Long-term temporal integration in the anuran auditory system

Todd B. Alder and Gary J. Rose

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA
Correspondence should be addressed to G.J.R. (gary.rose@m.cc.utah.edu)

Analysis of the temporal structure of acoustic signals is important for the communication and survival of a variety of animals including humans. Recognition and discrimination of particular temporal patterns in sounds may involve integration of auditory information presented over hundreds of milliseconds or seconds. Here we show neural evidence for long-term integration in the anuran auditory system. The responses of one class of auditory neurons in the torus semicircularis (auditory midbrain) of frogs reflect the integration of information, gathered over approximately 45–150 ms, from a series of stimulus pulses, not stimulus energy. This integration process is fundamental to the selective responses of these neurons for particular call types.

Biologically relevant information in acoustic signals often is present in its temporal structure. For example, the temporal relationships of particular acoustic elements are important for identifying human speech signals[1], the proximity of a target in bat sonar[2], songs in birds[3], and calls in anuran amphibians[4] (frogs and toads) and crickets[5]. In the sensory periphery, these signals are represented by unique spatiotemporal patterns of activity. It is still unclear to what extent these spatiotemporal patterns are read by higher-order 'detectors' of temporal features[6].

Several decades of work have provided substantial evidence for temporal feature detectors. Neurons have been recorded that respond best when tonal or frequency-modulated stimuli are of a particular duration[7–9]. 'Combination-sensitive' neurons have been found that respond best when at least two acoustic elements are presented in the appropriate temporal order and spacing[10–17]. In these cases, the response magnitude to the temporally appropriate combination of acoustic elements is much greater than the sum of their responses to each element presented alone.

Ultimately, however, recognition and discrimination of biologically important sounds might entail integration over many acoustic elements, and over hundreds of milliseconds or seconds. In this scenario, each pair of elements in the appropriate temporal relationship would activate combination-sensitive neurons, which then provide inputs to a long-term integrator. This stage might represent the top of a neuronal hierarchy, where the number of consecutive correct 'bits' of information would be evaluated. Such integration processes would complete the formation of selective filters for biologically relevant signals.

To investigate such long-term integration processes, we have conducted neurophysiological studies of the anuran auditory system. Because anurans are able to discriminate between calls that differ primarily in temporal pulse density[18,19], it is likely that they use temporal integration. We tested this hypothesis by making recordings from single units in the torus semicircularis (auditory midbrain) of two species of frogs, *R. pipiens* and *H. regilla*. *Rana pipiens* has been used in many of the previous studies of temporal processing in the auditory system; *H. regilla* was studied because these frogs are able to discriminate behaviorally between calls that differ almost exclusively in pulse repetition rate. The torus is a logical place to look for neurons that integrate information over multiple pulses because neurons in this area respond selectively to particular rates of amplitude modulation (AM)[20,21], and the distribution of 'best rates of AM' is species specific[22,23]. Here we show that long-term integration occurs in the auditory system of anurans. As a result, neurons show strong selectivity for temporal patterns of acoustic signals that represent particular call types.

Results

Recordings were made from 109 neurons in the torus semicircularis of 25 male frogs. Sixty seven of these cells were tuned for AM rate (that is, responded most strongly to a particular AM rate) and these cells were studied further. Forty-four of these band-pass cells had response latencies less than 40 ms and, in most cases, were excited best by AM rates less than approximately 60 Hz. The remaining 23 band-pass cells were of particular interest with respect to processes of integration over long time scales. These neurons responded best to AM rates above approximately 60 Hz and had response latencies that ranged from 45–150 ms. The integration processes that operate in these latter cells are the subject of this study. These neurons were clustered in the medial torus; multiunit recordings revealed little response in this area to AM rates below approximately 50 Hz. These cells, in *H. regilla*, would respond to advertisement calls, but not to aggressive calls, whereas the opposite relation would hold for *R. pipiens*. In a representative single cell, this selectivity was largely independent of whether the AM was sinusoidal or of more natural shape (Fig. 1a). This selectivity cannot stem simply from integration of stimulus intensity because these cells failed to respond to tone bursts or to slow rates of sinusoidal AM, even when stimulus energy was constant across AM rates.

Alternatively, the tuning to high rates of AM could result either from underlying sensitivities to the duration and rise-times of pulses, or from a process by which neural activities that stem from individual pulses are combined when in an appropriate temporal pattern. The first hypothesis can be ruled out, because stimulus pulses that were 10 ms in duration and had natural shapes (rise/fall characteristics) effectively excited these neurons

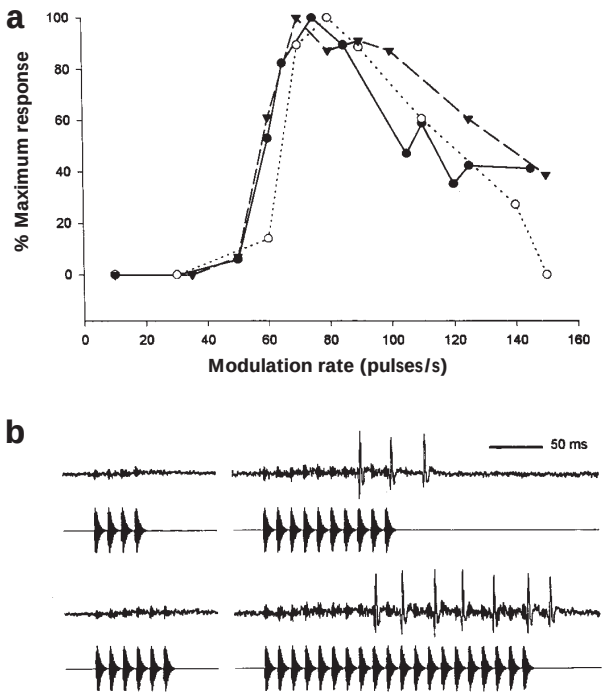

Fig. 1. Pulse-integrating properties of a neuron in the torus or auditory midbrain of the frog *H. regilla*. **(a)** Normalized response level versus the rate of modulation of the amplitude of a 1.2 kHz carrier signal. Responses were normalized with respect to the maximum spike rate for each form of amplitude modulation. Signal amplitude was modulated sinusoidally (closed circles) or so that the relative rise and fall characteristics of pulses resembled those found in natural calls. In the latter case, pulse duration was either held constant at 10 ms (triangles) or varied with pulse repetition rate to maintain a duty cycle of 1.0 (open circles). Stimulus duration was 400 ms, except when pulse duration and number (15 pulses) were held constant, in which case stimulus duration varied with pulse repetition rate. **(b)** Recordings from this neuron in response to stimuli that consisted of 4, 6, 10 and 20 pulses per presentation. Pulses were 12.5 ms in duration. A faint microphonic potential, reflecting the stimulus, can be seen in these recordings.

when presented at high repetition rates, but were ineffective when presented at low repetition rates (Fig. 1a). To test the hypothesis that the selectivity of these cells for high pulse repetition rates might stem from an integrating process, the number of pulses per stimulus presentation was varied (Fig. 1b). In this case, at least eight pulses, presented at approximately 80 pulses/s, were required to elicit spiking. The long response latency of this neuron, therefore, resulted primarily from an integration process with a time constant longer than the time required to conduct signals to this area of the brain. The minimum number of pulses, delivered at the optimal repetition rate, that was sufficient for eliciting at least one spike per stimulus presentation (pulse train) ranged from 4–15 (median, 8.5, n = 18) for this group of neurons. The responses of these cells, therefore, are evidence of an integrating process, but what is being integrated? The system may integrate either stimulus intensity that is distributed in a specific temporal pattern or information relating to the number and temporal density of pulses. In the latter case, the cell's response level would be largely independent of the amplitude of each pulse.

To distinguish between these two possibilities, pulse amplitude was varied while holding the repetition rate of pulses constant. If the responses of neurons reflect the integration of stimulus intensity that is presented in a temporally appropriate pattern, then, as the stimulus intensity is increased, fewer pulses should be required to elicit spiking. This was not the case. For seven of the eight neurons that were tested across a 6–12 dB range in stimulus amplitude, a constant number of stimulus pulses, delivered at the optimal repetition rate, was required to elicit a response. In a representative neuron (Fig. 2), more than 10 pulses, delivered at the optimal rate of 100 pulses/s, were required to elicit spikes, regardless of whether the stimulus amplitude was 4, 10 or 16 dB above threshold. For example, when stimulus amplitude was set at a level 16 dB above the threshold required for a series of 16 pulses to elicit spikes, a stimulus consisting of 8 pulses failed to evoke spikes; this difference represents approximately a 40-fold increase in stimulus power and a 20-fold increase in stimulus energy. The latency of response to the sequence of 16 pulses was similar across a wide range of stimulus amplitudes (Fig. 2, inset). These findings indicate that the activity of this neuron was not simply a function of the amount of stimulus energy that was presented in a temporally appropriate pattern. Instead, the cell responded only when a threshold number of stimulus pulses occurred within a particular time window. This conclusion is also supported because this neuron's selectivity for the repetition rate of pulses was similar across these stimulus amplitudes.

Because of compressive nonlinearities (that is, dynamic range limits) in the auditory system, however, increases in stimulus peak amplitude may not translate into proportionately greater activity levels in afferents to these neurons. We therefore varied the duty cycle (pulse duration divided by interpulse interval) of pulses as a means of varying the energy in stimuli while holding the peak amplitude and repetition rate of pulses constant; halving the duty cycle of pulses (Fig. 3b, inset) results in halving stimulus energy. For this class of neurons, the number of pulses that were required to elicit a threshold level of response was largely independent of the duty cycle of the pulses (Fig. 3a). If neural thresholds (number of pulses) for different pulse duty cycles were related to the amount of stimulus energy, values would fall along the dotted and dashed lines. Instead, data points fall along the solid line, indicating that the critical feature for exciting these neurons was the number of pulses that occurred within a particular time window. In addition, the selectivity of neurons for pulse repetition rate was almost completely independent of the pulse duty cycle (Fig. 3b). These data, therefore, support the hypothesis that these auditory neurons integrate information concerning the number and temporal density of pulses, not simply stimulus intensity that is presented in a temporally appropriate pattern.

Discussion

Our results indicate that one class of neurons in the frog torus or auditory midbrain derives its selectivity for fast PRRs by integrating information gathered over a time period of approximately 45–150 ms; these cells only begin to respond after the appropriate temporal density of pulses has been maintained for a sufficient length of time. As a result, individual cells that we term 'pulse-integrator' (PI) neurons code the number of consecutive correct interpulse intervals in the stimulus. The biological significance of integrating information across a series of pulses is clear for anurans. Behavioral studies have shown that these animals are able to discriminate between calls that differ almost

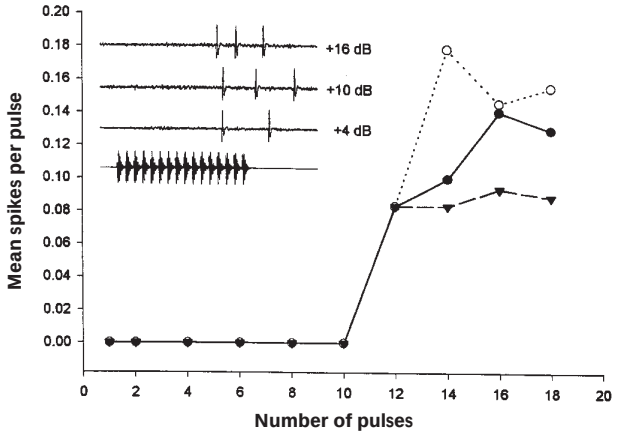

Fig. 2. Energy independence of the pulse-integration process: pulse amplitude variations. The normalized response levels of a single representative neuron in *R. pipiens* for stimulus amplitudes 4 (triangles), 10 (open circles) and 16 (closed circles) dB above threshold, as a function of the number of pulses in the stimulus. Pulse repetition rate was 100 pulses/s. The carrier frequency was 600 Hz. Recording traces (inset) show that response latency was largely independent of stimulus amplitude.

exclusively in pulse repetition rate[18,19] or number of pulses[24]. In these cases, pulse shape and spectral structure are highly similar between calls. Pulse-integrating processes of the kind shown here form the basis of selective neural responses to the calls of high AM rate.

Long-term integration and detection thresholds

In addition to its role in temporal pattern recognition, long-term integration across pulses might influence the thresholds for detection of calls. Consider how behavioral threshold might be related to the activity of a population of neurons, some of which are PI neurons. Assume that PI and nonPI neurons have similar thresholds; for neurons of either type, thresholds were between 25 and 75 dB SPL. Presentation of single pulses will only excite neurons that do not integrate across pulses (nonPI neurons). As the number of pulses per stimulus is increased, PI neurons will progressively be recruited. The aggregate activity therefore will increase in a nonlinear fashion with pulse number.

As the stimulus intensity is progressively decreased, the level of activity required for behavioral detection will only be met for stimuli consisting of enough pulses to recruit the PI neurons. For stimuli of lower pulse number, only nonPI neurons will fire; the pulse-integrating neurons will be silent. Thus when only one or a few pulses are delivered, the stimulus intensity will have to be raised to elicit a threshold level of activity in the pool of nonPI neurons. Behavioral thresholds, therefore, should be lower for calls that consist of many pulses (maintaining correct interpulse intervals) than for calls that are repeated more frequently but have fewer consecutive pulses.

The tradeoff between stimulus pulse number and stimulus intensity for detection is fundamental to 'multiple-looks' models[25] of temporal integration in the auditory system. Multiple-looks models of temporal integration in humans primarily account for the finding that subjects detect stimuli at lower amplitudes as the duration of the stimulus is increased, even out to hundreds of milliseconds. Two types of models have been described. A statistical formulation[25] postulates a tradeoff between the number of sampling opportunities and stimulus intensity. Each sampling, corresponding to the minimum integration time of the auditory system, represents an opportunity that the threshold for detection will be met. Because the number of sampling opportunities increases with stimulus duration, the probability of detection increases[26]. To achieve behavioral threshold when presenting only one or a few pulses per stimulus presentation, the intensity of the stimulus would have to be increased; that is, a tradeoff occurs between intensity and the duration of a pulse train.

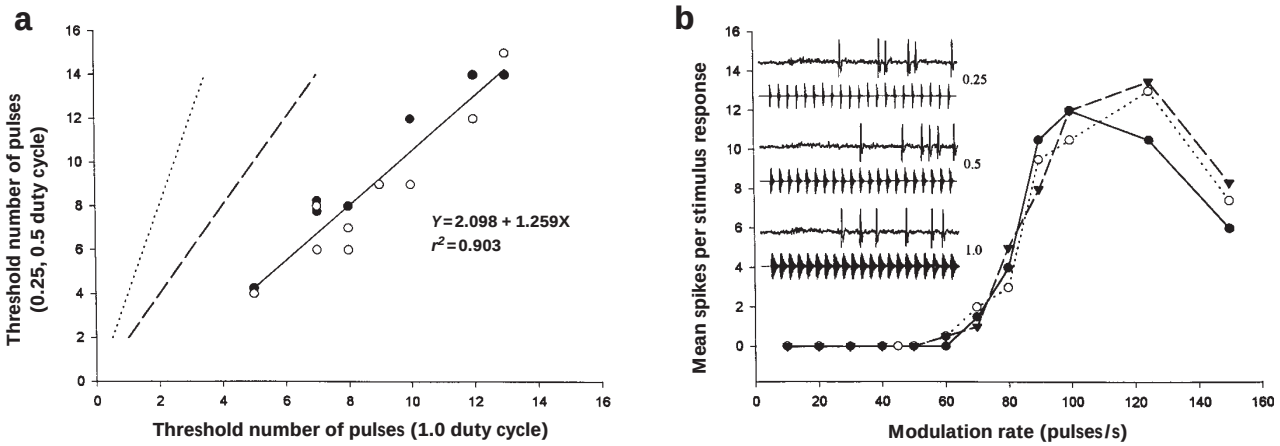

Fig. 3. Energy independence of the pulse-integration process: variations of pulse duty cycle. **(a)** Number of stimulus pulses required to elicit a threshold level of neuronal response, for pulse duty cycles of 1.0 versus 0.5 (filled symbols) and 1.0 versus 0.25 (open symbols). The rate of repetition of stimulus pulses was held at the optimal value for each neuron; this value was independent of the duty cycle of pulses, as shown in (b). Dotted lines signify equal energy conditions for stimuli of 1.0 versus 0.25 duty cycle, and dashed lines show equal energy for 1.0 versus 0.5 duty cycle. **(b)** Response level of a single neuron versus stimulus AM rate for pulse duty cycles of 1.0 (closed circles), 0.5 (open circles) and 0.25 (triangles). Inset, examples of stimuli and responses at these duty cycles.

A second, nontrivial, multiple-looks model posits that information from individual looks might be held in memory and combined later; stimulus intensity *per se* is not integrated over hundreds of milliseconds[25]. Pulse-integrating (PI) neurons provide evidence supporting this memory- and information-based multiple-looks model. These neurons respond in a highly nonlinear fashion to stimuli of different pulse number. For example, the number of spikes that are evoked by a stimulus consisting of sixteen pulses greatly exceeds that evoked by four repetitions of a stimulus consisting of four pulses. The number of 'looks' are identical for these two stimulus conditions, and a similar number of spikes would be evoked from neurons that do not integrate over many pulses.

The multiple-looks hypothesis functionally relates information integration to the tradeoff between pulse number and stimulus intensity for behavioral threshold. As shown in this paper, however, pulse-integrating processes also seem to be important in the ability of organisms to discriminate between different patterns of amplitude modulation. In humans, for example, the ability to detect changes in the AM rate of stimuli does not peak until at least five pulses (or cycles of modulation) are present[27]. For threshold of detection, a 'look' may constitute the minimum integration time of the auditory system. For pattern recognition, however, the temporal complexity of 'looks' may depend on the level of temporal pattern analysis that is being considered; each 'look' would generate an input to integrators at the next level of processing. In birdsong, for example, a look could represent a frequency sweep at one level, a syllable at an intermediate level, or a series of syllables called a 'phrase'.

Relation to temporal integration in other systems

The long-term integration process that is demonstrated in this paper analyzes the number of consecutive 'correct' interpulse intervals. By firing only after at least eight (on average) consecutive correct intervals have occurred, the responses of these neurons signify with a high statistical reliability that the species-specific call type has indeed occurred. Long-term integration processes might also be important in other systems. Neurons that show facilitated responses to particular temporal configurations of information-bearing elements[17] have been recorded in the auditory systems of songbirds[10–15], bats[16,17] and primates[28–30]. In songbirds, neurons that respond best to the bird's own song generally are selective for particular pairs of syllables/notes that are presented with a particular temporal order and spacing. Neurons have been recorded in bats that respond selectively to frequency-modulated signals that are separated by a particular interval of time and code target range information. Such 'combination-sensitive' neurons represent detectors for these temporal features of sounds. In principle, any of these temporal feature detectors could provide inputs to a long-term integrator. Take, for example, neurons that respond selectively to particular temporal pairings of syllables in birdsong and rapidly habituate to repeated presentation of the stimulus. The correct temporal sequence of syllables in a song phrase could activate sequentially an array of combination-sensitive neurons, which would then provide the requisite number of inputs within a particular time window to elicit spiking from a subsequent long-term auditory integrator. To the best of our knowledge, pulse-integrating neurons have only been found in anurans. Future work is needed, therefore, to determine whether similar processes of integrating information and their underlying mechanisms are present in other animals, including mammals.

Methods

Recording procedure. Extracellular recordings were made from single neurons in the torus semicircularis of 19 *R. pipiens* and 6 *H. regilla* using glass micropipettes. The frog's body temperature was 17–18º C. Animals were prepared for recording *in vivo* according to methods previously published[20]; criteria for establishing threshold also can be found in this article. Many neurons that were excited by AM stimuli failed to respond to pure tones, regardless of duration. The frequency tuning characteristics of these neurons, therefore, could not be determined by conventional methods. Instead, the carrier frequency in the modulation was varied, holding the rate of AM at the optimal value, until the maximal response was obtained.

Stimulus generation and delivery. Acoustic stimuli were generated using Tucker Davis Technologies (TDT) System II hardware and custom software on a Pentium 90 computer. Amplitude-modulated stimuli were generated by multiplying a white noise or pure tone carrier with a modulating waveform, which contained a DC offset equal to half its peak to peak amplitude. Tones and noise were created using a TDT AP2 card. The sampling rate for these carriers and all modulating waveforms was 25 kHz.

Stimuli that consisted of pulses of natural shape were generated by multiplying the carrier signal by a modulating envelope that was a mathematical representation of the natural pulse envelope. Based upon analysis of field-recorded calls, a single pulse envelope was generated using the following equations:

$$V = k\,[e^{-t/\tau 1} - e^{-t/\tau 2}]$$
$$\tau 2 = (\tau 1)/2$$

where $\tau 1$ and $\tau 2$ defined the relative rising and falling phases of the envelope and k was a constant. The pulse envelope was then repeated to produce the modulating waveform. Stimulus duration was 400 ms. Total stimulus energy was held constant (within 0.1 dB) with changes in PRR. Neurons were typically tested using modulation duty cycles of 1.0, 0.5, 0.25 and 0.1.

In another stimulus set, pulse duration and shape were held constant at natural dimensions while the pulse repetition rate was varied. As rates increased, pulses moved closer together, thereby increasing the pulse duty cycle and decreasing total stimulus duration. Because pulse shape, duration and number did not vary with pulse repetition rate, total energy remained constant.

While searching for auditory units, we delivered a sinusoidally amplitude modulated acoustic stimulus. Stimuli were amplified and presented free field in an audiometric room. The speaker was situated 0.5 meters from the frog, contralateral to the recording site. For frequencies greater than 500 Hz, reflections in the booth were attenuated by at least 30 dB relative to the stimulus. Correspondingly, stimuli were presented at levels not exceeding 30 dB above each unit's threshold. A microphone situated above the frog was used to measure stimulus levels via a sound level meter. Stimuli were presented once every 2.5 seconds, a rate that is sufficient for producing a consistent response across stimulus repetitions. Sound level was varied using a programmable attenuator. The modulation rate and carrier frequency were varied to cover the range represented in the vocalizations of these species. Upon encountering a single unit that was excited by any of these search stimuli, threshold was determined[20], and tests of the neuron's temporal selectivity were performed at approximately 10 decibels above this threshold value.

Analysis. Neurophysiological data were analyzed off-line using SPIKE-2 software and the 1401 data acquisition interface from Cambridge Electronic Design, Cambridge UK. Pulse integration thresholds were determined from smoothed response histograms, where bin number equaled the pulse number; threshold was defined as the first bin to have an average of 1.0 spikes per stimulus repetition. Histograms were smoothed according to the function, $Ri = [R(i-1) + Ri + R(i+1)]/3$.

Acknowledgements

This work was supported by a grant from the National Organization for Hearing Research to G.J.R. We thank J. Canfield, E. Fortune and D. Carrier for critically reading this manuscript.

Received 24 April: accepted 13 August 1998

1. Olive, J. P., Greenwood, A. & Coleman, J. *Acoustics of American English Speech. A Dynamic Approach* (Springer, New York, 1993).
2. Moss, C. F. & Schnitzler, H. U. in *Hearing by Bats* (eds Popper, A. N. & Fay, R. R.) 87–145 (Springer, New York, 1995).
3. Konishi, M. Birdsong: from behavior to neuron. A*nnu. Rev. Neurosci.* 8, 125–170 (1985).
4. Gerhardt, H. C. in *The Evolution of the Amphibian Auditory System* (eds Fritzsch, B. Ryan, M. J., Wilczynski, W., Hetherington, T. E. & Walkowiak, W.) 455–484 (Wiley, New York, 1988).
5. Huber, F. & Thorson, J. Cricket auditory communication. *Sci. Amer.* 253, 60–68 (1985).
6. Bullock, T. H. Some principles in the brain analysis of important signals: mapping and stimulus recognition. *Brain Behav. Evol.* 28, 145–156 (1986).
7. Piheiro, A. D., Wu, M. & Jen, P. H. S. S. Encoding repetition rate and duration in the inferior colliculus of the big brown bat *Eptesicus fuscus*. *J. Comp. Physiol*. 169, 69–85 (1991).
8. Gooler, D. M. & Feng, A. S. Temporal coding in the frog midbrain: The influence of duration and rise-fall time on the processing of complex amplitude-modulated stimuli. *J. Neurophysiol*. 67, 1–22 (1992).
9. Casseday, J. H., Ehrlich, D. & Covey, E. Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. *Science* 264, 847–850 (1994).
10. Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci*. 3, 1039–1057 (1983).
11. Doupe, A. J. & Konishi, M. Song-selectivity in auditory circuits in the vocal control system of the zebra finch. *Proc. Natl. Acad. Sci. USA* 88, 11339–11343 (1991).
12. Margoliash, D. & Fortune, E. S. Temporal and harmonic combination-sensitive neurons in the zebra finch. *J. Neurosci*. 12, 4309–4326 (1992).
13. Lewicki, M. S. & Konishi, M. Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. *Proc. Natl. Acad. Sci. USA* 92, 5582–5586 (1995).
14. Lewicki, M. S. Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. *J. Neurosci.*16, 5854–5863 (1996).
15. Doupe, A. J. Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J. Neurosci.* 17, 1147–1167 (1997).
16. Suga, N., O'Neill, W. E. & Manabe, T. Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. *Science* 200, 778–781 (1978).
17. Suga, N. Principles of auditory information-processing derived from neuroethology. *J. Exp. Biol.* 146, 277–286 (1989).
18. Brenowitz, E. A. & Rose, G. J. Behavioural plasticity mediates aggression in choruses of the Pacific treefrog. *Anim. Behav*. 47, 633–641 (1994).
19. Rose, G. J. & Brenowitz, E. A. Plasticity of aggressive thresholds in *Hyla regilla*: discrete accommodation to encounter calls. *Anim. Behav*. 53, 353–361 (1997).
20. Rose, G. & Capranica, R. R. Temporal selectivity in the central auditory system of the leopard frog. *Science* 219,1087–1089 (1983).
21. Walkowiak, W. in *Evolution of the Amphibian Auditory System* (eds Fritszch, B. Wilczynski, W., Ryan, M. J. & Walkowiak, W.) 275–294 (Wiley, New York, 1988).
22. Rose G. J. & Capranica, R. R. Processing amplitude-modulated sounds by the auditory midbrain of two species of toads: matched temporal filters. *J. Comp. Physiol*. 154, 211–219 (1984).
23. Rose, G. J., Brenowitz, E. A. & Capranica, R. R. Species specificity and temperature dependency of temporal processing by the auditory midbrain of two species of treefrogs. *J. Comp. Physiol*. 157, 763–769 (1985).
24. Klump, G. M. & Gerhardt, H. C. Use of non-arbitrary acoustic criteria in mate choice by female gray tree frogs. *Nature* 326, 286–288 (1987).
25. Viemeister, N. F. & Wakefield G. H. Temporal integration and multiple looks. *J. Acoust. Soc. Am*. 90, 858–865 (1991).
26. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
27. Lee, J. Amplitude modulation rate discrimination with sinusoidal carriers. *J. Acoust. Soc. Am*. 96, 2140–2147 (1994).
28. Wollberg, Z. & Newman, J. D. Auditory cortex of squirrel monkey: response patterns of single cells to species-specific vocalizations. *Science* 175, 212–214 (1972).
29. Wang, X., Merzenich, M. M., Beitel, R. & Schreiner, C. E. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol*. 74, 2685–2706 (1995).
30. Rauschecker, J. P. Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol. Suppl. (Stockh.)* 532, 34–38 (1997).