Computational principles underlying the recognition of acoustic signals in insects

Jan Clemens · R. Matthias Hennig

Abstract Many animals produce pulse-like signals during acoustic communication. These signals exhibit structure on two time scales: they consist of trains of pulses that are often broadcast in packets—so called chirps. Temporal parameters of the pulse and of the chirp are decisive for female preference. Despite these signals being produced by animals from many different taxa (e.g. frogs, grasshoppers, crickets, bushcrickets, flies), a general framework for their evaluation is still lacking. We propose such a framework, based on a simple and physiologically plausible model. The model consists of feature detectors, whose time-varying output is averaged over the signal and then linearly combined to yield the behavioral preference. We fitted this model to large data sets collected in two species of crickets and found that Gabor filters—known from visual and auditory physiology—explain the preference functions in these two species very well. We further explored the properties of Gabor filters and found a systematic relationship between parameters of the filters and the shape of preference functions. Although these Gabor filters were relatively short, they were also able to explain aspects of the preference for signal parameters on the longer time scale due to the integration step in our model. Our framework explains a wide range of phenomena associated with female preference for a widespread class of signals in an intuitive and physiologically plausible fashion. This approach thus constitutes a valuable tool to understand the functioning and evolution of communication systems in many species.

Keywords Perceptual decision making · Insect · Song · Linear-nonlinear model · Gabor filter

J. Clemens (✉) · R. M. Hennig
Behavioral Physiology Group, Department of Biology,
Humboldt-Universität zu Berlin, 10115 Berlin, Germany
e-mail: clemensjan@googlemail.com

R. M. Hennig
e-mail: matthias.hennig@biologie.hu-berlin.de

J. Clemens
Bernstein Center for Computational Neuroscience Berlin,
10115 Berlin, Germany

Present Address:
J. Clemens
Princeton Neuroscience Institute, Princeton University,
Princeton, NJ 08544, USA

1 Introduction

Throughout the animal kingdom acoustic signals are used for intra-specific communication. In many insects, acoustic signals play a key role for mate finding and selection. Sexual selection and speciation are the major factors shaping acoustic signals. Hence, understanding the computational principles underlying the processing of these courtship signals is a key to understand the principles underlying the evolution of communication systems (Phelps and Ryan 1998; Akre et al. 2011).

The temporal patterns of many communication signals convey information on multiple time scales. For instance, in human speech, the voicing and the transition between formants occupy power between 30 and 50 Hz while syllables are produced at a rate of 4–7 Hz (Giraud and Poeppel 2012). Although the songs of many insects and frogs exhibit a much simpler, pulse-like pattern, they too have a multi-scale

temporal structure (Otte 1992; Gerhardt and Huber 2002) (Fig. 1a). Trains of stereotyped pulses at a rate between 10–200 Hz are produced in packets—so-called chirps or bouts—at a rate of up to 5 Hz. This principal structure is shared by many insect groups like crickets (Alexander 1962; Otte 1992), bush crickets (Schul 1998), grasshoppers (von Helversen 1972; Safi et al. 2006), fruit flies (Hoy et al. 1988) as well as frogs (Gerhardt and Huber 2002).

Insects display a stereotyped, easily quantifiable and highly selective behavior for parameters of the pulse and the chirp (Fig. 1b and c). Behavioral experiments,
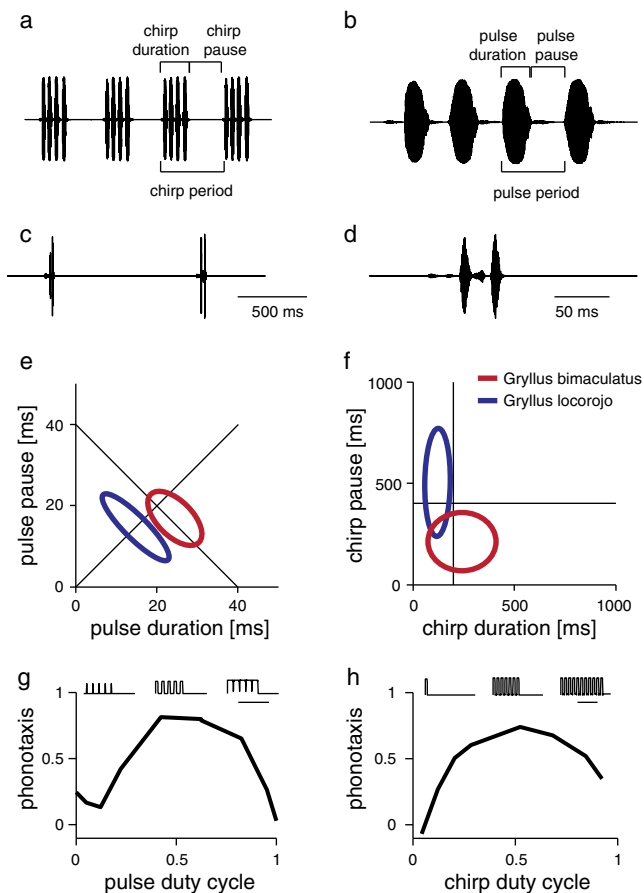
combined with electrophysiological studies have provided insights to song pattern recognition in some species (Webb et al. 2007; Creutzig et al. 2010; Hennig 2003; Schmidt et al. 2008; Kostarakos and Hedwig 2012; Zorovic and Hedwig 2011). However, we still lack a general understanding of the computational principles underlying the evaluation of this widespread class of signals. Most progress is confined to the short time scale of the pulse, while only few studies have addressed the processing of the chirp structure on the long time scale (but see Grobe et al. 2012). Here, we propose a computational framework for processing in sensory pathways that is able to explain behavioral selectivity on both time scales.

We applied a simple and generic model of a perceptual decision making system to the problem of evaluating pulse-like patterns (Fig. 2a). We employed linear-nonlinear models as feature detectors that are widely used to model the input-output function of sensory neurons in insects and vertebrates (Clemens et al. 2012; Machens et al. 2001; Nagel and Doupe 2006). The output of these linear-nonlinear models is integrated over the stimulus and linearly weighted to yield a scalar value equivalent to the behavioral response. We successfully fitted this model using preference data for two species of crickets. Generalizing the structure of these models revealed that it can reproduce the majority of principal preference functions known in insects.
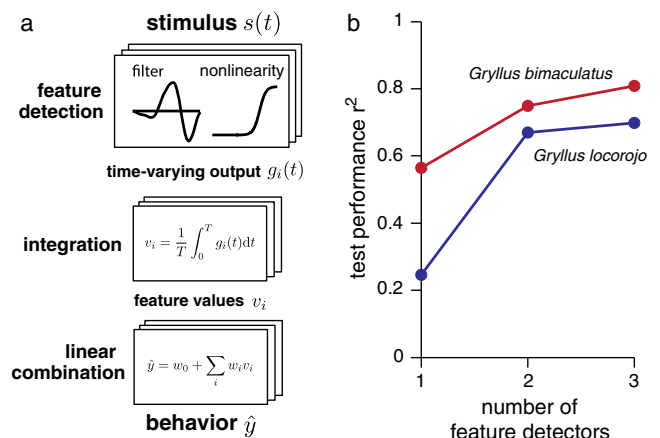

Fig. 1 Song patterns and behavioral preference for two species of crickets **a–d** Song of *Gryllus bimaculatus* (**a, b**) and *Gryllus locorojo* (**c, d**) with the chirp structure (**a, c**) and a close-up of one chirp showing the pulse structure (**b, d**). Horizontal bars in (**c**) and (**d**) indicate the time scale for plots (**a, c**) and (**b, d**), respectively. **e, f** Schematic representation of the preferred parameters of the pulse (**e**) and the chirp (**f**) for *Gryllus bimaculatus* (*red ellipses*) and *Gryllus locorojo* (*blue ellipses*). The diagonal line in (**e**) corresponds to a duty cycle of 0.5 (equal pulse duration and pause). The anti-diagonal line in (**e**) marks patterns for which the sum of the pulse and the pause is 40 ms and thus correspond to a pulse period of 40 ms. Lines in (**f**) correspond to a chirp duration of 200 ms and a chirp pause of 400 ms respectively. **g, h** Phonotaxis values to songs with varying pulse and chirp duty cycle for *Gryllus bimaculatus*. Stimuli with small, intermediate and high pulse (**g**) or chirp (**h**) duty cycles are shown on top (small horizontal bars in (**g**) and (**h**) correspond to 200 ms)

Fig. 2 Model structure and dependence of test performance on the number of feature detectors. **a** The model consisted of three stages. First, the signal's envelope $s(t)$ was processed by a bank of feature detectors (i = 1–3) which were modeled as linear-nonlinear units. A short linear filter (duration 64 ms) was followed by a static, sigmoidal nonlinearity. Second, the output of each feature detector, $g_i(t)$, was integrated over time and thereby reduced to a feature value v_i. Third, the set of feature values—one for each feature detector per stimulus—was linearly weighted to yield a scalar behavioral response score $\hat{y}$. **b** Dependence of model performance on the number of feature detectors in the model. For both species, two features saturated test performance

2 Methods

2.1 Preference data

Behavioral preferences were estimated on a walking compensator (for details see Hennig 2009). The behavioral response was measured as the directionality of the walk of the female towards the loudspeaker and was expressed as a normalized phonotaxis value (Schul 1998). A phonotaxis value of 0.8–1.0 corresponded to the female's directionality in response to a pattern similar to the song of a conspecific male (positive control). Scores of 0.0 corresponded to undirected phonotaxis as it is usually observed during silence or when a pure tone with the conspecific carrier frequency was presented (negative control). For fitting the model, we averaged the phonotaxis values of 20–90 animals. The standard deviation of phonotaxis values between animals was around 0.2.

The data sets consisted of 489 stimuli for *Gryllus bimaculatus* and 148 stimuli for *Gryllus locorojo*. The carrier frequencies were chosen to match the carrier frequency of conspecific calling songs (4.5 kHz for *Gryllus bimaculatus*, 5.0 kHz for *Gryllus locorojo*). The envelopes covered a large range of attractive and unattractive patterns. The majority of them were pulse-like patterns with a binary amplitude distribution and exhaustively sampled the space of pulse and chirp parameters (see e.g. Fig. 1g, h). Some patterns were constructed by superposing sinusoids with different frequencies yielding patterns with complex modulations. The stimuli and corresponding behavioral responses are described in detail in Hennig (2009), Schneider and Hennig (2012) and Grobe et al. (2012) for *Gryllus bimaculatus*, and in Rothbart and Hennig (2012a) for *Gryllus locorojo*.

For fitting the models we restricted the data sets to reduce computation time and improve performance of the fitting algorithm. We chose stimuli with a chirp duty cycle between 0.1 and 0.67 and chirp periods shorter than 600 ms. For *Gryllus locorojo* we imposed no constraint on the chirp duty cycle but excluded stimuli with a chirp period longer than 1000 ms. The stimulus sets used for fitting the models had a size of 316 for *Gryllus bimaculatus* and 100 for *Gryllus locorojo*.

While the calling songs of crickets may last for many hours, the great stereotypy and temporal redundancy of the songs allowed us to consider only a single chirp period for fitting. By setting the boundary conditions for the convolution appropriately, we emulated a situation were this chirp was embedded in the middle of a long stretch of song.

2.2 Model structure

The envelope of the song formed the input to the model. The envelope was extracted from the raw waveform by the root mean square method, transformed to a logarithmic decibel scale and thresholded at 35 dB.

This input, $s(t)$, was processed by up to three parallel feature detectors, which were implemented as linear-nonlinear models. The signal was filtered $f_j(t) = \int_{-\infty}^{\infty} s(\tau)h_j(t-\tau)d\tau$ and subsequently transformed with a sigmoid nonlinearity $g_j(t) = 1/(1 + \exp(-a_j f_j(t) + b_j))$. The subscript j indexes the feature detector, h_j is the linear filter, f_j the filtered stimulus, a_j is the steepness of the sigmoidal, b_j is the input for which the nonlinearity assumes its half-maximal value, g_j the output of the feature detector.

The output of each feature detector was integrated over the whole stimulus and thereby reduced to one feature value, v_j, per feature detector: $v_j = 1/T \int_0^T g_j(t)$. The feature values were then linearly weighted to yield a predicted preference value: $\hat{y} = w_0 + \sum_j w_j v_j$ for the stimulus.

2.3 Model fitting and evaluation

The envelopes in each data set were normalized such that the full stimulus set had zero-mean and unit-variance. Parameters for the feature detectors—the linear filter h_j and the parameters for the nonlinearity a_j and b_j—were learned using a genetic algorithm (Mitchell 1998). Filters h_j had a duration of 64 ms and were represented as a weighted sum of 16 raised-cosine basis functions (Pillow et al. 2008). This sped-up training time by reducing stimulus dimensionality and enforced smooth filters. The weights w_j for the individual features were determined by standard linear regression. Thus, each feature detector was described by 18 parameters—16 coefficients for the filter and two parameters for the non-linearity—and one feature weight in the weighting stage. The models containing two feature detectors discussed below thus have of $2 \cdot 19$ parameters plus a bias term w_0.

Model performance was quantified as the coefficient of correlation between the behavioral responses and the model output by leave-one-out cross-validation. For analysis of the model structure (Fig. 3a–d, f), we trained the model on the full data set. This yielded a structure and performance similar to that obtained in the majority of cross-validation runs.

2.4 Visualization of 2D response profiles

To produce the response surfaces in Fig. 5a and b, we used the Matlab function `TriScatteredInterp` with natural neighbor interpolation. The temporal stimulus parameters pulse period and global duty cycle were determined from the envelopes of the stimuli used for fitting the model. Pulse period was given by the interval between the onset of subsequent pulses in a chirp. Global duty cycle was the integral of the envelope, normalized by its length. Chirp period
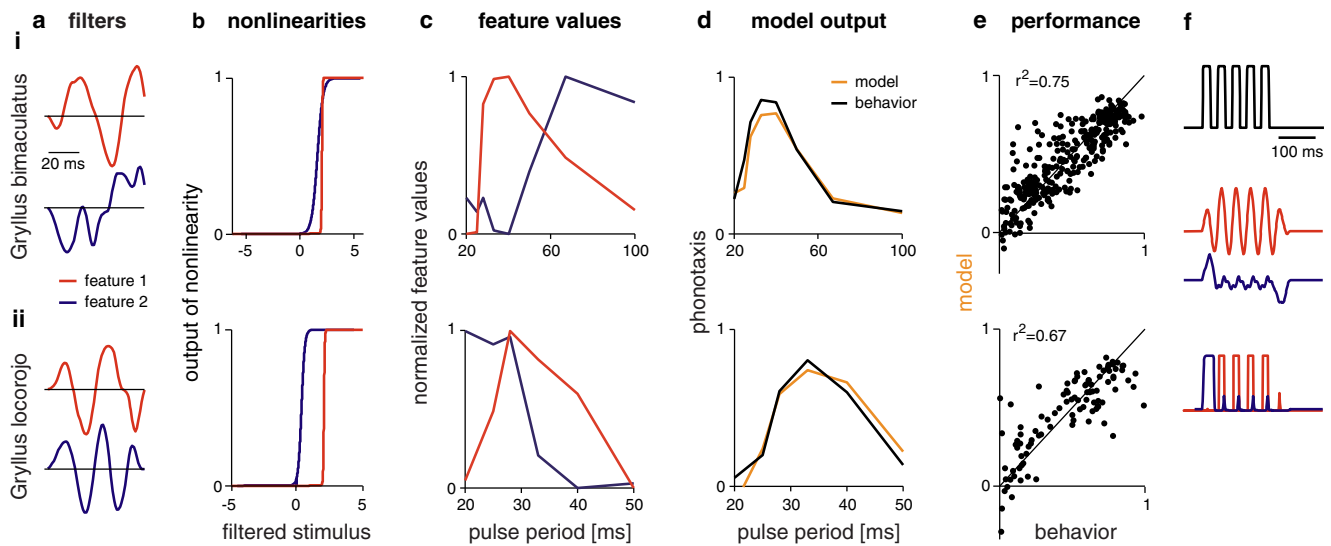
Fig. 3 Structure and processing properties of the models fit to experimental data. Rows correspond to the different species: **i** *Gryllus bimaculatus*, **ii** *Gryllus locorojo*. **a** Shape of the filters associated with the first (*red*) and second (*blue*) feature detector. **b** Nonlinearities associated with the first (*red*) and second (*blue*) feature detector. **c** Transfer functions for the two features for pulse period. Feature values are normalized to a range between 0 and 1. Pulse duty cycle of the signals was 0.5. The chirp parameters were chosen to be optimal for the respective species. **d** Transfer functions for behavior (*black*) and model predictions (*orange*). Same stimuli as in (**c**). **e** Behavioral and modeled phonotaxis values for all stimuli in the data set. **f** Processing of an exemplary stimulus by the model for *Gryllus bimaculatus*. top: stimulus, middle: stimulus filtered by the first (*red*) and the second (*blue*) filter, bottom: filtered stimulus after nonlinearity. Averaging these last traces over the full stimulus yields feature values of 0.18 and 0.11, respectively. The feature values are then linearly combined according to: $0.55 + 2.38 \cdot 0.18 - 1.89 \cdot 0.11$, to yield a preference value of 0.77 which is close to the measured one of 0.81. Note the negative weight for the second feature

was taken as the interval between the onset of adjacent chirps. We excluded stimuli for which temporal parameters could not be unequivocally defined, e.g. due to complexly modulated envelopes or small modulation depths.

3 Results

The preference data of two species of crickets, *Gryllus bimaculatus* and *Gryllus locorojo* (Weissman et al. 2012), were used for fitting the linear-nonlinear models. The calling song of both species consists of pulses with a characteristic duration and pause. These pulses are produced in packets, called chirps, which themselves have a species-specific duration and pause (Fig. 1a–d). This stereotyped structure can be described compactly in a four dimensional parameter space, comprising the duration and pause of the pulse and the duration and pause of the chirp. From these basic parameters, two further important parameters can be derived. First, the period which is given by the sum of the duration and the pause and which corresponds to the pulse or chirp rate; and second the duty cycle, which measures the duration of the pulse or chirp relative to the period and is related to the signal's energy.

The selectivity for the pulse parameters in the two species considered here was most compactly described as a preference for a narrow range of pulse periods (Fig. 1c, Rothbart and Hennig 2012a; Hennig 2009; Grobe et al. 2012). In addition to pulse period, pulse duty cycle was also important (Fig. 1g). While *Gryllus bimaculatus* preferred songs with a duty cycle greater than 0.5, *Gryllus locorojo* was less selective for duty cycle. Likewise, the two species exhibited specific preferences for parameters on the time scale of the chirp. (Fig. 1c, h). *Gryllus bimaculatus* mainly preferred short chirp pauses, while *Gryllus locorojo* responded best to short chirp durations and longer chirp pauses.

3.1 Model for song recognition in two species of crickets

To derive principles underlying the preference for parameters of the songs on both time scales, we fitted a simple model of a perceptual decision-making system (Fig. 2a) to preference data for *Gryllus bimaculatus* and *Gryllus locorojo*.

In both species, models with two feature detectors clearly outperformed models with a single filter, while inclusion of a third filter barely increased performance (Fig. 2b). For further analysis, we used models with two feature detectors,

which explained the preference functions very well (Fig. 3e, $r^2 = 0.75$ and 0.67, respectively). We will now briefly canvas the structure of the model for each species, show the structure of the two feature detectors and describe their role in reproducing the behavioral tuning for pulse.

3.1.1 Linear-nonlinear models for Gryllus bimaculatus

The model for the first species, *Gryllus bimaculatus*, exhibited two differently shaped filters (Fig. 3a (i)). The first filter resembled a Gabor function, which can be constructed by multiplying a sinusoidal modulation and a Gaussian envelope. The underlying sine had a frequency of approximately 2 periods per 64 ms and responded strongest to a pulse period of 30–40 ms (Fig. 3c (i), *red*). The second filter's main mode was slower (1 period per 64 ms) and this filter accordingly preferred longer pulse periods (longer than 60 ms, Fig. 3c (i), *blue*). At the animal's preferred pulse period, the second filter produced a single peak at the onset of the chirp (Fig. 3f, *blue*). The nonlinearities for both filters were relatively similar, though that of the second filter was slightly shallower and more sensitive (Fig. 3b (i)).

In the model, the signal envelope was processed by these two filter-nonlinearity pairs in parallel (Fig. 3f). The time-averaged output of each of these feature detectors yielded a scalar feature value, which was linearly combined to the phonotaxis value. Consequently, the weights of each feature indicated whether a given feature was associated with increases or decreases in response and hence, whether that feature was "excitatory" or "suppressive". The first feature detector was associated with a positive weight of 2.38 and was excitatory. The second feature detector had a negative weight of -1.89 and was hence suppressive. Due to the higher pass band of the second filter (Fig. 3c (i), *blue*), responses to pulse periods longer than 60 ms were suppressed, yielding the narrow tuning seen in the behavior (Fig. 3d (i)).

Thus, in *Gryllus bimaculatus*, an excitatory feature established the basic band-pass tuning for pulse period (compare Fig. 1c). The second filter sharpened this tuning by suppressing responses to longer pulse periods.

3.1.2 Linear-nonlinear models for Gryllus locorojo

The model for the second species, *Gryllus locorojo*, exhibited two similar, Gabor-like filters (Fig. 3a (ii)). Both filters differed in their modulation frequency, with the first filter having a smaller modulation frequency than the second (2.0 periods and 2.5 periods per 64 ms, respectively). Accordingly, the tuning of both filters for pulse period differed: the first filter responded stronger to longer periods than the second filter (Fig. 3c (ii)). The nonlinearity associated with the

first filter had a higher threshold and a steeper slope than that of the second one (Fig. 3b (ii)).

As in *Gryllus bimaculatus*, the first filter was excitatory while the second filter was suppressive (weights 11.45 and -9.3, respectively). However, in contrast to the first species, in *Gryllus locorojo* the suppressive feature preferred shorter periods and hence sharpened the basic tuning established by the first feature by suppressing responses to short periods (Fig. 3c (ii) and d (ii)).

3.1.3 General features of linear-nonlinear models for both species

The results on the two species have shown that the simple model developed here reproduced behavior well (Fig. 3e). Tuning for pulse period was the outcome of the interaction of an excitatory and a suppressive feature (Fig. 3c). While one feature established the basic behavioral tuning, the other feature usually sharpened it. We found two principal types of filters (Fig. 3a): Gabor-like filters with a band-pass tuning for pulse period (both filters for *Gryllus locorojo*, first filter for *Gryllus bimaculatus*) and unimodal, low-pass filters with a clear preference for long pulse periods (second filter for *Gryllus bimaculatus*).

3.2 Modeling the diversity of preference functions for pulse in insects with Gabor filters

The majority of filters derived from the preference data in the two species of crickets were Gabor-like (Fig. 3a) and exhibited a band-pass tuning for pulse period (Fig. 3c, *red*). Gabor filters are known in many sensory systems and are the basic building block of models of the visual and auditory pathways of vertebrates and invertebrates (e.g. Priebe and Ferster 2012; Smith and Lewicki 2006). We sought to generalize the results obtained for *Gryllus bimaculatus* and *Gryllus locorojo* to explain the diversity of preference functions for pulse in other species. Many orthopterans exhibit preference functions similar to those found in the two species of crickets used for fitting the initial models but display different preferred periods and duty cycles (Fig. 4a (i), Hennig 2009; Rothbart and Hennig 2012a, b; Hennig 2003; Schul 1998), or resonant behavior (Fig. 4a (ii), Schul 1998; Bush and Schul 2005). Other species have preference functions with a fundamentally different shape, e.g. band-pass filters for pulse duration or pause duration (Fig. 4a (iii), Schul 1998; Hennig 2003), or high-pass filters for duty-cycle (Fig. 4a (iv), Schul 1998).

To better link the structure of the filters to that of the preferred stimulus as well as to explore whether Gabor filters can explain the biological diversity of preference functions we built a simplified version of our perceptual
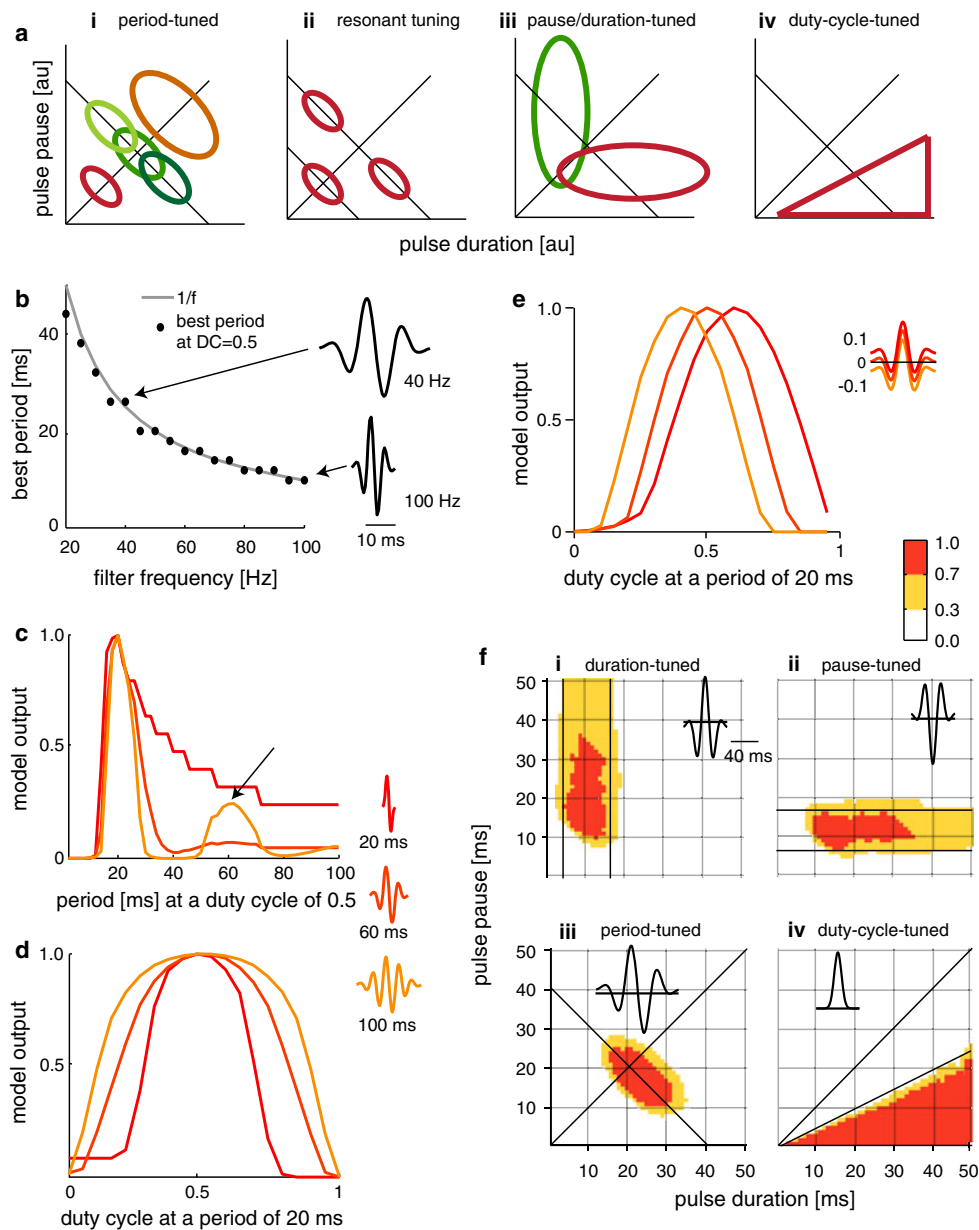
Fig. 4 Tuning properties of models based on Gabor-filters. **a** Diversity of preference functions for pulse parameters found in different species of insects. Ellipses depict schematically the range of preferred stimuli. Axes are in arbitrary units. (i) Variations in preferred pulse period (ellipses along the diagonal) and preferred duty cycle (green ellipses) among different species, (ii) resonant preference function (ellipses depict the range of preferred stimuli for a single species) (iii) pulse duration (*green*) and pulse pause (*red*) filters, (iv) high-pass filter for duty cycle. **b** Relation between modulation frequency of the filter and preferred period of the model. See pictograms to the right of the plot for example filters (phase 0 π, offset 0, duration $3/f$). Black dots depict simulation results, the gray line shows the theoretical value for the best frequency. The stimuli in all subplots exhibited a chirp period of 500 ms with a chirp duration of 250 ms. **c** Relation between filter duration and selectivity for pulse period. Shown are the phonotaxis values for stimuli with a pulse duty cycle of 0.5 and varying pulse periods for Gabor filters of 20, 60 and 100 ms duration (modulation frequency 50 Hz, phase 0 π, offset 0). The arrow indicates elevated responses to a multiple of the best period, indicating resonant selectivity. **d** Relation between filter duration and selectivity for pulse duty cycle at the best period for Gabor filters of 20, 60 and 100 ms duration (modulation frequency f=50 Hz, phase π, offset 0). **e** Influence of the offset of the Gabor filter on model output (modulation frequency f=50 Hz, phase $\pi/2$, duration $3/f$). **f** Model output for stimuli with different pulse durations and pauses. Output is color coded (see color bar). The filters producing the model output are shown as pictograms in each subplot with identical temporal scaling (see color bar)

decision-making system. It consisted of a single Gabor filter, and a piece-wise linear nonlinearity approximating a sigmoid with a threshold and a saturation. The integral output of such a feature detector was taken as being proportional to the phonotaxis value. Note that the reduced, single–filter model would exhibit reduced performance when probed with the full stimulus set used to fit the models in Fig. 3 (Fig. 2b). However, since in this section we were exclusively interested in reproducing the selectivity for pulse pattern in different species, we used stimuli with a fixed chirp structure.

We examined the impact of the Gabor filter's parameters on tuning for pulse duration, pause, period and duty cycle. The filters were parameterized as $exp(-(t/\sqrt{2}/\sigma)^2) \cdot \sin(2\pi f t + \phi) + \omega$. f was the frequency of the sinusoidal modulation. ϕ was the phase of the sinusoid. σ corresponded to the width of the Gaussian envelope and controlled the duration of the filter. ω was an offset. These reduced models thus consisted of 4 parameters for the filter plus 3 parameters for the piece-wise linear nonlinearity.

The modulation frequency f determined the preferred pulse period of the model, with slower modulations leading to longer preferred periods (Fig. 4b). Filter duration σ determined the sharpness of tuning for pulse period (Fig. 4c for f=50 Hz). For longer filters, resonant tuning arose in the form of additional peaks in the tuning curve (Fig. 4c, arrow). Resonant phenomena are known in song recognition of a species of bush crickets, *Tettigonia cantans* (Bush and Schul 2005). Note that these peaks need not necessarily be visible if the threshold of the nonlinearity following the filter is sufficiently high. In addition, there existed a trade-off between the sharpness of tuning along the period and the duty cycle axes (Fig. 4d). That is, while longer filters led to sharper tuning for pulse period, they yielded broader tuning for duty cycle (compare Figs. 4c and d).

Offset ω and phase ϕ also had an impact on the response of Gabor filters for pulse-like signals (Figs. 4e, f). Offset changed the preferred duty cycle from 0.5 for no offset ($\omega = 0$, Fig. 4e) to higher values for positive ($\omega = 0.1$) and smaller preferred duty cycles for negative offsets ($\omega = -0.1$). Thus, Gabor filters were not only able to reproduce the species-specific differences for pulse period (via the frequency f, Fig. 4b) but also for pulse duty cycle (via the offset ω). That is, the preference for duty cycles between 0.5 and 0.75 in *Gryllus bimaculatus* can be created by equipping the Gabor filter with a positive offset (compare Figs. 1c and 4f (iii)).

While offset changed the duty cycle preference, phase tended to change the extension of the preferred stimuli along the principal axes for pulse duration ($\phi = 0$) and pause ($\phi = \pi$) (Fig. 4f, (i) and (ii)). In conjunction with large absolute offsets, this yielded preference functions

which were tuned to a small range of durations or pauses. Similar response profiles are known in crickets (*Teleogryllus commodus*, Hennig 2003) and bushcrickets (*Tettigonia viridissima*, Schul 1998).

Gabor filters with a slow modulation relative to the filter duration acted as low-pass filters and created preference functions which corresponded to duty-cycle filters (Fig. 4f, (iv)). Such principal forms of preference are known in a species of bushcrickets, *Tettigonia caudata* (Schul 1998).

Our modeling results show that a simple Gabor-like filter is able to reproduce many tuning properties for parameters of the pulse's time scale found in crickets and other insects, including the preference for pulse period and duty cycle, resonant phenomena, and preference for pulse duration and pause. Furthermore, we have shown that there exists a simple and systematic relationship between parameters of the Gabor filter and the shape of the preference function.
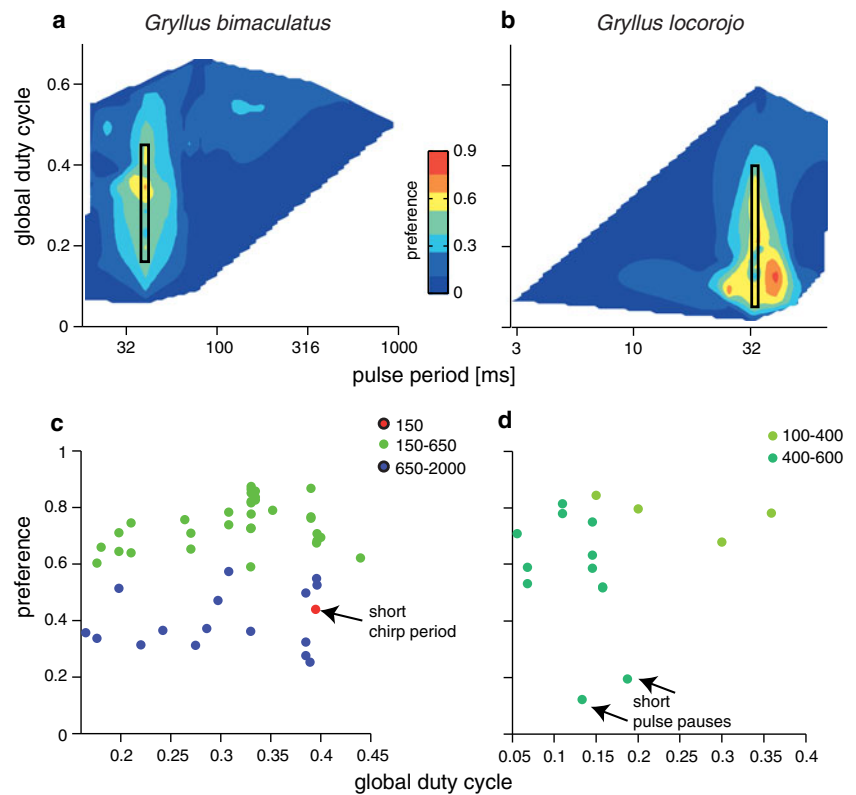
3.3 A new space for song signals

Above, we have shown that Gabor filters well reproduce most forms of preference for parameters of the short time scale of the pulse. Are Gabor filters also sensitive to aspects of the longer time scale of the chirp? Although the filters themselves were relatively short (64 ms) when compared to the time scale of the chirp, the integrating stage of our model also led to encoding of information about long time scales.

The nature of this information depended on whether the underlying filter was mainly differentiating or integrating. An ideal differentiating filter responded stereotypically at the onset of each pulse. Since its output was integrated over the full signal in our model, the resulting feature value was proportional to the number of pulses in the signal, irrespective of the duty cycle. The Gabor filters exhibited a mainly differentiating character through the modulating sinusoid. In contrast, a purely integrating filter responded proportionally to the stimulus' integral and thereby encoded the total energy or the global duty cycle of the song.

Thus, in addition to reproducing the selectivity for pulse period and pulse duty cycle conveyed by the Gabor filter stage (Fig. 4) the feature values also encoded the global duty cycle or overall pulse rate over the signal through the integration stage. Can this reduced signal space comprehensively and parsimoniously explain the preference in both species? We determined the pulse period and global duty cycle of our stimuli to visualize female phonotaxis values in this parameter space (Fig. 5a–b).

Selectivity for pattern in both species was described by the combination of two band-passes which aligned well with the principal axes "pulse period" and "global duty cycle" (Fig. 5a, b). However, some stimuli with optimal pulse period and global duty cycle displayed low phonotaxis values (Fig. 5c, d). For *Gryllus locorojo*, two outliers exhibited

Fig. 5 Filters for chirp duty cycle. **a**, **b** Response profiles in the feature space spanned by pulse period and global duty cycle for *Gryllus bimaculatus* (a) and *Gryllus locorojo* (b). Preference is color coded (see colorbar in (**a**)). **c–d** Preference values for stimuli with optimal pulse period and overall duty cycle (black outline in a and b) for both species. The chirp period is color coded (see legend right to each panel). Blue dots in (**c**) indicate stimuli with very long chirp periods (> 650 ms). Black arrows indicate outliers with pulse pauses below the temporal resolution of the animal (**d**)

a very short pulse pause (3 ms) (Fig. 5d, arrows) which was likely below the temporal resolution of the cricket's auditory system and hence prevented the detection of the correct pulse period. The outliers for *Gryllus bimaculatus* exhibited either very short or very long chirp periods (Fig. 5c, *blue* and *red* dots). Chirp period was a parameter which could not be extracted from the output of the filters as they only discriminated between different chirp duty cycles or pulse rates. Apparently there is a limitation for the correct prediction of patterns on very long time scales. Nevertheless, preference functions in the two species of crickets can be explained by selectivity for pulse period and global duty cycle over a wide range of patterns limited only by very short and very long time constants on either time scale.

4 Discussion

In this study we proposed a simple framework to understand the computational principles underlying the recognition of pulse-like songs as found in many species of insects and frogs. The framework based behavioral preference on the integral output of linear-nonlinear models. Starting by fitting this framework to preference data from two species of crickets (Figs. 1–3), we found that Gabor filters were surprisingly powerful in reproducing the principal shapes of the preference functions for pulse patterns

in many species of insects (Fig. 4). Adjusting the parameters of the Gabor filters allowed us to shift the preference function in a systematic way. In addition, the integrating nature of the framework introduced selectivity for the longer time scale of the chirp, for example by creating a dependence on the global duty cycle of the song. The resulting feature space spanned by pulse period and global duty cycle explained most of the preference in our data sets (Fig. 5).

4.1 Extensions of the model

Given that the model was relatively simple, it is surprising how well it was able to explain behavioral preference in two large data sets (Figs. 3, 5) and a great diversity of principal preference functions (Fig. 4). Extending the model to allow for more nonlinear feature detectors and integration of feature values might further increase the power of our framework.

First, there has been a great variety of modifications to the simple one-filter-one-nonlinearity model used as a feature detector in this study. Extensions to two-filter models with a full, two-dimensional nonlinearity have allowed to explain common, nonlinear response properties of sensory neurons like adaptation or sparseness (e.g. Fairhall et al. 2006; Atencio et al. 2008; Clemens et al. 2012). Incorporating these computations in the feature detection

step might allow for a more robust and more nonlinear feature detection.

Second, we assume that the integral output of the feature detectors is combined linearly to yield the predicted behavioral response. However, there are cases where cue combination is clearly nonlinear, e.g. if the presence of one feature vetoes the response to a stimulus (Ronacher and Stumpner 1988). Such interaction could be reproduced by nonlinear transformations of the feature values (e.g. $v_i \rightarrow v_i^2$), by including an interaction term ($v_i \cdot v_j$), or by a boolean operations like (v_i AND v_j) or (v_i XOR v_j).

Note that increasing model complexity comes at the price of requiring more data when fitting. However, incorporation of prior knowledge about the system under study could help reduce the amount of data needed, e.g. through parameterizing the filters (see Fig. 4) or the shape of the nonlinearity (Pillow and Simoncelli 2006).

4.2 Linear-nonlinear models, temporal selectivity and fusion of time scales

The building block of our framework for song recognition was the linear-nonlinear model. It served as a feature detector, the integral output of which was linearly related to behavioral preference.

Linear filters can be described as having the principal properties of integration or differentiation. An ideal integrating filter consists of a single lobe and acts as a low-pass filter on the signal (see Fig. 4f (iv)). Its integral output corresponds to the global duty cycle of the stimulus. In contrast, an ideal differentiating filter has a bimodal structure and acts as a high-pass filter. As it most strongly responds to transients in the signal, its integral output reflects the amount of onsets or offsets in a stimulus. For pulse-like signals, this corresponds to the pulse rate.

The filters found for the preference data (Fig. 3a) as well as the Gabor filters (Fig 4) displayed a mixture of integrating and differentiating properties, and thus often had band-pass like properties. On the short time scale of the pulse, this yielded pulse period filters (Figs. 3d and 4b, b). Extending the duration of the Gabor filter to span multiple periods (Fig. 4c) resulted in a resonant preference function known from a species of bush crickets (Bush and Schul 2005; Webb et al. 2007). Changing the offset of the Gabor filter—thereby rendering it more integrating—created filters for pulse pause and pulse duration that were relatively invariant to pulse period (Fig. 4f (i) and (ii)). Thus, in combining two principal types of filters, Gabor functions allowed the reproduction of the principal preference functions found in many insects and also in frogs (Fig. 4f).

The integration step in our framework added a long time scale and hence selectivity for aspects of the chirp to the model. This led to a fusion of the two time scales in the integral output of the linear-nonlinear models. Depending on whether the filter was mainly differentiating or integrating, the number of pulses or the global duty cycle of the signal was encoded by the feature detectors. Our analysis of the feature space spanned by these types of filters explained most aspects of preference for song parameters on the short and the long time scale (Fig. 5).

4.3 Predictions

The fusion of time scales through the integrating step in the model has consequences for the role of temporal aspects of song. Feature values depend both on local properties of the song—e.g pulse period and duty cycle (Fig. 4b–d)—and on average signal characteristics over the whole song—e.g. the global duty cycle (Fig. 5). The short filters used as feature detectors in the model (Fig. 3a) well matched features on the time scale of the pulse pattern and could thereby impose the necessity for a precise song pattern matching the shape of the filter. Hence, the model could be highly selective for parameters of the pulse like period or duty cycle (Fig. 4). In contrast, selectivity for global song parameters was the outcome of an averaging process and could not demand a precise pattern on this time scale. Accordingly, the model was much less selective for the precise structure on this long time scale—e.g. chirp duration or chirp pause (Fig. 5).

A surprising consequence of the integration is that the specific timing of pulse periods within a song is of little relevance. For instance, the species *Teleogryllus oceanicus* switches between two different pulse periods during a song. Experiments with artificial songs have tested whether the precise arrangement of these two pulse periods as found in natural songs is required by females (Pollack and Hoy 1979; Hennig and Weber 1997). Interestingly, females responded well to such shuffled songs as long as the duration and frequency of occurrence of the two pulse periods matched that of the natural songs. This is consistent with a model in which the occurrence of each of the two pulse periods is detected and counted by a dedicated Gabor filter.

Another corollary of the fusion of time scales by the integrating step is that female preference can be independent of the precise chirp structure. That is, in the absence of a preference for chirp period, the chirp structure could serve mainly the purpose of keeping the global duty cycle within the preference range of the receivers. Consistent with this hypothesis, there exist species which produce irregular chirps (Alexander 1957, 1962; Desutter Grandcolas and Robillard 2003). The fact that chirp structure is regular in most species could be an epiphenomenon of the oscillatory nature of the central pattern generators underlying song

production and need not necessarily reflect adaptation to receiver preference.

The new feature space relies on the fact that the global duty cycle is a relevant parameter (Fig. 5a–b). That is, preference did depend on the product of the duty cycle of the pulse and that of the chirp and not on their individual values. Then, a long duty cycle on one time scale could be compensated for by a short duty cycle on the other time scale, as long as the product of the two falls within the preferred range. For instance a male singing with a short pulse duty cycle could compensate for that by producing a longer chirp duty cycle.

4.4 Neural implementation

Our framework provides a phenomenological description of how the structure of pulse-like patterns determines behavior. That is, it is agnostic to a specific biophysical implementation. However, the framework of linear-nonlinear models makes basic but realistic assumptions about sensory pathways. Hence an implementation is conceivable, as the individual building blocks correspond to canonical computations which are found in many systems (Carandini and Heeger 2012). A linear filter well describes the temporal selectivity of sensory neurons. A threshold and saturation as implemented by the sigmoidal nonlinearity corresponds well to the nonlinear properties of the neuronal membrane. Integration of neuronal inputs over long time scales is known to be implemented through the accumulation of intracellular calcium or other mechanisms.

The filters themselves, namely their positive and negative lobes can be constructed with excitatory and inhibitory synaptic inputs, respectively. The delay between excitation and inhibition governs the frequency of the modulating sinusoid and therewith the preferred pulse period (compare Fig. 4b). The balance between both kinds of inputs determines the filter's offset: stronger relative excitation yields a filter with a positive offset while excess inhibition yields a negatively offset filter (compare Fig. 4f (i) and (ii)). A strong influence of inhibition on shaping the temporal tuning of auditory neurons in the brain of the cricket *Gryllus bimaculatus* has recently been proposed (Kostarakos and Hedwig 2012). The long, multi-cyclic Gabor filters producing resonant phenomena (Fig. 4c) are most likely implemented either in a recurrent network or at the single-cell level with resonant conductances (Webb et al. 2007; Schreiber et al. 2004).

In principle the observed models reflect the overall number of properties and operations in the auditory pathway of crickets. Therefore, the calculated preference values likely correspond to the result of network computations. However, the small–sized auditory pathway of most insects renders it feasible that there exist single cells that correspond to the output of our feature detector(s). Indeed, Kostarakos and Hedwig (2012) have found an auditory neuron in the brain of *Gryllus bimaculatus*, whose integral output was highly correlated to female preference. Interestingly, its response pattern resembled the output generated by the Gabor filter fit to behavioral data (Fig. 3f, *red*). This suggests that the Gabor filter in association with a sigmoidal nonlinearity well describes this cell's output, which in turn is likely the result of the network feeding into this cell. Along this line, Zorovic and Hedwig (2011) have described a cell which responds to the onset of the chirp, resembling the output of the second filter for *Gryllus bimaculatus* in our model (Fig. 3f, *blue*).

4.5 Conclusion

Our framework has provided a surprisingly simple computational principle for many aspects of female preference to pulse-like songs. Gabor filters well explained preference for parameters of the pulse while a preference for the chirp emerged from the integrating step of the model. The model revealed a direct relation between the parameters of Gabor filters and the shape of preference functions. Our approach thus provides a valuable tool that could explain the changes underlying female preference for communication signals during speciation events.

These results are likely to generalize to other species groups as well. Frogs produce similarly structured signals and preference functions. They may thus implement similar algorithms for their evaluation. The songs of birds as well as the speech of humans is much more complex and exhibits rich spectro-temporal structure. While the recognition systems is much more parallel and the filters are more complex, similar principles—feature detection and integration—might apply here as well.

References

Akre, K.L., Farris, H.E., Lea, A.M., Page, R.A., Ryan, M.J. (2011). Signal perception in frogs and bats and the evolution of mating signals. *Science*, *333*(6043), 751–752.

Alexander, R.D. (1957). The song relationships of four species of ground crickets (Orthoptera: Gryllidae: Nemobius). *Ohio Journal of Science*, *57*(3), 153–163.

Alexander, R.D. (1962). Evolutionary change in cricket acoustical communication. *Evolution*, *16*, 443–467.

Atencio, C.A., Sharpee, T.O., Schreiner, C.E. (2008). Cooperative Nonlinearities in Auditory Cortical Neurons. *Neuron*, *58*, 956–966.

Bush, S.L., & Schul, J. (2005). Pulse-rate recognition in an insect: evidence of a role for oscillatory neurons. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *192*, 1–9.

Carandini, M., & Heeger, D.J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62.

Clemens, J., Wohlgemuth, S., Ronacher, B. (2012). Nonlinear computations underlying temporal and population sparseness in the auditory system of the grasshopper. *Journal of Neuroscience*, *32*(29), 10,053–10,062.

Creutzig, F., Benda, J., Wohlgemuth, S., Stumpner, A., Ronacher, B., Herz, A.V.M. (2010). Timescale-Invariant Pattern Recognition by Feedforward Inhibition and Parallel Signal Processing. *Neural Computation*, *22*(6), 1493–1510.

Desutter Grandcolas, L., & Robillard, T. (2003). Phylogeny and the evolution of calling songs in Gryllus (Insecta, Orthoptera, Gryllidae). *Zoologica Scripta*, *32*(2), 173–183.

Fairhall, A.L., Burlingame, A.C., Narasimhan, R., Harris, R.A., Puchalla, J.L., Berry, M.J. (2006). Selectivity for multiple stimulus features in retinal ganglion cells. *Journal of Neurophysiology*, *96*, 2724–2738.

Gerhardt, C.H., & Huber, F. (2002). *Acoustic Communication in Insects and Anurans*. Chicago: University of Chicago Press.

Giraud, A.L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Grobe, B., Rothbart, M.M., Hanschke, A., Hennig, R.M. (2012). Auditory processing at two time scales by the cricket Gryllus bimaculatus. *Journal of Experimental Biology*, *215*(10), 1681–1690.

Hennig, R.M. (2003). Acoustic feature extraction by cross-correlation in crickets? *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *189*(8), 589–598.

Hennig, R.M. (2009). Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket Gryllus bimaculatus. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *195*(10), 971–987.

Hennig, R.M., & Weber, T. (1997). Filtering of temporal parameters of the calling song by cricket females of two closely related species: a behavioral analysis. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *180*(6), 621–630.

Hoy, R., Hoikkala, A., Kaneshiro, K. (1988). Hawaiian courtship songs: evolutionary innovation in communication signals of Drosophila. *Science*, *240*(4849), 217–219.

Kostarakos, K., & Hedwig, B. (2012). Calling song recognition in female crickets: temporal tuning of identified brain neurons matches behavior. *Journal of Neuroscience*, *32*(28), 9601–9612.

Machens, C.K., Stemmler, M., Prinz, P., Krahe, R., Ronacher, B., Herz, A.V.M. (2001). Representation of acoustic communication signals by insect auditory receptor neurons. *Journal of Neuroscience*, *21*(9), 3215–3227.

Mitchell, M. (1998). *An introduction to genetic algorithms (complex adaptive systems)* (3rd printing ed.). A Bradford Book.

Nagel, K.I., & Doupe, A.J. (2006). Temporal processing and adaptation in the songbird auditory forebrain. *Neuron*, *51*(6), 845–859.

Otte, D. (1992). Evolution of cricket songs. *Journal of Orthoptera Research*, *1*(1), 25–49.

Phelps, S.M., & Ryan, M.J. (1998). Neural networks predict response biases of female túngara frogs. *Proceedings of the Royal Society of London Series B*, *265*(1393), 279–285.

Pillow, J.W., & Simoncelli, E.P. (2006). Dimensionality reduction in neural models: An information-theoretic generalization of spike-triggered average and covariance analysis. *Journal of vision*, *6*, 414–428.

Pillow, J.W., Shlens, J., Paninski, L., Sher, A., Litke, A.M., Chichilnisky, E.J., Simoncelli, E.P. (2008). Spatio-temporal correlations and visual signaling in a complete neuronal population. *Nature*, *454*(7207), 995–999.

Pollack, G.S., & Hoy, R. (1979). Temporal pattern as a cue for species-specific calling song recognition in crickets. *Science*, *204*(4391), 429–432.

Priebe, N.J., & Ferster, D. (2012). Mechanisms of neuronal computation in mammalian visual cortex. *Neuron*, *75*(2), 194–208.

Ronacher, B., & Stumpner, A. (1988). Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *163*, 517–523.

Rothbart, M.M., & Hennig, R.M. (2012a). The Steppengrille (Gryllus spec./assimilis): Selective filters and signal mismatch on two time scales. *PLoS ONE*, *7*(9), e43975.

Rothbart, M.M., & Hennig, R.M. (2012b). Calling song signals and temporal preference functions in the cricket Teleogryllus leo. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *198*(11), 817–825.

Safi, K., Heinzle, J., Reinhold, K. (2006). Species recognition influences female mate preferences in the common European grasshopper (Chorthippus biguttulus Linnaeus, 1758). *Ethology*, *112*(12), 1225–1230.

Schmidt, A., Ronacher, B., Hennig, R.M. (2008). The role of frequency, phase and time for processing of amplitude modulated signals by grasshoppers. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *194*(3), 221–233.

Schneider, E., & Hennig, R.M. (2012). Temporal resolution for calling song signals by female crickets, Gryllus bimaculatus. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *198*(3), 181–191.

Schreiber, S., Erchova, I., Heinemann, U., Herz, A.V.M. (2004). Subthreshold resonance explains the frequency-dependent integration of periodic as well as random stimuli in the entorhinal cortex. *Journal of Neurophysiology*, *92*(1), 408–415.

Schul, J. (1998). Song recognition by temporal cues in a group of closely related bushcricket species (genus Tettigonia). *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *183*(3), 401–410.

Smith, E.C., & Lewicki, M.S. (2006). Efficient auditory coding. *Nature*, *439*(7079), 978–982.

von Helversen, D. (1972). Gesang des M'annchens und Lautschema des Weibchens bei der Feldheuschrecke Chorthippus biguttulus (Orthoptera, Acrididae). *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *81*(4), 381–422.

Webb, B., Wessnitzer, J., Bush, S.L., Schul, J., Buchli, J., Ijspeert, A. (2007). Resonant neurons and bushcricket behaviour. *Journal of Comparative Physiology A: Sensory Neural, and Behavioral Physiology*, *193*(2), 285–288.

Weissman, D.B., Gray, D.A., Pham, H.T., Tijssen, P. (2012). Billions and billions sold: Pet-feeder crickets (Orthoptera: Gryllidae), commercial cricket farms, an epizootic densovirus, and government regulations make for a potential disaster. *Zootaxa*, *3504*, 67–88.

Zorovic, M., & Hedwig, B. (2011). Processing of species-specific auditory patterns in the cricket brain by ascending, local and descending neurons during standing and walking. *Journal of Neurophysiology*, *105*, 2181–2194.