TEMPORAL ASPECTS OF THE PROCESSING OF CALLING SONGS IN ORTHOPTERA

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Gundula Meckenhäuser
from Celle

2014

The research presented in this dissertation was carried out from August 2010 until March 2014 at the Theoretical Neuroscience & Neuroinformatics group, Freie Universität Berlin, under the supervision of Prof. Dr. Martin Nawrot.

1st Reviewer:	Prof. Dr. Martin P. Nawrot - Freie Universität Berlin
2nd Reviewer:	Prof. Dr. R. Matthias Hennig - Humboldt-Universität zu Berlin

Date of defense: 19. 06. 2014

Acknowledgements

First of all, I would like to thank my supervisor Martin Nawrot for the opportunity to work in his interdisciplinary and multicultural research group, for his continuous support, encouragement and openness towards new projects during the past years.

I express my gratitude to my experimental collaborators: Matthias Hennig for introducing me to the fascinating world of crickets, for collecting the data for the first project and his clear thoughts for the manuscript; Stefanie Krämer for confiding me to her data from grasshoppers and her patience; Bernhard Ronacher for his ideas and support for finishing the second manuscript. Thanks for providing me with real neuroscientific data!

Thanks to all my colleagues. In particular, I want to thank Farzad Farkhooi without whom I would have been stuck in the last year: thanks for the numerous encouraging discussions and helpful remarks on my second project. I also thank Jan Soelter for discussing artificial neural networks over and over again, Thomas Rost for installing packages I could not install, Chris Häusler for his relaxing mood, Michael Schmuker for taking care of Bommel, Evren Pamir for his critical view, Rinaldo Betkiewicz for not loosing control in the polish Milchbar, Joachim Haenicke for his handcreme, and Tara Dezhdar for being a great office mate.

I thank Yulia Oganian and Joachim Haenicke how helped to polish this thesis during the final weeks and Florian Rau who never got tired of correcting my writings about crickets and grasshoppers.

Many thanks go to my “ladies” from Berlin and friends from Hamburg for sharing laughter and tears. Thank you so much for your moral support and love.

I dedicate this thesis to my parents Helga Meckenhäuser and Wolfram Seidel to whom I am so grateful for their endless support and encouragement.

This dissertation is based on the following two manuscripts:

Critical song features for auditory pattern recognition in crickets

Authors:
Gundula Meckenhäuser[1], R. Matthias Hennig[2], Martin P. Nawrot[1]
Author contributions:
GM, RMH, MPN conceived the reseach idea. RMH designed the experiment.
GM analyzed the data. GM, MPN wrote the manuscript. RMH revised the manuscript.
Manuscript status:
Manuscript has been published in PLoS ONE: doi: 10.1371/journal.pone.0055349

Decoding of calling songs and their behavioral relevance from grasshopper auditory neurons

Authors:
Gundula Meckenhäuser[1,*], Stefanie Krämer[2,*], Farzad Farkhooi[1], Bernhard Ronacher[2], Martin P. Nawrot[1]
Author contributions:
SK, BR designed the experiments. SK carried out the experiments. GM, FF, MPN developed ideas for data analysis. SK analyzed behavioral data. GM analyzed behavioral and electrophysiological data. GM, SK, BR, MPN wrote the manuscript. FF revised the manuscript.
Manuscript status:
Manuscript has been submitted to Frontiers in Systems Neuroscience on March, 14th, 2014.

Author affiliations:
1 Theoretical Neuroscience & Neuroinformatics, Institute of Biology, Department of Biology, Chemistry and Pharmacy, Freie Universität Berlin
2 Behavioural Physiology Group, Department of Biology, Humboldt-Universität zu Berlin
$*$ equal contribution

Zusammenfassung

Zentraler Untersuchungsgegenstand von *computational neuroscience* sind die Verrechnungsprinzipien, welche der neuronalen Informationsverarbeitung zugrunde liegen. In dieser Doktorarbeit werden experimentelle Daten mit Methoden des maschinellen Lernens analysiert um zu einem besseren Verständnis der Verarbeitung von Balzgesängen in Grillen und Grashüpfern beizutragen.

Welche Charakteristika der Balzgesänge von Grillen sind ausschlaggebend für die Erkennung und Beurteilung potentieller Paarungspartner? Kapitel 2 untersucht diese Fragestellung durch Analyse von Verhaltensdaten mit künstlichen neuronalen Netzen. Es werden Modelle präsentiert die das phonotaktische Verhalten der Weibchen auf Basis beschreibender Größen eines Gesanges quantitativ vorhersagen. Diese Vorhersagen ermöglichen Teilmengen von mehreren Charakteristika zu identifizieren welche die Verhaltensdaten am besten beschreiben.

Wie sind Informationen eines Balzgesanges, z.B. die Identität oder die Attraktivität, in den neuronalen Antworten auditorischer Neurone in Grashüpfern kodiert? Um diese Frage zu beantworten werden in Kapitel 3 Verhaltensdaten und elektrophysiologische Daten mit Bayes Klassifikatoren untersucht. Es wird gezeigt, dass Informationen über den Gesang in der Anzahl der Aktionspotentiale von Populationen auditorischer Neurone kodiert ist.

Summary

The central aim of computational neuroscience is to understand the computational principles underlying the neuronal processing of information. In this thesis experimental data is analyzed with methods from machine learning to contribute to a better understanding of the processing of calling songs in crickets and grasshoppers.

Which features of cricket calling songs are critical for the recognition and evaluation of a potential mating partner? Chapter 2 investigates this question by analyzing a large body of behavioral data recorded from female crickets during phonotactic experiments with artificial neural networks. Models are presented that quantitatively predict the experimental measure of phonotactic behavior for a given set of feature values from calling songs. The model predictions allow to identify minimal feature sets that best describe the behavioral data.

How is information about a calling song, i.e. the song identity and its attractiveness, encoded in the neuronal pattern of auditory neurons in grasshoppers? To answer this question behavioral as well as electrophysiological data of auditory neurons are analyzed with naïve Bayes classifiers in chapter 3. It is shown that information about a stimulus is encoded in the spike count of populations of neurons.

Keywords:
insect acoustic communication, neural information processing, pattern recognition, artificial neural network, naïve Bayes classifier

Contents

Chapter 1

General introduction

As in many insects, acoustic communication plays a key role for mating in Orthoptera [Bradbury and Vehrencamp, 1998]. In the cricket species *Gryllus bimaculatus* males produce calling songs by rubbing their wings and females use these songs to localize potential mating partners. If a female rates a song, and thereby the singer, as attractive she approaches the male, which eventually leads to the act of copulation [Hedwig, 2006]. In the grasshopper species *Chorthippus biguttulus* acoustic communication is bidirectional. Males produce calling songs by rubbing their hind legs across a vein on the forewings and a calling song is answered by a female who is then approached by the male [von Helversen and von Helversen, 1997]. Orthoptera offer several advantages as model systems to study how information about acoustic signals is processed in the brain. Research of the last decades provides convincing evidence that the decisive cues for the recognition of calling songs reside in the temporal pattern of amplitude modulations, i.e. the song's envelope, and not the spectral content (see von Helversen [1972]; von Helversen and von Helversen [1998] for grasshoppers and Grobe et al. [2012]; Hennig [2009] for crickets). The aim of this thesis is to highlight temporal aspects of the processing of calling songs in crickets and grasshoppers by analyzing different types of experimental data with methods of machine learning.

1.1 Methods of machine learning for analyzing neuroscientific data

In order to analyze so-called big data, machine learning algorithms [Bishop, 2006] make use of the increase in computational power that has taken place over the last years. In neuroscience, modern experimental methods generate large bodies of data that require novel computational approaches to fully grasp their complexity. This thesis shows that machine learning algorithms provide effective means to complement and guide such experiments.

Methodically, this thesis focuses on supervised learning. Supervised learning algorithms infer a function based on a known set of input data and known labels that allows to predict labels of new data. In chapter 2 a large set of calling songs labeled with the phonotactic score that reflects the behavioral responses of female crickets is analyzed with artificial neural networks (ANNs) [Rosenblatt, 1958]. ANNs are related to their biological counterparts in the way that they are represented as networks of interconnected units. The units can be thought of as neurons and the connections as synaptic weights. In feed-forward ANNs, where the units are ordered in layers with connections between the layers, the connections are typically learned with the supervised backpropagation algorithm [Riedmiller and Braun, 1993; Rumelhart et al., 1986]. After learning, new data is then forward propagated through the network to predict unknown labels. Although ANNs were originally inspired by the brain [McCulloch and Pitts, 1943], the degree to which they mirror functions of real neural circuits is not clear [Hinton, 2011].

The aim of chapter 2 is to predict the phonotactic score of untested calling songs, whereas chapter 3 tackles the reverse problem of decoding which stimulus evoked a particular neuronal response in grasshoppers. To this end, a set of neuronal spike patterns labeled with stimulus identity is analyzed with the naïve Bayes classifier (for reviews see Pouget et al. [2000]; Quiroga and Panzeri [2009]). In contrast to ANNs, naïve Bayes classifiers require a small amount of data to learn the necessary parameters and are extremely fast.

Unsupervised learning techniques are used to find structures in unlabeled data or as ways to preprocess data before classification or regression. In chapter 2, before learning the ANN, a principal component analysis (PCA) [Ringnér, 2008] was applied to remove linear correlations across the temporal features of the calling songs.

1.2 Crickets and grasshoppers as model systems

Orthoptera are ideal model systems to study the neuronal processing of acoustic signals, because the stimulus space is clearly defined and their auditory system is relatively small and, to a certain level, easily accessible.

The auditory system of crickets and grasshoppers evolved mainly for the detection of echolocation sounds of predatory bats and for the recognition and localization of conspecific signals from potential mating partners [Stumpner and Helversen, 2001]. The information content of an acoustic signal can be represented by its spectral content and the temporal pattern of amplitude modulations, i.e. the envelope. The animals make use of the clear difference between the carrier frequencies of conspecific signals and ultrasonic sounds of predators to discriminate between good and evil. Some species also use the spectral content to distinguish between male and female [von Helversen and von Helversen, 1997] or to estimate the distance to the singer [Römer, 1987]. The focus of this thesis is on the recognition of conspecific songs, for which the decisive features reside in the temporal pattern of amplitude modulations (see von Helversen [1972]; von Helversen and von Helversen [1998] for grasshoppers and Grobe et al. [2012]; Hennig [2009] for crickets). Using artificial song models that mimic natural patterns of calling songs allows to test behavioral (cf. data in chapters 2 and 3) and neuronal (cf. data in chapter 3) responses under laboratory conditions.

Another reason that supports Orthoptera as model systems is that their auditory systems are small in comparison to vertebrates, such as mammals that have billions of neurons [Williams and Herrup, 1988] for the processing of sensory information. The auditory network of crickets, which are subject of chapter 2, is organized in a feed-forward structure: The tympanal organs of crickets, which are located on the two front legs, are set into vibration by sound. Approximately 70 receptor neurons [Young and Ball, 1974] transduce these vibrations into spike trains that travel into the prothoracic ganglion. Here, the omega-neuron [Marsat and Pollack, 2005; Selverston et al., 1985] and two ascending neurons (ANs) [Schildberger, 1984] were identified as auditory interneurons which constitute the first processing level for auditory signals. These three neurons exist in each hemisphere and receive excitatory input from the receptor neurons. Ascending neurons transmit information to the brain, whereas the omega-neuron inhibits the contralateral ANs. The ascending neuron AN1 is most sensitive to low carrier frequencies [Schildberger, 1984] associated with conspecific songs whereas the AN2 is most sensitive to ultrasonic frequencies of predators [Marsat and Pollack, 2007]. In the brain, Schildberger [1984] and Kostarakos and Hedwig [2012] identified auditory neurons in the protocerebrum that are involved in the processing of conspecific songs.

The auditory pathway of grasshoppers also exhibits the feed-forward motif [Vogel and Ronacher, 2007]: At each side of the first abdominal segment resides an ear from where information about acoustic signals is transduced by more than 60 auditory receptor neurons into spike trains traveling into the metathoracic ganglion [Jacobs et al., 1999]. There, approximately 15 local interneurons constitute the first processing level and provide input to approximately 20 ascending neurons that project information to the brain [Römer and Marquart, 1984; Stumpner and Ronacher, 1991]. *In vivo* intracellular recordings of ascending neurons in grasshoppers are analyzed in chapter 3. The next processing level is the brain, where Hedwig and Heinrich [1997] identified auditory brain neurons in the mediolateral protocerebrum.

In cricket and grasshoppers, the decision centers located in the female brain must evaluate whether a perceived song follows the conspecific pattern and whether it is attractive enough to trigger a behavioral response. However, in comparison to ascending neurons, less is known about the processing of auditory information by brain neurons [Stumpner and Helversen, 2001].

1.3 Outline of the thesis

This thesis is divided into two self-contained manuscripts that both address different temporal aspects of the processing of calling songs in Orthoptera.

In chapter 2 I investigated which temporal features of calling songs critically determine the phonotactic performance of female crickets. The natural pattern of cricket calling songs consists of repetitive pulses that are grouped into pulse trains called chirps. Thus, they carry essential information on a short and a long time scale, defined by the pulses and chirps, somewhat in analogy to the words and phrases of human speech. The song pattern can be described with four independent parameters, for example the pulse duration, pulse pause, chirp duration, and chirp pause. However, it is not clear whether these four cues are analyzed independently in the cricket brain. The period, that is the sum of duration and pause, as well as the duty cycle, that is the ratio of duration and period, for both pulses and chirps have also been implicated as relevant descriptors [Doherty, 1985; Grobe et al., 2012; Hennig, 2009]. I applied ANNs to analyze a large set of behavioral data that consists of 218 different artificial calling songs and a phonotactic score that reflects the walking behavior of female crickets *Gryllus bimaculatus*. In a first step, I used four temporal features of calling songs to predict the phonotactic score of untested songs. The model predictions showed a high correlation with the experimentally measured results. I used this model to investigate the interaction of the two time scales defined by pulses and chirps. In a further step, I investigated all feature sets, each one consisting of different combinations of calling song features, and identified the most important temporal cues for the phonotactic performance of females.

In chapter 3 I studied how information about an acoustic song is encoded in the neuronal pattern of auditory neurons in grasshoppers. Typically, a calling song consists of a repetition of syllables with species-specific amplitude modulations of a broad carrier frequency band. In nature, female grasshoppers are often confronted with songs whose temporal structure is degraded, e.g. as a result of the overlap with other songs [Lang, 2000; Römer et al., 1989]. Behavioral experiments with *Chorthippus biguttulus* revealed that females accept songs with a perturbation in the beginning of a syllable, whereas syllables that were degraded in the middle or at the end lead to a rejection. This finding allowed to divide the stimulus set into attractive and unattractive songs. Applying naïve Bayes classifiers to *in vivo* intracellular recordings from ascending neurons from *Locusta migratoria*, I investigated how information about acoustic signals, i.e. the stimulus identity and its attractiveness, is encoded in the spike pattern of auditory neurons that project to the brain. Focusing on the simplest coding scheme, I put emphasis on spike count coding. First, I investigated the effect of the integration time of the neuronal activity in single neurons on the performance of the classifier. As the grasshopper brain receives input not only from one ascending neuron but from up to 20, I also decoded the stimulus identity and its attractiveness from the counts of populations of different sizes. Finally, I used a time-resolved firing rate profile for decoding the stimulus identity and its attractiveness in single neurons.

Chapter 2

Critical song features for auditory pattern recognition in crickets[1]

Many different invertebrate and vertebrate species use acoustic communication for pair formation. In the cricket *Gryllus bimaculatus*, females recognize their species-specific calling song and localize singing males by positive phonotaxis. The song pattern of males has a clear structure consisting of brief and regular pulses that are grouped into repetitive chirps. Information is thus present on a short and a long time scale. Here, we ask which structural features of the song critically determine the phonotactic performance. To this end we employed artificial neural networks to analyze a large body of behavioral data that measured females' phonotactic behavior under systematic variation of artificially generated song patterns. In a first step we used four non-redundant descriptive temporal features to predict the female response. The model prediction showed a high correlation with the experimental results. We used this behavioral model to explore the integration of the two different time scales. Our result suggested that only an attractive pulse structure in combination with an attractive chirp structure reliably induced phonotactic behavior to signals. In a further

---

[1] This chapter is the reprint version of Meckenhäuser et al. [2013]. Please refer to page v in this dissertation for the detials of authors contributions.

step we investigated all feature sets, each one consisting of a different combination of eight proposed temporal features. We identified feature sets of size two, three, and four that achieve highest prediction power by using the pulse period from the short time scale plus additional information from the long time scale.

2.1 Introduction

Acoustic communication plays a key role for mating behavior in many different species, most prominently in birds [Weiss et al., 2012], fish [Kastenhuber and Neuhauss, 2011], amphibians [Woolley et al., 2004], and insects [Hedwig, 2006; von Helversen and von Helversen, 1998; von Philipsborn et al., 2011]. In the cricket species *Gryllus bimaculatus* males produce calling songs by rubbing their wings and females use these songs to localize the potential partner. If females recognize the conspecific song and rate it as attractive they approach the singing male, a behavior called phonotaxis (for an overview see Hedwig [2006] and Pollack [2000]). The natural pattern of a calling song consists of repetitive pulses that are grouped into pulse trains called chirps [Huber et al., 1989]. The attractiveness of different patterns can be easily tested under laboratory conditions by monitoring the phonotactic behavior of females toward artificial signals [Hennig, 2009; Weber et al., 1981]. Extensive phonotaxis experiments suggested that the brain processes the patterns in the temporal domain [Grobe et al., 2012; Hennig, 2009] rather than in the spectral domain as has been proposed earlier [Thorson et al., 1982]. Schneider and Hennig [2011] provided evidence that females evaluate only the coarse temporal structure of a pattern. Consequently, the abstract song pattern of *Gryllus bimaculatus* can be described with four independent parameters [Doherty, 1985], for example the pulse duration, pulse pause, chirp duration, and chirp pause (see Figure 2.1). However, it is not clear whether these four cues are analyzed independently in the cricket brain. The period, that is the sum of duration and pause, as well as the duty cycle, that is the ratio of duration and period, for both pulses and chirps have also been implicated as relevant descriptors [Doherty, 1985; Grobe et al., 2012; Hennig, 2009]. Behavioral experiments [Hennig, 2009] show that a pulse period of 40 ms at a pulse duty cycle of 0.5 elicits highest phonotactic scores. For the organization of the chirps Grobe et al. [2012] observed optimal ranges between 200 and 500 ms for the chirp period, provided that the chirp duty cycle lies between 0.3 and 0.7. However, the relative importance of each of these song features is as yet unclear.

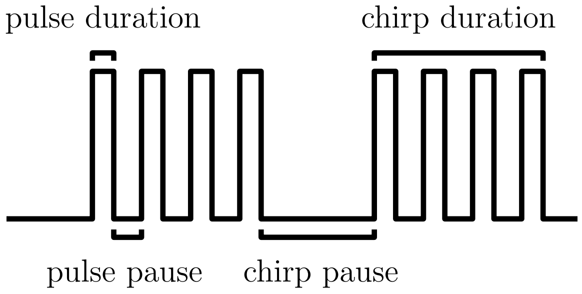

Figure 2.1: **Artificial song pattern of the cricket *Gryllus bimaculatus* and its temporal features.** Typically, a calling song consists of repetitive pulses that are grouped into chirps. The temporal structure of an artificial song pattern is fully determined by four descriptors, e.g. the duration and pause for both pulses and chirps. Four additional descriptors are frequently used to characterize cricket songs, namely the period (the sum of duration and pause), and the duty cycle (the ratio of duration and period) for both, the short and the long time scale.

Here, we employ artificial neural networks, which are also known as multilayer perceptrons, to analyze a large body of behavioral data obtained in phonotaxis experiments. We provide a detailed investigation of the relevance of individual song parameters on a quantitative measure that rates phonotactic behavior. Our models provide quantitative predictions for the attractiveness of hitherto untested song parameters, which helps guiding future phonotaxis experiments. Finally, we carefully interpret our results with respect to the underlying neural processing employed for acoustic pattern evaluation in the cricket brain.

2.2 Materials and Methods

2.2.1 Behavioral experiments and data

We used behavioral tests to measure the phonotactic score of the cricket *Gryllus bimaculatus* as explained in detail in Hennig [2009]. In brief, female crickets were placed on top of a trackball system that records their 2D walking trace. The females were presented with song patterns that mimic natural calling songs. These were constructed by amplitude modulated sinusoidal signals with a carrier frequency of 4.5 kHz. The amplitude was modulated to construct a periodical series of rectangular sound pulses that are grouped into chirps (see Figure 2.1). As a measure for the attractiveness of a particular song pattern, we computed the phonotactic score according to the formula in Schul [1998]. The phonotactic score is an integral measure that involves the walking length, the accuracy of the course maintenance, and the orientation of the female. It assumes values between -1 and 1, whereat a value close to one indicates a high level of attractiveness of the tested song pattern. For this study, we grouped data from experiments of 218 song patterns differing in their temporal parameters each of which was presented to several female crickets (mean: 31, range: $8-225$). For each song the phonotactic score was averaged across individual animals.

The data set was preprocessed as follows. First, we examined the distribution of the response values of the song patterns: 35% of the patterns were unattractive with a phonotactic score smaller than 0.2, 48% were intermediate between 0.2 and 0.6, and 17% were attractive with a value greater than 0.6 [Grobe et al., 2012; Hennig, 2009]. Then we split the data set into a training data set and a test data set of 200 and 18 data points, stratified according to the above allocation of unattractive, intermediate and attractive songs. This method is known as stratified sampling and was applied whenever data sets were divided into subsets. Then, we whitened the temporal calling song features of the training data set and applied the obtained transformation to the features of the test data set as well. In the whitening process, the features are first projected onto their principal components which removes linear correlations across features and then each feature is normalized to zero mean and unit variance. This linear coordinate transformation is widely used to preprocess the data before applying regression methods such as artificial neural networks [Bishop, 2006].

2.2.2 Model

Artificial neural networks

Artificial neural networks are commonly employed for regression tasks [Bishop, 2006], that is in our case to predict the phonotactic score from untested patterns. Figure 2.2A shows an example of a network diagram with four input variables that represent the features of a calling song, ten neurons in the hidden layer and one output neuron that represents the corresponding phonotactic score. In detail, the information about the features is forward propagated as follows: input variables x_i that represent calling song features are linearly combined to activations $\alpha_j = \sum w_{ji} x_i$ of hidden neuron j, where w_{ji} denotes the synaptic weight between input neuron i and hidden neuron j. Then, the activations of each hidden neuron are transformed with a nonlinear sigmoidal function $f(\alpha_j) = \frac{1}{1+exp(-\alpha_j)}$. Finally, the output variable $y = \frac{1}{2} \sum w_{\text{out},j} f(\alpha_j)$ is computed, where $w_{\text{out},j}$ indicates the synaptic weight between hidden neuron j and the output neuron. Thus, in artificial neural networks the temporal calling song features are nonlinearly processed to predict the phonotactic score. We implemented artificial neural networks in the Python programming language, using the Fast Artificial Neural Network Library [Nissen, 2003].

Training and validation

For training the synaptic weights, we chose the RProp algorithm which is a well-established supervised learning technique for multilayer feed-forward networks [Riedmiller and Braun, 1993]. The algorithm uses a training data set to update the randomly initialized weights in each training cycle such that the mean squared error between the model's prediction and the experimentally observed phonotactic score is minimized. We used the whitened training data set to perform a stratified 5-fold cross validation for training and validating networks. The training was stopped after $10,000$ cycles. This stopping criterion enabled us to compare the performance of networks with different architectures. To produce a single error estimation, the mean squared errors for validation (MSE_{val}) and training ($\text{MSE}_{\text{train}}$) were averaged over folds. In order to account for random initialization of the weights, the 5-fold cross validation was repeated for 100 times and we calculated the mean and standard deviation of the MSE_{val} and $\text{MSE}_{\text{train}}$.

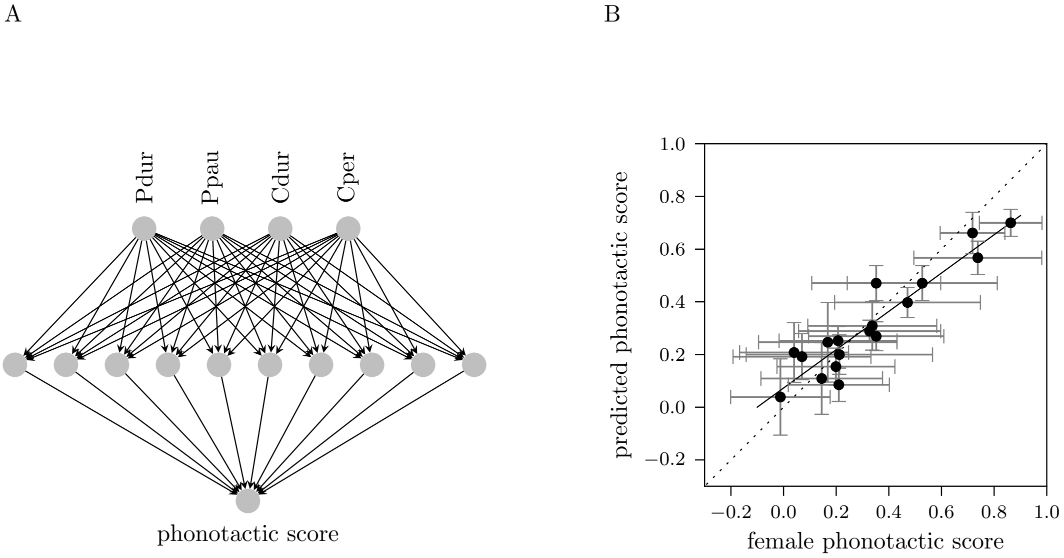

Figure 2.2: **Network diagram and predictive performance of the best 4-feature model.** (A) The network diagram consists of four input neurons representing temporal calling song features, which project to input-evaluating neurons in the hidden layer. These in turn project to the output neuron mimicking the relative phonotactic score; abbreviations: Pdur - pulse duration, Ppau - pulse pause, Cdur - chirp duration, Cper - chirp period. (B) Correlation between the phonotactic score of 18 test samples predicted by the best 4-feature model and the experimentally measured scores. Each dot shows the mean phonotactic score for a given song pattern that was presented to on average 31 females and tested for 100 times with the model. The errorbars indicate standard deviation across individual females (horizontal) and across 100 repeated model simulations (vertical). The solid regression line has a slope of 0.73. The performance: $\text{MSE}_{\text{test}} = 0.017$ and $r = 0.93$.

Model selection

In a first step we determined the appropriate number of neurons in the hidden layer by comparing the validation errors of networks with $n = 1$ to $n = 20$ hidden neurons. In detail, for each $n > 1$, we calculated the percent change of the validation error with respect to the network consisting of $n = 1$ neuron. Then, we chose the smallest n such that networks with $n + 1$ hidden neurons lead to an improved performance of no more than 1% as compared to networks with n neurons. This criterion ensured to select a network with high predictive power on the one hand and a simple model architecture on the other hand.

Performance

To obtain an unbiased estimate of a network's ability to generalize we used the test data set of 18 song patterns to test the network's performance. Therefore, we trained a network with the whitened training data for $10,000$ cycles and ran it with the test data set. Again, we repeated this for 100 times and averaged the network's prediction. Then we calculated the mean squared test error ($\mathrm{MSE}_{\mathrm{test}}$) as well as the linear Pearson correlation coefficient between the averaged network's predictions and the mean phonotactic scores averaged over females.

Prediction

To predict the phonotactic score of an untested song pattern, we first trained a chosen network over $10,000$ cycles with whitened features and the corresponding phonotactic score of the initial feature set of 218 data points. Then, we transformed the features of the untested song pattern with the transformation obtained in the whitening process of the features belonging to the initial data set. Next, we run the trained model with the transformed features of the untested song pattern. Finally, we repeated this training and prediction procedure for 100 times and averaged the phonotactic scores across the repetitions.

Feature selection

We considered in total eight different temporal features of a song pattern that have been previously used as descriptors. This is a redundant set of descriptors as four features, two on the short and two on the long time scale, are sufficient to fully define the song pattern. However, it is not a priori known, which set of features will best describe the behavioral data. Thus, we investigated all 255 feature sets, each one consisting of a different combination of the eight temporal features. For each feature set, we trained and validated models for a different number of hidden neurons followed by the selection of the appropriate model, as described above.

2.3 Results

Our analyses comprised a large body of behavioral data from experiments in which artificial calling songs were presented to female crickets under systematic variation of the song parameters. The phonotactic behavior was monitored with a single quantity, the phonotactic score. The acoustic pattern of an artificial song is shown in Figure 2.1. We trained artificial neural networks that receive as input the values of a particular set of song features to predict the phonotactic score. First, we considered feature sets made up by two features on the short pulse time scale and two on the long chirp time scale and analyzed how well an artificial neural network trained on parts of the experimental data can predict the phonotactic score on the remaining test data. In order to investigate the interplay of pulse and chirp information with respect to the phonotactic score we systematically varied pulse period and chirp period. Finally, we compared feature sets, each one consisting of a different combination of temporal features, in order to determine those features that are most efficient in correctly predicting average phonotactic behavior.

2.3.1 Predictive performance of models using full temporal pattern information

How well can we predict the behavioral outcome in an experimental trial based on the particular song pattern that was presented? To answer this question we trained and validated different artificial neural networks on non-redundant input features using a cross-validation procedure. From a total of eight potential features we investigated all combinations made up by two features on the short pulse time scale and two on the long chirp time scale that together fully determine the temporal song structure (see Figure 2.1). The best performing 4-feature model was selected based on the validation error. It used pulse duration, pulse pause, chirp duration, and chirp period as input features and comprised $n = 10$ hidden neurons. The network diagram is shown in Figure 2.2.A.

The average performance of this 4-feature model was quantified on the test data set as shown in Figure 2.2.B, where each point corresponds to one song pattern and the model prediction is plotted against the average phonotactic score computed from the animals' behavior. The predicted response values for the test data set were highly correlated with the experimentally measured responses, that is with a linear correlation coefficient of $\mathrm{r} = 0.93$. The mean squared error between the predictions and the experimental measurements was $\mathrm{MSE}_{\mathrm{test}} = 0.017$. The vertical errorbars indicate standard deviation, indicating the prediction variability of the best 4-feature model that was simulated for 100 times toward the same calling song. The main source for this variability is that before training the weights were initialized randomly, which resulted in slightly different predictions for one song pattern. The horizontal errorbars indicate inter-individual response variability of different females toward the same song.

2.3.2 Fusion of the short and long time scale

Female crickets use information from both, the pulse pattern and the chirp pattern to recognize and evaluate the conspecific song. How is this information on the short pulse and the long chirp time scales combined by female crickets during auditory processing? We hypothesize two basic models as sketched in Figure 2.3.A: in case of a logical AND-operation only an attractive pulse structure in combination with an attractive chirp structure generates highest phonotactic scores. This would indicate a synergistic processing. In contrast, a logical OR-operation requires either a suitable pulse or an attractive chirp structure to drive high phonotactic scores, thus optimal parameters for both time scales do not transmit extra information. The latter behavior is known as hypo-additive effect [Duchamp-Viret et al., 2003]. We evaluated the best 4-feature model (pulse duration, pulse pause, chirp duration, and chirp period) for different patterns by systematic variation of chirp and pulse periods. While varying the periods we fixed the duty cycles at 0.5, which ensured that one parameter of

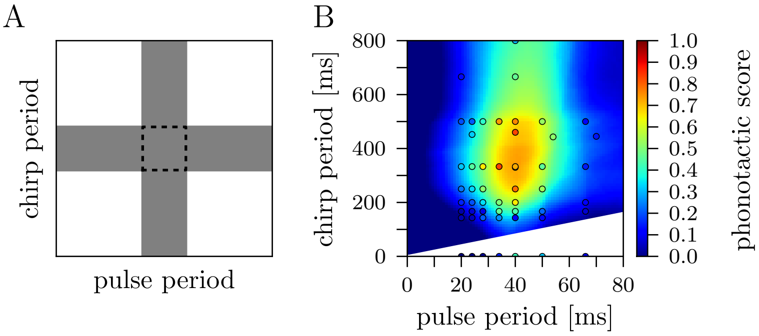

Figure 2.3: **Interaction of the short and long time scale.** (A) Sketch of a logical AND-operation (central square) and an OR-operation (gray shading). (B) Chirp period - pulse period response field predicted by the best 4-feature model (pulse duration, pulse pause, chirp duration, chirp period). The dominant circular area of highest response values suggests an AND-operation. Circles indicate experimentally measured phonotactic scores.

each time scale was in an attractive range [Grobe et al., 2012; Hennig, 2009]. The plane spanned by the chirp period and the pulse period in Figure 2.3.B shows highest response values for patterns with a chirp period between 250 and 500 ms and pulse periods from 35 to 45 ms. The maximal response value was obtained for a pattern with a pulse period of 40 ms and a chirp period of 340 ms. The dominant circular shape of highest responses suggested that the model approximates a logical AND-operation for high phonotactic scores.

2.3.3 Selection of the most informative song features

Which are the critical temporal song features that carry the most information for phonotaxis? A number of different song parameters have previously been tested experimentally and several have been suggested to be of particular importance. We considered a total of eight temporal features, namely duration, pause, period, and duty cycle for both pulses and chirps, as introduced in Figure 2.1. Above we already presented a model that uses two features from the short time scale (pulse duration and pulse pause) and two features from the long time scale (chirp duration and chirp period). However, it is not clear, which set of features will best describe the behavioral data. Thus, we investigated all possible feature sets, each one consisting of a different combination of the eight temporal features and compared the prediction accuracy of the corresponding models. Figure 2.4 shows the ten best models. The overall best performance with respect to the validation error was obtained for the 3-feature model that uses pulse period, chirp duration, and chirp duty cycle as input and $n = 10$ neurons in the hidden layer. All models using the pulse period plus two features from the long time scale were among the ten best performing networks that use three features as input. Surprisingly, the best 2-feature model that only uses pulse period and chirp pause and $n = 7$ hidden neurons did not perform significantly different from the best 3-feature model (p-value $=$ 0.028 for a two-sided Wilcoxon rank-sums test; significance level of 0.01). The feature combinations of pulse period plus one chirp feature are the four best in the class of models that only use two features as input. In contrast, the best 4-feature model that uses pulse duration, pulse pause, chirp duration, and chirp period as input and $n = 10$ hidden neurons performed significantly worse than the best 3-feature model (p-value $=$ 0.005 for a two-sided Wilcoxon rank-sums test; significance level of 0.01). Models with only one or more than four features as input were not ranked top ten.

2.3.4 Model predictions for pulse and chirp response fields

We investigated pulse and chirp response fields predicted by the best 4-feature, 3-feature, and 2-feature models. Pulse response fields describe two-dimensional subspaces spanned by the pulse duration and pulse pause of the eight dimensional feature space in which the attractiveness is color coded. To this end, we trained the models using all data of 218 songs and their phonotactic scores. For the best 4-feature model, we predicted the phonotactic scores for patterns with different pulse durations and pulse pauses but with a fixed chirp duration of 200 ms and a fixed chirp period of 333 ms that construct an attractive chirp structure [Hennig, 2009]. The pulse response field of this model, as shown in Figure 2.5.A, reveals an oval structure: song patterns with high phonotactic scores are displayed in an area bounded by pulse periods of 30 and 45 ms and pulse duty cycles of 0.4 and 0.7. For the best 3-feature and best 2-feature model we predicted responses toward patterns with different pulse periods but with a fixed chirp duration of 200 ms and a chirp duty cycle of 0.6 (best 3-feature model), and a fixed chirp pause of 133 ms (best 2-feature model). The pulse

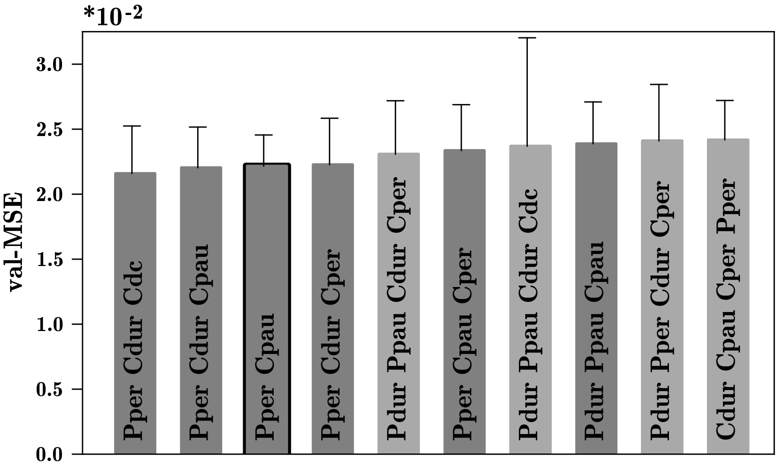

Figure 2.4: **Ten best performing models.** Model of size four (light gray), three (dark gray), and two (black edging) are ranked top ten. The overall best performing model uses the pulse period, chirp duration, and chirp duty cycle. The best 2-feature model (pulse period and chirp pause) did not perform significantly different (p-value $= 0.028$ for a two-sided Wilcoxon rank-sums test; significance level of 0.01). The best 4-feature model (pulse duration, pulse pause, chirp duration, and chirp period) performed significantly worse than the best 3-feature model (p-value < 0.01 for a two-sided Wilcoxon rank-sums test). Abbreviations: Pdur - pulse duration, Ppau - pulse pause, Pper - pulse period, Pdc - pulse duty cycle, Cdur - chirp duration, Cpau - chirp pause, Cper - chirp period, Cdc - chirp duty cycle. The models were validated 100 times and errorbars indicate standard deviation.

response fields of the best 3-feature model (Figure 2.5.B) and the best 2-feature model (Figure 2.5.C) were highly similar: due to the fact that only a single parameter on the short time scale was used we obtained a 1-dimensional structure where the phonotactic score varied along the diagonal defined by the pulse period. Particularly, the phonotactic scores were invariant under different pulse duty cycles. Higher phonotactic scores were in the range of 40 ± 10 ms pulse period, which was consistent with predictions of the best 4-feature model.

Next, we analyzed the chirp response fields. We predicted the response values for song patterns with different chirp durations and chirp periods but a fixed pulse duration of 20 ms and a fixed pulse pause of 20 ms for the best 4-feature model. In case of the 3-feature model we predicted responses toward patterns with different chirp durations and chirp duty cycles but a fixed pulse period of 40 ms. The response fields of the best 4-feature (Figure 2.5.D) and 3-feature model (Figure 2.5.E) revealed highest phonotactic scores for chirp durations and pauses between 100 and 300 ms. The chirp response field of the 2-feature model, obtained by varying the chirp pause at a fixed pulse period of 40 ms, revealed a 1-dimensional structure in which the scores only vary for different chirp pauses, irrespective of the chirp duration, see Figure 2.5.F. Here, highest phonotactic scores were predicted for chirp pauses between 100 and 250 ms.

2.4 Discussion

In this study we trained artificial neural networks to predict the attractiveness of calling songs of the cricket *Gryllus bimaculatus*. We studied the dependence of the model performance on the parameters of the calling song and aimed to identify minimal subsets of temporal features that carried sufficient information to describe the experimentally measured behavioral performance.

2.4.1 The most relevant song features for behavior

A number of different song parameters, namely the duration, pause, period, and duty cycle for both pulses and chirps are commonly used in cricket studies [Doherty, 1985; Grobe et al., 2012; Hedwig and Poulet, 2005; Hennig, 2009; Thorson et al., 1982]. But,

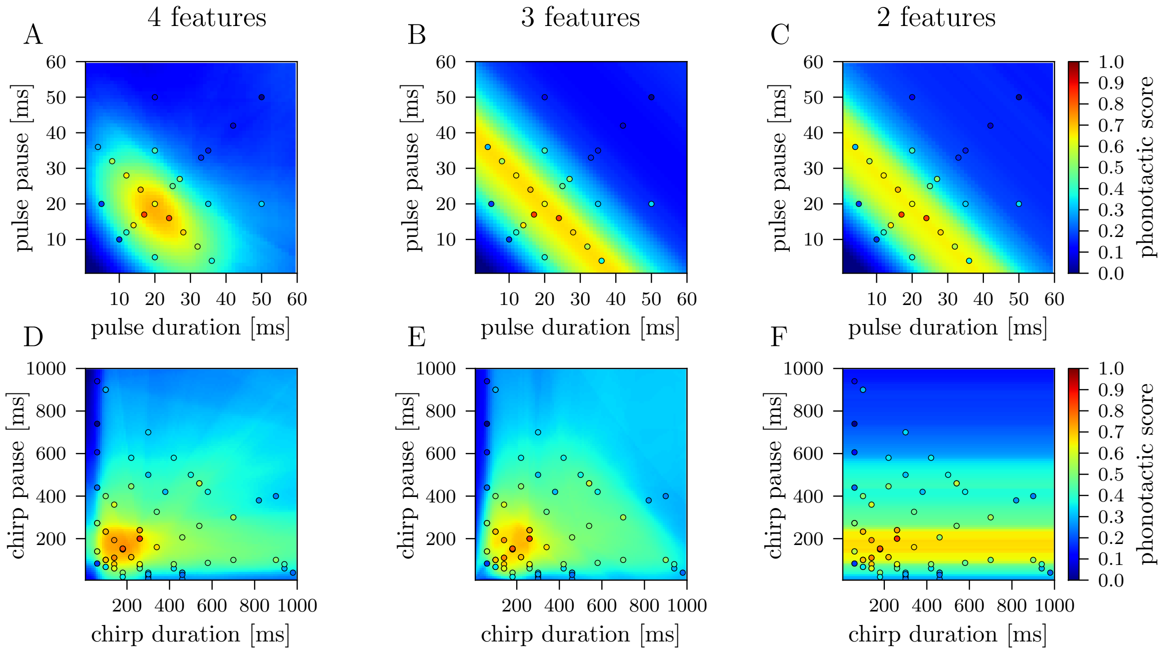

Figure 2.5: **Pulse and chirp response fields predicted by the best 4-feature, 3-feature and 2-feature model.** (A) The pulse response field of the best 4-feature model shows highest phonotactic scores for patterns that are accumulated in an oval bounded by pulse periods of 30 and 45 ms and pulse duty cycles of 0.4 and 0.7. The pulse response fields of the best 3-feature model (B) and the best 2-feature model (C) are clearly independent of the pulse duty cycle and show an extension on the diagonal defined by a pulse period of 40 ms. The chirp response field of the best 4-feature model (D) and the best 3-feature model (E) are qualitatively similar and reveal best scores for patterns with chirp durations and pauses between 200 and 300 ms. The best 2-feature model predicts highest scores for patterns with a chirp pause between 100 and 250 ms, irrespective of the chirp duration (F). Circles indicate experimentally measured phonotactic scores.

this set is overcomplete in the following sense: four features, two from each time scale, are sufficient to describe the artificial calling song. Thus, we investigated the performance of in total 255 models each one using a different set of song features. We identified three feature sets of different sizes that are best describing the behavioral data. The best 4-feature model, which used pulse duration, pulse pause, chirp duration and chirp period was ranked top ten (see Figure 2.4). The overall best model uses three features, the pulse period, chirp duration, and chirp duty cycle. Remarkably, the six combinations consisting of the pulse period plus two features from the long time scale are among the ten best performing sets of three features. Also the best 2-feature model uses the pulse period from the short time scale plus the chirp pause as input and regarding only models with two input features, the pulse period plus one feature from the long chirp time scale are the best four models. These findings suggest that the pulse period is the most crucial feature from the short time scale. For optimal prediction information on the short time scale (pulses [Schildberger, 1984; Zorovic and Hedwig, 2011]) and information on the long time scale (chirps [Doherty, 1985; Grobe et al., 2012]) are equally important. Also, in a taxonomic study [Desutter-Grandcolas and Robillard, 2003] temporal features on both time scales (number of pulses per second, number of pulses per chirp, number of chirps per minute) were relevant for relating phylogeny to the species-specific song patterns.

2.4.2 Logical AND-operation of the time scales

Calling songs of crickets carry information on short and long time scales, somewhat in analogy to words and phrases of human speech. How does the female cricket fuse auditory information that is present on the two distinct time scales? The response profile (Figure 2.3.B) for different combinations of pulse and chirp periods showed a synergistic effect, that is only attractive pulse structures combined with attractive chirp structures drove highest phonotactic scores. This provided evidence for a logical AND-operation of the time scales and was in line with results from Grobe et al. [2012] who interpolated behavioral measurements in the plane spanned by chirps per second and pulses per second, that is in the frequency domain. Notably, the combination of attractive pulse periods between 35 and 45 ms and unattractive chirp periods (greater than 500 ms) already caused intermediate responses. This again underlined the importance of the pulse period which we determined as the behaviorally most important feature of the short time scale.

The finding that the time scales are fused in an AND-operation can be interpreted with respect to the neuronal processing in the cricket brain. If our results had indicated a logical OR-operation of short and long time scales, then an independent, that is parallel processing of both time scales in the brain would have been likely. The result of the interdependence indicates that processing could be either parallel or serial. In the former case we expect from physiological experiments to find neuronal responses in the central brain that are independently tuned to either the short [Kostarakos and Hedwig, 2012; Schildberger et al., 1986] or the long time scale. The fusion of both information streams would happen only at a late stage of the brain network. Alternatively, in the latter case of serial processing we expect neural representations to be dependent on both time scales at an earlier stage of the brain network.

2.4.3 Song pattern complexity in crickets versus grasshoppers

Acoustic communication is also widely studied in grasshoppers. In mating behavior, male *Chorthippus biguttulus* grasshoppers produce courtship songs consisting of syllables that are grouped into phrases which in comparison with the songs of crickets exhibit a more complex song structure [Balakrishnan et al., 2001; Klappert and Reinhold, 2003; von Helversen and von Helversen, 1994]. If females rate the song as attractive, they produce response signals that direct the male toward her [van Helversen and van Helversen, 1983]. Wittmann et al. [2010] employed an approach similar to ours and analyzed courtship songs of the grasshopper *Chorthippus biguttulus* with artificial neural networks. Seven structural features of courtship songs were introduced and served as input to artificial neural networks. The linear correlation of $r = 0.93$ between the model's predictions and the experimentally measured response probabilities was in a similar range as for our best 4-feature model. Wittmann et al. [2010] also investigated the features that affect a female's assessment of a male's quality by excluding each song parameter once. In their case, none of the excluded features led to an increased performance of the corresponding reduced model. This indicates that the employed features are non-reducible and results in a feature space of at least seven dimensions. Thus, the processing of auditory information in female grasshoppers is more complex than in crickets.

2.4.4 Non-linear extension improves performance

A closer inspection of the best 4-feature model's performance as shown in Figure 2.2.B revealed a systematic mismatch between the behavioral measurements and the model prediction. For small experimental phonotactic scores (≤ 0.1) the model overestimated the attractiveness of the corresponding song patterns. Likewise, the model underestimated the attractiveness of models that were experimentally found to be highly attractive (≥ 0.7). The same systematic bias was observed in the behavioral predictions by Wittmann et al. [2010]. What could be the reason for this result and how could we improve model predictions? The classical artificial neural networks devised non-linear elements in the hidden layer while the output neuron computes a linear sum. We additionally applied a non-linear transformation of sigmoidal shape to the predictions of the model. In detail, we first used the training data set to choose $n = 10$ hidden neurons. Then, we chose the parameters $y = 0.5$ and $b = 5.23$ of the sigmoidal $s_{b,y}(x) = \frac{1}{1+\exp(-b(x-y))}$ as they minimized the mean squared error between the experimentally measured phonotactic scores of the training data set and the sigmoidal transformed predictions. This improved the predictive power: on the test data set the error measure reduced to $\mathrm{MSE}_{\mathrm{test}} = 0.008$ (as compared to $\mathrm{MSE}_{\mathrm{test}} = 0.017$) and the linear correlation coefficient was $\mathrm{r} = 0.94$. A possible interpretation of this result in a biological context involves two-step processing. In a first processing stage, the attractiveness of the stimulus pattern is evaluated. In a second stage, the outcome of this evaluation is non-linearly translated into behavior analog to a behavioral decision. To investigate this possibility it would be of interest to study in detail behavioral thresholds in individual animals [Beckers and Wagner, 2011].

2.4.5 Towards future models of neural network processing

We presented artificial neural networks that are suitable for predicting phonotactic scores of untested song patterns and thus for complementing behavioral as well as guiding electrophysiological studies. However, artificial neural networks do not attempt to model the natural neural processing of auditory information in the cricket brain. To improve our understanding of the underlying neuronal mechanisms during pattern recognition computational models of neural function are required that incorporate our anatomical, morphological, and physiological knowledge. Any such model should attempt to reproduce female phonotactic behavior and to provide testable hypotheses at the biophysical level.

Insects in general are well suited because they achieve the required tasks of pattern recognition and evaluation of the fitness parameters with relatively small brains. The cricket *Gryllus bimaculatus* is a well suited insect model for studying the neural basis of the processing of auditory information and the generation of choice behavior due to its highly limited neuronal resources. In the auditory pathway receptor neurons converge to two ascending neurons that project to a small number of neurons in the brain. Much is already known about the physiological properties [Schildberger, 1984; Wohlers and Huber, 1982] of the ascending interneurons and ongoing work investigates the connectivity and physiological properties of the brain neurons. It has been shown that for varying pulse patterns some neurons match the average behavioral tuning [Kostarakos and Hedwig, 2012; Schildberger, 1984]. Several modeling approaches that use the cricket as a model system exist. Webb [1995] investigates sound-seeking in crickets with robots. Mhatre and Balakrishnan [2007] used a stochastic model to simulate the walking path of crickets. But, only few attempts have been made to model the neural mechanisms for pattern recognition in crickets. Benda and Hennig [2008] showed that spike-frequency adaptation can generate intensity invariance in ascending neurons. In a preliminary study short term depression and short term facilitation in central brain synapses were suggested as plausible mechanisms for the parametric tuning on the short and long time scale [Rost, 2011]. Recently, based on their physiological investigation of central brain neurons, Kostarakos and Hedwig [2012] suggested a network scheme that includes mutual excitation and inhibition of central brain neurons as a plausible alternative that awaits testing in a future neural network study.

Chapter 3

Decoding of calling songs and their behavioral relevance from grasshopper auditory neurons[1]

Acoustic communication plays a key role for mate attraction in grasshoppers. Males use songs to advertise themselves to females. Females evaluate the song pattern to recognize a conspecific male and as proxy to its fitness. The song pattern, a repetitive structure of sound syllables separated by short pauses, provides a primary cue to identify conspecifics. In their natural habitat females often receive songs of degraded temporal structure. Perturbations may, for example result from the overlap with other songs. We studied the response behavior of females to songs that show different signal degradations. A perturbation of an otherwise attractive song at later positions in the syllable diminished the behavioral response, whereas the same perturbation at the onset of a syllable did not affect song attractiveness. Using intracellular recordings from auditory neurons we explored how information about the acoustic stimulus and its attractiveness is encoded in the neural responses. We find that populations of

---

[1] This chapter has been submitted to Frontiers in Systems Neuroscience. Please refer to page v in this dissertation for the detials of authors contributions.

three or more neurons were sufficient to reliably decode the acoustic stimuli and to predict its behavioral relevance from the single-trial integrated firing rate. Decoding accuracy increased with neural population size but not with integration time. We discuss our results in light of the time persistence and trial-to-trial variability of auditory neuronal responses with respect to potential mechanisms of pattern recognition and decision making in the brain.

3.1 Introduction

Acoustic communication of grasshoppers has become a prominent model system to investigate principles of neuronal processing of acoustic stimuli and they provide the opportunity to study perceptual decision making in a comparatively simple nervous system. Grasshoppers produce acoustic signals, termed “songs”, to attract a mating partner. Natural songs consist of a repetition of stereotyped subunits (syllables) with species-specific amplitude modulations of a broad carrier frequency band that are produced by moving the hind legs against the forewings [von Helversen and von Helversen, 1997]. Due to characteristic differences between grasshopper species the songs constitute an important barrier against hybridization. Both the song production and the song recognition are innate behaviors, and therefore we can be confident that the corresponding neuronal circuits are “hard-wired”. In behavioral tests one can use artificial song models that mimic and vary certain song features, and thereby explore which cues are crucial for song recognition [von Helversen, 1972; von Helversen and von Helversen, 1997, 1998]. These experiments demonstrated that the decisive cues for song recognition reside in the temporal pattern of amplitude modulations, i.e. in a song’s envelope. In the grasshopper *Chorthippus biguttulus*, the subject of this investigation, a very simple but highly attractive song model consists of a series of sound “syllables” separated by pauses (see Figure 3.1.A). Using song models we can reduce the signal’s complexity and compare the behavioral responses directly with the processing capacities of neurons at different stages of the auditory pathway.

The nervous system of grasshoppers offers an important advantage: many neurons can be identified as individuals on the basis of their characteristic morphology [Stumpner and Ronacher, 1991]. Thus specific processing properties can be assigned to groups of individual neurons in the auditory pathway. The first stage of auditory processing comprises three neuron classes: auditory receptor neurons, local neurons (LN) and ascending neurons (AN). The ears of grasshoppers are located on the sides of the first abdominal ganglion. A total of approximately 60 receptor neurons transduce the vibrations of the tympanum into series of action potentials that travel via the axons into the metathoracic ganglion complex, which houses the first auditory processing stage. There, axons make contact to various types of local neurons (LNs) – about 15 individual LNs have been identified so far. The LNs then contact a set of about 20 ANs, the axons of which ascend to the animal's head and constitute the sole auditory input to upstream processing circuits and decision centers located in the brain [Bauer and von Helversen, 1987; Ronacher et al., 1986]. Since the population of ANs constitutes a bottleneck for the information that is available to the brain, they will be in the focus of the present study. Remarkably, the auditory pathway including the ANs is highly conserved between different grasshopper species [Neuhofer et al., 2008; Ronacher and Stumpner, 1988]. Not only are the neurons' morphologies extremely similar in two not related species (*C. biguttulus* and the locust *Locusta migratoria*), but homologous neurons also exhibit the same physiological properties and processing capacities. For this reason we can compare neuronal properties of e.g. LN and AN of the locust with behavioral data obtained with *C. biguttulus*.

The decision centers located in the female brain must evaluate whether a heard song follows the con-specific pattern and whether it is attractive enough to trigger a response song as the appropriate behavior. This task appears simple under ideal conditions, since the song patterns of different species differ considerably [Gottsberger and Mayer, 2007; Stumpner and von Helversen, 1994]. However, in nature there are many factors that may degrade the acoustic signal on its way from sender to receiver. This aggravates the classification problem. Here we introduced perturbations of the signal envelope that strongly influenced behavioral decisions. Applying perturbations to the pattern of an attractive song model affected the signal's attractiveness as measured by the female response rates differently, depending on the specific position of a perturbation within a song syllable (Figure 3.1.A). Presenting the same stimuli while performing intracellular recordings from identified neurons allowed to investigate the neural representation of the stimulus identity and of its behavioral relevance.

Using naïve Bayes classifiers (for review see Quiroga and Panzeri [2009]; Pouget et al. [2000]) we specifically asked to what degree the acoustic stimulus can be decoded and whether the behavioral stimulus category can be predicted from the single-trial responses of single neurons and neuron populations.

3.2 Materials and Methods

3.2.1 Animals

The behavioral tests were performed with females of *Chorthippus biguttulus*. The animals were reared as the filial generation (F1) from eggs of individuals collected as adults near Göttingen, Germany. After adult moult females and males were held separately in plastic cages to ensure virginity. In this species the females respond to a male's song with a song of their own, thereby indicating their readiness to mate. This response song is an ideal criterion showing that a female has identified a song as belonging to a potential conspecific mating partner.

Electrophysiological experiments were performed on locusts, *Locusta migratoria*, that were bought from a commercial supplier (for details of the breeding and keeping procedures see Schmidt et al. [2008]; Stange and Ronacher [2012]). We can homologize identified neurons between the two species on the basis of their characteristic morphology [Stumpner and Ronacher, 1991]. The homologous auditory neurons of the thoracic ganglia show quantitatively similar response patterns in both species [Neuhofer et al., 2008; Ronacher and Stumpner, 1988; Sokoliuk et al., 1989]. On the basis of this strong homology we can use recordings from *L. migratoria* neurons and compare their spike patterns with behavioral responses of *C. biguttulus*.

3.2.2 Acoustic stimuli

A digitally generated song envelope consisting of rectangular syllables (subunits) of 72 ms duration separated by 12 ms pauses served as an attractive standard stimulus (Figure 3.1.A). In order to systematically screen the detrimental effect of degradation at different syllable positions, we inserted perturbations of 24 ms either in the first, or in the middle, or in the last part of each syllable (Figure 3.1.A). A perturbation

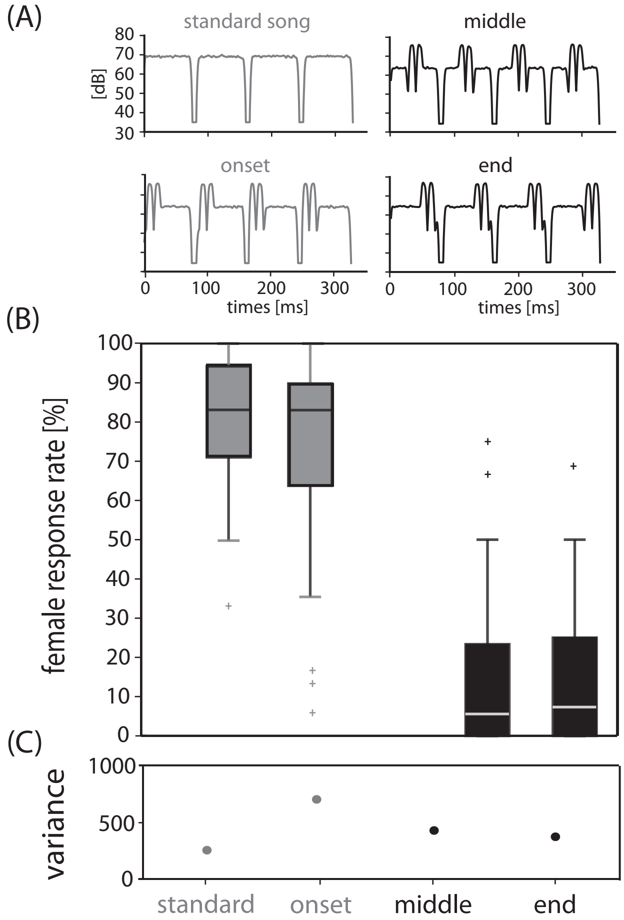

Figure 3.1: **Perturbation of the standard song affects attractiveness when placed at later syllable positions.** (A) Envelopes of song models used for behavioral and neurophysiological tests. An attractive standard song consisted of 72 ms syllables and 12 ms pauses. The other stimuli had the same syllable and pause durations but exhibited perturbations at different positions within a syllable (onset, middle, end). (B) The median response rate of 33 *C. biguttulus* female responses for the stimulus with onset perturbation was 83%, thus very similar to the response to the standard stimulus. In contrast, stimuli with perturbation in the middle and end were mostly rejected (median response rate 6%). The median is displayed as the central mark in the box plot. The edges of the box are the 25th and 75th percentiles. (C) Note the high variance in female responses, especially when perturbation is at syllable onset.

consisted of 2 alternating accents and gaps, each of 6 ms duration and 12 dB higher or lower relative to the syllable plateau. Earlier experiments had revealed that gaps within a syllable do markedly reduce the stimulus attractiveness; accentuations that occur at the end of a syllable have similar detrimental effects [Ronacher and Stumpner, 1988; von Helversen, 1972; von Helversen and von Helversen, 1997; von Helversen, 1979]; for reviews see Ronacher et al. [2004]; Ronacher and Stange [2013].

The envelopes of all song models were convolved with the same carrier frequency (a broad band noise spectrum of $5 - 40$ kHz). Sound intensity was calibrated with a half inch microphone (type 4133; Brüel & Kjær, Nærum, Denmark) and a measuring amplifier (type 2209, Brüel & Kjær) at the position of the animal. All four test patterns were presented with the same effective intensity (RMS) of 70 dB SPL; therefore the peak and plateau intensities differed between stimuli (syllable plateau 70 dB for the standard stimulus and 65 dB for perturbed stimuli, Figure 3.1.A). Yet, these intensities fall into the intensity range well accepted by *C. biguttulus* females [von Helversen and von Helversen, 1997, 1994]. The songs presented in the behavioral and electrophysiology tests comprised the same envelope structure but differed in length: 2772 ms (33 subunits; behavior) and 756 ms (9 subunits for electrophysiology), respectively.

3.2.3 Behavioral experiments

Virgin *C. biguttulus* females were tested in a sound proof chamber at a constant temperature of 30 ± 2°C. The experiments were automatically conducted by a custom made program (written by M. Hennig in Labview 7.1, National Instruments) presenting songs in a pseudo-randomized order while recording the females' responses (for details of the apparatus and testing procedures see Schmidt et al. [2008]). Each song was iterated 18 times. As a measure of stimulus attractiveness we used the percentage of responses normalized to the 18 presentations for each female. Out of these individual responses median response rates were calculated. Additionally, a negative control was presented, comprising the same carrier frequency and length as the standard signal, but lacking any syllable pause structure. In applying this negative control stimulus those females indicating a not discriminative behavior for song patterns could be detected. From further analysis we therefore excluded 11 from 44 tested females as they responded more than twice to the negative control.

3.2.4 Electrophysiological experiments

Intracellular recordings were obtained with conventional methods from adult *Locusta migratoria*. For a detailed description of the experimental settings see Vogel et al. [2005]; Wohlgemuth and Ronacher [2007]. Temperature during the experiments was kept constant at $30 \pm 2°$C. At the end of an experiment the cells were stained with Lucifer Yellow and identified according to their characteristic morphology and physiology. For this study ascending neurons (ANs) were analyzed which represent the third processing stage in the metathoracic ganglion and transmit the auditory information to the grasshopper's brain. Recordings were obtained from AN1 ($n = 9$), AN3 ($n = 10$), AN4 ($n = 4$) and AN12 ($n = 2$) neurons (for the terminology see Römer and Marquart [1984]; Stumpner and Ronacher [1991]). Each song was presented within a looped order: standard stimulus, onset-perturbation, perturbation in the middle, then perturbation in the end, and starting again with the standard stimulus. Stimulus iteration was 8 times.

3.2.5 Data analysis

Estimation of firing rates

Neuronal firing rates were estimated from spike trains by convolution with a Gaussian kernel with width σ ranging from $1, \ldots, 30$ ms and support $[-4\sigma, 4\sigma]$ [Nawrot et al., 1999]. The kernel was normalized to unit area such that the time integral of the estimated rates equals the number of spikes.

Naïve Bayes classification

Naïve Bayes classifiers are statistical classifiers that are based on Bayes' theorem together with naïve independency assumptions. We applied Bayesian classifier to decode which stimulus class evoked a particular neural response (see Pouget et al. [2000]; Quiroga and Panzeri [2009] for reviews). Let $P(s)$ denote the probability of presentation of stimulus class s and $P(x_1, \ldots, x_n|s)$ the conditional probability of observing spike train features $x_1, \ldots, x_n$ given s. The posterior probability that stimulus class s was presented given $x_1, \dots, x_n$ is according to Bayes' theorem

$$P(s|x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n|s)}{P(x_1, \dots, x_n)} \cdot P(s)$$

with

$$P(x_1, \dots x_n) = \sum_{s \in S} P(x_1, \dots, x_n|s) \cdot P(s).$$

The naïve independency assumption that each feature x_i is conditionally independent of feature x_j given s simplifies to

$$P(s|x_1, \dots, x_n) = \frac{\prod_{i=1}^{n} P(x_i|s)}{P(x_1, \dots, x_n)} \cdot P(s).$$

From this posterior probability distribution the stimulus class $\hat{s}$ that maximizes the probability that $x_1, \dots, x_n$ was observed is chosen:

$$\hat{s} = \mathrm{argmax}_{s \in S} \left\{ P(s|x_1, \dots, x_n) \right\}.$$

Since $P(x_1, \dots, x_n)$ is constant given stimulus class s, the classification rule can be written as

$$\hat{s} = \mathrm{argmax}_{s \in S} \left\{ \prod_{i=1}^{n} P(x_i|s) \cdot P(s) \right\}.$$

Different decoding approaches

First, we decoded stimulus classes based on the spike count of single neurons which can be considered as a very simple descriptor of a neural spike response pattern. For each stimulus of 756 ms duration we counted the number of spikes for each of the eight trials. For every count c we decoded the stimulus class as

$$\hat{s} = \mathrm{argmax}_{s \in S} \{ P(c|s) \cdot P(s) \}.$$

Based on the remaining counts we performed a kernel density estimation to compute the probability density functions $P(c|s)$. The estimation was implemented with `scipy.stats.gaussian_kde` [Oliphant, 2007]. As the procedure includes automatic bandwidth determination, the probability density functions were estimated with different bandwidths. To account for the nonnegativity of the counts, we restricted the support to positive values and normalized the probability density function to unit area. For the very rare case that not more than two counts had different values we assumed a Poisson distribution with mean of the counts.

Second, we investigated a "voting algorithm": for a fixed trial the stimulus class was decoded as explained above but starting with the spike count measured over the first period of 84 ms (syllable plus following pause). We then successively decoded the stimulus class based on the count restricted to each period separately which yielded nine estimations of the stimulus class. After all nine periods the stimulus was chosen that was most often predicted in the nine periods.

Third, for decoding from a pseudo-population of neurons, we used the counts $c_1, \ldots, c_n$ of n neurons of different type recorded in different females and calculated

$$\hat{s} = \mathrm{argmax}_{s \in S} \left\{ \prod_{i=1}^{n} P(c_i|s) \cdot P(s) \right\}$$

to decode which stimulus class triggered the counts $c_1, \ldots, c_n$.

We further chose time-resolved firing rates, which capture the temporal dynamics of a neural pattern, for decoding stimulus classes in single neurons. To this end, we estimated the firing rates for each millisecond which yielded rates $r_0, \ldots, r_{756}$ from $t = 0, \ldots, 756$ ms (see 3.2.5 Estimation of firing rates). Based on the estimated firing rates we decoded the stimuli according to

$$\hat{s} = \mathrm{argmax}_{s \in S} \left\{ \prod_{i=0}^{756} P(r_i|s) \cdot P(s) \right\}.$$

Here, $P(r_t|s)$ was estimated as explained for above.

Grouping of stimuli into classes

We followed the decoding approaches to first decode the four stimuli. In this case the set S of stimulus class consists of the standard stimulus, onset perturbation, middle-perturbed song and end-perturbed song, i.e. each song forming a single class. As all four songs were equally often presented we applied the classification rules with $P(s) = 1/4$ for all $s \in S$. However, we may also define stimulus classes that consist of grouped stimuli. For example, decoding whether or not a song shows degradation yields two classes, one consisting of the standard stimulus and the other one of the three perturbed songs. The prior of these two classes is

$$P(s) = \begin{cases} 1/4 & \text{for } s = \text{standard stimulus} \\ 3/4 & \text{for } s = \text{perturbed stimulus.} \end{cases}$$

Performance of the classifier

To validate the performance of the classifier we performed a leave-one-out cross validation in which each single trial response was used once for decoding based on the distribution of the remaining trials. The results were stored in a confusion matrix [Jurman et al., 2012] whose entry (i, j) represents the number of times that a presentation of stimulus class i was predicted to be stimulus class j. Based on the confusion matrix we quantified the decoding performance with the Matthews correlation coefficient (MCC) as it is defined in Jurman et al. [2012]. The MCC assumes values between -1 and 1, where 0 indicates chance level classification and 1 perfect prediction. In case of binary classification (e.g. decoding the standard stimulus against the three perturbed stimuli) the formula reads

$$\mathrm{MCC} = \frac{\mathrm{TP} \cdot \mathrm{TN} - \mathrm{FP} \cdot \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP}) \cdot (\mathrm{TP} + \mathrm{FN}) \cdot (\mathrm{TN} + \mathrm{FP}) \cdot (\mathrm{TN} + \mathrm{FN})}}$$

where TP, TN, FP and FN denote true positives, true negatives, false positives and false negatives, respectively. This measurement has the advantage that it can be applied in multiclass problems even if the classes are of different sizes [Gorodkin, 2004; Jurman et al., 2012].

To test whether a classifier decodes significantly better than chance we performed a leave-one-out cross validation based spike train features that were randomly reassigned to the stimuli, followed by a calculation of the MCC. We repeated this procedure $1,000$ times and calculated the p-value as the percentage of MCCs that are larger than or equal to the actual MCC. A significance level of 0.05 was chosen.

3.3 Results

3.3.1 Behavioral decisions reveal two stimulus classes

In behavioral tests we investigated how degradation at specific positions within the signal does affect signal recognition. We compared the responses of *C. biguttulus* females to four stimulus types (Figure 3.1.A): (i) standard stimulus without perturbation, (ii) with perturbation during the first third of the syllable ("onset"), (iii) during the second third ("middle"), and (iv) during the last third ("end"). Figure 3.1.B shows the distribution of response rates across individual females to these stimuli (see 3.2.3 Behavioral experiments). The standard stimulus was highly attractive (median: 83%), although individual females differed considerably in their response rate (compare quartile ranges and see variance in Figure 3.1.C). Females showed similar high response rates towards the stimulus with onset perturbation, whereas the same perturbation in the middle or the end of a syllable led to a behavioral rejection (median response levels of $< 10\%$). Only 3 out of 33 females responded to the latter stimuli in more than 50% of the stimulus presentations.

In order to further analyze differences in attractiveness we pairwise compared stimulus responses in individual females. For each female, the response rates for any two stimuli (see left column in Figure 3.2) were subtracted. Thus it could be shown that the responses to the onset stimulus did not differ significantly from the responses to the standard (top row, Figure 3.2); the same is true for the comparison of the stimuli perturbed in the second and third part of the syllable (lowest row, Figure 3.2). In contrast, the responses to the unperturbed song and the song with middle and end perturbations differed significantly (p-value < 0.001; Friedman and Dunn's Multiple Comparison Test), and in both cases the median difference was about 60%. Similar results were found for the comparison between the onset perturbed stimulus and the other two perturbed stimuli (median differences $> 50\%$, p-value < 0.001).

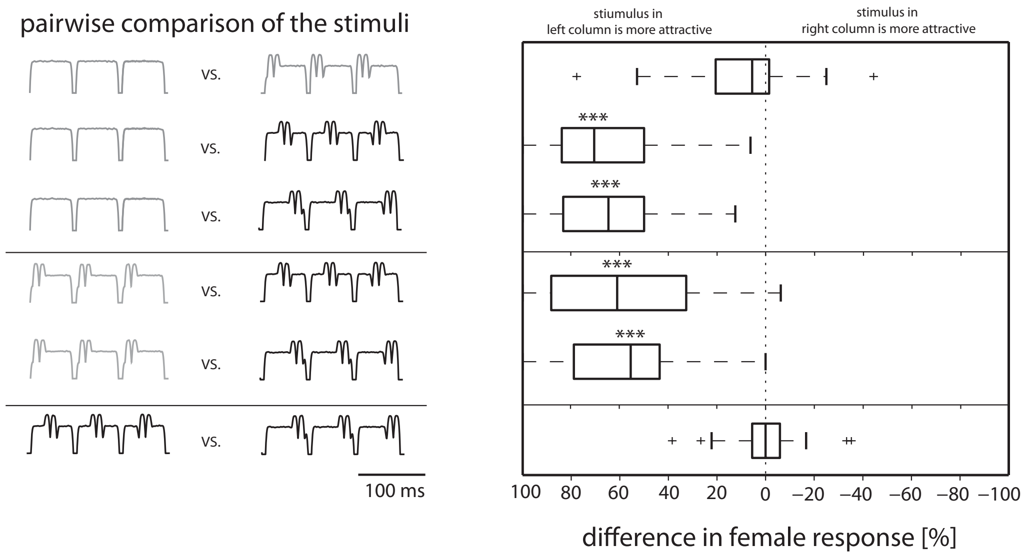

Figure 3.2: **Pairwise comparison of individual female responses allows distinction in attractive and unattractive stimulus classes.** Box plots show medians of response differences in individual females for stimulus comparisons shown in the left. Whereas there is no difference in response between stimuli with onset perturbation and the standard song, they are both significantly more attractive than stimuli with perturbation at middle and end (p-value < 0.001, Dunn's post hoc test after Friedman).

3.3.2 Decoding stimulus identitiy and behavioral class from the neuronal spike counts

Grasshoppers have to make their decisions based on the information about the environment provided by the sensory and downstream neurons of the auditory pathway. The clear separation into two behavioral stimulus classes raises the question of how the different stimuli and these different behavioral classes are represented and discriminated within the grasshopper's nervous system. We address this question in intracellular *in vivo* recordings of identified ANs during repeated presentations of all four songs. To quantify the encoded information we apply a single-trial decoding approach to the neural spiking activity using a Bayesian classifier. We start out with decoding the identity of the auditory stimulus before we predict the behavioral class (attractive versus non-attractive). Finally, we compare the decoding performance based on the simple measure of the spike count – equivalent to the time-integral of the firing rate – to the decoding performance when taking into account the full time-resolved firing rate profile.

Stimulus classification based on single neurons

How is information about a stimulus, such as the stimulus type or its attractiveness, represented in the spike responses of the ANs? Figure 3.3 shows example voltage traces of *in vivo* intracellular recordings from two individual ANs, and the corresponding spike raster plots. The example AN3-neuron responded with a burst of spikes to the stimulus onset and with smaller bursts at syllable onsets. In the two unattractive stimuli, however, additional spike bursts occurred in the middle or at the end of the syllables. The AN1-neuron marked the syllable onsets of the standard stimulus, whereas the perturbations evoked additional spikes within the syllables. The trial-averaged firing rates (Figure 3.3, color coded) of all recorded neurons indicate that neuronal response patterns vary for the four different song patterns. Also, neurons that are of the same morphological type (AN1, AN3, AN4, AN12) show variations in their response patterns across individual animals.

We use a Bayesian approach to classify the acoustic stimulus based on the neural activity (see 3.2.5 Naïve Bayes classification). To this end we counted the number of spikes in each single trial and for each of the four stimuli during the complete stimulus duration of 756 ms, comprising 9 syllables and the respective pauses. Based on the counts we decoded the stimulus identity according to the classification rules in 3.2.5 Different decoding approaches. We measured the classification performance by the Matthews correlation coefficient (MCC).

Figure 3.4 shows the results for decoding the four stimuli from single neuron activity. The MCC was higher than chance level for all but two neurons (see Figure 3.4) and 11 out of 25 decoded the stimuli significantly better than on basis of randomized counts (black dots in Figure 3.4, p-value < 0.05). Averaging across all 25 neurons yielded a mean MCC of 0.32. The decoding results were best for the standard song (not shown). As shown in Figure 3.1.A the standard song had a higher syllable plateau than the perturbed songs which is a consequence of our constraint that all stimuli have the

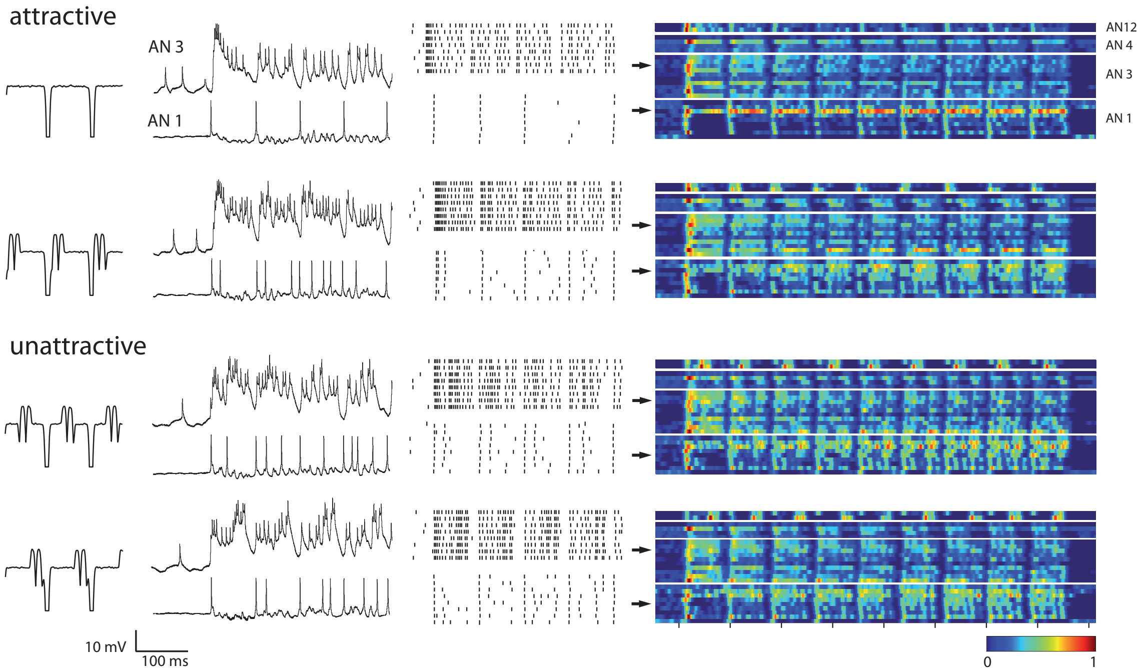

Figure 3.3: **Neuronal responses to all four songs categorized by their behavioral relevance.** Voltage traces and spike raster plots (8 trials) in the second and third columns show responses to the first four syllable–pause subunits for two example neurons AN3 and AN1. The fourth column shows trial-averaged firing rates estimated with a Gaussian kernel of width $\sigma = 4$ms during the whole stimulus presentation. Each row within a block of a neuron type represents the response of a single neuron (from top to bottom AN12 ($n = 2$), AN4 ($n = 4$), AN3 ($n = 10$), AN1 ($n = 9$)). Color denotes the amplitude of the estimated firing rates normalized to the maximum rate within each neuron class. Arrows point out the firing rates of the shown examples.

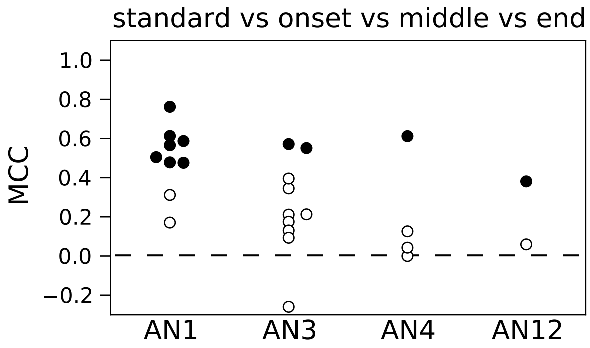

Figure 3.4: **Count based decoding of stimuli in single neurons.** A classification of the four stimuli is in 11 (filled circles) out of 25 neurons significantly better than a classification based on randomized counts. The distribution of MCC values of all 25 neurons differs significantly from the MCC distribution of the classifiers that are based on randomized counts (p-value < 0.05, one-sided Wilcoxon rank-sum test). Dashed line represents chance level based on randomized counts.

same effective intensity. A closer look showed that the trial-averaged spike count elicited by the standard syllables differed from the spike counts evoked by the perturbed syllables. However, this is not consistent across neurons. For some neurons the spike count evoked by the standard stimulus is considerably larger than the spike count evoked by any of the perturbed stimuli, for other neurons this relation is reversed. This difference between the spike count triggered by the standard and the perturbed stimuli is reflected in a higher performance in decoding the standard stimulus against the class of perturbed stimuli (SI Figure 3.8: averaged MCC is 0.78; 22 neurons decode significantly better than by chance). To avoid a bias of the decoding performance due to the higher syllable plateau of the unperturbed standard stimulus, we reduced the stimulus set to the three perturbed songs throughout the rest of the manuscript. This reduced stimulus set yielded only 5 neurons that allowed for a successful decoding of the three stimuli, and the average MCC dropped sharply to 0.08 (Figure 3.5.A).

So far, the spike count was measured during complete stimulus presentation which consists of nine periods (syllable plus pause). Next, we asked how good we can decode the stimuli based on the spike count extracted over shorter time windows. To

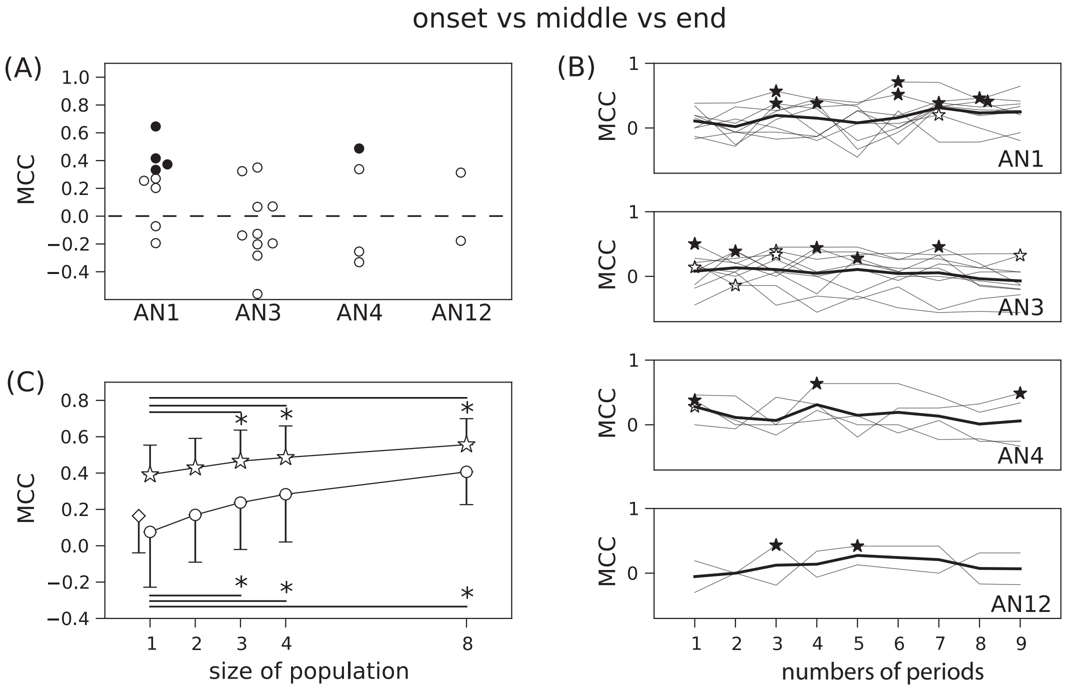

Figure 3.5: **Count based decoding of the three perturbed stimuli in single neurons and populations.** (A) Only in 5 neurons the three perturbed stimuli are decoded significantly better than a classification based on randomized counts. The distribution of MCC values of all 25 classifiers does not differ significantly from the MCC distribution of the classifiers that are based on randomized counts (p-value $= 0.23$, one-sided Wilcoxon rank-sum test). Dashed line represents chance level. (B) Averaged time course of the MCC is not increasing with stimulus time (thick black line). Stars depict the number of periods where the MCC peaks. Markers are filled if the classifier decodes significantly better than classifiers based on randomized counts. (C) Decoding performance increases with population size. Circles denote MCC values of classifiers that are based on spike count measured over all nine periods (cf. circles in (A)). Stars depict MCCs of decoder based on counts over population-individual optimal number of periods (cf. stars in (B)). Diamonds show the MCC of classifier that use the voting algorithm. MCCs are averaged across neurons and vertical errorbars depict standard deviation. The mean performance increases significantly from single neurons to populations of size three, four, and eight (p-value < 0.05, one-sided Wilcoxon rank-sum test).

this end, we investigated the MCC as a function of the number of periods starting at stimulus onset (Figure 3.5.B). Surprisingly, the MCC, averaged across neurons within one class, did not increase but stayed constant over stimulus time (see thick lines in Figure 3.5.B). For single neurons the MCC fluctuated without apparent increase or decrease (thin lines in Figure 3.5.B). Restricting the counts from onset to the period where the MCC was maximal (stars in Figure 3.5.B) remarkably improved the decoding results: now 18 neurons performed significantly better than by chance (black stars in Figure 3.5.B) and the MCC averaged across all neurons increased from 0.08 to 0.39. At least three periods were needed for successful stimulus reconstruction in 15 out of the 18 neurons.

In addition, we decoded the stimuli using a "voting" algorithm (see 3.2.5 Different decoding approaches). We first decoded the stimuli on the basis of each single period separately. In a second step we chose the stimulus which was most often predicted in the nine periods. In comparison to the decoding results that were based on counts over the whole stimulus length the averaged MCC increased moderately from 0.08 to 0.16 (diamond in Figure 3.5.C).

3.3.3 Population decoding

As the grasshopper brain receives input from several ascending neurons (up to twenty at each side [Stumpner and Ronacher, 1991]) we next decoded the three perturbed songs from neuronal populations (see 3.2.5 Different decoding approaches). We constructed neuronal populations up to size four with each neuron from a different type, representing a subpopulation of ANs in one hemisphere. Additionally, we decoded on a basis of populations of size eight, consisting of two different neurons of each available type, which reflects the input from both ears. As to be expected the averaged decoding performance is increasing with population size up to 0.41 for 8 neurons if counts were extracted over the complete stimulus duration (circles in Figure 3.5.C). In case counts were extracted over increasing stimulus duration for each population individually, the MCC is again increasing with population size up to 0.56 for population size 8 (stars in Figure 3.5.C). This improvement was significant between populations of size 3 or larger and single neurons (p-value < 0.05, one-sided Wilcoxon rank-sum test).

3.3.4 Discrimination of behaviorally relevant classes by single neurons and populations

In our behavioral experiments stimuli fell into two behaviorally relevant classes: the standard song and the onset-perturbed song were attractive whereas songs with middle- and end-perturbed syllables were rejected (Figure 3.1.B). Here we asked: is it possible to predict whether a song belongs to the accepted or rejected class based on the single trial neuronal responses? We again used a Bayesian decoder and evaluated the success of correct predictions in single trials with the MCC (3.2.5 Naïve Bayes classification). We first considered the total spike count over all nine periods in single neurons. Only half of all MCC values were larger than zero and the number of neurons that decoded significantly better than by chance was reduced to 3 (Figure 3.6.A). The MCC averaged across all 25 neurons was 0.19 and the distribution of the MCC did not differ significantly from the distribution of the performance values based on randomized counts (p-value = 0.45, one-sided Wilcoxon rank-sum test). As above, this result can be improved by individually choosing the number of periods over which the counts are calculated for each neuron: then single-trial classification was significantly better than chance in 12 neurons (see black stars in Figures 3.6.B). In this case the average MCC was 0.49 and the distribution differed significantly from decoders based on randomized counts ($p\text{-value} < 0.05$, one-sided Wilcoxon rank-sum test). Again, the averaged time course of the MCC was not accumulating and the optimal numbers of periods were widely spread among neurons (Figure 3.6.B).

If information was used from populations, the performance improved remarkably up to an average MCC of 0.69 (counts over all nine periods; circles in Figure 3.6.C) and 0.79 (counts over population-individual stimulus length; stars in Figure 3.6.C) for populations of size eight. The increase differed significantly between single neurons and populations of size three or larger for counts measured over the complete stimulus duration (circles in Figure 3.6.C) and between single neurons and population of size two and larger if counts were measured over population-individual numbers of periods (stars in Figure 3.6.C)

In contrast, the performance of a decoder that used the voting algorithm was small (0.07) for single neurons (diamond in Figure 3.6.C).

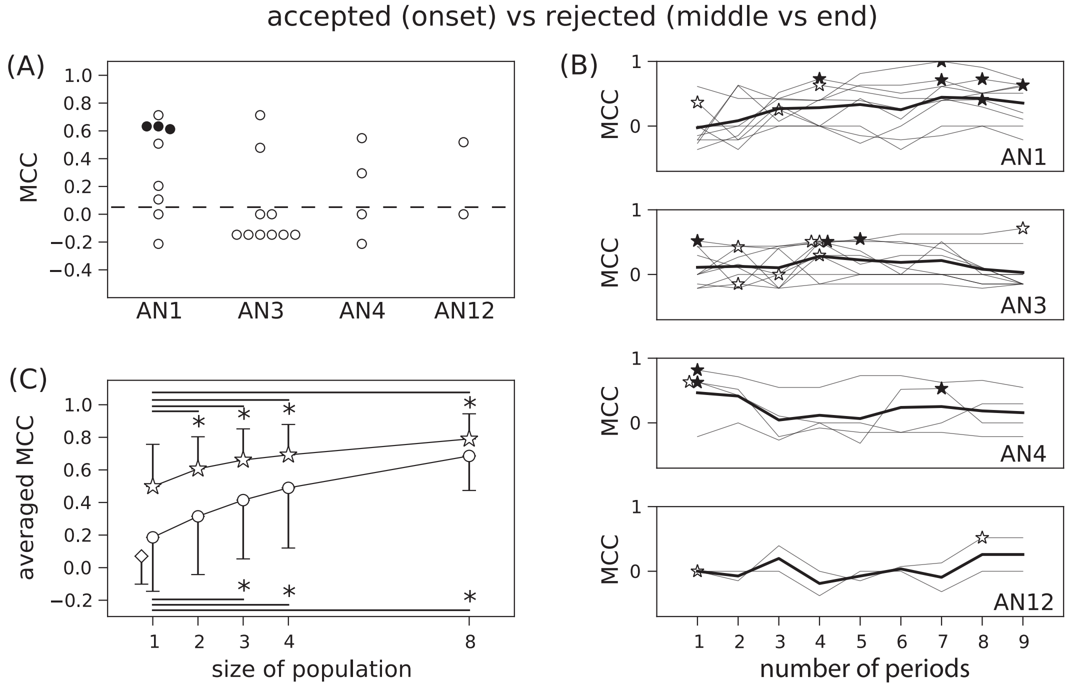

Figure 3.6: **Count based decoding of behaviorally relevant classes in single neurons and populations.** (A) Decoding the class of accepted versus the class of rejected stimuli is in only 3 neurons successful. The distribution of the 25 MCC values does not differ significantly from the MCC distribution of the classifiers that are based on randomized counts (p-value = 0.45, one-sided Wilcoxon rank-sum test). (B) Averaged time course of the MCC is not increasing with stimulus time (thick black line). (C) Decoding performance increases with population size. The increase differs significantly between single neurons and populations of size three and larger for counts measured over complete stimulus duration (circles) and between single neurons and population of size two and larger if counts were measured over population-individual numbers of periods (stars) ($p\text{-value} < 0.05$, one-sided Wilcoxon rank-sum test).

3.3.5 Precise timing information improves coding in single neurons

How much information does the temporal structure of the neural spike pattern carry about the evoked stimulus type or its attractiveness? To this end, we investigated naïve Bayes classifiers based on time-resolved firing rate estimates using a Gaussian kernel and for different kernel widths σ (see 3.2.5 Different decoding approaches). The precise timing of spikes is reflected if we estimate firing rates with very small kernel widths (see Figure 3.3 for a detailed structure of the rates that were estimated with kernel width of $\sigma = 4$ ms), whereas, at the other extreme, rates estimated with very large kernel widths correspond to the spike count. Figure 3.7 shows the performance of single-trial classification averaged across neurons as a function of the Gaussian kernel width. Classifying the three perturbed stimuli yields good performances for rate estimates with kernels widths σ in the range of $1, \ldots, 12$ ms (continuous line in Figure 3.7.A). The average MCC reaches a maximum of 0.89 at $\sigma = 4$ ms (see filled circle in Figure 3.7.A). This result underlines the differences in the temporal structure of the spike patterns evoked by the three stimuli (see Figure 3.3). Decoding behaviorally relevant classes (accepted versus rejected songs) performed best (MCC $= 0.68$) with $\sigma = 1$ ms, and $\sigma = 1, \ldots, 5$ ms yielded classifiers with similar performance (not significantly different; Figure 3.7.B). However, the range of MCC covered by individual neurons (grey area in Figure 3.7.B) was much larger than in Figure 3.7.A, indicating that some neurons performed well in this task whereas others performed badly. We conclude that the detailed structure of the estimated rates evoked by the middle- and end-perturbed songs bear similarities and thus can be discriminated from the attractive pattern evoked by the onset perturbation, at least by some neurons. Note that the temporal resolution of the rates had to be finer for classifying behavior-relevant classes than for decoding stimuli.

3.4 Discussion

We applied the framework of naïve Bayes classifiers to investigate how information about the stimuli and their respective attractiveness is encoded in the responses of auditory neurons of grasshoppers. Naïve Bayes classifiers have frequently been used to quantify encoded information in neural spike trains. Hoare et al. [2011] applied naïve Bayes classifiers to predict odors from a population of 15 olfactory sensory neurons in

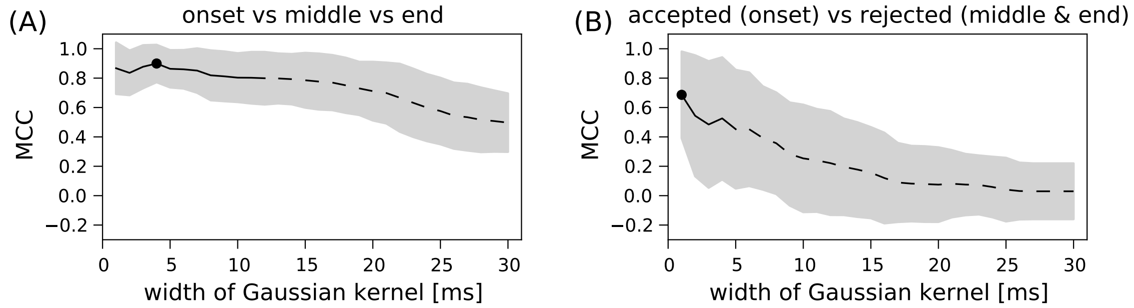

Figure 3.7: **Time-resolved firing rate based decoding in single neurons.** Averaged performance is shown as a function of the width of the Gaussian kernel that was used for estimation of the firing rates. Grey area shows standard deviations. Classifiers are best performing with firing rates determined with high temporal resolutions (Gaussian kernel of widths $\sigma = 1, \dots, 12$ ms (A) and $\sigma = 1, \dots, 5$ ms (B) and within these ranges they do not differ significantly from the best ones (black circle; p-value < 0.05, two-sided Wilcoxon rank-sum test).

Drosophila larvae based on their spike count. Using modeled neuronal responses of 10 visual interneurons in the blowfly *Calliphora vicina* Karmeier et al. [2005] decoded optic flow fields and found that the decoding performance increases with integration time. In behaving monkeys Rickert et al. [2009] applied naïve Bayes classifiers to classify the cueing stimulus and to predict the direction of voluntary movements from multiple single-unit recordings in the motor cortex. Here, we followed three different approaches: First, we decoded stimulus classes based on the spike count of single neurons and studied how the decoding performance changes with the integration time of neurons. We also investigated a voting algorithm, which bears similarity to the multiple looks model proposed for psychophysical experiments by Viemeister and Wakefield [1991]. Second, we studied count based decoding in populations of up to 8 neurons. Third, we used information of time-resolved firing rates for decoding in single neurons.

3.4.1 Rate coding in the grasshopper auditory system

Remarkably, the voting algorithm did not significantly improve the performance when compared to the spike count decoding in single neurons (see diamond in Figures 3.5.C, 3.6.C). Therefore we restrict our discussion to the other two approaches. Decoding the stimulus class from the single neuron spike count resulted in an average performance as quantified with the Matthews correlation coefficient (MCC) of 0.39 (see star in Figure 3.5.C), provided that the counts were measured from stimulus onset to a neuron-specific optimal number of periods. When predicting the behavioral category of the stimulus we obtained a MCC of 0.49 (Figure 3.6.C). As to be expected, the performance improved considerably when decoding from the time-resolved firing rates of single neurons yielding average performances of MCC = 0.89 (Figure 3.7.A) and MCC = 0.68 (Figure 3.7.B) for stimulus classification and classification of accepted versus rejected songs, respectively. However, in the realistic scenario of decoding the spike counts from a population of neurons the performance increased significantly as compared to the single neuron case. For the maximum population size of 8 ANs we obtained MCC = 0.56 for stimulus decoding (Figure 3.5.C) and MCC = 0.79 for predicting the behavioral class (Figure 3.6.C). We grouped maximally 8 neurons, two of each of the morphological types that had been recorded in our experiments. This represents a realistic subpopulation of ANs from an individual animal. We can expect that the intact population of at least 20 morphologically distinct ANs per hemisphere in the grasshopper will reach considerably higher performances that might approach hundred percent correct classification. Moreover, our decoding results are hampered by the fact that we have only 8 experimental trials at hand for training and testing the decoding algorithm. Thus our performance estimates are suboptimal due to this methodological restriction. Interestingly, the behavioral experiments showed that individual females do not show a consistent 100% “performance” across repeated presentations of the same song (Figure 3.1.B) which might be due to variability in the sensory computation and decision making process or to motivational state. We conclude that the simple measure of the population spike count, reflecting the average stimulus-induced neuronal firing rates, reliably encodes identity and behavioral relevance of a particular song, without need for a time-resolved evaluation of firing rates.

These results are particularly interesting in view of recent papers investigating different aspects of the grasshopper’s auditory pathway. Clemens et al. [2011] provided evidence that between the local and ascending neurons, i.e. between the second and third processing stage, the coding principle changes from a summed population code to a labeled-line population code where the population’s information is maximal if a decoder takes into account neuronal identity. At the level of the AN population, the temporal sparseness as well as the population sparseness increases [Clemens et al., 2012]. At the same time, integrated spike rate information gains in significance compared to spike timing information [Clemens et al., 2011, 2012; Creutzig et al., 2009; Ronacher, 2014; Wohlgemuth and Ronacher, 2007]. The connection to behavioral classifications into attractive and not attractive songs is provided by a recent study that used abstract LN (linear-nonlinear) models of feature extraction, and was able to predict behavioral responses with high reliability ($r^2 = 0.87$), with a set of only two feature detectors [Clemens and Ronacher, 2013]. This cascade model involves the extraction of local stimulus features followed by a long term temporal integration over the whole stimulus duration. Thus the model explicitly neglects the exact temporal position of specific features and assumes a comparison of the detectors' output in form of long-term averaged spike counts. The high performance of the spike count classifiers in the behavior-related classification task found in the present study (Figure 3.7.B) thus fits nicely to the structure of the model proposed by Clemens and Ronacher [2013]. In addition, the use of a spike count code would also explain why the remarkable sloppiness of spike timing found in ascending neurons [Vogel et al., 2005] does not impair the precise evaluation of song features in the ms range which was observed in behavioral tests [Ronacher, 2014; Ronacher and Stange, 2013; Ronacher and Stumpner, 1988; von Helversen, 1979].

3.4.2 Constancy of the encoded information and trial-by-trial variability

Another interesting result is revealed by Figures 3.5.B and 3.6.B: against the naïve expectation that longer stimuli might provide more information, the average MCC did not rise with longer stimulus durations, and for several neurons the classification was best after only a few periods. Most classifiers needed more than 3 or 4 periods to reach a significance level of 0.05. On average significant decoding was reached at approximately 5 periods (4.8 ± 2.4 in Figure 3.5.B, 5.2 ± 2.9 in Figure 3.6.B). A number of 3 to 5 periods corresponds to a duration in the range of 250 to 420 ms. These values fit well to the minimal number of periods and song durations of 250 to 450 ms needed by *C. biguttulus* males to respond to a female song [Ronacher and Krahe, 1998; Ronacher et al., 2000].

The decoding results depend on the encoding signal as well as on the spike train variability, which increases the uncertainty of the decoder model. We estimated the single neuron trial-by-trial spike count variability using the Fano factor [Nawrot, 2010; Nawrot et al., 2008]. As shown in the supplementary Figure 3.9 the variability either remained constant or increased with increasing the stimulus time. This fits the result of the constant decoding performance independent of stimulus duration. The absolute values of Fano factor fit previous results showing that variability of spike trains increases from receptor neurons to the ascending neurons [Neuhofer et al., 2011; Ronacher et al., 2004; Vogel et al., 2005; Vogel and Ronacher, 2007]. The trial-to-trial variability is particularly high in ascending neurons, and on average these showed a reduced performance in stimulus classification compared to local neurons [Wohlgemuth and Ronacher, 2007]. Using song models that were progressively degraded, Neuhofer et al. [2011] could estimate the respective contributions of external signal degradation and the trial-to-trial variability of spike trains caused by intrinsic neuronal noise. Intrinsic neuronal noise had a very strong impact on the spike train variability, in particular in ANs, thus likely affecting the representation of acoustic signals along the auditory pathway, and thus also the discrimination and recognition of grasshopper songs [Ronacher, 2014].

3.4.3 Integrating evidence for behavioral decisions – a hypothetical brain algorithm in the grasshopper

At the level of ascending neurons we found a steady representation of information about the stimulus and its behavioral relevance in the spike count. Based on this observation we hypothesize a two-stage process for the recognition of conspecifics and behavioral decision making in the grasshopper's brain. A first stage process is postulated which evaluates population activity in a manner similar to our Bayesian decoder. For the central brain we suggest an algorithm that, based on the population input, continuously monitors the behavioral relevance throughout the song. If the song is built of attractive elements this provides steady evidence throughout the song that a conspecific and attractive male is present. A second stage process integrates this evidence over time and accumulating evidence is transformed into an appropriate behavioral response through thresholding. This second stage of our model hypothesis is reminiscent to a class of decision-making models that assume evidence over time to be integrated in order to reach a threshold. These models have previously been formulated for alternative choices in sensory decision tasks [Beck et al., 2008; Drugowitsch and Pouget, 2012; Gold and Shadlen, 2007]. In the grasshopper, recognition and evaluation of a conspecific calling song simplifies to the female's decision between showing or not showing her response behavior depending on whether and when the evidence reaches a threshold. In a neuroethological context as well as in controlled behavioral experiments animals can modulate their behavioral response level [von Helversen and von Helversen, 1997, 1994; Wirmer et al., 2010]. In our model hypothesis this could be realized by a modulation of response threshold, e.g. through neuromodulators in the relevant brain circuit [Heinrich et al., 2001; Wirmer et al., 2010].

If we assume a time-integrating algorithm in the grasshopper brain, what could be the underlying mechanism? The time span is indicated by the duration of the reported response times in the range of several hundreds of milliseconds (see above). One cellular mechanism that could subserve this task is short-term synaptic plasticity. Fascilitation and depression at such synapses are governed by processes with typical time constants in the right order of magnitude and they have repeatedly been suggested to be involved in decision making processes [Martínez-García et al., 2011; Mongillo et al., 2008] including a suggested algorithm for auditory pattern recognition in the cricket's central brain [Rost et al., 2013].

In summary, our results support the hypothesis of a population rate code in ANs that project the acoustic information to the central brain (see Clemens et al. [2011, 2012]. The information about the behavioral relevance of a stimulus is well represented in the population rate and this information is constant throughout the stimulus presentation. We suggest a computational process located within the grasshopper brain that infers the behaviorally relevant information and integrates this evidence over time to reach a behavioral decision based on accumulated evidence.

3.5 Supplementary Information

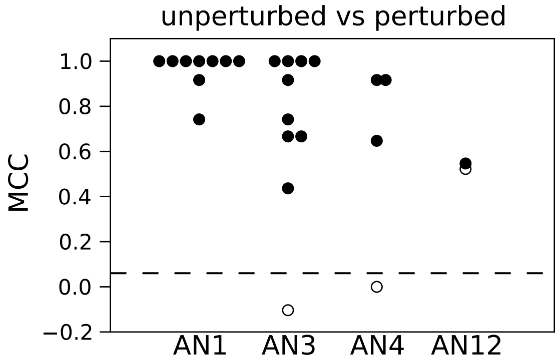

Figure 3.8: **Count based decoding of standard versus perturbed stimuli.** A classification of the standard against the perturbed stimuli is in 22 (filled circles) out of 25 neurons significantly better than a classification based on randomized counts. The distribution of MCC values of all 25 neurons differs significantly from the MCC distribution of the classifiers that are based on randomized counts ($p\text{-value} < 0.05$, one-sided Wilcoxon rank-sum test). Dashed line represents chance level based on randomized counts.

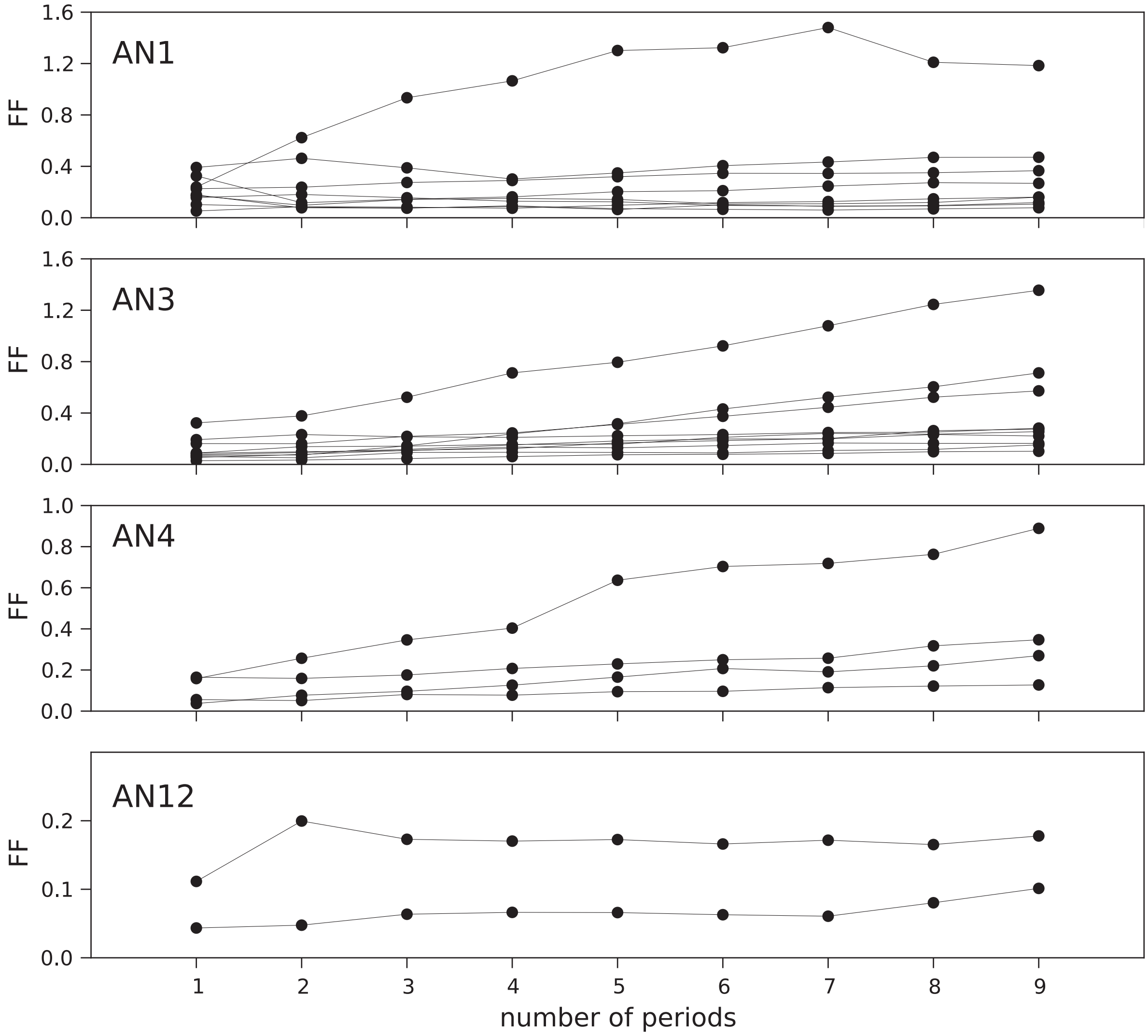

Figure 3.9: **Spike count variability in single neurons.** Each point depicts the Fano factor as calculated from stimulus onset to the number of periods, averaged across the four stimuli. The Fano factors either remained constant or increased with increasing the stimulus time.

Chapter 4

General discussion

This thesis comprises temporal aspects of the neuronal processing of calling songs in crickets and grasshoppers. In chapter 2 I identified minimal subsets of temporal features of calling songs that carry sufficient information to describe the experimentally measured behavioral performance of female crickets. Chapter 3 provides evidence that information about a stimulus is encoded in the spike count of populations of ascending neurons in grasshoppers and describes the effect of the integration time on the decoding performance. The integration of the results into the current understanding of the processing of acoustic signals in crickets and grasshoppers has been discussed in each of the chapters. Here, I will first discuss a common theme, namely the inter-individual response variability, that reoccurred in chapters 2 and 3. Then, I will point out alternative methods that provide means to analyze the data presented in this thesis.

4.1 Inter-individual response variability

The behavioral data set analyzed in chapter 2 consists of 218 different artificial calling songs and the corresponding phonotactic scores that reflect the walking behavior of female crickets *Gryllus bimaculatus* (see Hennig [2009] for a detailed description of the experiment). The phonotactic score [Schul, 1998] takes values between -1 and 1 whereas a value close to 1 indicates a high level of attractiveness of the tested song. In Figure 2.2.B the model prediction is plotted against the average phonotactic score computed from the walking behavior of females. The horizontal errorbars depict the standard deviation across females and show a considerable inter-individual behavioral response variability.

The beauty - and also challenge - of the data set I analyzed in chapter 3 is that it involves behavioral as well as electrophysiological experiments in grasshoppers. The behavioral response to four different artificial calling songs was measured in female *Chorthippus biguttulus*. Figure 3.1 shows a considerable variability in the response rates across females. A similar response variability is also present in the *in vivo* intracellular recordings of ascending neurons in *Locusta migratoria*. These neurons share the same physiological properties and processing capacities as ascending neurons of *Chorthippus biguttulus*. The trial-averaged, kernel estimated firing rates of the ascending neurons in Figure 3.3 show distinct codes toward the same stimulus. In chapter 3 the behavioral data from *Chorthippus biguttulus* and electrophysiological data from *Locusta migratoria* was pooled. Being aware of the associated difficulties, I suggest to perform experiments not only in the same species, but in the same animal.

The observed variabilities occurred across females under laboratory conditions, i.e. the very same song was presented to numerous females. In nature, there are also processes that cause signal variability [Ronacher et al., 2004]. Slight differences between calling songs from individual senders might carry important information about the sender, e.g. its health. Many factors may degrade an acoustic signal on its way from sender to receiver (cf. stimuli in chapter 3): scattering and reverberations when a signal travels through dense vegetation, atmospheric turbulences, chorus noise produced by other singers.

4.2 Alternatives to artificial neural networks

The data set from chapter 2 is predestined for building a model that predicts the phonotactic score of untested calling songs, because it consists of more than 200 experimentally tested calling songs labeled with the real-valued phonotactic score. I deliberately chose artificial neural networks to compare the results with those of

Wittmann et al. [2010] (see 2.4.3 Song pattern complexity in crickets versus grasshoppers). My model prediction showed high correlation with the experimental results (see Figure 2.2.B in chapter 2) and allowed investigating the interaction of the two time scales given by the pulses and chirps. The predictions provide evidence that the two time scales are integrated in an AND-like operation, i.e. only an attractive pulse structure in combination with an attractive chirp structure generates high phonotactic scores (see Figure 2.3 in chapter 2). There are two major points of criticism related to the chosen approach. First, ANNs are (still) computationally costly. For a speed-up, I recommend to analyze the data set with support vector regression machines [Drucker et al., 1996]. Support vector machines were, for example, successfully applied by Yovel et al. [2008] for a classification of plants according to their echoes to ultrasonic signals emitted by bats. Second, an assumption underlying the application of ANNs is that they assume that all features are available at the same time. To overcome this issue, I suggest to apply the framework of temporal restricted Boltzmann machines [Sutskever et al., 2008].

4.2.1 A structural method for identifying the most informative calling song features with ANNs

In chapter 2 I asked which temporal features of a calling song critically determine the phonotactic performance of female crickets. I built all 255 possible subsets of the overcomplete feature set consisting of the eight features, i.e. duration, pause, period and duty cycle for both pulses and chirps. By comparing the performance of ANNs using all different combinations of feature sets, I identified the set of pulse period, chirp duration, and chirp duty cycle as best performing. However, learning and testing 255 ANNs is computationally very costly. The method of greedy backward elimination [Guyon, 2003] offers an alternative way to minimize the dimension of the feature space. I will now present the results of the greedy backward elimination in comparison with the result of the brute-force method of comparing all 255 models as provided in chapter 2.

Starting with all eight temporal calling song features, in each step I eliminated the one feature whose removal yielded the smallest validation error until all validation errors increased if yet another feature is removed. The results (cf. Table 4.1) showed that the overcomplete feature set can be reduced to a feature set of size (pulse duration, pulse pause, chirp pause) three without any loss of prediction power. However, greedy backward elimination is risky in the sense that a removed feature might later be important in another combination of features. To account for this, I trained and tested all 56 networks, each one using a different combination of three features as input. Among these networks, the one that uses the pulse period, chirp duration and chirp duty cycle as input was best performing. This is the very same feature set I identified with the brute-force method of investigating all 255 feature sets.

Table 4.1: Error reduction due to backward elimination of features. The first column shows the difference between the validation errors of networks using all but one feature and the validation error of the full model. The following columns state the differences between the validation errors of networks obtained in the corresponding elimination step and the validation error of the networks investigated in the step before. All values were multiplied by 10^3.

removed feature	step 1	step 2	step 3	step 4	step 5	step 6
pulse period	-1.08	–	–	–	–	–
chirp duration	-0.25	-0.11	–	–	–	–
pulse duty cycle	1.46	0.71	-3.02	–	–	–
chirp duty cycle	-0.65	2.3	0.003	-5.34	–	–
chirp period	-0.24	-0.003	-2.00	-3.86	-0.72	–
pulse duration	1.54	4.63	5.22	8.37	6.28	4.77
pulse pause	-0.39	3.85	2.93	11.18	10.24	9.10
chirp pause	-0.42	0.19	-0.71	-0.54	1.01	19.04

4.3 Alternatives to naïve Bayes classifiers

In chapter 2 I asked how information about acoustic signals, i.e. the stimulus identity and its attractiveness, is encoded in the spike pattern of auditory neurons that project to the brain. Naïve Bayes classifiers provide a well-established statistical framework to answer this question. The underlying rule for classifying stimuli or behavior allows to investigate decoding based on different spike train features, such as the spike count or time-resolved firing rate vectors. One point of criticism relates to the estimation of the distributions of the spike train features. The data set comprises only eight trials per neuron and an estimation of the parameters based on more trials would have improved the decoding results. Karmeier et al. [2005] solved this issue by first modeling neuronal responses of ten visual interneurons in the blowfly *Calliphora vicina*, which allowed for a large trial number, followed by a Bayesian decoding of optic flow fields.

The data set of chapter 2 is also suitable for a classification based on similarities of spike trains, as explained in [Machens et al., 2003]: first, for each acoustic signal one spike train is arbitrarily chosen as a template and the remaining spike trains are classified to a signal by assigning each one to the closest of the eight templates. To this end, the spike train metric [van Rossum, 2001] is used to calculate the distance between two spike trains. Then, averages are computed by permuting all possible template choices, yielding a confusion matrix. The width τ of the kernel which is used to convolve the spike trains when applying the spike train metric allows to investigate the effect of spike timing and of spike count differences on the discrimination performance. Good classification performances for small values of τ indicate the importance of spike timing, whereas large values of τ point to the importance of the spike count. Wohlgemuth and Ronacher [2007] used this classification algorithm based on the spike train metric to investigate how well amplitude modulations of acoustic signals can be distinguished in auditory neurons of *Locusta migratoria*. In receptor neurons and local interneurons the discriminability of amplitude modulations of acoustic signals depends on the timing of spikes. In contrast, differences in the spike count of ascending neurons become important for the discrimination, but the discrimination performance at higher processing levels is less successful than at early stages. One advantage of this geometrical algorithm is that it is not limited by small trial numbers. Although the parameter τ accounts for effects of spike timing or spike count, statistical naïve Bayes classifiers are more flexible, because the classification rules allow for testing numerous spike train features. For example, next to the spike count and time-resolved firing rate vectors (cf. chapter 3), they also allow for a classification of the stimulus based on the latency, i.e. the time the very first spike occurred after stimulus onset.

Another way of analyzing the data of chapter 3 is to apply generalized linear models (GLMs) [Pillow et al., 2008], a generalization of linear–nonlinear–Poisson cascade models [Chichilnisky, 2001]. They serve to model the input-output function of sensory neurons [Clemens and Hennig, 2013; Tripathy et al., 2013]. In such a framework, the stimulus is described by a stimulus filter, a post-spike filter, and a coupling filter.

The sum of these filter responses is evaluated with an exponential nonlinearity that defines an instantaneous spike rate and is used to draw spikes. To further investigate the response variability across ascending neurons presented in chapter 3, I suggest to fit their spiking responses to GLMs and compare the GLM parameters across neurons and neuron types.

4.3.1 PCA-based classification of grasshopper ascending neurons

I also analyzed the data set of chapter 3 with a PCA-based cluster analysis, similar to the one applied in Meyer et al. [2013], who classified antennal lobe neurons in the honeybee based on physiological response features. For each ascending neuron, I calculated the coefficient of variation, the Fano factor [Nawrot et al., 2008], the trial-averaged latency, the trial-averaged spike count, the lifetime sparseness [Willmore and Tolhurst, 2001], the adaptation from the first syllable of a song to the last one [Hildebrandt et al., 2009], and the pairwise spike train distances [van Rossum, 2001] averaged across all trial pairings. Thus, the response of each neuron to an acoustic song was described by feature vector with seven elements. Is it possible to discriminate between stimuli or their attractiveness based on these features? To investigate this question, I performed a principal component analysis. Figure 4.1 shows all 100 (25 neurons times 4 stimuli) feature vectors projected to the first two principal components, which capture 57% of the variance in the data. A classification into stimulus classes (Figure 4.1.A) or behaviorally relevant classes (Figure 4.1.B) failed. Only clusters of neuron types (Figure 4.1.C) appeared, though with areas of overlap. But, as the type of ascending neurons can be easily identified according to their characteristic morphology, this classification is not helpful.

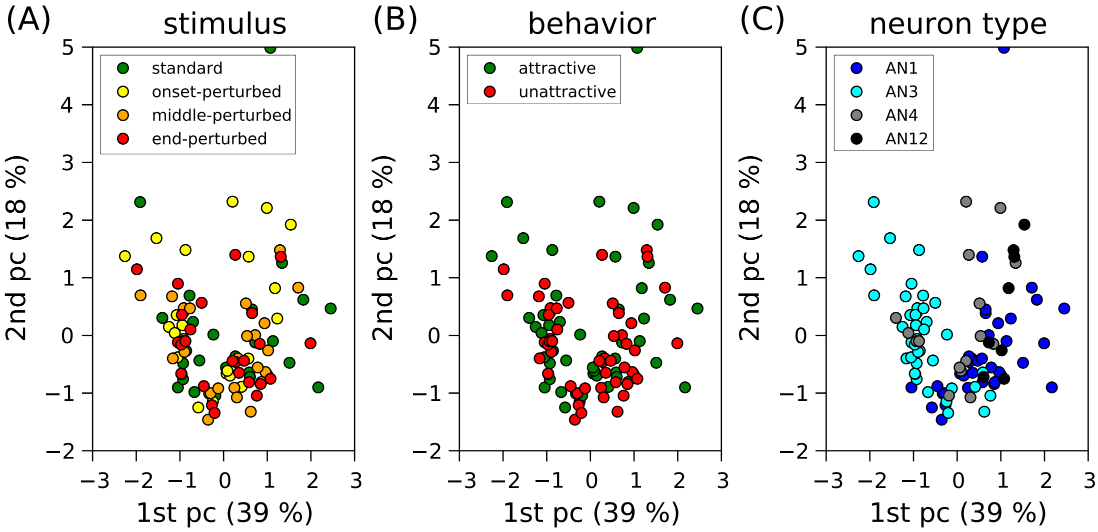

Figure 4.1: **PCA-based cluster analysis of stimuli, behavior related classes and neuron types from spike trains of grasshopper ascending neurons.** Each dot corresponds to a feature vector that describes the response of a single neuron, projected to the first two principal components. Neither stimulus classes (A) nor behaviorally relevant classes (B) appear; only classes of neuron types (C).

Bibliography

Balakrishnan, R., von Helversen, D., and von Helversen, O. (2001). Song pattern recognition in the grasshopper Chorthippus biguttulus : the mechanism of syllable onset and offset detection. *J. Comp. Physiol. A*, 187(4):255–264.

Bauer, M. and von Helversen, O. (1987). Separate localization of sound recognizing and sound producing neural mechanisms in a grasshopper. *J. Comp. Physiol. A*, 161(1):95–101.

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, 60(6):1142–1152.

Beckers, O. M. and Wagner, W. E. (2011). Mate sampling strategy in a field cricket: evidence for a fixed threshold strategy with last chance option. *Anim. Behav.*, 81(3):519–527.

Benda, J. and Hennig, R. M. (2008). Spike-frequency adaptation generates intensity invariance in a primary auditory interneuron. *J. Comput. Neurosci.*, 24(2):113–136.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer, New York.

Bradbury, J. and Vehrencamp, S. (1998). *Principles of animal communication*. Sinauer Associates, Sunderland.

Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses. *Netw. Comput. Neural Syst.*, 12(2):199–213.

Clemens, J. and Hennig, R. M. (2013). Computational principles underlying the recognition of acoustic signals in insects. *J. Comput. Neurosci.*, 35(1):75–85.

Clemens, J., Kutzki, O., Ronacher, B., Schreiber, S., and Wohlgemuth, S. (2011). Efficient transformation of an auditory population code in a small sensory system. *PNAS*, 108(33):13812–13817.

Clemens, J. and Ronacher, B. (2013). Feature extraction and integration underlying perceptual decision making during courtship behavior. *J. Neurosci.*, 33(29):12136–12145.

Clemens, J., Wohlgemuth, S., and Ronacher, B. (2012). Nonlinear computations underlying temporal and population sparseness in the auditory system of the grasshopper. *J. Neurosci.*, 32(29):10053–10062.

Creutzig, F., Wohlgemuth, S., Stumpner, A., Benda, J., Ronacher, B., and Herz, A. V. M. (2009). Timescale-invariant representation of acoustic communication signals by a bursting neuron. *J. Neurosci.*, 29(8):2575–2580.

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

Doherty, J. A. (1985). Trade-off phenomena in calling song recognition and phonotaxis in the cricket, Gryllus bimaculatus (Orthoptera, Gryllidae). *J. Comp. Physiol. A*, 156(6):787–801.

Drucker, H., Burges, Christopher J. C. Kaufman, L., Smola, A. J., and Vapnik, V. N. (1996). Support vector regression machines. *Adv. Neural Inf. Process. Syst.*, 9:155–161.

Drugowitsch, J. and Pouget, A. (2012). Probabilistic vs. non-probabilistic approaches to the neurobiology of perceptual decision-making. *Curr. Opin. Neurobiol.*, 22(6):963–969.

Duchamp-Viret, P., Duchamp, A., and Chaput, M. (2003). Single olfactory sensory neurons simultaneously integrate the components of an odour mixture. *Eur. J. Neurosci.*, 18(10):2690–2696.

Gold, J. I. and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, 30:535–574.

Gorodkin, J. (2004). Comparing two K-category assignments by a K-category correlation coefficient. *Comput. Biol. Chem.*, 28(5):367–374.

Gottsberger, B. and Mayer, F. (2007). Behavioral sterility of hybrid males in acoustically communicating grasshoppers (Acrididae, Gomphocerinae). *J. Comp. Physiol. A*, 193(7):703–714.

Grobe, B., Rothbart, M., Hanschke, A., and Hennig, R. M. (2012). Auditory processing at two time scales by the cricket, Gryllus bimaculatus.

Guyon, I. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182.

Hedwig, B. (2006). Pulses, patterns and paths: neurobiology of acoustic behaviour in crickets. *J. Comp. Physiol. A*, 192(7):677–689.

Hedwig, B. and Heinrich, R. (1997). Identified descending brain neurons control different stridulatory motor patterns in an acridid grasshopper. *J. Comp. Physiol. A*, 180(3):285–294.

Hedwig, B. and Poulet, J. F. A. (2005). Mechanisms underlying phonotactic steering in the cricket Gryllus bimaculatus revealed with a fast trackball system. *J. Exp. Biol.*, 208(5):915–927.

Heinrich, R., Wenzel, B., and Elsner, N. (2001). A role for muscarinic excitation: control of specific singing behavior by activation of the adenylate cyclase pathway in the brain of grasshoppers. *PNAS*, 98(17):9919–9923.

Hennig, R. M. (2009). Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket Gryllus bimaculatus. *J Comp Physiol A*, 195(10):971–987.

Hildebrandt, K. J., Benda, J., and Hennig, R. M. (2009). The origin of adaptation in the auditory pathway of locusts is specific to cell type and function. *J. Neurosci.*, 29(8):2626–2636.

Hinton, G. E. (2011). Machine learning for neuroscience. *Neural Syst. Circuits*, 1(1):12.

Hoare, D. J., Humble, J., Jin, D., Gilding, N., Petersen, R., Cobb, M., and McCrohan, C. (2011). Modeling peripheral olfactory coding in Drosophila larvae. *PLoS One*, 6(8):e22996.

Huber, F., Moore, T. E., and Loher, W. (1989). *Cricket behavior and neurobiology*. Cornell University Press.

Jacobs, K., Otte, B., and 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(3):270–285.

Jurman, G., Riccadonna, S., and Furlanello, C. (2012). A comparison of MCC and CEN error measures in multi-class prediction. *PLoS One*, 7(8):e41882.

Karmeier, K., Krapp, H. G., and Egelhaaf, M. (2005). Population coding of self-motion: applying Bayesian analysis to a population of visual interneurons in the fly. *J Neurophysiol*, 94(3):2182–2194.

Kastenhuber, E. and Neuhauss, S. (2011). Acoustic communication: Sound advice from piranhas. *Curr. Biol.*, 21(24):R986–R988.

Klappert, K. and Reinhold, K. (2003). Acoustic preference functions and sexual selection on the male calling song in the grasshopper Chorthippus biguttulus. *Anim. Behav.*, 65(1):225–233.

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

Lang, F. (2000). Acoustic Communication Distances of a Gomphocerine Grasshopper. *Bioacoustics*, 10(4):233–258.

Machens, C. K., Schütze, H., Franz, A., Kolesnikova, O., Stemmler, M. B., Ronacher, B., and Herz, A. V. M. (2003). Single auditory neurons rapidly discriminate conspecific communication signals. *Nat. Neurosci.*, 6(4):341–342.

Marsat, G. and Pollack, G. S. (2005). Effect of the temporal pattern of contralateral inhibition on sound localization cues. *J. Neurosci.*, 25(26):6137–6144.

Marsat, G. and Pollack, G. S. (2007). Efficient inhibition of bursts by bursts in the auditory system of crickets. *J. Comp. Physiol. A*, 193(6):625–633.

Martínez-García, M., Rolls, E., Deco, G., and Romo, R. (2011). Neural and computational mechanisms of postponed decisions. *PNAS*, 108(28):11626–11631.

McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 5:115–133.

Meckenhäuser, G., Hennig, R. M., and Nawrot, M. P. (2013). Critical song features for auditory pattern recognition in crickets. *PLoS One*, 8(2):e55349.

Meyer, A., Galizia, C. G., and Nawrot, M. P. (2013). Local interneurons and projection neurons in the antennal lobe from a spiking point of view. *J. Neurophysiol.*, 110(10):2465–2474.

Mhatre, N. and Balakrishnan, R. (2007). Phonotactic walking paths of field crickets in closed-loop conditions and their simulation using a stochastic model. *J. Exp. Biol.*, 210(20):3661–3676.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, 319(5869):1543–1546.

Nawrot, M. P. (2010). Analysis and interpretation of interval and count variability in neural spike trains. In Grün, S. and Rotter, S., editors, *Anal. parallel spike trains*, pages 37–58.

Nawrot, M. P., Aertsen, A., and Rotter, S. (1999). Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. *J. Neurosci. Methods*, 94(1):81–92.

Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *J. Neurosci. Methods*, 169(2):374–90.

Neuhofer, D., Stemmler, M., and Ronacher, B. (2011). Neuronal precision and the limits for acoustic signal recognition in a small neuronal network. *J. Comp. Physiol. A*, 197(3):251–265.

Neuhofer, D., Wohlgemuth, S., Stumpner, A., and Ronacher, B. (2008). Evolutionarily conserved coding properties of auditory neurons across grasshopper species. *Proc. R. Soc. B*, 275(1646):1965–1974.

Nissen, S. (2003). Implementation of a Fast Artificial Neural Network Library (fann). Technical report, Department of Computer Science University of Copenhagen.

Oliphant, T. E. (2007). Python for scientific computing. *Comput. Sci. Eng.*, 9(3):10–20.

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

Pollack, G. (2000). Who, what, where? Recognition and localization of acoustic signals by insects. *Curr. Opin. Neurobiol.*, 10(6):763–767.

Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci.*, 1(2):125–132.

Quiroga, R. Q. and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.*, 10(3):173–185.

Rickert, J., Riehle, A., Aertsen, A., Rotter, S., and Nawrot, M. P. (2009). Dynamic encoding of movement direction in motor cortical neurons. *J. Neurosci.*, 29(44):13870–13882.

Riedmiller, M. and Braun, H. (1993). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. *IEEE*, 1:586–591.

Ringnér, M. (2008). What is principal component analysis? *Nat. Biotechnol.*, 26(3):303–304.

Römer, H. (1987). Representation of auditory distance within a central neuropil of the bushcricket Mygalopsis marki. *J. Comp. Physiol. A*, 161(1):33–42.

Römer, H., Bailey, W., and Dadour, I. (1989). Insect hearing in the field III. Masking by noise. *J. Comp. Physiol. A*, 164(5):609–620.

Römer, H. and Marquart, V. (1984). Morphology and physiology of auditory interneurons in the metathoracic ganglion of the locust. *J. Comp. Physiol. A*, 155(2):249–262.

Ronacher, B. (2014). Processing of species-specific signals in the auditory pathway of grasshoppers. In Hedwig, B., editor, *Insect Hear. Acoust. Commun.*, chapter 11, pages 185–204.

Ronacher, B., Franz, A., Wohlgemuth, S., and Hennig, R. M. (2004). Variability of spike trains and the processing of temporal patterns of acoustic signals-problems, constraints, and solutions. *J. Comp. Physiol. A*, 190(4):257–77.

Ronacher, B., Helversen, D., and Helversen, O. (1986). Routes and stations in the processing of auditory directional information in the CNS of a grasshopper, as revealed by surgical experiments. *J. Comp. Physiol. A*, 158(3):363–374.

Ronacher, B. and Krahe, R. (1998). Song recognition in the grasshopper Chorthippus biguttulus is not impaired by shortening song signals: implications for neuronal encoding. *J. Comp. Physiol. A*, 183(6):729–735.

Ronacher, B., Krahe, R., and Hennig, R. M. (2000). Effects of signal duration on the recognition of masked communication signals by the grasshopper Chorthippus biguttulus. *J. Comp. Physiol. A*, 186(11):1065–1072.

Ronacher, B. and Stange, N. (2013). Processing of acoustic signals in grasshoppers - A neuroethological approach towards female choice. *J. Physiol.*, 107(1):41–50.

Ronacher, B. and Stumpner, A. (1988). Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. *J. Comp. Physiol. A*, 163(4):517–523.

Rosenblatt, F. (1958). The perceptron: A probalistic model for information storage and organization in the brain. *Psychol. Rev.*, 65(6):386–408.

Rost, T. (2011). *Modelling pattern recognition in cricket phonotaxis*. PhD thesis, Freie Universität Berlin.

Rost, T., Ramachandran, H., Nawrot, M. P., and Chicca, E. (2013). A neuromorphic approach to auditory pattern recognition in cricket phonotaxis. *IEEE*, pages 1–4.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(9):533–536.

Schildberger, K. (1984). Temporal selectivity of identified auditory neurons in the cricket brain. *J. Comp. Physiol. A*, 155(2):171–185.

Schildberger, K., Wohlers, D. W., Schmitz, B., Kleindienst, H. U., and Huber, F. (1986). Morphological and physiological changes in central auditory neurons following unilateral foreleg amputation in larval crickets. *J. Comp. Physiol. A*, 158(3):291–300.

Schmidt, A., Ronacher, B., and Hennig, R. M. (2008). The role of frequency, phase and time for processing of amplitude modulated signals by grasshoppers. *J. Comp. Physiol. A*, 194(3):221–33.

Schneider, E. and Hennig, R. M. (2011). Temporal resolution for calling song signals by female crickets, Gryllus bimaculatus. *J. Comp. Physiol. A*, 198(3):181–191.

Schul, J. (1998). Song recognition by temporal cues in a group of closely related bushcricket species (genus Tettigonia). *J. Comp. Physiol. A*, 183(3):401–410.

Selverston, A., Kleindienst, H.-U., and Huber, F. (1985). Synaptic connectivity between cricket auditory interneurons as studied by selective photoinactivation. *J. Neurosci.*, 5(5):1283–1292.

Sokoliuk, T., Stumpner, A., and Ronacher, B. (1989). GABA-Iike immunoreactivity suggests an inhibitory function of the thoracic low-frequency neuron (TN1) in acridid grasshoppers. *Naturwissenschaften*, 76:223–225.

Stange, N. and Ronacher, B. (2012). Grasshopper calling songs convey information about condition and health of males. *J. Comp. Physiol. A*, 198(4):309–318.

Stumpner, A. and Helversen, D. V. (2001). Evolution and function of auditory systems in insects. *Naturwissenschaften*, 88(4):159–170.

Stumpner, A. and Ronacher, B. (1991). Auditory interneurones in the metathoracic ganglion of the grasshopper Chorthippus biguttulus I. Morphological and physiological characterization. *J. Exp. Biol.*, 158(1):391–410.

Stumpner, A. and von Helversen, O. (1994). Song production and song recognition in a group of sibling grasshopper species (Ch. dorsatus, Ch. dichorus and Ch. loratus: Orthoptera, Acrididae). *Bioacoustics*, 6(1):1–23.

Sutskever, I., Hinton, G., and Taylor, G. (2008). The Recurrent Temporal Restricted Boltzmann Machine. In *NIPS*.

Thorson, J., Weber, T., and Huber, F. (1982). Auditory behavior of the cricket II. Simplicity of calling-song recognition in Gryllus, and anomalous phonotaxis at abnormal carrier frequencies. *J. Comp. Physiol. A*, 146(3):361–378.

Tripathy, S. J., Padmanabhan, K., Gerkin, R. C., and Urban, N. N. (2013). Intermediate intrinsic diversity enhances neural population coding. *PNAS*, 110(20):8248–8253.

van Helversen, D. and van Helversen, O. (1983). Species recognition and acoustic localization in acridid grasshoppers: a behavioral approach. In Huber, F. and Markl, H., editors, *Neuroethol. Behav. Physiol.*, pages 253 – 284. Springer.

van Rossum, M. C. (2001). A novel spike distance. *Neural Comput.*, 13(4):751–763.

Viemeister, N. F. and Wakefield, G. H. (1991). Temporal integration and multiple looks. *J. Acoust. Soc. Am.*, 90(2):858–865.

Vogel, A., Hennig, R. M., and Ronacher, B. (2005). Increase of neuronal response variability at higher processing levels as revealed by simultaneous recordings. *J. Neurophysiol.*, 93(6):3548–3559.

Vogel, A. and Ronacher, B. (2007). Neural correlations increase between consecutive processing levels in the auditory system of locusts. *J. Neurophysiol.*, 97(5):3376–3385.

von Helversen, D. (1972). Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke Chorthippus biguttulus (Orthoptera, Acrididae). *J. Comp. Physiol. A*, 81(4):381–422.

von Helversen, D. and von Helversen, O. (1997). Recognition of sex in the acoustic communication of the grasshopper Chorthippus biguttulus (Orthoptera, Acrididae). *J. Comp. Physiol. A*, 180(4):373–386.

von Helversen, D. and von Helversen, O. (1998). Acoustic pattern recognition in a grasshopper: processing in the time or frequency domain? *Biol. Cybern.*, 79(6):467–476.

von Helversen, O. (1979). Angeborenes Erkennen akustischer Schlüsselreize. In Rathmayer, W., editor, *Verhandlungen der Dtsch. Zool. Gesellschaft*, pages 42–59. Gustav Fischer Verlag, Stuttgart.

von Helversen, O. and von Helversen, D. (1994). Forces driving coevolution of song and song recognition in grasshoppers. In Schildberger, K. and Elsner, N., editors, *Fortschr. Zool.*, pages 253–283. Gustav Fischer Verlag.

von Philipsborn, A. C., Liu, T., Yu, J. Y., Masser, C., Bidaye, S. S., and Dickson, B. J. (2011). Neuronal control of Drosophila courtship song. *Neuron*, 69(3):509–522.

Webb, B. (1995). Using robots to model animals: a cricket test. *Rob. Auton. Syst.*, 16(2):117–134.

Weber, T., Thorson, J., and Huber, F. (1981). Auditory behavior of the cricket I. Dynamics of compensated walking and discrimination paradigms on the Kramer treadmill. *J. Comp. Physiol. A*, 141(2):215–232.

Weiss, M., Kiefer, S., and Kipper, S. (2012). Buzzwords in females' ears? The use of buzz songs in the communication of nightingales (Luscinia megarhynchos). *PLoS One*, 7(9):e45057.

Williams, R. W. and Herrup, K. (1988). The control of neuron number. *Annu. Rev. Neurosci.*, 11(1):423–453.

Willmore, B. and Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Netw. Comput. Neural Syst.*, 12(3):255–70.

Wirmer, A., Faustmann, M., and Heinrich, R. (2010). Reproductive behaviour of female Chorthippus biguttulus grasshoppers. *J. Insect Physiol.*, 56(7):745–53.

Wittmann, J. P., Kolss, M., and Reinhold, K. (2010). A neural network-based analysis of acoustic courtship signals and female responses in Chorthippus biguttulus grasshoppers. *J Comput Neurosci*, 31(1):105–115.

Wohlers, D. W. and Huber, F. (1982). Processing of sound signals by six types of neurons in the prothoracic ganglion of the cricket, Gryllus campestris L. *J. Comp. Physiol.*, 146(2):161–173.

Wohlgemuth, S. and Ronacher, B. (2007). Auditory discrimination of amplitude modulations based on metric distances of spike trains. *J. Neurophysiol.*, 97(4):3082–3092.

Woolley, S., Sakata, J., and Crews, D. (2004). Evolutionary insights into the regulation of courtship behavior in male amphibians and reptiles. *Physiol. Behav.*, 83(2):347–360.

Young, D. and Ball, E. (1974). Structure and development of the auditory system in the prothoracic leg of the cricket Teleogryllus commodus (Walker). *Zeitschrift für Zellforsch. und Mikroskopische Anat.*, 147(3):293–312.

Yovel, Y., Franz, M. O., Stilz, P., and Schnitzler, H.-U. (2008). Plant classification from bat-like echolocation signals. *PLoS Comput. Biol.*, 4(3):e1000032.

Zorovic, M. and Hedwig, B. (2011). Processing of species-specific auditory patterns in the cricket brain by ascending, local, and descending neurons during standing and walking. *J. Neurophysiol.*, 105(5):2181–2194.