162.
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.
Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained by using natural sounds. J Neurosci 20:2315–2331.
von Helversen D (1972) Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke *Chorthippus biguttulus*. J Comp Physiol 81:381–422.
von Helversen D, von Helversen O (1998) Acoustic pattern recognition in a grasshopper: processing in the time or frequency domain? Biol Cybern 79:467–476.
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.
Warland DK, Reinagel P, Meister M (1997) Decoding visual information from a population of retinal ganglion cells. J Neurophysiol 78:2336–2350.
Wessel R, Koch C, Gabbiani F (1996) Coding of time-varying electric field amplitude modulations in a wave-type electric fish. J Neurophysiol 75:2280–2293.