Storage of auditory temporal patterns in the songbird telencephalon

Patrick D. Roberts, [a] Roberto Santiago, [b] Tarciso Velho, [a] Claudio V. Mello [a]

[a] *Neurological Sciences Institute, Oregon Health & Science University 505 N.W. 185th Avenue, Beaverton, OR 97006*

[b] *Portland State University, Portland, OR 97201*

Abstract

A quantitative model of auditory learning is presented to predict how auditory patterns are stored in the songbird auditory forebrain. This research focuses on the caudomedial nidopallium (NCM) in the songbird telencephalon, a candidate site for song perception and the formation of song auditory memories. The objective is to introduce simplified features of bird song that could be used by the auditory forebrain to identify and distinguish memorized songs. The results elucidate which biological mechanisms are sufficient for temporal pattern prediction and the storage of higher-order patterns, where by higher-order, we mean the specific arrangement of syllables into song motifs (phrases) to reveal neural mechanisms of syntax.

Key words: auditory, learning, avian forebrain, synaptic plasticity, habituation

Auditory processing and auditory memories are essential to animals that use vocal communication in social behavoir (see reviews in [1]). In songbirds, vocal learning and retention of vocalizations depends on the ability to hear a song model and to perfect vocal performance through auditory feedback [2–4]. Songbirds also rely on auditory processing to identify and to discriminate among individuals of their species [5,6]. Although song auditory processing and memory play central roles, their underlying neuronal mechanisms remain largely unknown.

Song production and learning depends on a well-characterized set of interconnected forebrain nuclei, but less is known about areas involved in *song perception and discrimination*. Furthermore, little is known about the formation and storage of song auditory memories. It is known that auditory information ascends along a brainstem pathway that is conserved in vertebrates, reaching the telencephalon through field-L. Field-L projections represent pathways for auditory information to reach higher-order areas [7–10], but the role of these targets in song auditory processing and leaning, and how song exposure affects song-evoked behaviors, is unclear.

1. NCM: a candidate site for song perception and the formation of song auditory memories.

The caudomedial nidopallium (NCM), a telencephalic field-L target, is analogous to supragranular layers of the mammalian auditory cortex [14,10]. One source of evidence is from activity-dependent genes, such as *zenk*, that are expressed in the brain upon activation of neurons. Analysis of gene expression has been very useful in revealing patterns of brain activation in response to specific stimuli or behavioral contexts. NCM has the most marked *zenk* induction response to song [15]. *Zenk* induction in NCM is rapid and transient [16,17], being highest for conspecific song, as compared to heterospecific song or tones [15]. *Zenk* induction decreases markedly upon repeated song presentations (song-specific "habituation"), it is re-elicited upon presentation of a novel song [17], and is abolished by deafening [18].

Additional evidence of NCMs role in song learning is from evoked electrophysiological responses to song

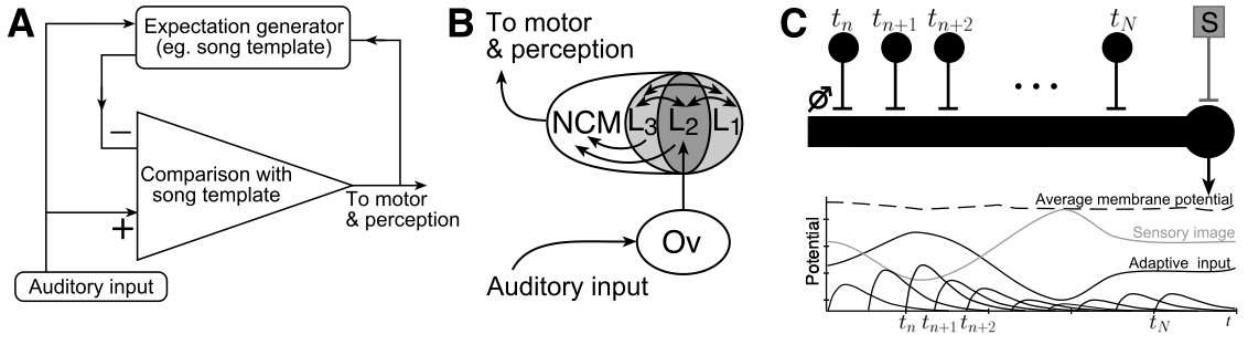

Fig. 1. (A) Temporal habituation occurs when a sensory (auditory) input is compared with a previously stored template that subtracts an expectation of the sensory image. The output is a constant unless the sensory input does not match the output of the expectation generation. When a mismatch is present, the sensory template is updated. (B) Schematic summary diagram representing connections of auditory structures (L1, L2 and L3 are subdivisions of field-L; NCM, caudomedial nidopallium; Ov, nucleus ovoidalis (thalamus)). (C) Model of temporal habituation [11,12] consists of 2 inputs: a series of inputs delayed $(t_n, t_{n+1} \ldots t_N)$ with respect to the beginning of a stimulus, and a non-adaptive input (S) that represents the stimulus. The expectation is generated by the series of delayed inputs and their weights are updated by an anti-Hebbian learning rule [13,11,12]. The bottom panel shows a simulation of the inputs to the membrane potential after habituation. The adaptive inputs sum (black, solid) to generate a negative image of the sensory input (grey, solid), and the average membrane potential (black, dashed) is approximate constant (arbitrary units).

recorded in NCM. These responses are less brisk and of longer latencies than in field-L, and show some selectivity for conspecific song [19–21]. NCM emerges as a major auditory processing station, and some insight to its organization and function is provided by studying these evoked electrophysiological responses.

The evoked electrophysiological responses of NCM neurons to song auditory stimuli decrease quickly upon repeated song presentations [19,20]. This habituation phenomenon is song-specific, since a high response level can be restored by a novel song stimulus, and habituation to specific songs is long-lasting. Multiunit responses have been shown to decrease following repeated presentations of the same conspecific song [19] when responses were recorded for several different songs presented sequentially at a single site. Each song elicited a different initial response, but all habituated during training. The multiunit responses were reduced by 60-80% during habituation. Immediately after training, each song was tested again and the habituated response level achieved for a given song was retained even after training with other songs [19,20].

Local injections of RNA and protein synthesis inhibitors show that the long-term maintenance of the NCM habituation to song is dependent on song-induced gene expression [19]. The *zenk* gene is a likely candidate for regulating long-lasting neuronal changes in NCM. *Zenk* exhibits a robust song regulation and has been implicated in long-term neuronal plasticity in other systems. To predict the expression of *zenk* in NCM, we developed an adaptive model of evoked responses, and used the model to demonstrate how a population of NCM neurons distinguish between two songs.

2. Model of auditory adaptation in NCM.

To investigate the principles of temporal habituation [13,11,12] (Fig. 1) in the songbird auditory system, we developed a model based on synaptic plasticity that could explain the habituation seen in the auditory forebrain of songbirds. The conceptual framework of temporal habituation is shown in Fig. 1A where a temporal pattern, such as an auditory input, is repeatedly presented to the bird. The temporal pattern triggers an expectation generator to predict the syllables of the song. This expectation is subtracted from the direct sensory input pattern and the output is an error between the expectation and the actual sensory temporal pattern. The error signal then projects to either motor regions to correct song production, or to perceptual regions to determine the familiarity of the song.

The location of NCM in the auditory pathway is shown in Fig. 1B. Auditory inputs enter the forebrain via a thalamic structure (nucleus ovoidalis, Ov), and acoustic features of stimuli are coded in spiking patterns of neurons in the subdivisions of field-L (L_1, L_2, L_3) [22]. This spike representation projects to NCM where the temporal patterns are

presumed to be combined in neurons of NCM that habituate to repeated auditory patterns.

A mechanistic model of the temporal habituation (Fig. 1C, [11,12])presupposes a time-ordered series of adapting synaptic inputs (such as from field-L) to a habituating neuron (in NCM). The spike-times of the adapting inputs are consistently correlated with a non-adapting input that generates an evoked response in the habituating neuron. The mechanism of habituation is an anti-Hebbian form of synaptic plasticity that depresses the synaptic efficacy when input spikes are correlated with output spikes, and otherwise potentiates the active synapses.

Modeling Methods. In our model of the NCM, we represented each model neuron as a single compartment, spike-response model [23] (see [11,12] for details). Response functions, or *kernels*, correspond to the postsynaptic potential (PSP), and represent the effects on the membrane potential caused by spikes. We separated fast, electrophysiological processes from slow, adaptive processes by representing time in 2 components, (t, n), where t denotes the time following the beginning of each stimulation cycle, and n represents the number of cycles. If T is the period of the stimulus cycle, then $(t + T, n) = (t, n + 1)$. In these variables, t parametrized the model's prediction of the evoked response, and n parametrized changes of the model neuron's response during habituation.

We represented the membrane potential, $V(t, n)$, as a random variable with a normal distribution function: a mean value of $V(t, n)$, and a variance of σ [24]. A probability function, $P_\theta(t, n) = (1 + \exp(-\sigma(V(t, n) - \theta)))^{-1}$, represented the probability that a spike occured at time (t, n). The contribution to the NCM-neuron model's membrane potential from synaptic inputs was computed using a weighted sum of excitatory response kernels, $\epsilon(t)$ where $t = 0$ represents the time that the presynaptic spike arrives at the synapse so that $\epsilon(t) = 0$ if $t \leq 0$. The contribution from an input neuron j was represented by presynaptic spike probability, convolved with the response kernel and weighted by the synaptic strength w_{ij}, with $w_{ij} > 0$. The average membrane potential of the i-th NCM neuron was $\langle V_i^{NCM}(t, n)\rangle = v_i(t) + \sum_j w_{ij} \sum_{t'} P_j(t', n)\epsilon(t - t')$, where $v_i(t)$ is a non-adapting input to NCM neuron$i$, and the last term is the contribution from the cells that make synaptic contact with i-th NCM neuron. The output of each NCM neuron was represented by the spike probability function, $P_i^{NCM}(t, n)$, that quantified the probability of a spike, where t denotes the time during the song presentation that begins at time n. The spike probability quantified the evoked response, and how the response changes during habituation. When a temporal pattern is first presented, the variance of the membrane potential is large, representing a high firing rate at the peaks of the input excitation. This variance of the response was used to predict *zenk* expression in each NCM neuron, because *zenk* expression has been shown to depend on the strength of the response. The result is a prediction of the percentage of NCM neurons that express *zenk* in response to repeated presentations of conspecific songs.

The biologically plausible assumptions for the NCM response model were (see Fig. 2A):

Assumption 1: A population of feature selective neurons. These neurons encode acoustic features based on linear response properties of field-L [22]. Each neuron responded with an increased spike probability to acoustic features extracted from the syllables of conspecific song. The responses were used to drive a pause-duration generator that encoded the pauses between the occurrence of each acoustic feature. The biophysical mechanism could be similar to delay-tuning found in the inferior colliculus of bats [25]; post-inhibitory rebound follows an inhibitory input from the first sound, and the second sound injects a subthreshold excitatory postsynaptic potential. Membrane properties of the cell determine the timing when the rebound and excitation add to generate a spike. Acoustic features separated by pauses of over 100 msec can be explained with this mechanism [26]. In our simulation, there are potentially 100 units for each feature, each selects a different pause (in msec) between each occurrence of the feature.

Assumption 2: Delay-lines from each feature selective neuron to NCM. The temporal structure of the song must provide all of the information for habituation. Thus, each syllable response must generate a series of inputs to the habituating cell so that a syllable that consistently occurs before a second syllable can generate a series of adaptive inputs to an NCM neuron that cancel the response of the NCM neuron to the second syllable. Two possible mechanisms for these delays would be either a strong rebound from an inhibitory input or a transient potassium current [27] following a subthreshold excitatory input. Varying membrane properties in a set of neurons could yield a range of delays following inputs from feature-duration selective cells.

Assumption 3: Spike-timing dependent plastic-

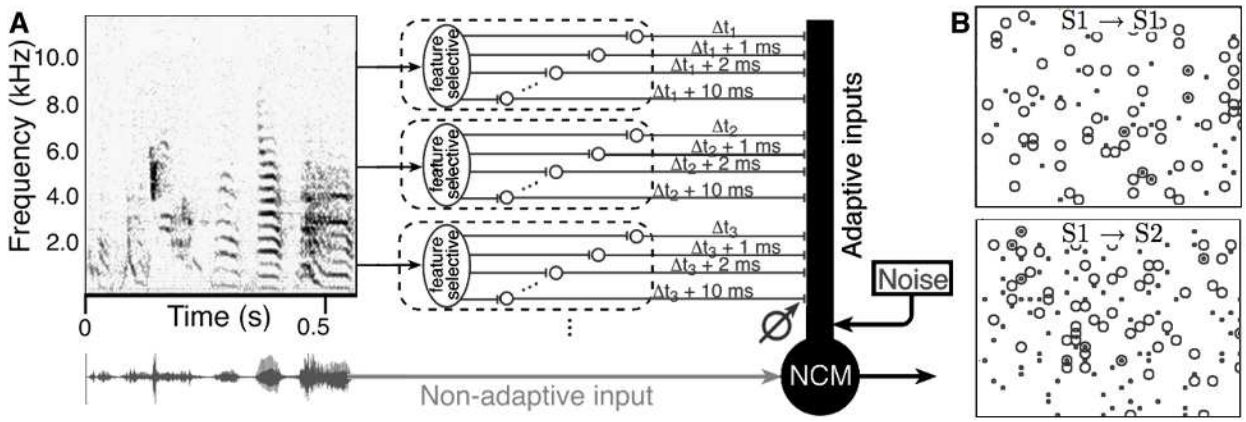

Fig. 2. *Model of an NCM neuron that habituates to temporal patterns.* (A) The acoustic features of syllables were detected by feature selective neurons that initiated spikes at different delays in a set of neurons ($\triangle t_1, \triangle t_2, \ldots$) that project to the NCM-neuron. The synaptic weights were allowed to adapt by an activity dependent learning rule that causes the evoked response to relax to a fixed point. The song syllables also had a direct, non-adaptive synaptic input to the NCM-neuron as a series of excitatory postsynaptic potentials. (B) The expression of *zenk* predicted by the model for 2 presentations of conspecific song. Top panel (S1 → S1): S1 is played for 50 song presentation cycles, then repeated another 50th cycles after a delay. Bottom panel (S1 → S2): S1 is played for 50 song presentation cycles, then S2 is played for 50th cycles. The responses of a grid of model cells are marked with a circle (∘) if they responded to the first presentation, and marked with a dot (·) if they responded to the second presentation. Model cells that responded to both presentations are double marked (⊙) to represent *zenk* expression in both the nucleus and the cytoplasm. The response to the second song was much greater for the novel song (bottom panel) than for the repeated song (top panel) because the number of cells expressing *zenk* in their nucleus was greater.

ity (STDP) with a learning rule that sculpts the adaptive input to cancel the sensory image from a non-adaptive input. The average change in synaptic weight per cycle was given by the non-associative weight change minus the average associative change, $\langle \triangle w_e(t,n) \rangle = \alpha - \sum_{t'} \beta L(t'-t) P_{NCM}(t,n)$. Although it is unknown whether such a synaptic learning rule exists at the input synapses onto NCM neurons, we are most interested in the learning dynamics of this particular learning rule; the *decorrelation* of inputs such that the output of the NCM neuron's evoked response approximates a constant during the song presentation. We have previously shown that this STDP leaning rule has a fixed point where the postsynaptic spike probability is $P_i^{NCM}(t, n \to \infty) = \alpha/\beta$ [11] that depends o the specifics on the learning rule for the adapting synapses. Due to the stability of the fixed point, the model correctly predicted the effects of habituation as would be caused by synaptic plasticity at the synapse from cells in field-L onto NCM-neurons.

Assumption 4: Non-adaptive response to auditory stimulus. To initiate a strong response to the song at the beginning of the learning protocol, we modeled the NCM neuron such that each neuron received a non-adaptive excitatory postsynaptic potential.

The results of our simulations provide an example to demonstrate the effects of habituation on NCM-neuron responses to song presentations. We quantified the probability of NCM-neurons responding to stimulation to estimate the percentage of cells that express *zenk* during each presentation. The result can be compared with the percentage of *zenk*-positive cells in the following stimulation paradigm. Birds are stimulated with 2 songs, for 5 min. each, separated by 25 min. to allow the zenk-expression evoked by the first song (S1) to be transported to the cytoplasm, while the zenk-expression evoked by the second song (S2) remains in the nucleus. Thus, this paradigm permits the identification of NCM neurons that respond to 2 song auditory stimuli, or to each of the stimuli alone, in the same bird.

3. Result: Habituation to multiple patterns of stimuli.

To investigate habituation to song structure, our numerical simulation used 2 finch songs to drive the model. Features were extracted from the song by convolving the pressure wave with a filter bank of gamma-tone filters ($te^{-t/\tau}cos(\omega t)$) to simulate primary auditory processing. Frequency bands were then combined with relative delays to generate a field-L response. The responses represented feature-selective cells in the model with an increased spike

rate, followed by a series of postsynaptic potentials in the NCM-neuron. An activity dependent synaptic learning rule altered the synaptic efficacy from the delay fibers onto the NCM model neurons (Fig. 2A). The resulting output of the NCM-neuron relaxed to its fixed-point [11]. The activity dependent plasticity reduced the variance of the membrane potential, and led to a prediction of a reduced expression of *zenk*.

The results of a simulation are shown in Fig. 2B. We tested the difference between the nuclear expression of *zenk* for two presentations of the same song (S1 → S1) compared with the presentation of two different songs (S1 → S2). After the first song is presented 50 times, a second song is presented 50 times following a delay. When the first song was was used in the second presentation, the expression of *zenk* in the nucleus was greatly reduced relative to the cytoplasmic expression, in agreement with experimental studies [17]. The interference between different songs, resulting in double expression, is partly due to similar features that are shared by both songs so that the NCM neuron re-adapts during each presentation.

We have shown that the features that are encoded in the neural activity of cells that project to NCM limits what song patterns can be distinguished. In the present model, we used a limited set of features encoded in the projection to NCM and, therefore, the model cannot presently identify complex arrangements of syllables that are distinguished by NCM. A more complete set of projection neurons to NCM, such as found in Field-L [22], would increase the number of auditory patterns that could be stored by this system, particularly selection for spectral features in addition to temporal features. In addition, the present model does not impose a spatial structure on the responses of cells in NCM, therefore no spatial map of *zenk* expression was predicted. In future studies, the model will be compared with *zenk* spatial patterns to determine whether acoustic features are projected to selective regions of NCM. Using the extended model, the model will predict how the spike probability of NCM-neurons changes during song presentation and generate a prediction comparable with the probability of *zenk*-expression in NCM that follows repeated presentations of conspecific songs.

References

[1] M. D. Hauser, E. M. Konishi, The design of animal communication, MIT Press, Cambridge, MA, 1999.

[2] W. H. Thorpe, The learning of song patterns by birds with special reference to the song of the chaffinch, Fringilla coelebs, Ibis 100 (1958) 535–570.

[3] M. Konishi, The role of auditory feedback in the control of vocalization in the white-crowned sparrow, Zeitung fr Tierpsychologie 22 (1965) 770–783.

[4] P. Marler, S. Peters, Selective vocal learning in a sparrow, Science 198 (1977) 519–521.

[5] C. K. Catchpole, P. J. B. Slater, Bird song: biological themes and variations, Cambridge University Press, U.K., 1995.

[6] D. E. Kroodsma, E. H. Miller, Ecology and evolution of acoustic communication in birds, Cornell University Press, Ithaca, NY, 1996.

[7] D. B. Kelley, F. Nottebohm, Projections of a telencephalic auditory nucleus-field l-in the canary, Journal of Comparative Neurology 183 (1979) 455–469.

[8] E. S. Fortune, D. Margoliash, Parallel pathways and convergence onto hvc and adjacent neostriatum of adult zebra finches (taeniopygia guttata), Journal of Comparative Neurology 360 (1995) 413–441.

[9] C. V. Mello, G. E. Vates, S. Okuhata, F. Nottebohm, Descending auditory pathways in the adult male zebra finch (taeniopygia guttata), Journal of Comparative Neurology 395 (1998) 137–60.

[10] G. E. Vates, B. M. Broome, C. V. Mello, F. Nottebohm, Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches, Journal of Comparative Neurology 366 (1996) 613–642.

[11] P. D. Roberts, Dynamics of temporal learning rules, Phys. Rev. E 62 (2000) 4077–4082.

[12] P. D. Roberts, Recurrent biological neural networks: The weak and noisy limit, Phys. Rev. E 69 (2004) 031910.

[13] M. E. Nelson, M. G. Paulin, Neural simulations of adaptive reafference suppression in the elasmobranch electrosensory system, J. Comp. Physiol. A 177 (1995) 723–736.

[14] H. J. Karten, T. Shimizu, The origins of neocortex: connections and lamination as distinct events in evolution, Journal of Cognitive Neuroscience 1 (1989) 291–301.

[15] C. V. Mello, D. S. Vicario, D. F. Clayton, Song presentation induces gene expression in the songbird forebrain, Proceedings of the National Academy of Sciences U S A 89 (1992) 6818–6822.

[16] C. V. Mello, D. F. Clayton, Song-induced zenk gene expression in auditory pathways of songbird brain and its relation to the song control system, Journal of Neuroscience 14 (1994) 6652–6666.

[17] C. Mello, F. Nottebohm, D. Clayton, Repeated exposure to one song leads to a rapid and persistent decline in an

immediate early gene's response to that song in zebra finch telencephalon, Journal of Neuroscience 15 (1995) 6919–6925.

[18] E. D. Jarvis, F. Nottebohm, Motor-driven gene expression, Proceedings of the National Academy of Sciences U S A 94 (1997) 4097–4102.

[19] S. J. Chew, C. Mello, F. Nottebohm, E. Jarvis, D. S. Vicario, Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain, Proc. Natl. Acad. Sci. USA 92 (1995) 3406–3410.

[20] S. J. Chew, D. S. Vicario, F. Nottebohm, A large-capacity memory system that recognizes the calls and songs of individual birds, Proc. Natl. Acad. Sci. USA 93 (1996) 1950–1955.

[21] T. A. Terleph, C. V. Mello, D. S. Vicario, Auditory topography and temporal response dynamics of canary caudal telencephalon, J Neurobiol 66 (3) (2006) 281–92.

[22] K. Sen, F. E. Theunissen, A. J. Doupe, Feature analysis of natural sounds in the songbird auditory forebrain, J Neurophysiol 86 (3) (2001) 1445–58.

[23] W. Gerstner, J. L. van Hemmen, Associative memory in a network of 'spiking' neurons, Network 3 (1992) 139–164.

[24] M. W. Levine, The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells, Biol. Cybern. 65 (1991) 459–467.

[25] E. Covey, J. H. Casseday, Timing in the auditory system of the bat, Annu. Rev. Physiol. 61 (1999) 457–476.

[26] E. W. Large, J. D. Crawford, Auditory temporal computation: interval selectivity based on post-inhibitory rebound, J. Comput. Neurosci. 13 (2) (2002) 125–142.

[27] J. Connor, D. Walter, R. McKown, Neural repetitive firing - modifications of the hodgkin-huxley axon suggested by experimental results from crustacean axons, Biophys. J. 18 (1977) 81–102.