LETTER **Communicated by Carlos Brody**

Timescale-Invariant Pattern Recognition by Feedforward Inhibition and Parallel Signal Processing

Felix Creutzig
felix.creutzig@tu-berlin.de
Department Economics of Climate Change, Technische Universität Berlin, 10623 Berlin, Germany

Jan Benda
benda@bio.lmu.de
Department of Biology and Bernstein Center for Computational Neuroscience München, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany

Sandra Wohlgemuth
sandra.wohlgemuth@fu-berlin.de
Abt. Verhaltensbiologie, Institut für Biologie, Freie Universität Berlin, 14195 Berlin, Germany

Andreas Stumpner
astumpn@gwdg.de
Johann-Friedrich-Blumenbach-Institut für Zoologie und Anthropologie, Abt. Zelluläre Neurobiologie, Georg-August Universität Göttingen, 37077 Göttingen, Germany

Bernhard Ronacher
bernhard.ronacher@rz.hu-berlin.de
Department of Biology and Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität zu Berlin, 10115 Berlin, Germany

Andreas V. M. Herz
herz@bio.lmu.de
Department of Biology and Bernstein Center for Computational Neuroscience München, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany

The timescale-invariant recognition of temporal stimulus sequences is vital for many species and poses a challenge for their sensory systems. Here we present a simple mechanistic model to address this computational task, based on recent observations in insects that use rhythmic acoustic communication signals for mate finding. In the model framework, feedforward inhibition leads to burst-like response patterns in one neuron of the circuit. Integrating these responses over a fixed time window by a readout neuron creates a timescale-invariant stimulus representation.

Only two additional processing channels, each with a feature detector and a readout neuron, plus one final coincidence detector for all three parallel signal streams, are needed to account for the behavioral data. In contrast to previous solutions to the general time-warp problem, no time delay lines or sophisticated neural architectures are required. Our results suggest a new computational role for feedforward inhibition and underscore the power of parallel signal processing.

1 Introduction

Sensory systems extract behaviorally relevant features from incoming stimuli, often in real time. A particular challenge arises when stimulus attributes need to be computed that extend over longer time intervals, such as the duration of sound elements within acoustic communication signals. In this situation, the stimulus time course must be represented such that all potentially relevant information remains transiently stored during the feature-extraction process.

Natural signal variations may strongly distort a temporal pattern sequence. In an acoustic communication system, for example, changes of the sender's biophysical properties can result in large variations of the frequency spectrum or even sequence speed. To correctly recognize a compressed or stretched sequence, stimulus invariances under local or global rescaling (time-warp and time-scale invariance, respectively) need to be exploited. Humans and various animals master this task, as shown by the ease with which we correctly identify spoken words even if the speaking rate is varied by a factor of two or three (Klatt, 1976; Port & Dalby, 1982). Similarly, many ectothermic animals use acoustic communication signals, although the overall temporal scale of these signals varies strongly with the surround temperature (Hauser & Konishi, 1999; Römer, 2001; Gerhardt & Huber, 2002). In contrast, technical approaches, such as hidden Markov models, which identify normal speech with great success, are often impaired under stimulus conditions with local time warp or global temporal rescaling (Juang & Rabiner, 1991).

Previous theoretical solutions to the time warp problem exploit subthreshold oscillations and coincidence detection (Hopfield, 1996), transient synchronization (Hopfield & Brody, 2001), and synfire chains with different sorts of inhibition (Jin, 2004). Some of these models achieve timescale-invariant sequence recognition at the expense of high computational costs or a large number of dedicated delay lines. We aim at a simpler yet general framework based on electrophysiological and behavioral data from a well-established model system: grasshopper acoustic communication.

Various grasshopper species generate sound signals to be detected by conspecific mates who can also use the song to assess the singer's genetic fitness (von Helversen & von Helversen, 1994). Grasshoppers generate these "songs" by rhythmically rasping their hind legs across their forewings. This

results in a periodic sound pattern that alternates between "syllables" and "pauses," that is, sound segments with high and low amplitudes, respectively. In species such as *Chorthippus biguttulus*, the behavioral response depends to a large extent on the ratio of syllable to pause duration (von Helversen, 1972). If this ratio is kept constant, the absolute length of one song unit (syllable plus pause) can vary by more than 300% without impairing the behavioral response (von Helversen & von Helversen, 1994).

On the receiver side, the communication signals are processed by a feedforward auditory system. Following mechanosensory signal transduction (Gollisch & Herz, 2005), receptor neurons encode the time-varying stimulus envelope with high reliability and temporal precision (Stumpner, Ronacher, & von Helversen, 1991; Machens et al., 2003; Rokem et al., 2006). These neurons project to local interneurons in the metathoracic ganglion whose responses are processed by a small set of ascending neurons (ANs) and then sent to the brain (Stumpner & Ronacher, 1994). As the highest neural processing stage, the brain integrates the AN signals with other, potentially multimodal information to generate behavioral responses (Ronacher, von Helversen, & von Helversen, 1986).

One particular ascending neuron (AN12) generates burst-like response patterns at syllable onset. In a recent study, we have shown that the intraburst spike count of this neuron scales linearly with the duration of the preceding sound pause (Creutzig et al., 2009; Creutzig, 2008). Integrating the number of spikes over an extended time interval results in an unexpected observation: the total spike count reflects the characteristic syllable-to-pause ratio of the species while being invariant to global temporal rescaling of the song. Simply counting the AN12 spikes within a fixed time window thus provides the behaviorally relevant information without any need to explicitly calculate ratios between the syllable and pause durations. We now address two key questions that are directly related to this finding. First, is there a biologically plausible mechanism to explain the AN12 burst patterns? Second, can this spike-counting strategy be applied to other song features, and, more generally, is there a simple neural circuit to account for the observed behavioral response pattern of grasshoppers?

2 Results

2.1 Feedforward Inhibition Explains Burst-like Discharge Patterns. To explain the burst-like responses of the AN12 neuron, we use a minimal model framework as depicted in Figure 1. In this model, the acoustic signal drives auditory receptor neurons whose firing rate adaptation emphasizes the initial part of the sound pulse. An excitatory channel forwards the receptor response directly to the AN12 neuron, whereas a second, inhibitory channel contributes with a low-pass-filtered version of the receptor output. This role could be played by the SN1 neuron, which may provide delayed inhibition to the AN12 neuron (Sokoliuk, 1992). The difference of

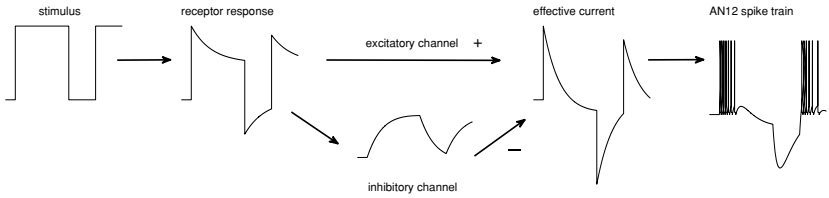

Figure 1: Signal processing of the AN12 neuron, a key player for timescale-invariant auditory sequence recognition in grasshoppers. The function of the feedforward inhibition circuit is demonstrated with a block stimulus that mimics the characteristic syllable-pause structure of grasshopper acoustic communication signals. Adaptation of the auditory receptor neurons emphasizes the initial part of the syllable. The reduced sound intensity during the pause still excites some of the receptor neurons that have, hence, a nonzero steady-state firing rate between syllables. Due to the adaptation dynamics, that asymptotic rate is higher than the receptor activity directly after stimulus offset. An excitatory channel forwards this signal directly to the AN12 neuron, whereas a second, inhibitory feedforward channel contributes with a low-pass-filtered version of the signal, as sketched in the lower central part of the figure. The difference of both activities drives the leaky integrate-and-fire dynamics of the AN12 model neuron and leads to the burstlike discharge patterns.

both activities drives the leaky integrate-and-fire dynamics of the AN12 neuron.

The mathematical model for this feedforward inhibition contains seven free parameters (see below and the appendix). The two parameters of the receptor neuron were calibrated by hand within the range of experimentally observed values (Benda, 2002): The relative adaptation level was set to $A_r = 0.5$, and the adaptation time constant was chosen as $\tau_r = 30$ ms. As shown by simulations, the circuit dynamics does not depend sensitively on the precise values of these parameters. The other five parameters (the two inhibitory channel parameters τ_{inh}, A_{inh} and the three parameters τ_V, V_{th}, and V_{reset} of the integrate-and-fire neuron) were optimized for all six measured AN12 neurons with respect to the average distance between experimental and model spike trains, as measured by the quality measure Γ (see the appendix). For all cells, the resulting parameter values for the inhibitory feedforward neuron (τ_{inh}, A_{inh}) were rather similar, $\tau_{inh} = [30 - 45$ ms$]$, $A_{inh} = [1.25 - 1.4]$. The model quality hardly varied within this range of parameters. We therefore fixed $\tau_{inh} = 40$ ms and $A_{inh} = 1.3$ for all cells and individually optimized the remaining integrate-and-fire parameters. Finally, the refractory period was set to 1.75 ms for all cells, corresponding to the shortest observed interspike interval. The optimization was done as follows. Optimization criterion was the Γ-value $\in [0, 1]$, a measure of similarity as defined in the appendix. All five parameters were optimized

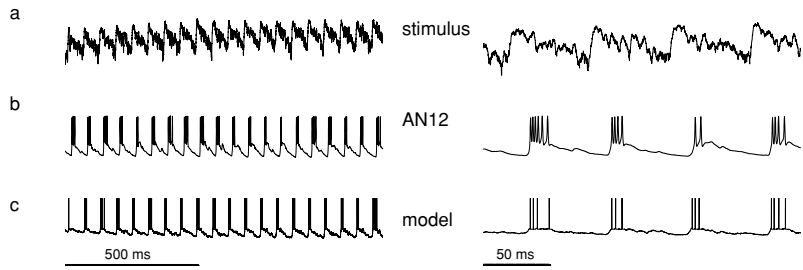

Figure 2: Comparison of experimental and simulated AN12 responses. (a) Amplitude envelope of a characteristic grasshopper song: left, longer sound stretch illustrating overall song characteristics; right, expanded time course emphasizing syllable-to-syllable variability. (b) Measured response of the AN12 neuron. Bursts follow syllable onset with high reliability and a response latency of about 12 ms. This latency is due to acoustic delays, response latencies of the receptor neurons, propagation delays along the auditory nerve, synaptic delays, and the intrinsic dynamics of the AN12 cell. (c) The feedforward inhibition circuit reproduces the experimental burst characteristics.

such that for seven songs, experimental and model spike trains were as similar as possible. With these parameter values, the Γ-value was calculated for the eighth song. This leave-one-out procedure was repeated for all songs to compute a mean Γ-value. Each cell was fitted with an individual parameter set. For the six cells, this led to a Γ-range of 0.33 to 0.75 with a median of 0.53. As a comparison, we also measured the intertrial similarity between repetitions of the same song and its average spike train for each cell (range: 0.49–0.84; median $= 0.60$) in the recorded AN12 data. This shows that the trial-to-trial variability is almost as large as the deviations between the experimental and simulated responses, underscoring the model quality.

The model dynamics agree well with measured AN12 responses, even if natural grasshopper songs are used as realistic input stimuli. This is demonstrated in Figure 2 for the raw spike patterns and in Figure 3 for two key response characteristics of the AN12 neuron: the burst-triggered average (BTA) of the model neuron, that is, the average stimulus preceding a burst with given spike count, displays the same qualitative features as the real neuron (see Figures 3a to 3c). Similarily, the correlation between the intraburst spike count (IBSC) and preceding pause duration closely follows a linear relation in both, experiment and model (see Figures 3b and 3d). Similar to physiological results from the AN12 neuron (Creutzig et al., 2009), the model spike count is independent from both onset slope and syllable duration and increases slightly with onset amplitude.

Without receptor adaptation, the model exhibits unrealistically high activity levels of up to 26 spikes per burst in response to block stimuli, far

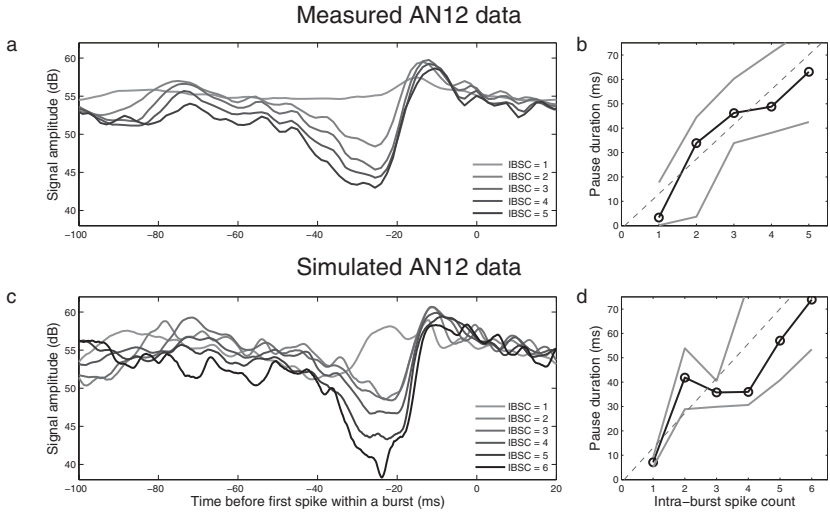

Figure 3: Encoding properties of the AN12 neuron. (a) Burst-triggered averages (BTAs) calculated from experimental AN12 responses. The BTA is the average stimulus preceding a burst with given intraburst spike count (IBSC). (b) Pause durations as a function of subsequent IBSCs; gray lines denote upper and lower quartiles. The best linear fit agrees well with a direct proportionality: the y-intercept is only $-1.1 + /- 2.1$ ms. (c) Burst-triggered averages from the modeled AN12 neuron resemble experimental BTAs. (d) Similarly, IBSCs of the model neuron reflect pause durations; the gray lines denote again upper and lower quartiles. Data in *a* and *b* redrawn from Creutzig et al. (2009).

more than observed in experiments (up to approximately 10 spikes). In accordance with the literature (Römer, 1976; Benda, 2002), we therefore kept receptor adaptation as an important part of the model.

Could the burst response of the AN12 neuron also be explained by a cell-intrinsic mechanism? It is well known that spike frequency adaption (Benda & Herz, 2003) leads to burst-like spike patterns at stimulus onset (Izhikevich, 2004). However, any adaptation mechanism caused by spike activity would imply that the intraburst spike count depended on the duration, and possibly also intensity, of the preceding syllable—but not on pause duration. Such a dependence cannot be inferred from statistical observations for AN12 neurons (Creutzig et al., 2009). The separation of cell-intrinsic adaptation from downstream processes has been investigated for some grasshopper interneurons by recording from the same cell applying both acoustic and intracellular current stimulation (Hildebrandt, Benda, & Hennig, 2009). However, this approach has not yet been applied to the AN12 neuron. We would expect that its phasic response is due to presynaptic processes. We conclude that feedforward inhibition as sketched in

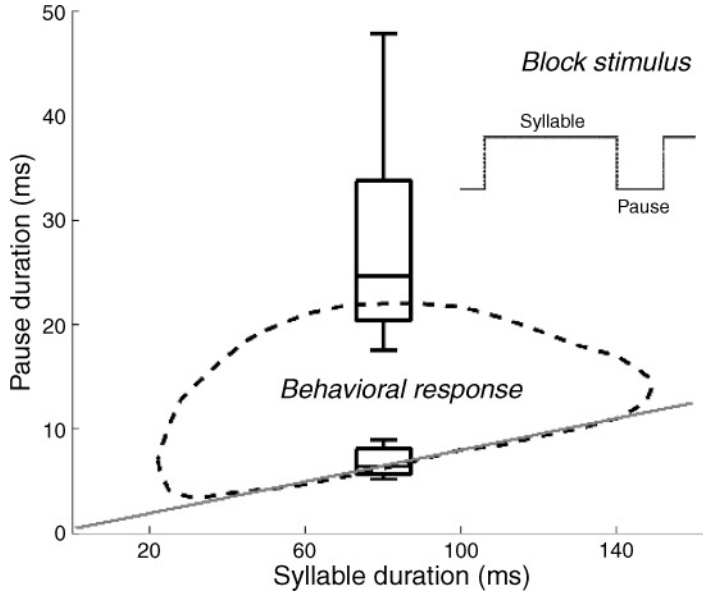

Figure 4: Behavioral response of *Chorthippus biguttulus* females. The dashed curve encircles the region of syllable and pause durations within artificial block stimuli (see inset) to which female grasshoppers respond (one animal, 20% response level). Boxes illustrate the response variability across 17 females for a syllable duration of 80 ms. The gray line indicates a timescale invariance: jointly, rescaling of pause and syllable durations does not change behavioral response, at least for syllable durations between about 40 and 140 ms. Modified from von Helversen and von Helversen (1994).

Figure 1 is sufficient to explain the burst-like discharge patterns observed in the electrophysiological experiments; no intrinsic burst mechanisms are required.

2.2 A Minimal Neural Circuit Accounts for Behavioral Responses. Female grasshoppers evaluate male songs and respond only to specific syllable-pause combinations. In particular, the behavioral experiments of von Helversen and von Helversen (1994) with artificial model songs show that equally preferred songs lie on oval curves in the syllable-pause plane (see Figure 4).

How is the information contained in AN12 spike trains related to these behavioral responses? On the lower border of the behavioral response curve in Figure 4 and for syllable durations between about 40 and 140 ms, pause duration is proportional to syllable duration. Within this parameter region, temporal rescaling of both syllable and pause duration thus does not

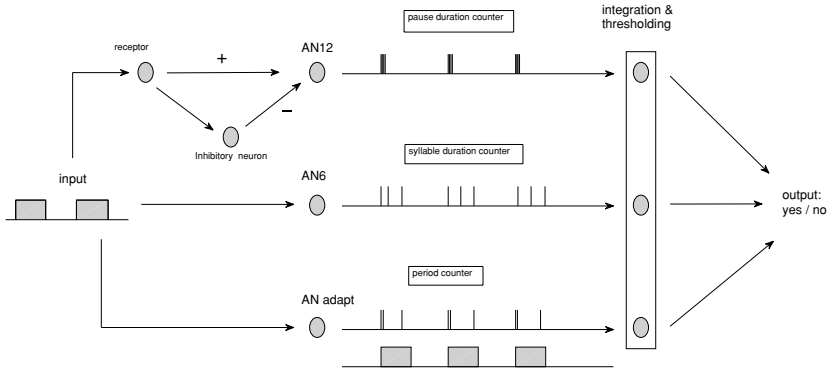

Figure 5: Full model circuit. A set of parallel feature detectors evaluates the input stimuli. Here, the AN12 neuron measures total pause duration, the AN6 encodes total syllable duration, and a third, rapidly adapting neuron counts the number of syllables.The outputs of these three neurons are integrated in time, thresholded, and then combined through an AND operation that could be implemented through a coincidence detector.

result in behavioral changes. Interestingly, the AN12 response patterns explain this timescale invariance. As discussed by Creutzig et al. (2009), the intraburst spike count of the AN12 neuron increases linearly with pause duration. Stretching both syllable and pause durations thus increases the number of spikes per burst. The number of bursts per unit time decreases by the same factor. Hence, the total number of spikes within a fixed readout window is constant under a global temporal rescaling of the song.

A simple mechanism could account for the observed behavior of Figure 4. If the total number of AN12 spikes per unit time is above a certain target value, the female responds to the song, and otherwise not. This would explain the almost straight lower boundary of the behavioral response curves. To explain the deviations at small and large syllable durations as well as the upper part of the iso-response oval, two additional mechanisms are required (see Figure 5). To avoid responses when the pause-to-syllable ratio is too large, the total duration of all syllables within a fixed time window could be measured and required to be above a certain threshold. The permitted songs would lie below the straight line denoted by "syllable-integration threshold" in Figure 6. Similarly, to avoid responses at high syllable and pause durations, the total number of syllables in a fixed time window could be measured and required to exceed another threshold. To satisfy this third criterion, songs would need to be to the left of the straight line denoted by the syllable-count threshold.

These threshold conditions can be formalized. For the pause-integration threshold, if spike count is proportional to pause duration and we inquire

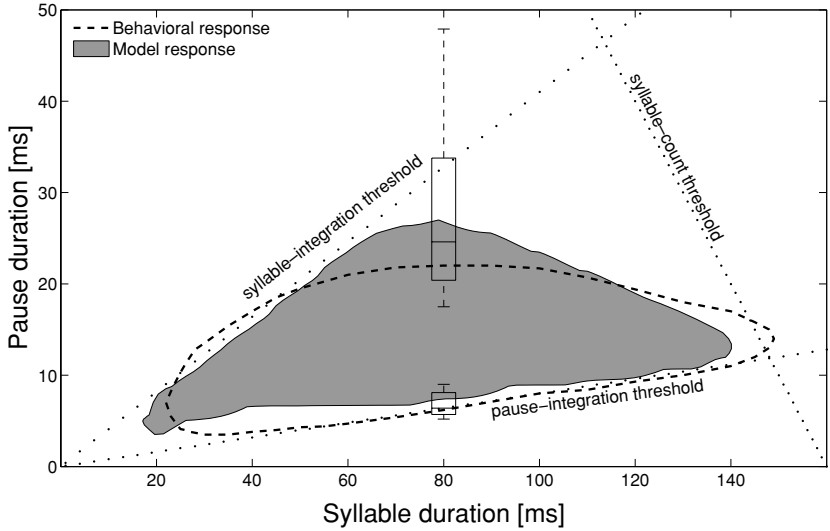

Figure 6: Model behavior. For artificial block stimuli, the circuit model sketched in Figure 5 generates a positive response if the syllable and pause durations of the input stimuli are within a triangle delimited by the dotted lines. Firing rate adaptation of the period-counter neuron bends the "syllable count" threshold and results in the shaded response region, which resembles the behavioral data from Figure 4, shown by the dashed lines. The zigzag contour is due to finite size effects of the numerical simulation.

about the total rate of the spikes per second, then if k is the spike count, p is pause duration, s is syllable duration, and α_p is a constant of proportionality,

$$k = \frac{\alpha_p p}{p + s}. \tag{2.1}$$

If k_p is the marginal spike count that elicits a response, then a necessary condition for behavioral reaction is given by $k > k_p$ which translates into

$$p > \frac{k_p s}{\alpha_p - k_p}. \tag{2.2}$$

Similarly for the syllable-integration threshold, if spike count is proportional to syllable duration and α_s is another constant of proportionality and k_s is the marginal spike count that elicits a response, then

$$s > \frac{k_s p}{\alpha_s - k_s}. \tag{2.3}$$

For the syllable-count threshold, the number of syllable-plus-pause units per time, that is, the syllable frequency, f_{s+p} must be above a certain

threshold:

$$f_{s+p} > f_0. \tag{2.4}$$

In the resulting circuit model, input information is transmitted by ascending neurons, whereas integration and thresholding are performed by downstream readout neurons. Pause integration is done by the summation of the AN12 spike train, as discussed before. Syllable integration is accomplished by a second neuron that fires tonically in response to sound inputs, resulting in a spike count that is proportional to syllable duration. The AN6 neuron fits this description (Stumpner et al., 1991). The AN6 neuron fires tonically throughout the syllables, with rather little adaptation during a song (Stumpner & Ronacher, 1991; Krahe, Budinger, & Ronacher, 2002). It is thus an ideal candidate for measuring syllable duration. A hypothetical readout neuron counts the number of spikes per time, and, assuming direct proportionality, this neuron infers information on the total syllable duration. A third neuron that responds phasically to syllable onsets provides the input for a syllable counter. This neuron spikes at the beginning of a syllable. A count of spikes by a readout neuron can then be interpreted as a syllable counter. A number of neurons respond phasically, such as the AN1, AN3, AN4, and AN12. However, the AN1 responds only to high intensities, and the AN3 and AN4 exhibit strong response variability. The AN12 is the most reliable candidate for syllable counting too. In this case, however, information is readout by taking a burst as a unit of information, that is, by interpreting the number of bursts as the number of syllables. As little is known about the readout in the brain ganglion, we chose to depict the syllable count as a separate process in Figure 6.

The ascending neurons, such as the AN12 and the AN6, transmit analog, continuous information on specific parameters, such as pause durations. However, the brain ganglion is not required to keep all this information. Rather, brain neurons can simply threshold the spike count and respond only with a positive output (e.g., a spike) when the specific spike count is high enough. This is a very simple neural architecture. As a final step, a logical AND operation could be implemented through coincidence detection—by another thresholding operation—which results in the triangular response field of Figure 6.

To obtain the desired response oval instead of the triangle, the third neuron adapts to a small, positive firing rate. Hence, in addition to its strong phasic response, this neuron is also active for long syllable duration. In effect, when pause duration increases while syllable-plus-pause duration is constant (i.e., along straight lines that are parallel to the syllable-count threshold in Figure 6), fewer spikes are elicited. Accordingly, a higher syllable count is required to reach the threshold so that the syllable count line bends leftward. With a time constant of $\tau = 3$ ms and a steady-state firing rate that equals 10% of the onset rate, the output of the model circuit comes

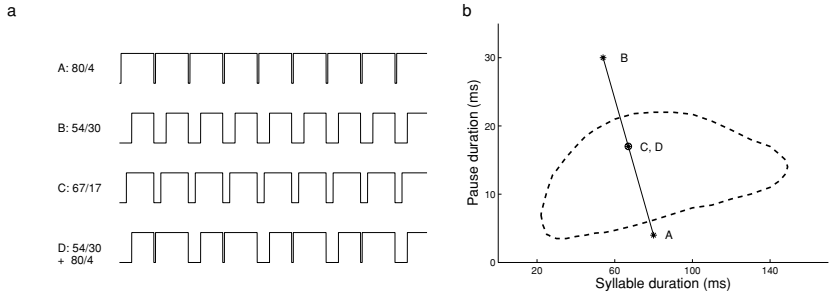

Figure 7: Prediction of the integration hypothesis. (a) Schematic drawing of four different block stimuli. Parameter values correspond to syllable and pause durations (in ms), respectively. Model songs A and B would not elicit a behavioral response because their pause and syllable durations lie outside the parameter region, resulting in positive behavioral responses. In contrast, model song C falls into that region. Song D combines elements of songs A and B and has the same syllable-pause ratio as song C. If syllable and pause durations are integrated over a timescale that is much longer than the song elements, mixture D should also elicit a behavioral response. (b) The mean syllable-to-pause combination of each model song together with the region of positive behavioral responses (redrawn from Figure 4).

close to the measured behavioral data, as shown in Figure 6. Note that the model predicts a positive behavioral response even for very short pause and syllable durations (the gray-shaded region in the lower left corner of the triangle of Figure 6). In experiments, *Chorthippus biguttulus* does not respond to these combinations of short syllables and short pauses (von Helversen & von Helversen, 1997). In this regime, however, additional processes such as gap detection (von Helversen, 1972; Ronacher & Stumpner, 1988) are operating, so that a proper extension of the model could account for the differences.

The predictive power of the model could easily be tested. Since temporal integration of ascending neurons is responsible for song recognition, the detailed song structure does not play any role within the model framework. All that matters are temporal averages of syllable and pause durations. Hence, alternating between two different syllable-pause combinations, each of which is not sufficient to elicit behavioral responses, should lead to a positive classification (see Figure 7).

3 Discussion

3.1 A New Computational Role of Feedforward Inhibition. As demonstrated in this letter, a small and rather simple neural circuit can be used to encode the duration of a stimulus pause (see Figure 1). A key

ingredient is a strong inhibitory channel parallel to the direct excitatory input to this neuron. The long time constant $\tau_{inh} = 40$ ms of the inhibitory neuron makes it sensitive for pause durations. For short pauses, the inhibitory activation r_{inh} is still near the maximum value reached during the preceding stimulus pulse so that it dominates the excitatory channel, and thus the AN12 receives a small input eliciting only one or two spikes. With increasing pause length, the inhibition decays away, and therefore the excitatory input to the AN12 gets stronger and stronger, resulting in more and more spikes evoked by the syllable onset. As a consequence, the output neuron generates a burst whose intraburst spike count directly reflects the pause duration (see Figure 3).

Feedforward inhibition can thus be used not only to enforce the temporal precision of downstream neurons, as has been reported for auditory cortex (Cruikshank, Rose, & Metherate, 2002; Elhilali, Fritz, Klein, Simon, & Shamma, 2004; Tan, Zhang, Merzenich, & Schreiner, 2004; Wehr & Zador, 2005), hippocampus (Pouille & Scanziani, 2001) and lateral geniculate nucleus (Blitz & Regehr, 2005). Furthermore, feedforward inhibition helps to tune bat auditory midbrain neurons to behaviorally important parameters, including signal duration (Covey & Casseday, 1999), and is used for gap detection in grasshoppers, another pattern detection circuit (Krahe et al., 2002). As shown by our results, feedforward inhibition can also help integrate information on long timescales and yet maintain a precisely timed response, thus providing a simple solution to the resolution-integration paradox (deBoer, 1985; Nelken, 2004). In addition, the grasshopper example demonstrates that feedforward inhibition can be used to compress stimulus information—a pause of 40 ms is represented by a short burst that may last for 4 ms only. This burst code allows the system to multiplex information about the time-of-occurrence and behavioral relevance of a particular sound pattern (Creutzig et al., 2009). Feedforward inhibition is thus a network motif with most interesting computational capabilities.

Similar to our model, an interplay of fast and slow dynamics in bursting neurons has been used to model time-warp invariance in Gollisch (2008). This model differs from our approach in relying on intrinsic bursting dynamics and a push-pull mechanism to drive the output neuron. Despite these fundamental cellular differences, the model generates similar output behavior for block stimuli. It would therefore be interesting to test the predictions of both models with more complicated inputs such as the mixture stimuli shown in Figure 7.

3.2 Parallel Signal Processing and Temporal Integration. Summation of AN12 spikes within a fixed time window can explain timescale invariance along one side of the behavioral response curve, as illustrated by the gray line in Figure 4. However, this moving average of AN12 activity cannot account for the full behavioral response curve. We have therefore assumed that three song features are processed independent of each other

(see Figure 5). Here, pause duration is measured by the AN12 neuron, syllable duration is measured by the AN6 neuron, and the syllable count is measured by a phasic neuron, potentially also the AN12. The spike (or burst) count of each neuron is thresholded by readout neurons. If each readout neuron responds positively, a behavioral response is elicited.

This model of parallel feature detection, integration, and subsequent thresholding is reminiscent of the pulse-integrator model proposed by Alder and Rose (1998) for sound-pattern recognition in acoustically communicating frogs. The behaviorally relevant timescale of integration for female grasshoppers is around 1 s. Figure 18 in von Helversen (1972) shows a steep decline of responsiveness if song models shorter than 1.2 s were presented. The required long integration times have been observed not only in frogs but also in various other species, including electric fish (Oestreich, Dembrow, George, & Zakon, 2006) and primates (Luna, Hernandez, Brody, & Romo, 2005). This makes us confident that a similar integration mechanism may exist in insects too.

In summary, we have shown how a seemingly complicated task such as timescale-invariant pattern recognition can be accomplished by a rather elementary circuit. The proposed combination of feedforward inhibition, parallel feature detection, and temporal integration is consistent with known neurophysiological data. The model framework explains behavioral data and makes quantitative predictions to be tested in future experiments. The simplicity of the circuit and its surprising computational capabilities make it an ideal building block for more general temporal-sequence processing tasks.

Appendix: Technical Notes on the Modeling Results

A.1 AN12 Model. Within the sound-preprocessing circuit (see Figure 1), the firing rate adaptation $a(t)$ of auditory receptor neurons is determined by the time-dependent acoustic stimulus $s(t)$ (Benda & Herz, 2003),

$$\tau_r \frac{da(t)}{dt} = -a(t) + s(t). \tag{A.1}$$

The adaptation time constants τ_r of grasshopper receptor neurons are in the range of 30 to 50 ms (Gollisch & Herz, 2004). They are thus at least one order of magnitude larger than the time constants relevant for electrical integration (Gollisch & Herz, 2005). On the phenomenological level of the current model framework, this separation of timescales allows us to approximate the receptors' time-dependent firing rate $r_r(t)$ by

$$r_r(t) = s(t) - A_r a(t), \tag{A.2}$$

where $A_r \in [0, 1]$ denotes the scaling factor for adaptation.

The AN12 neuron receives a mixed input typical for a feedforward inhibition circuit. The first component is the receptor output $r_r(t)$. The second component is a sign-inverted low-pass-filtered copy of $r_r(t)$ generated by an inhibitory interneuron with firing rate $r_{inh}(t)$,

$$\tau_{inh}\frac{dr_{inh}(t)}{dt} = -r_{inh}(t) + r_r(t).$$

To capture the discrete spike output of the AN12 neuron, it is described as a leaky integrate-and-fire neuron,

$$\tau_{RC}\frac{dV(t)}{dt} = -V(t) + r_r(t) - A_{inh}r_{inh}(t), \tag{A.3}$$

where A_{inh} denotes the relative strength of the inhibitory input. Whenever the voltage $V(t)$ reaches the threshold V_{th}, the model neuron spikes and $V(t)$ is reset to V_{reset}. After a refractory period determined by the minimum interspike interval of the experimental data (1.75 ms), the cell again integrates the feedforward inputs. Optimizing parameter values led to $\tau_{\mathrm{RC}} = 6.9 \pm 1.2$ ms, $V_{th} = 0.01 \pm 0.004$, and $V_{reset} = 0.006 \pm 0.004$.

A.2 Model Quality. From the digitized recording signal, spike times were determined using a voltage-threshold criterion. To account for neural fatigue, bursts were defined by a recursive formula: a spike belongs to a burst if it follows the preceding nth spike by no more than $(3 + n)$ ms. The first interspike interval within a burst is thus at most 4 ms, the second at most 5 ms, and so on. This definition reflects the observation that interspike intervals within an AN12 burst typically increase from spike to spike. Applying a fixed interspike interval (ISI) criterion would either miss the later spikes within a long burst (for short ISI cutoffs) or misclassify isolated spikes as belonging to a burst (for longer ISI cutoffs). For notational simplicity, isolated spikes were treated as bursts with one spike. The intraburst spike count (IBSC), that is, the number of spikes within a given burst, was assigned to the time of the first spike of that burst (bin size: 2 ms). This procedure defines a reduced spike train representation that takes the IBSC time course into account but neglects the temporal fine structure within bursts. Mean IBSCs were then calculated from eight repetitions of the recorded AN12 responses. Finally, bursts in the recorded AN12 data and in the model spike train were compared and considered coincident when the respective burst times were within one bin. The smaller of the two IBSCs contributes to n_{coinc}, the number of coincidence spikes.

Model parameters were optimized by maximizing the similarity Γ between the mean responses of model and recorded data. Let n_{AN12} and n_{model} denote the total mean recorded and modeled spike counts, respectively, and n_{coinc} the number of coincidence spikes, as defined above. Then the

coincidence measure is defined as $\Gamma = \frac{2n_{coinc}}{n_{AN12}+n_{model}}$. This quantity equals one for spike trains that are identical within this definition and vanishes if no coincident spikes occur.

A.3 Full Model. The AN12 neuron is specified as above. For the subsequent threshold operation, the AN12 output R_{AN12} is defined as the sum of all spikes within a readout window with length 1 s.

The firing rate of the AN6 neuron is assumed to be $r_{AN6}(t) = 1\ s^{-1}$ when the model is presented with a syllable and $r_{AN6}(t) = 0$ between syllables. For the subsequent threshold operation, the output of AN6 is integrated over the same readout window: $R_{AN6}(t) = \int_0^{1s} r_{AN6}(t-\tau)\,d\tau$.

The firing rate $r_{adapt}(t)$ of the rapidly adapting neuron in the circuit's third branch follows a dynamics that is that of the receptor neurons,

$$r_{adapt}(t) = s(t) - A_{adapt}\, a(t)$$
$$\tau_{adapt}\frac{da(t)}{dt} = -a(t) + s(t),$$

where $\tau_{adapt} = 3$ ms and $A_{adapt} = 0.9$. For the subsequent threshold operation, $r_{adapt}(t)$ is also integrated over a 1 s time window: $R_{adapt}(t) = \int_0^{1s} r_{adapt}(t-\tau)d\tau$. The threshold values are

$$\theta_{AN12} = 8$$
$$\theta_{AN6} = 0.72$$
$$\theta_{adapt} = 0.13.$$

Note that θ_{AN6} can be interpreted as the minimum syllable-to-period ratio needed to elicit a response. Only if all three thresholds are exceeded—$R_{AN12} > \theta_{AN12}$ and $R_{AN6} > \theta_{AN6}$ and $R_{adapt} > \theta_{adapt}$—is the binary behavioral output set to 1.

Acknowledgments

This work was supported by the German Research Foundation (through Sonderforschungsbereich 618), the Federal Ministry for Education and Research (through the Bernstein Centers for Computational Neuroscience in Berlin and Munich), the Boehringer Ingelheim Fonds, and the German National Merit Foundation.

References

Alder, T. B., & Rose, G. J. (1998). Long-term temporal integration in the anuran auditory system. *Nat. Neurosci., 1*, 519–522.

Benda, J. (2002). *Single neuron dynamics—Models linking theory and experiment*. Unpublished doctoral dissertation, Humboldt-Universität zu Berlin.

Benda, J., & Herz, A. V. M. (2003). A universal model for spike-frequency adaptation. *Neural Computation, 15*, 2523–2564.

Blitz, D. M., & Regehr, W. G. (2005). Timing and specificity of feed-forward inhibition within the LGN. *Neuron, 45*, 917–928.

Covey, E., & Casseday, J. H. (1999). Timing in the auditory system of the bat. *Annual Review of Physiology, 61*, 457–476.

Creutzig, F. (2008). *Sufficient encoding of dynamical systems*. Unpublished doctoral dissertation, Humboldt-Universität zu Berlin.

Creutzig, F., Wohlgemuth, S., Stumpner, A., Benda, J., Ronacher, B., & Herz, A. V. M. (2009). Time-scale invariant representation of acoustic communication signals by a bursting interneuron. *Journal of Neuroscience, 29*, 2575–2580.

Cruikshank, S. J., Rose, H. J., & Metherate, R. (2002). Auditory thalamocortical synaptic transmission in vitro. *J. Neurophysiol., 87*, 361–384.

deBoer, E. (1985). Auditory time constants: A paradox? In A. Michelsen (Ed.), *Time resolution in auditory systems* (pp. 141–158). Berlin: Springer.

Elhilali, M., Fritz, J. B., Klein, D. J., Simon, J. Z., & Shamma, S. A. (2004). Dynamics of precise spike timing in primary auditory cortex. *J. Neurosci., 24*, 1159–1172.

Gerhardt, C. H., & Huber, F. (2002). *Acoustic communication in insects and anurans*. Chicago: University of Chicago Press.

Gollisch, T. (2008). Time-warp invariant pattern detection with bursting neurons. *New Journal of Physics, 10*, 015012.

Gollisch, T., & Herz, A. V. M. (2004). Input-driven components of spike-frequency adaptation can be unmasked in vivo. *J. Neurosci., 24*, 7435–7444.

Gollisch, T., & Herz, A. V. M. (2005). Disentangling sub-millisecond processes within an auditory transduction chain. *PLoS Biol., 3*, e8.

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

Hauser, M. D., & Konishi, M. (1999). *The design of animal communication*. Cambridge, MA: MIT Press.

Hopfield, J. J. (1996). Transforming neural computations and representing time. *Proc. Natl. Acad. Sci., 93*, 15440–15444.

Hopfield, J. J., & Brody, C. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *Proc. Natl. Acad. Sci., 98*, 1282–1297.

Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks, 15* (5), 1063–1070.

Jin, D. (2004). Spiking neural network for recognizing spatiotemporal sequences of spikes. *Phys. Rev. E, 69*, 021905.

Juang, B. H., & Rabiner, L. R. (1991). Hidden Markov models for speech recognition. *Technometrics, 33*, 251–272.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am., 59*, 1208–1221.

Krahe, R., Budinger, E., & Ronacher, B. (2002). Coding of a sexually dimorphic song feature by auditory interneurons of grasshoppers: The role of leading inhibition. *J. Comp. Physiology A, 187*, 977–985.

Luna, R., Hernandez, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary sensory cortex. *Nat. Neurosci., 8*, 1210–1219.

Machens, C. K., Schutze, H., Franz, A., Kolesnikova, O., Stemmler, M. B., Ronacher, B., et al. (2003). Single auditory neurons rapidly discriminate conspecific communication signals. *Nat. Neurosci., 6*, 341–342.

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Curr. Op. Neurobiol., 14*, 474–480.

Oestreich, J., Dembrow, N. C., George, A. A., & Zakon, H. H. (2006). A "sample-and-hold" pulse-counting integrator as a mechanism for graded memory underlying sensorimotor adaptation. *Neuron, 49*, 577–588.

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysics, 32*, 141–152.

Pouille, F., & Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science, 293*, 1159–1163.

Rokem, A., Watzl, S., Gollisch, T., Stemmler, M., Herz, A. V. M., & Samengo, I. (2006). Spike-timing precision underlies the coding efficiency of auditory receptor neurons. *J. Neurophysiol., 95*, 2541–2552.

Römer, H. (1976). Die Informationsverarbeitung tympanaler Rezeptorelemente von *Locusta migratoria* (Acrididae, Orthoptera). *Journal of Comparative Physiology, 109*, 101–122.

Römer, H. (2001). Ecological constraints for sound communication: From grasshoppers to elephants. In F. G. Barth & A. Schmid (Eds.), *Ecology of sensing* (pp. 59–77). New York: Springer.

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

Ronacher, B., von Helversen, D., & von 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*, 363–374.

Sokoliuk, T. (1992). *Neuroanatomische Untersuchungen der Hörbahn von* Chorthippus biguttulus. Unpublished doctoral dissertation, Friedrich-Alexander-Universität Erlangen Nürnberg.

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

Stumpner, A., & Ronacher, B. (1994). Neurophysiological aspects of song pattern recognition and sound localization in grasshoppers. *Am. Zool., 34*, 696–705.

Stumpner, A., Ronacher, B., & von Helversen, O. (1991). Auditory interneurones in the metathoracic ganglion of the grasshopper *Chorthippus biguttulus*. II. Processing of temporal patterns of the song of the male. *J. Exp. Biol., 158*, 411–430.

Tan, A. Y., Zhang, L. I., Merzenich, M. M., & Schreiner, C. E. (2004). Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. *J. Neurophysiol., 92*, 630–643.

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

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

von Helversen, O., & von Helversen, D. (1994). Forces driving coevolution of song and song recognition in grasshoppers. In K. Schildberger & N. Elsner (Eds.), *Neural basis of behavioural adaptation* (pp. 253–284). Jena: Fischer.

Wehr, M., & Zador, A. (2005). Visual adaptation as optimal information transmission. *Neuron, 47*, 437–445.

Received May 11, 2009; accepted October 5, 2009.