

Information Transfer Along the Ventral Auditory Processing Stream in the Awake Macaque

Elliot H. Smith, Spencer S. Kellis, Paul A. House, and Bradley Greger, *Member, IEEE*

Abstract— Few studies have examined the physiology of the auditory cortical processing streams in the context of information transfer among cortical areas. This study examines information transfer in two cortical areas in the ventral auditory processing stream in an awake macaque. We show conditional information examined over different durations of neural responses provides insight into the time scale and direction of cortical hierarchical processing.

I. INTRODUCTION

Neocortical sensory processing is thought to occur simultaneously in multiple cortical streams as the brain transforms sensory information into more cognitive representations. Support for this hierarchical processing in visual and auditory cortex stems from neuroanatomical studies outlining projection patterns through the cortical circuit [1-4], and from lesion studies [5, 6]. These studies have identified two processing streams: spatial elements of a stimulus are processed in a “dorsal” stream extending along the parietal lobe, and stimulus identity or quality is processed in a “ventral” stream extending rostrally along the temporal lobe. While there is compelling neurophysiological evidence for a “dorsal” stream in audition [7], physiological studies of a “ventral” stream are limited.

Researchers have examined selectivity of neurons along the ventral auditory processing stream using differences in firing rate [8], statistical classifiers [9], and mutual information [10]. None of these measures, however, examines cortical physiology in the context of information transfer along the processing stream. Modeling work, and studies in other sensorimotor modalities have used, cross-correlation [11], Granger causality [12], and directed mutual information [14] to inform the understanding of cortical connectivity and information flow in cortical circuits. These measures all infer causality from a time delay.

Conditional mutual information (CMI) is the mutual information between two random variables given a third random variable is known. This measure has been used to infer connectivity from resting fMRI data [13]. For sensory cortex, CMI allows one to calculate the information between a sensory stimulus and a neural response, given that the neural response in another cortical area is known. CMI is therefore a relatively simple way of analyzing information transfer among multiple, interconnected cortical areas. The

advantages of CMI in neural responses are that CMI does not rely on Gaussian statistics, or infer causality from a time delay.

The goal of the present study was to characterize the physiology of the two cortical areas in the auditory ventral stream in the context of a processing stream. We therefore examined information transfer of neural responses in two cortical areas in an awake primate. Specifically we looked at action potential (AP) and local field potential (LFP) responses to species-specific vocalizations from primary auditory cortex (AI) and rostral parabelt auditory cortex (PBr).

II. METHODS

A. Research Subject

The neural data examined in this study were recorded from two microelectrode arrays chronically implanted in an adult male rhesus macaque. Each microelectrode array contained 96 electrodes and was implanted in AI (figure 1a) and PBr (Figure 1b). All experiments were performed according to NIH guidelines for animal care and use and with approval from the University of Utah Institutional Animal Care and Use Committee.

B. Experimental Setup

Each of seven macaque vocalization exemplars was randomly presented free field through a piezoelectric speaker (ES1, Tucker Davis Technologies, Alachua, FL). There was at least one second in time between each stimulus presentation. Stimuli were presented while the monkey sat in an acoustically, optically, and electromagnetically shielded chamber (Acoustic Systems, Austin, TX).

C. Data Acquisition and Preprocessing

Neural data were recorded with a Cerebus System (Blackrock Microsystems, Salt Lake City, UT). The broadband data were recorded at 30 kHz and later filtered

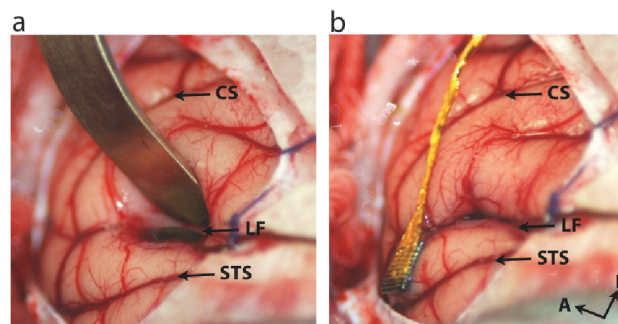


Figure 1. Electrode arrays implanted in AI (a) and PBr (b). CS, Central Sulcus; LF, Lateral Fissure; STS, Superior Temporal Sulcus; A, Anterior; D, Dorsal.

E. H. Smith and B. Greger are with the Department of Bioengineering, University of Utah, Salt Lake City, UT, 84112, USA (phone: 206-349-1249; e-mail: e.h.smith@utah.edu).

S. S. Kellis is with the Division of Biology, California Institute of Technology, Pasadena, CA 91125 USA.

P. A. House is with the Department of Neurosurgery, University of Utah, Salt Lake City, UT, 84112, USA

and down-sampled to 2 kHz for analysis of the LFP. High-frequency AP data were digitally filtered (eight pole, high-pass 250Hz, low-pass 7.5kHz) and sampled at 30 kHz. Individual spikes were detected offline using t-distribution E-M sorting [15]. Only units that were well isolated from the noise cluster in 3-D PCA space were retained for further analysis (12 in PBr, and 24 in AI).

LFPs were recorded on 96 channels for 144 trials per class in AI (1008 total trials), and on 95 channels for 112 trials per class in PBr (784 total trials) over multiple weeks.

Multitapered spectrograms and peri-stimulus time histograms (PSTHs) were generated using an open source neural analysis package (www.chronux.org) [16]. Single-trial PSTHs and LFP spectrograms were generated separately for four durations after the beginning of each vocalization (200, 400, 800, and 1600 msec). PSTHs were calculated using Gaussian kernels of five widths (5, 10, 25, 50, and 100 msec). Spectrograms were calculated using multitapered analysis with a time-bandwidth product of 5, 9 leading tapers, 200 msec windows, 10 msec step sizes, and using 1024-point fast Fourier transforms.

D. Statistical Classification

We extended the method of Kellis et al. [17] to simultaneously incorporate features representing dynamics in time, space, and frequency for both AP data and LFP data. We use the term *class* to refer to one type of species-specific-vocalization (e.g., “grunt”), and the term *trial* to refer to one instance of a vocalization being played for the monkey.

To select training features for LFP responses, two-dimensional spectrograms were calculated for a subset of trials from each class and each recording channel. These multidimensional data were unwrapped to produce a two-dimensional matrix in which each row contained all the time and frequency features from all channels for a single trial. The feature matrix was z-scored, orthogonalized using principal component analysis (PCA), and projected into the principal component space using a sufficient number of leading principal components to retain 95% of the variance in the data. A different subset of trials was selected for testing the classifier, and spectrograms were calculated for each class and each recording channel. These data were unwrapped into a two-dimensional matrix, z-scored, and projected into the same principal component space calculated for training features.

To produce training and testing sets from AP data, we selected consecutive days for which at least four common channels had units with similar firing responses (6 pairs of days for AI, 4 units for each day; 2 pairs of days for PBr, 6 units for each day). For each pair of consecutive days, the data from the first day were used for training, and data from the second day were used as testing. For both training and testing, AP data were collected into a large two-dimensional matrix where each row represented all firing rate data from each unit for a given trial. Then, the data were z-scored, orthogonalized, and projected into the principal component space using the same process described above for LFP data.

The classifier was evaluated for all pair-wise combinations of vocalizations using linear discriminant

analysis. Pairwise classifications were examined in order to determine probability distributions to be used in information theoretic analyses.

E. Information Theoretic Analyses

To examine cortical information transfer we examined CMI between a stimulus and a response in one cortical area given that there was a response in the other cortical area. Let R_x and R_y be two random variables representing neural response classes in two different cortical areas, we calculated CMI as follows:

$$I(S;R_x | R_y) = \sum_{s \in S} \sum_{x \in R_x} \sum_{y \in R_y} p(s,x,y) \log_2 \left(\frac{p(y)p(s,x,y)}{p(s,y)p(x,y)} \right). \quad (1)$$

Where probability distributions $p(y)$, $p(s,y)$, $p(x,y)$, and $p(s,x,y)$ were taken from confusion matrices of pair-wise classification frequencies for all AP and LFP responses. Equation (1) was evaluated for R_x being responses in AI and R_y being responses in PBr to calculate CMI for stimulus and response in AI given that the response in PBr is known ($I(S;R_{AI} | R_{PBr})$). CMI for stimulus and response in PBr given a known response in AI ($I(S;R_{PBr} | R_{AI})$) was calculated in the same way. Mean information values across data durations are reported as mean \pm standard deviation.

III. RESULTS

A. Transfer Information Varies for Different Time Scales of Auditory Processing

To examine cortical information transfer between AI and PBr for the seven species-specific vocalization stimuli, we calculated the mutual information between an auditory stimulus and the neural response in one cortical area given the response in the other cortical area was known; the CMI. For AP responses, when CMI is averaged over all kernel-widths, information in AI given PBr was greater than the information in PBr given AI for all 4 data durations ($I(S;R_{AI} | R_{PBr}) - I(S;R_{PBr} | R_{AI}) = 1.68 \pm 0.51$ bits) (Figure 2a). For LFP responses, the information in AI given PBr was also greater than the information in PBr given AI for the 400, 800 and 1600 millisecond data durations ($I(S;R_{AI} | R_{PBr}) - I(S;R_{PBr} | R_{AI}) = 0.49 \pm 0.40$ bits). However, for the 200 millisecond data duration for LFP responses information in PBr given AI was greater than the information in AI given PBr ($I(S;R_{PBr} | R_{AI}) - I(S;R_{AI} | R_{PBr}) = 3.33$ bits) (Figure 2b). The highest CMI observed across all LFP responses was for the 200-millisecond data duration for responses in PBr, given responses in AI (6.64 bits) (figure 2b).

B. Transfer Information for Short Time-Scale AP Responses is Dependent on PSTH Temporal Resolution

While the bias of information transfer reverses for 200-millisecond versus 400, 800 and 1600 millisecond LFP responses, this change in direction is not evident for the AP responses averaged over the different kernel-widths. However, this change in bias is apparent for the 25 and 50 millisecond kernel widths (red and blue lines are higher for the 200 millisecond data duration in Figure 3b, whereas all

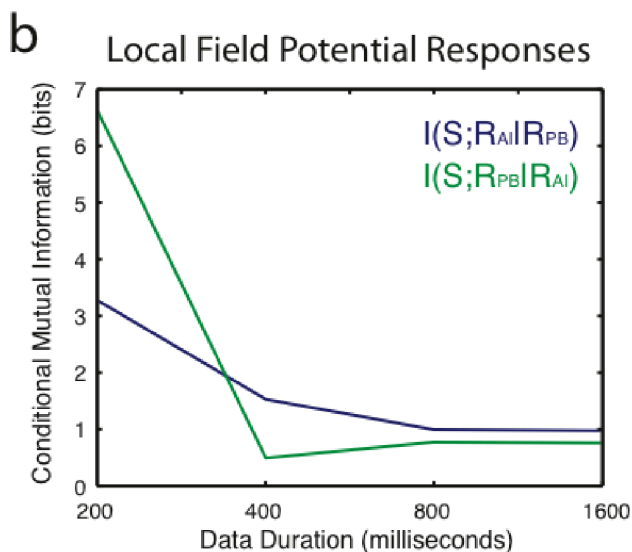
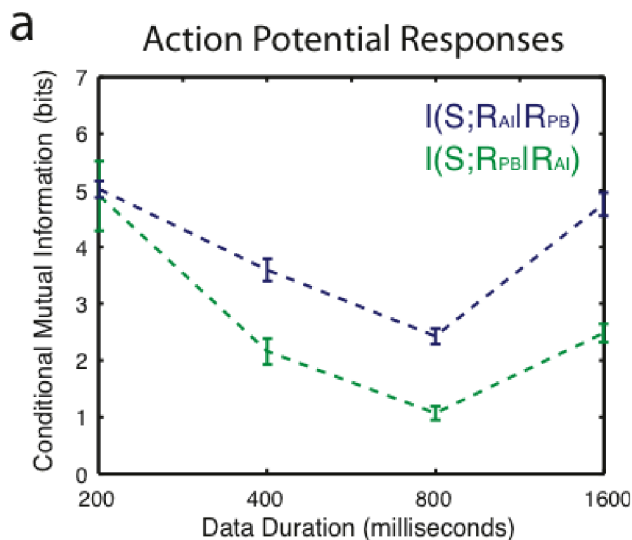


Figure 2. CMI in LFP responses and average AP responses. (a) Average CMI for AP responses. Average CMI over kernel-widths from CMI in Figure 2 is shown. Bars represent standard error. (b) CMI for LFP responses. In both (a) and (b), blue lines indicate CMI between stimulus and responses in AI given responses in PBr are known and green lines indicate CMI between stimulus and AP responses in PBr given the response in AI is known.

colored lines are lower in Figure 3a). The highest CMI observed for AP responses was for the 200 millisecond data duration for responses in PBr, given responses in AI for the 50-millisecond kernel width (8.77 bits) (figure 3b).

IV. DISCUSSION

At the beginning of a cortical response to an auditory stimulus, information transfer is largely feed-forward, or dependent on lower auditory nuclei to filter and propagate the transduced neural information to the cerebral cortex. Then, as the identity of a stimulus is recognized, information transfer feeds back from higher cortical areas. The current results reflect this schema by showing the first 200 milliseconds of a response in PBr being informed by the first 200 milliseconds of a response in AI, and longer responses in AI being informed by longer responses in PBr. In examining CMI

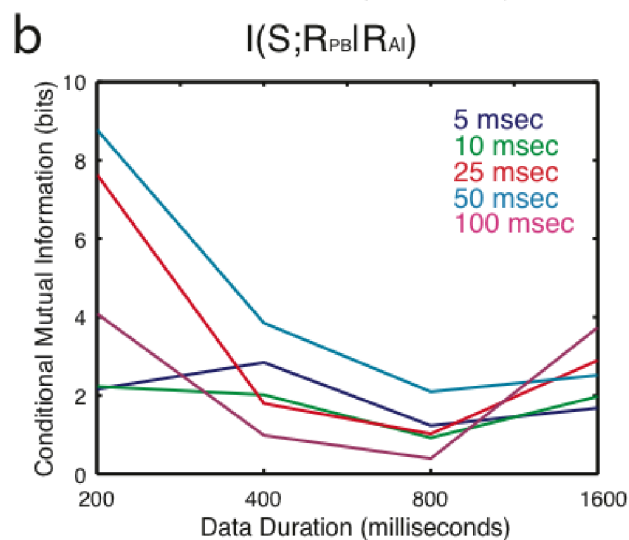
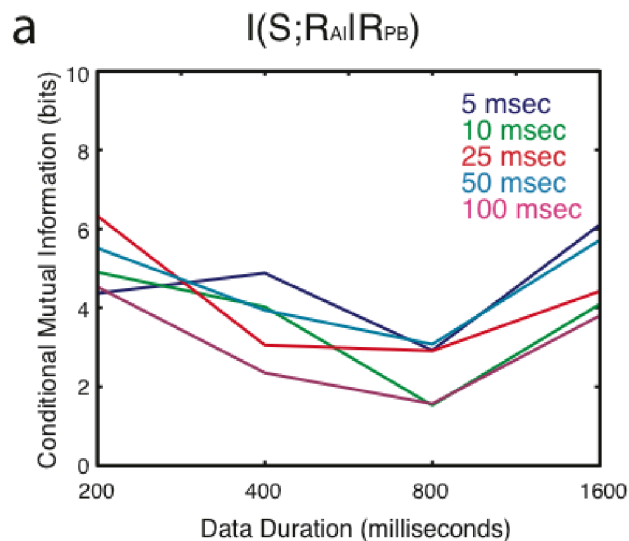


Figure 3. CMI in AP responses over kernel widths. (a) CMI between stimulus and AP responses in AI given the response in PBr is known. Shows CMI for action potential responses for different data durations and PSTH kernel-widths (colored lines). (b) CMI between stimulus and AP responses in PBr given the response in AI is known. Shows CMI for action potential responses for different data durations and PSTH kernel widths (colored lines), as in (a).

between AI and PBr, we found that information in AI given responses in PBr was higher for the 400, 800 and 1600 millisecond data durations. Interestingly, information in PBr given responses in AI was highest for the 200 millisecond data duration for LFP responses. Information from the 200 millisecond data duration in PBr given responses in AI was also higher for two of the five kernel widths (25 and 50 milliseconds).

Responses to full-length vocalizations in AI contained more information about the stimulus given that responses in PBr were known, when compared with information in PBr given responses in AI were known. The duration of the vocalizations we played ranged from 0.23 to 1.27 seconds (average vocalization length was 0.53 ± 0.37 seconds). The 400 through 1600 millisecond data durations therefore each contain the full duration of at least one vocalization, whereas

the 200 millisecond data duration doesn't contain the full duration of any of the seven vocalizations. These results accord with the theory that cortico-cortical feedback is evident in offset, or at least later, neural responses.

The output of our classifier provided more information about the first 200 milliseconds of responses in PBr given that the first 200 milliseconds of responses in AI were known. Further examination of the CMI from AP responses shows that CMI from the 5, 10 and 100 millisecond kernel-widths is lower in PBr given AI than that in AI given PBr. The lack of information in PBr given AI for the 5 and 10 millisecond kernel-widths points to a particular window of temporal integration in PBr. A recent study examining temporal integration in macaque auditory cortex shows that neurons from the belt region of auditory cortex (between AI and parabelt in the processing stream) have difficulty synchronizing to stimulus modulations above and below 5 Hz [18]. PBr may have a similar window of temporal integration, as the high information for the 25 and 50 millisecond kernel-widths corresponds to a temporal integration window of between 20 and 40 Hz. The 100 millisecond kernel-width is wide enough to be interpreted as an average firing rate for the 200 millisecond data duration. Therefore we can conclude that temporal fluctuations in the neural responses, in addition to average firing rate, provide information about a stimulus in the first 200 milliseconds in PBr. Together these results suggest that coding in PBr relies on information from overall firing rate and precise temporal fluctuations of responses in AI, though neurons in PBr might not be able to synchronize to rapid temporal modulations, like those in AI.

V. CONCLUSION

We have presented a study of neural responses in two cortical areas along the ventral processing stream in the context of information transfer along the cortical processing stream. We have shown that responses in AI contain more information about species-specific vocalizations when late responses in PBr are known. In contrast, early responses in PBr contain more information about stimuli when responses in AI are known. These results elucidate the time scale and directionality of processing along the ventral auditory processing stream.

References

- [1] M. Mishkin and L. G. Ungerleider, "Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys," *Behav Brain Res*, vol. 6, pp. 57-77, Sep 1 1982.
- [2] D. J. Felleman and D. C. Van Essen, "Distributed hierarchical processing in the primate cerebral cortex," *Cereb Cortex*, vol. 1, pp. 1-47, Jan 1 1991.
- [3] T. A. Hackett, "Information flow in the auditory cortical network," *Hear Res*, vol. 271, pp. 133-46, Jan 1 2011.
- [4] J. A. Winer and C. C. Lee, "The distributed auditory cortex," *Hear Res*, vol. 229, pp. 3-13, Jul 1 2007.
- [5] H. E. Hefner and R. S. Hefner, "Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques," *Journal of Neurophysiology*, vol. 56, pp. 683-701, Sep 1 1986.
- [6] A. U. Turken and N. F. Dronkers, "The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses," *Front. Syst. Neurosci.*, vol. 5, pp. 1-20, Jan 1 2011.
- [7] T. M. Woods, *et al.*, "Effects of Stimulus Azimuth and Intensity on the Single-Neuron Activity in the Auditory Cortex of the Alert Macaque Monkey," *Journal of Neurophysiology*, vol. 96, pp. 3323-3337, Aug 9 2006.
- [8] X. Wang, "Neural coding strategies in auditory cortex," *Hear Res*, vol. 229, pp. 81-93, Jul 1 2007.
- [9] J. W. H. Schnupp, *et al.*, "Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex," *Journal of Neuroscience*, vol. 26, pp. 4785-95, May 3 2006.
- [10] L. M. Romanski, *et al.*, "Neural representation of vocalizations in the primate ventrolateral prefrontal cortex," *Journal of Neurophysiology*, vol. 93, pp. 734-47, Feb 1 2005.
- [11] A. Adhikari, *et al.*, "Cross-correlation of instantaneous amplitudes of field potential oscillations: a straightforward method to estimate the directionality and lag between brain areas," *Journal of Neuroscience Methods*, vol. 191, pp. 191-200, Aug 30 2010.
- [12] A. K. Seth, "A MATLAB toolbox for Granger causal connectivity analysis," *Journal of Neuroscience Methods*, vol. 186, pp. 262-273, Mar 15 2010.
- [13] R. Salvador, *et al.*, "Conditional mutual information maps as descriptors of net connectivity levels in the brain," *Frontiers in neuroinformatics*, vol. 4, p. 115, 2010.
- [14] K. So, *et al.*, "Assessing functional connectivity of neural ensembles using directed information," *Journal of Neural Engineering*, vol. 9, p. 026004, Mar 13 2012.
- [15] S. Shoham, *et al.*, "Robust, automatic spike sorting using mixtures of multivariate t-distributions," *J Neurosci Methods*, vol. 127, pp. 111-22, Aug 15 2003.
- [16] H. Bokil, *et al.*, "Chronux: a platform for analyzing neural signals," *J Neurosci Methods*, vol. 192, pp. 146-51, Sep 30 2010.
- [17] S. Kellis, *et al.*, "Decoding spoken words using local field potentials recorded from the cortical surface," *J. Neural Eng.*, vol. 7, p. 056007, Oct 1 2010.
- [18] B. H. Scott, *et al.*, "Transformation of temporal processing across auditory cortex of awake macaques," *Journal of Neurophysiology*, pp. 1-75, Nov 24 2010.