

Processing of Natural Sounds in the Auditory Forebrain of Songbirds.

Frédéric Theunissen. Department of Psychology. UC Berkeley.

The long-term goal of our research is to understand the neural mechanisms underlying complex sound recognition and the formation of perceptual acoustical memories. Our methods combine neuroethology, classical auditory neuro-physiology and computational neuroscience. More specifically, we are studying the avian homologue of the mammalian auditory cortex in songbirds to understand its role in processing species specific vocalizations including the specific sounds that play a special role in the song learning behavior: the tutor song and the bird's own song (BOS). Songbirds learn to sing and this learning behavior is critically dependent on a sensory phase during which the young birds must recognize and memorize the song of a tutor. In this respect song learning is similar to speech learning in infants. Moreover, both young and adult songbirds are capable of recognizing the songs of their own species (conspecific songs) as well as songs of particular individuals within their own species. By understanding how these natural sounds are represented in the auditory cortex of songbirds, we hope to advance our understanding of how the mammalian auditory cortex processes complex sounds, including speech sounds. We are also comparing the analysis of the functional organization of the auditory forebrain based on this neuroethological approach to the one that would be deduced from neuronal data obtained solely in response to synthetic sounds. Our computational techniques provide the link between the two analyses. Although the use of ethological sounds has many advantages, it has been difficult to use such stimuli to investigate systematically which aspects of the sound are actually being encoded in the neural responses. To address this issue, we have developed a methodology that allows us to estimate the stimulus-response function of a neuron from data obtained with any type of stimulus, including natural sounds (Theunissen et al., 2000). By comparing the stimulus-response function obtained from ensembles of natural sounds to that obtained from synthetic sounds, we can compare the two approaches and ultimately attempt to combine them. This effort is crucial because most of our current knowledge of the auditory system is based on experiments performed with simple synthetic sounds.

We recorded the activity of single neurons in the primary auditory and secondary auditory forebrain of zebra finches, Field L and cHV. In the anatomical chain of acoustical processing stages of the avian brain, the field L region lies between the thalamic auditory relay nucleus ovoidalis (Ov) and higher-level auditory areas such as song nuclei HVC and the caudal hyperstriatum ventrale, cHV (Figure 1). This location is analogous to the location of auditory cortex in mammals. As in the primary auditory areas of many other animals, field L in zebra finches and other birds displays a tonotopic organization. Based on Nissl and Golgi staining studies, the field L region has been divided into 5 sub-regions called L2a, L2b, L1, L3 and L. Neuro-anatomical tracer studies have shown that the thalamic input from Ov projects strongly to area L2a and L2b and more weakly to L1 and L3. L2a projects strongly to L1 and L3, and all field L regions project to cHV (Figure 1). We have obtained neural responses to sounds chosen from a large ensemble that includes simple synthetic sounds (such as tones), complex synthetic sounds (such as combination of tones and static ripples) and conspecific songs. Our neuroethological hypothesis is that that we will observe a hierarchical processing of complex sounds leading to neurons or brain areas that are particularly selective to conspecific songs or more specifically to the BOS and the tutor song. Our computational work would then expose the features of the complex sounds that are crucial for such neural recognition.

We have recently just finished the first systematic mapping of the auditory areas in the forebrain of adult zebra finches. We found that neurons in all of these brain areas showed stronger responses to natural sounds than to *matched* complex synthetic sounds (Figure 2). The complex sounds were matched to the natural sounds in the sense that some of their characteristic average acoustical properties (such as their power spectra) were identical. We also discovered that the *difference* in the strength of the response to natural sounds versus synthetic sounds was similar in primary and secondary auditory areas in apparent

contradiction with our hypothesis. However, we also found that the secondary auditory areas exhibited weaker responses to all sounds such that in some cases the response to the synthetic sounds was close to 0 while the response to conspecific song was still significantly positive. In this manner, a hierarchical selectivity for complex natural sounds could be achieved in the secondary areas by filtering out the weaker auditory responses of the primary auditory areas.

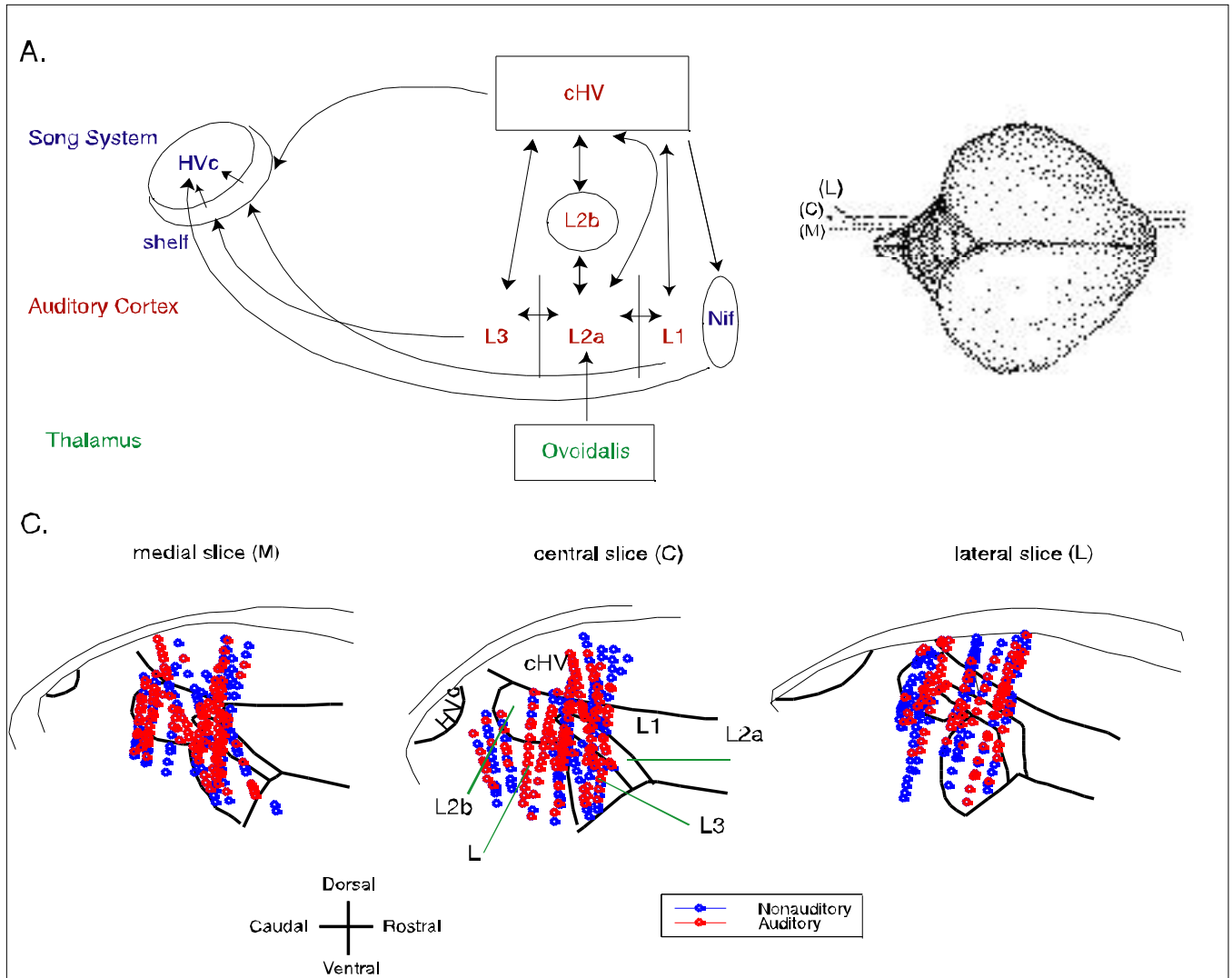


Figure 1. A. Schematic of the anatomical subdivisions and connectivity of the avian auditory forebrain and relation with the song system. The diagram is based on the results and schematics found in B and C. Location of all our recording sites projected onto one of three sagittal slices of the auditory forebrain. B shows the location relative to the midline of the 3 slices (medial slice 1200 μ , central slice 1400 μ , lateral slice 1600 μ). The three panels in C show the actual location within a given slice of the recording site. The boundaries separating the different auditory sub-regions are shown in bold (Nif is not shown). Units that were determined to be auditory are shown in red.

We also compared the neural activity obtained in response to unrelated conspecific songs to the one obtained in response to the BOS. Surprisingly, we found that the BOS elicited weaker responses on average than conspecific song. We postulate that these weaker responses might reflect a perceptual memory, a weaker response to an expected or common stimulus. This same neural circuitry would also be able to participate in song learning by selectively responding to the bird's vocalizations that deviate from a stored template. We are currently recording from young birds to directly test this hypothesis.

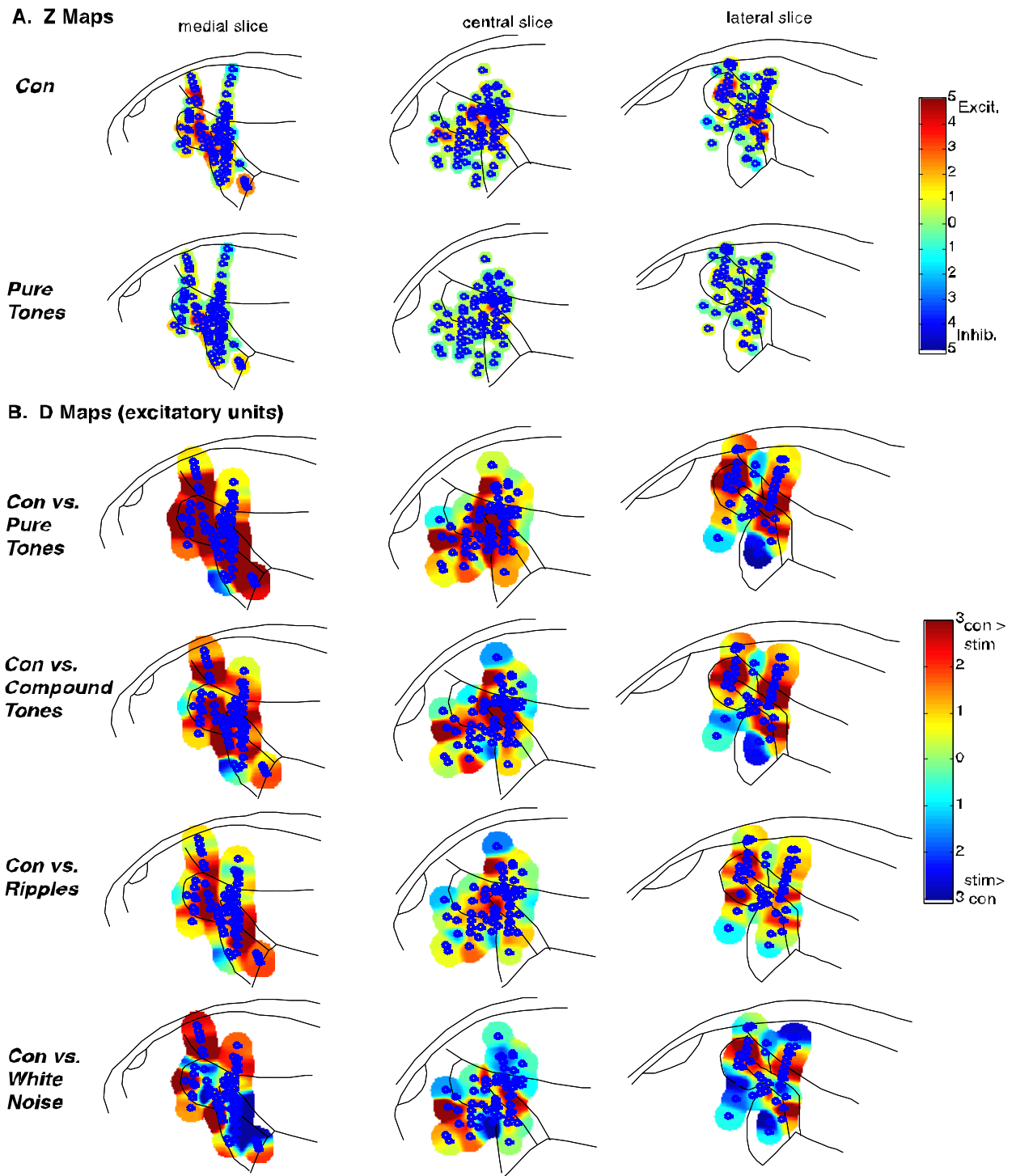


Figure 2. Comparison of the response to conspecific song (con) to the response to the synthetic sounds shown in figure 1. A. Maps of the response strength (z-score) to conspecific (con) and pure tone stimuli in the auditory forebrain regions described in figure 2. The strength of the response is color coded. Inhibitory responses are in blue and excitatory responses in green-yellow and red as shown in the color bar. B. Selectivity maps (d values) for the con versus synthetic comparison for the excitatory units. The color illustrates the strength of the selectivity as shown in the color bar. Note how the first row of panel B can be obtained approximately by subtracting the two response maps shown in panel A from each other.

Using our computational methods, we have also generated functional maps of response properties based on the spectral-temporal receptive fields (STRF) of the recorded neurons. The STRF shows temporal succession of acoustical features that would yield the highest response from a linear neuron. Our analytical methods allow us to estimate STRFs both from response to natural complex stimuli and to simple synthetic stimuli. From these STRFs, we have extracted classical auditory response parameters, such as the best frequency and the best amplitude modulation of sound, as well as more complex response parameters such as frequency modulation tuning. We are currently mapping these parameters to better understand how sounds are analyzed in the auditory cortex. Preliminary results are consistent with a hierarchical processing of sounds: the STRFs in the primary auditory areas have shorter latency and higher BMF (Figure 3). Also the STRFs in the primary auditory areas are better models of the stimulus-response functions of the neurons in the sense that they are better at predicting responses to novel stimuli. Therefore the stimulus response functions in the primary auditory areas are more linear than those in higher auditory areas.

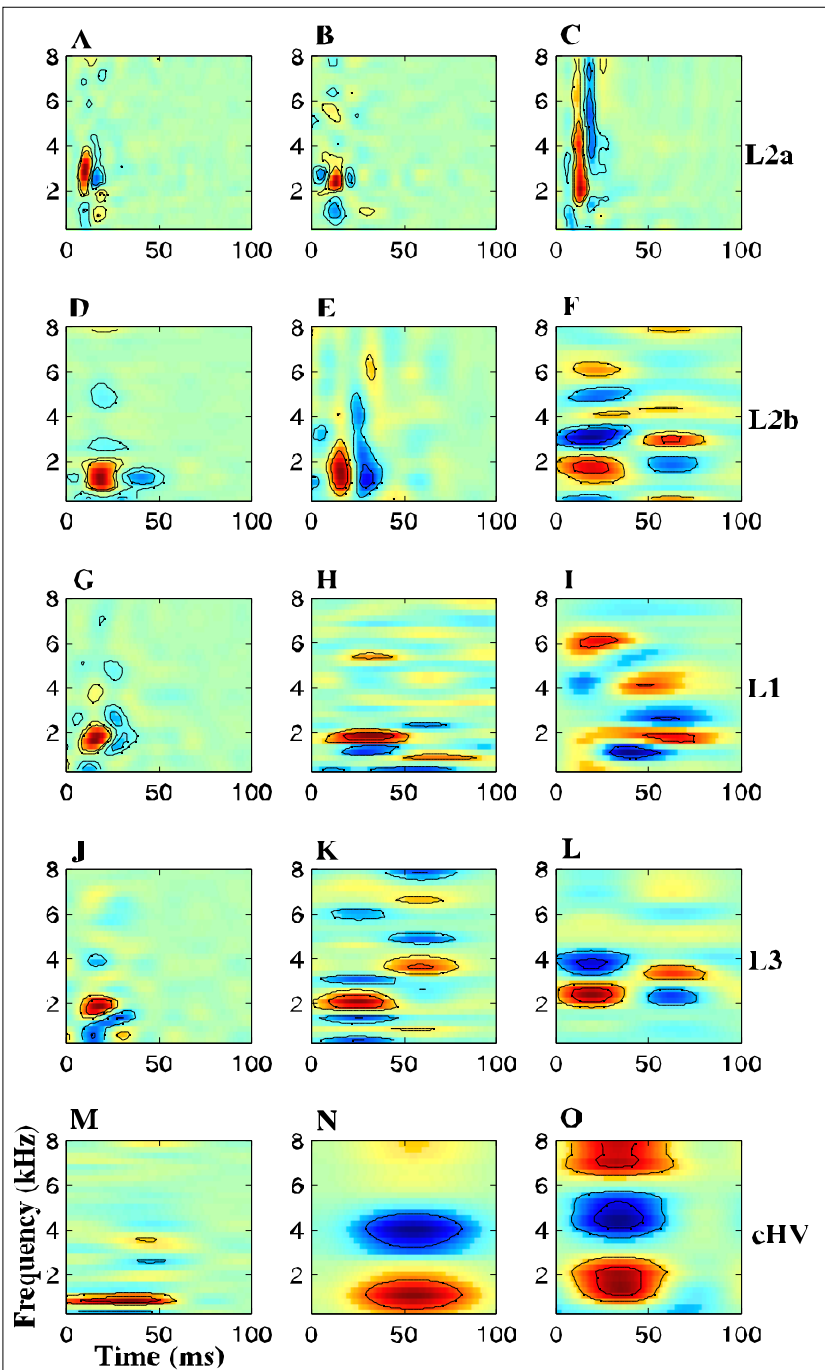


Figure 3. Examples of STRFs from different regions in the auditory forebrain. 3 examples from each sub-region are shown. A-C. Examples from sub-region L2a (sites 21.2, 14.11 and 27.4A). D-F. Examples from sub-region L2b (sites 20.4, 26.2B and 23.3B). G-I. Examples from sub-region L1 (sites 18.6, 18.7 and 14.9). J-L. Examples from sub-region L3 (sites 25.2, 27.5B and 27.4B). M-O. Examples from sub-region cHV (sites 14.2, 27.1A and 27.2A). The examples in A, D, K and L were from single units (see Methods). The STRFs display the spectral-temporal features of songs to which auditory forebrain neurons respond. This figure illustrates the diverse range of features to which neurons responded in the auditory forebrain. These range from simple features showing narrow-band components of song, as in the examples in 5A from sub-region L2a and 5D from sub-region L2b, to more complex multi-peaked features as in 5L from sub-region L3 and 5O from sub-region cHV. The figure also shows the wide range of time-scales of the features that can be found in the auditory forebrain.

Suprisingly, we also found that the STRF obtained with natural stimuli and synthetic stimuli can be significantly different. In other words, the response properties obtained from synthetic sounds does not describe properties of the neuron that only become apparent when natural sounds are used, and vice-versa. More recently, we have attempted to obtain a generalized non-linear model of the stimulus-response properties of the neurons, with the idea that some of the observed differences were due to non-linear properties of the cells. The non-linear receptive fields were somewhat better models of the stimulus-response but exhibit the same phenomena: the models did generalize across stimulus ensembles. We are tentatively explaining this phenomenon by stating that the neuron response properties can rapidly adapt to a different states. A more complete description of high-level sensory neurons will therefore involve multiple receptive fields to accommodate for these different processing states.

Relevant Publications:

*Theunissen, F.E., Sen, K. and Doupe, A.J. (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neuroscience*, **20**(6):2315-2331

*Theunissen. F. E., David S., Singh N., Hsu A., Vinje W. and Gallant J. (in press) Estimating Spatial Temporal Receptive Fields of Auditory and Visual Neurons From Their Responses to Natural Stimuli. *Network: Computation in Neural Systems*.