Behavioral/Systems/Cognitive

The Influence of Natural Scene Dynamics on Auditory Cortical Activity

Chandramouli Chandrasekaran,[1,2] **Hjalmar K. Turesson,**[1,2] **Charles H. Brown,**[4] **and Asif A. Ghazanfar**[1,2,3]

[1]Neuroscience Institute and Departments of [2]Psychology and [3]Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08540, and [4]Department of Psychology, University of South Alabama, Mobile, Alabama 36688

The efficient cortical encoding of natural scenes is essential for guiding adaptive behavior. Because natural scenes and network activity in cortical circuits share similar temporal scales, it is necessary to understand how the temporal structure of natural scenes influences network dynamics in cortical circuits and spiking output. We examined the relationship between the structure of natural acoustic scenes and its impact on network activity [as indexed by local field potentials (LFPs)] and spiking responses in macaque primary auditory cortex. Natural auditory scenes led to a change in the power of the LFP in the 2–9 and 16–30 Hz frequency ranges relative to the ongoing activity. In contrast, ongoing rhythmic activity in the 9–16 Hz range was essentially unaffected by the natural scene. Phase coherence analysis showed that scene-related changes in LFP power were at least partially attributable to the locking of the LFP and spiking activity to the temporal structure in the scene, with locking extending up to 25 Hz for some scenes and cortical sites. Consistent with distributed place and temporal coding schemes, a key predictor of phase locking and power changes was the overlap between the spectral selectivity of a cortical site and the spectral structure of the scene. Finally, during the processing of natural acoustic scenes, spikes were locked to LFP phase at frequencies up to 30 Hz. These results are consistent with an idea that the cortical representation of natural scenes emerges from an interaction between network activity and stimulus dynamics.

Introduction

Auditory communication usually takes place in acoustic scenes with myriad biotic and abiotic sources generating sound energy in a wide range of spectral frequencies and with a variety of time constants, repetition rates, and durations. These scenes contain vocalizations and other species-typical sounds produced by conspecifics and numerous other species sharing a common habitat, as well as the noises produced by the physical features of the environment such as wind and water (Waser and Brown, 1986; Slabbekoorn, 2004; Brumm and Slabbekoorn, 2005). Primates and other animals must efficiently encode the complex spectrotemporal structure of these scenes. Typically, the study of cortical representations of natural auditory scenes emphasized spiking activity (Nelken et al., 1999; Rotman et al., 2001), but recent efforts suggest a role for local field potentials (LFPs) (Belitski et al., 2008; Montemurro et al., 2008; Panzeri et al., 2010). LFPs are network-level signals thought to reflect the input and intracortical processing in a cortical area. They are made up of synchronized populations of synaptic potentials (Mitzdorf, 1987; Juergens et al., 1999) and other types of slow activity unrelated to

synaptic events, including voltage-dependent membrane oscillations and spike afterpotentials (Logothetis, 2003).

Understanding the role that LFPs may have in encoding natural scenes is tricky for at least three reasons. First, both natural scenes (Singh and Theunissen, 2003) and LFPs (Buzsáki and Draguhn, 2004; Buzsáki, 2006) evolve on multiple temporal scales with similar frequency bands. Second, the LFP is sensitive to other factors not related to the stimulus structure, such as the behavioral state of the animal (Fontanini and Katz, 2005, 2008). It is therefore unclear how scene structure and ongoing network activity would interact with one another or how this interaction would impact spiking activity. Third, most previous studies focusing on the role of LFPs in sensory coding focused on higher-frequency oscillations, such as the gamma band (Brosch et al., 2002; Henrie and Shapley, 2005; Liu and Newsome, 2006), and essentially ignored the lower-frequency bands.

To bridge these epistemic gaps, we addressed two fundamental questions: (1) what is the influence of natural scene structure on LFPs in auditory cortex? and 2) how does this influence affect spiking activity? To address these questions, we first characterized the spectrotemporal structure of natural auditory scenes recorded from three different primate habitats: savanna, riverine forest, and rainforest. We then recorded LFPs and spiking activity in the primary auditory cortex of macaque monkeys while they listened to these natural scenes. We used spectral analysis and phase coherence to understand the relationship between stimulus dynamics, network activity, and spiking responses during the processing of natural acoustic scenes. Our results suggest that the cortical encoding of natural scenes by spiking activity emerges through an interac-

tion between ongoing activity in auditory cortex and the temporal structure of the natural scene.

Materials and Methods

Subjects and surgery

Two adult male long-tailed macaques (*Macaca fascicularis*) were used in the experiments. For each monkey, we used preoperative whole-head magnetic resonance imaging (3 T magnet, 500 μm slices) to identify the stereotaxic coordinates of auditory cortex. The monkeys underwent sterile surgery for the implantation of a head post and recording chamber. The surgery was done with the animal under general anesthesia. The chamber was vertically oriented to allow an approach to the superior surface of the superior temporal gyrus (Pfingst and O'Connor, 1980; Recanzone et al., 2000). All experiments were performed in compliance with the guidelines of the Princeton University Institutional Animal Care and Use Committee.

Stimuli

We used natural auditory scenes from three different habitats common to Old World monkeys (including *Macaca* spp.). These habitats were the savanna, riverine forest, and rainforest. These scenes were used not because our captive macaque monkey subjects were familiar with them (they were not) but because these scenes contain the spectrotemporal complexity that the primate auditory system evolved to process. These scenes were recorded during early mornings (by C.H.B.) between the years 1975 and 1990. Geographical details about the source for these acoustic scenes, along with the animal species living in these habitats, have been published previously (Waser and Brown, 1986). All three types of natural scenes variously contained vocalizations by birds, insects, and mammals, as well as abiotic sounds. For our purpose, we chose nine (3 categories $\times$ 3 exemplars) 60-s-long segments from our database.

Paradigm and data collection

Subjects sat in a primate restraint chair placed in a darkened, sound-attenuating radio frequency enclosure. The enclosure was lined with echo-attenuating foam. The sound amplitude ranged from 55 to 68 dB sound pressure level measured at the location of the subject's head. Each "trial" lasted for 1 min. Intertrial intervals were 10 s in duration, and each scene was repeated five times. To ensure that the subject was awake and listened to the scene, juice was administered at random intervals throughout the duration of the experiment, including during auditory stimulation.

Recordings were made from the left primary auditory cortex using standard electrophysiological techniques. We used an eight-channel microdrive (NAN Instruments) that allowed us to move multiple electrodes independently. Guide tubes were used to penetrate the overlying tissue growth and dura. Electrodes (FHC Inc.) were glass-coated 125-μm-thick tungsten wire with impedances between 2 and 4 MΩ (measured at 1 kHz). Grounding was achieved by connecting the microdrive to the ground provided by the head stage and by using an additional grounding wire connected to the circuitry in the preamplifier. This eliminated or minimized 60 Hz line noise. All signals were acquired using the Plexon multichannel acquisition system (MAP; Plexon Instruments). Electrodes were lowered until multiunit cortical responses could be driven reliably by auditory stimuli. Search stimuli included pure tones, frequency-modulated (FM) sweeps, noise bursts, clicks, and vocalizations. Our recordings were performed with the head stages provided by Plexon Instruments. In one of the subjects, we used the HST/8o50-G20 head stage, which is known to have some issues regarding low-frequency LFP phase (Nelson et al., 2008). The FPalign utility from Plexon Instruments was used to correct for any frequency-dependent phase shifts for data recorded using this head stage (Nelson et al., 2008).

Local field potentials. Mean extracellular field potential signals were low-pass filtered at 100 Hz. Power spectra of the LFP signal were computed to determine whether there was any corruption by 60 Hz line noise. We discerned a small 60 Hz peak in a few cortical sites and removed this line noise using a notch filter with cutoffs between 59.9 and 60.1 Hz. We overlaid the filtered and unfiltered spectra to ensure that notch filtering did not lead to distortions of the signal. We then applied the notch filter to all cortical sites to ensure that they were all manipulated in the same manner.

Spiking activity: multiunits. To extract spiking activity from multiunits, we adapted methods published previously (Quiroga et al., 2004; Montemurro et al., 2008; Nauhaus et al., 2009). For the extraction of spike times, the raw neural signal was filtered in the high-frequency range of 500–5000 Hz using a four-pole Butterworth filter. We then used a spike detection threshold of 3.5 SDs. A spike was recognized as such only if the last spike occurred >1 ms earlier. This method is excellent for detecting spiking events reflecting the activity from one or more neurons around the tip of the electrode; it is not as effective for isolating single neurons.

Spectral tuning and tonotopy. For every cortical site, the frequency tuning was measured by using one of two approaches: (1) 40-tone pips of 200 ms each spaced logarithmically between 125 Hz and 20 kHz and presented at three different attenuation levels, or (2) 100 ms tone pips repeated five times per second with five frequencies logarithmically spaced between 100 Hz and 1 kHz and 20 frequencies logarithmically spaced between 1 and 20 kHz. Tonotopy was determined by measuring both the maximal deflection of the LFP along with the maximal spike discharge rate as a function of spectral frequency. Based on our stereotaxic coordinates, tonotopy, and the spectral selectivity of neurons, our recordings were in the primary auditory cortex. In both monkey subjects, we discerned a high-to-low transition of frequency selectivity in the caudal-to-rostral direction (Hackett et al., 1998).

Data analysis

All analyses were performed in MATLAB (MathWorks) using a combination of custom-made scripts and the Chronux toolbox (http://chronux.org).

Wideband envelope of the auditory scenes. To compute the wideband envelope of the natural scenes, we first bandpass filtered the sound into spectral frequency bands spaced in 1 kHz intervals. Choosing such a large spacing between cutoffs ensured that we minimized the overlap in energy between adjacent bands while designing filters so that the bands could be summed to create a wideband envelope. Narrowband elliptic filters with a 40 dB stop band and a 1 dB pass-band ripple (fdatool; MATLAB) were designed and applied to the scenes. We then computed the envelope for each of these bands by computing the absolute value of the Hilbert transform of this signal. Summing these individual narrow bands provided us with a wideband envelope.

Narrow-band envelopes of the auditory scenes. To identify the precise spectral frequencies that may have led to a response, we split the natural scenes into narrow frequency bands and then computed narrow-band envelopes. Specifically, we designed narrow-band elliptic filters with a 40 dB stop band and 1 dB pass-band ripple as above, but cutoff frequencies in this case were linearly spaced at 300 Hz. We then computed the envelope by taking the absolute value of the Hilbert transform of this signal. We could then use these narrow bands to either compute modulation spectra or, in phase coherence analysis, identify the precise spectrotemporal structure that drove LFP responses and spiking activity.

Temporal modulation spectra. To compute a temporal modulation spectrum, we first computed the narrow-band envelopes as described above. We then subjected each narrow-band envelope to a Fourier analysis using the multitaper methods outlined below to compute the temporal structure of each of these envelopes. For every narrow spectral frequency band, this analysis provided us with power as a function of temporal frequency. Previous approaches for computing the modulation spectrum have used the two-dimensional Fourier transform of the spectrogram to get the modulation spectrum because it provides additional characterization of sounds such as direction of FM sweeps, etc. (Singh and Theunissen, 2003; Cohen et al., 2007).

Spectral analysis parameters for both neural and acoustic data. Spectral analysis was performed using multitaper Fourier methods as performed in several previous studies (Pesaran et al., 2002; Ghazanfar et al., 2007; Womelsdorf et al., 2007; Maier et al., 2008). All spectra were computed over 4 s windows moved over either the natural scene stimuli or the neural responses. The 4 s window provided enough cycles to ensure that we were not undersampling the structure in higher-frequency bands (Belitski et al., 2010). There was no overlap between successive windows, and they were therefore independent. This allowed us to average across windows to summarize spectra. We computed Fourier spectra over the

frequency range from 1 to 90 Hz. Spectra were estimated on this 4 s window using nine Slepian tapers. According to the taper–window relationship, $k = 2NW - 1$, where k is the number of tapers, W is the bandwidth, and N is the time window; we get a bandwidth of 1.25 Hz. The bandwidth of 1.25 Hz was sufficient for us to discern the peaks in both Fourier spectra and the phase coherence. Varying the number of tapers did not affect the results in any significant manner.

Stimulus related power in the LFP. We averaged the absolute value of the spectrum on a trial-by-trial basis to arrive at an estimate of the stimulus-related power. All comparisons were made relative to the baseline by expressing the power during the scene period as percentage enhancement relative to baseline. An 8 s period during the 10 s intertrial interval was used as the baseline for all comparisons. We chose not to use the 1 s period before and after sound stimulation because of our relatively large window for spectral analysis, which could potentially be contaminated by on and off responses. We used one-sample t tests to determine whether the enhancement above baseline was significantly different from zero at every frequency. A Bonferroni's correction was applied for multiple comparisons.

Phase coherence between neural response and wideband stimulus envelope. To compute the relationship between the dynamics of the stimulus and neural activity, we computed the phase coherence (Womelsdorf et al., 2007; Ghazanfar et al., 2008; Maier et al., 2008). We chose phase coherence as a method to relate the structure of natural acoustic scenes to the neural responses. Phase coherence offers considerable advantages to methods such as cross-correlation because it avoids some normalization problems (Jarvis and Mitra, 2001). We first computed the multitapered Fourier transform for each trial of the LFP/spiking response $[X(f)]$ and the multitapered Fourier spectrum for the wideband stimulus envelope $[Y(f)]$. We next computed the cross-spectrum between the neural response and the stimulus envelope defined as $X^*(f)Y(f)$, where * denotes the complex conjugate. The argument of this complex number, termed the cross-spectral phase angle, provides the difference in phase between the LFP/spiking response and that of the acoustic scene across trials. The circular mean of these phase angles is the phase coherence. When the phase difference angles are consistent, the phase coherence is high, and when they are randomly distributed, the phase coherence is low.

To compute the baseline level of the phase coherence, we used a shuffling method. We took the LFP (or spiking activity) response to one natural scene exemplar and computed the phase coherence between it and the wideband envelope from the two other exemplars from the same scene category. To illustrate, this would mean taking the savanna 01 LFP (or spiking activity) response and then computing the phase coherence between it and the savanna 02 sound envelope. We repeated this for all mismatched pairs. This ensured that, for a given site, we could compute a baseline phase coherence given the excitability of the site. We term this as the "shuffled phase coherence."

Phase coherence between spiking activity and the local field potential. We computed the phase coherence between the spiking activity and the corresponding LFP recorded from the same site to investigate how well the spikes phase-locked to different frequencies of the LFP. We also computed shuffled phase coherence between the LFP and the spiking activity by taking spikes in response to one natural scene and the LFP response to another natural scene of the same category. To test whether the spike–field phase coherence was enhanced above the baseline, we expressed the phase coherence during the scene period as a percentage enhancement relative to the baseline spike–field phase coherence. We then used one-sample t tests to determine whether the enhancement above baseline was significantly different from zero at every frequency. A Bonferroni's correction to an α of 0.05 was applied for multiple comparisons.

Phase coherence as a function of spectral frequency. To compute the phase coherence as a function of spectral frequency, we used the narrowband envelopes instead of using the wideband stimulus envelope. We then computed phase coherence between it and the narrow-band envelopes of the neural response. This gave us a two-dimensional matrix of phase coherence between stimulus and neural response as a function of both spectral and temporal frequency.

SDs and inference for phase coherence. For single cortical site exemplars, we computed errors for the phase coherence by using a bootstrap proce-

dure. We resampled the cross-spectral phase angle 1000 times and computed the bootstrap SE as the SD across all these 1000 repetitions. For all subsequent inferences, we compared the intact versus shuffled phase coherences. Within a stimulus category such as savanna, riverine forest, or rainforest, "intact" refers to congruent neural activity–stimulus pairs ($n = 3$; e.g., savanna 01 LFP with savanna 01 scene), whereas "shuffled" refers to all incongruent pairs ($n = 6$; e.g., savanna 01 LFP with savanna 02 or savanna 03 scene). We then used two-sample t tests to identify whether there were significant differences between the intact and shuffled coherences on a site-by-site basis for both spike-scene and LFP-scene phase coherence.

Results

We recorded spiking activity and LFP responses from 60 primary auditory cortical sites over 11 sessions in two macaque monkeys (29 in one monkey, 31 in the other) while they listened to multiple exemplars of 1-min-long natural auditory scenes. These scenes were recorded in three different habitats of Old World monkeys—savanna, riverine forest, and rainforest (Waser and Brown, 1986; Brown, 2003)—and three exemplars from each of these scenes were presented to our monkey subjects (five repetitions each/random order). We chose such scenes not because our captive monkey subjects had extensive experience with them (they did not) but because they contain the spectrotemporal complexity that would allow us to probe how cortical circuits represent complex natural scenes. After characterizing the statistics of these natural scenes, we then used spectral analysis and phase coherence to explore the relationship between (1) network activity (in the form of the LFP signals) and the natural scenes, (2) spiking activity and the natural scenes, and finally (3) between LFPs and spiking activity. For the purposes of pithiness, we only focus on one natural scene, the savanna scene, in the main text. Identical analyses for the other two scenes—riverine forest and rainforests—are presented as supplemental data (available at www.jneurosci.org as supplemental material) and referred to when appropriate.

Spectral and temporal modulations of scenes from three primate habitats

We first analyzed the three natural acoustic scenes in our stimulus set to identify their statistical properties. Figure 1*A* shows a 5 s waveform and spectrogram of a natural scene exemplar recorded in the savanna habitat. The fluctuations present in the waveform are attributable to biotic sources, such as birds, insects, amphibians, and other mammals, as well as abiotic sources. The dominant fluctuations, visible as tall peaks in the waveform (gray arrows), are attributable to insect chirps. Figure 1*B* shows the waveform and the corresponding spectrogram of a segment between 0.5 and 1 s of the waveform shown in Figure 1*A* (gray shaded box). Dominant spectral modulations found in these chirps are observed between 5 and 6 kHz (Fig. 1*B*, downward FM sweeps), with smaller modulations in the 8–10 kHz band. These spectral frequencies are consistent with previous characterizations of "biotic noise" attributable to orthoptera and cicadas in similar habitats (Slabbekoorn, 2004). This spectral structure in the 5–6 and 8–10 kHz regions also seems to possess a regular pulse train-like structure (Fig. 1*C*), similar to those observed for many insect species (Pollack, 2000). The interpulse interval in these vocalizations is ~150–170 ms, suggesting a temporal modulation frequency in the range of 5.5–7 Hz.

To examine this more quantitatively, we generated a temporal modulation spectrum by taking the raw sound waveform and then bandpass filtering the sound into narrow frequency bands in the frequency range from 0 to 10 kHz. Figure 1*D* shows the

temporal modulation spectrum computed over the whole 60 s segment of this savanna acoustic scene averaged over 15 bins of 4 s each. This temporal modulation spectrum confirms our visual analyses of the spectrogram, demonstrating that energy in the 5–6 kHz band and higher spectral frequencies are temporally modulated. Three selected frequencies labeled in Figure 1*D* are shown in Figure 1*E*. The temporal frequency spectra for the 5.85 and 7.35 kHz spectral frequency regions are multipeaked, with a dominant temporal modulation at ~6 Hz and other peaks extending up to 40–50 Hz. Such multiple peaks in the temporal spectrum may be partially attributable to the Fourier transform of a regular signal train having modulations at multiples of the dominant frequency, in our case ~6 Hz (Bracewell, 1999). To illustrate this further, supplemental Figure S1*A* (available at www.jneurosci.org as supplemental material) shows an artificial click train with a dominant frequency of 6 Hz, with its corresponding Fourier spectrum (supplemental Fig. S1*B*, available at www.jneurosci.org as supplemental material) also showing modulations at multiples of this dominant frequency. Figure 1*F* summarizes the global temporal frequency structure by summing across all the spectral frequencies. As expected, the temporal structure is dominated by the 5–6 kHz spectral frequency band and the dominant temporal modulation frequency seems to be ~6.3 Hz, which agrees well with our estimates from Figure 1*C*. Because these scenes were recorded at approximately the same time and in the same place and were dominated by insect vocalizations (for the other two savanna scene exemplars, see supplemental Fig. S2*A*,*B*, available at www.jneurosci.org as supplemental material), the spectrotemporal structure was almost identical across all three exemplars with spectral peaks at ~5–6 kHz and temporal peaks at ~6 Hz. Indeed, such consistency in spectral structure is a hallmark of primate acoustic environments in the wild (Slabbekoorn, 2004).

We performed similar analyses on the riverine forest and rainforest acoustic scenes (results are shown in supplemental Fig. S2*C–F,G–J*, respectively, available at www.jneurosci.org as supplemental material). Dominant temporal modulations in the riverine forest acoustic scene were at 25 Hz as a result of insect vocalizations. The rainforest acoustic scene had a 1/*f* temporal structure with dominant modulations at <10 Hz. These analyses suggest that natural scenes recorded directly from an actual habitat contain considerable temporal structure (particularly the savanna and riverine forest scenes), and such structure, at least for the lengths of the
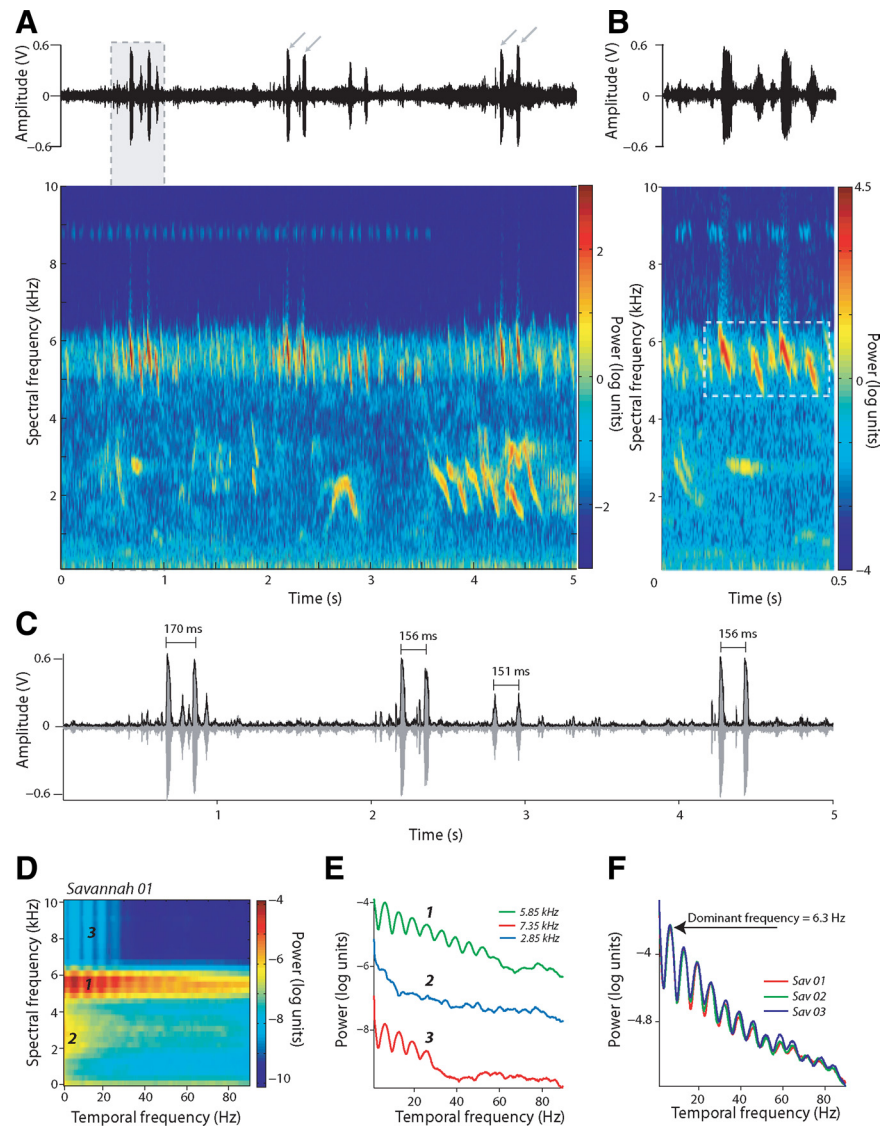
Figure 1. Spectrotemporal structure of the savanna acoustic scene. *A*, Top, A 5 s segment of the sound waveform from the savanna acoustic scene. *x*-axes depict time in seconds; *y*-axes depict the amplitude of the waveform in volts. Bottom, A spectrogram representation of the same sound shown in the top. *x*-axes depict time in seconds; *y*-axes depict spectral frequency in kilohertz. Color bar denotes power in log units. *B*, A segment between 0.5 and 1 s of the sound shown in *A* along with corresponding spectrogram. Figure conventions as in *A*. *C*, The waveform of the sound segment shown in *A* along with the envelope and hand-measured interpulse intervals. *x*-axes depict time in seconds; *y*-axes depict amplitude in volts. *D*, Temporal modulation spectrum of a savanna acoustic scene as a function of both temporal and spectral frequency. *x*-axes depict temporal frequency in hertz; *y*-axes depict spectral frequency in kilohertz. Color bar denotes the power in log units. *E*, Power in the modulation spectrum as a function of temporal frequency for three different selected spectral frequencies marked as 1, 2, and 3 in the modulation spectrum for the savanna acoustic scene (*C*): red, 7.35 kHz; green, 5.85 kHz; blue, 2.85 kHz. *x*-axes depict temporal frequency in hertz; *y*-axes depicts power in log units. *F*, Global temporal structure in the savanna acoustic scenes pooled by summing over all spectral frequencies for the three different savanna scenes. *x*-axes depict temporal frequency in hertz; *y*-axes depict log power. For statistics of the two other acoustic scenes, see also supplemental Figure S2 (available at www.jneurosci.org as supplemental material). For the statistics of periodic click trains, see supplemental Figure S1 (available at www.jneurosci.org as supplemental material).

scenes we analyzed here, seems similar to previous characterizations of natural sounds (Attias and Schreiner, 1997; Lewicki, 2002; Singh and Theunissen, 2003). We note, however, that the upper limit was ~25 Hz rather than the 5–10 Hz reported in previous studies. More importantly, because these temporal modulations also overlap with known structure of the EEG/LFP signal (Buzsáki and Draguhn, 2004; Buzsáki, 2006), we can now examine the effect of such scene structure on the LFP and spiking activity during the representation of such scenes.
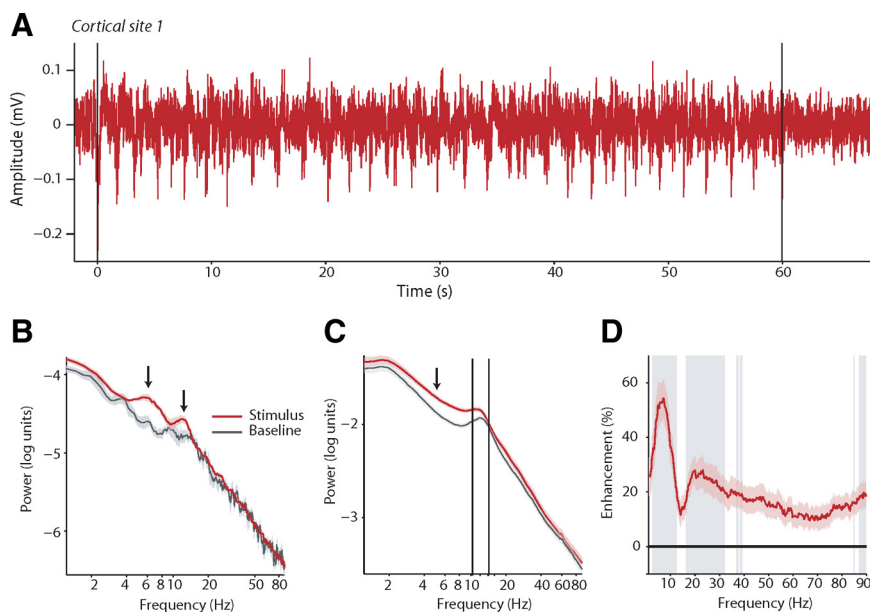
Figure 2. Single exemplars of LFP responses to natural savanna scenes. **A**, LFP response to a single savanna exemplar as a function of time for a single cortical site averaged over five repetitions. x-axes depict time in seconds; y-axes depict amplitude in millivolts. Black lines denote onset and offset of acoustic stimulation. **B**, Power spectrum of the LFP response shown in **A** during the presentation of the scene and during baseline. x-axes depict frequency in hertz; y-axes the power in logarithmic units. Error bars denote the SEM. **C**, Power spectrum of the LFP averaged over all cortical sites during the presentation of the savanna scenes and during baseline. Conventions as in **B**. Solid black lines are provided to guide the eye to a region of similarity between the stimulus-related LFP power spectrum and the baseline LFP power spectrum. **D**, Average enhancement of the power in the LFP over all cortical sites during the presentation of the savanna scenes relative to baseline. x-axes depicts frequency in hertz; y-axes depicts the enhancement in percentage. For additional details on the change in the LFP power for the riverine forest and rainforest acoustic scenes, see supplemental Figure S3 (available at www.jneurosci.org as supplemental material).

Power in the LFP is a function of the scene and ongoing network activity

We wanted to understand how the presentation of natural acoustic scenes modified the structure of the LFP in primary auditory cortex relative to its baseline state. We therefore compared the power spectrum of the LFP during the presentation of the natural scene with the power spectrum during the baseline period. We computed total stimulus-related power because it would include both phase-locked and non-phase-locked responses to the scene. Figure 2A shows the LFP from a single cortical site averaged over five repetitions in response to a savanna scene exemplar. The corresponding power spectrum during the stimulus and 8 s baseline periods is shown in Figure 2B. Stimulus onset evokes regular repetitive motifs in the LFP signal, and this is borne out in the spectra. Two peaks are clearly visible in the spectra during the stimulus period. The first peak is at ~6 Hz, which suggests that the stimulus induces power in this frequency band (paired $t_{(4)} = 6.33$, $p = 0.003$). The second peak was centered on 12 Hz, but this was not significantly different from the baseline (paired $t_{(4)} = 2.44$, $p = 0.07$). Similar exemplars for the riverine forest acoustic scene, showing peaks at 25 Hz in the spectra, are found in supplemental Figure S3A (available at www.jneurosci.org as supplemental material).

Across the population of cortical sites, the stimulus-related power in the LFP was enhanced relative to the baseline over a range of frequencies. Figure 2C plots the average spectrum of the LFP averaged over all cortical sites and savanna exemplars during the stimulus period and the baseline. There is a strong peak around the 11–16 Hz region in the baseline LFP that remains essentially unchanged during the stimulus period. In contrast, the presentation of the scene leads to an enhancement in power in

a lower-frequency range (2–11 Hz) and perhaps in a higher-frequency range (16–30 Hz). Figure 2D plots the percentage enhancement of the average power spectrum relative to the baseline across all cortical sites in response to the savanna acoustic scenes. Gray bars show the regions where the average enhancement in the power spectrum across all cortical sites for that particular frequency was significantly different from 0 relative to the baseline (one-sample t test relative to 0, $p < 0.05$). The largest enhancement was observed for the lower-frequency (2–11 Hz) region, and this was significantly different from 0 ($p < 0.05$); a second band of smaller enhancement was observed for the other frequencies >16 Hz but <30 Hz. Finally, only a weak modulation was observed beyond 30 Hz, and this was not significantly different from the baseline ($p > 0.05$). We repeated the analysis for the riverine forest and rainforest acoustic scenes and observed very similar results (supplemental Fig. S3A–C, available at www.jneurosci.org as supplemental material). However, there were differences between the exact frequencies that were modulated by each scene type (supplemental Fig. S3D, available at www. jneurosci.org as supplemental material). For the riverine forest scene, enhancements were observed in the 4–9 and 18–27 Hz regions; however, the region between 9 and 18 Hz was not modulated by the scene relative to the baseline. LFP frequencies at ~14 Hz were the least modulated relative to the baseline by the presentation of the scene. In response to the rainforest acoustic scene, LFP frequencies between 3–9 and 16–23 Hz were modulated by the scene, but LFP frequencies in the region between 9 and 16 Hz were not modulated by the scene, with the least modulated region relative to baseline again occurring at ~14 Hz.

To summarize, the power spectrum of the LFP in response to natural acoustic scenes involves two components. First, the presentation of the scene induces a reliable change in two frequency regions of the LFP. The first region is at ~2–9 Hz, and the second is at ~16–30 Hz. In addition, ongoing network activity in auditory cortex, as manifested by power in the 9–16 Hz region, was essentially unchanged relative to baseline. This suggests that a natural scene induces changes in the network dynamics at certain frequencies but leaves certain ongoing components of the activity unchanged.

Phase locking to natural scenes is a partial contributor to the change in LFP spectral structure

What is the source of such changes in the spectrum of the LFP in the frequency ranges between 2–9 and 16–30 Hz? One putative source is the temporal structure of the natural scene itself. Our analysis of the statistical structure of the three natural scenes suggested that there was considerable energy concentrated in the frequency range from 0 to 25 Hz. We therefore systematically analyzed the relationship between the temporal structure of the scene and the temporal structure of the LFP. Figure 3A shows a snippet between 10 and 15 s of the LFP response shown in Figure

2*A* along with the envelope of the savanna scene at the corresponding time points. There is a temporal correspondence between the LFP response and the envelope of the acoustic scene, perhaps a tracking of the temporal dynamics in the scene. To quantify the degree of tracking as well as the frequencies at which such tracking occurred, we computed the phase coherence between the raw LFP response and the wideband envelope of the sound as a function of temporal frequency. Phase coherence measures whether, on a trial-by-trial basis, a reliable phase relationship exists between two signals. The phase coherence allows us to identify modulation by the stimulus over a whole range of frequencies without being constrained by variations in power.

Because the three savanna scene exemplars had similar temporal modulation spectra (Fig. 1*C*,*E*) (supplemental Fig. S2*A*, available at www.jneurosci.org as supplemental material), we subdivided the different pairings of the LFP response–sound envelope into intact and shuffled conditions. In the intact condition, we computed coherence between corresponding pairs (e.g., savanna 01 LFP and savanna 01 envelope). For the shuffled conditions, we computed the phase coherence between noncorresponding pairs (e.g., savanna 01 LFP with savanna 02 envelope). The three intact phase coherences were essentially similar (supplemental Fig. S4, available at www.jneurosci.org as supplemental material) as were the shuffled coherence estimates. Figure 3*B* shows the average phase coherence for all intact pairs of savanna scene–LFP response and all shuffled pairs of savanna scene–LFP response for the cortical site shown in Figure 3*A*. We used average phase coherence in the frequency range of 5–7 Hz as a measure of the phase locking for the savanna scene. For this cortical site, average ± SD intact savanna scene–LFP phase coherence in the 5–7 Hz band was 0.93 ± 0.01 and significantly stronger than the average shuffled savanna scene–LFP phase coherence (0.47 ± 0.03; $t_{(7)}$ = 21.98, $p < 0.05$). The intact savanna scene–LFP coherence was enhanced over the shuffled savanna scene–LFP coherence by 95%. Figure 3*C*–*E* shows intact and shuffled phase coherence from three other cortical sites. Again, intact phase coherence was enhanced in the 5–7 Hz band over the shuffled phase coherence for all three exemplars (*t* tests, all $p < 0.05$). These results suggest that the stimulus modulations in the scene lead to a very robust phase locking in the LFP at the same frequency range.

Having observed the locking of LFP to the temporal dynamics of the stimulus, we next examined whether spiking activity also showed a similar pattern of locking to the temporal modulations of the acoustic scene. We focused on multiunit activity because phase coherence between spikes and continuous signals (such as the LFP or stimulus envelope) is more robust for multiunits than single units (Fries et al., 2001; Zeitler et al., 2006; Womelsdorf et al., 2007). Figure 4*A* shows the peristimulus time histogram from a multiunit cluster over five repetitions of a savanna scene. The response is smoothed with an adaptive 20 ms Gaussian window.
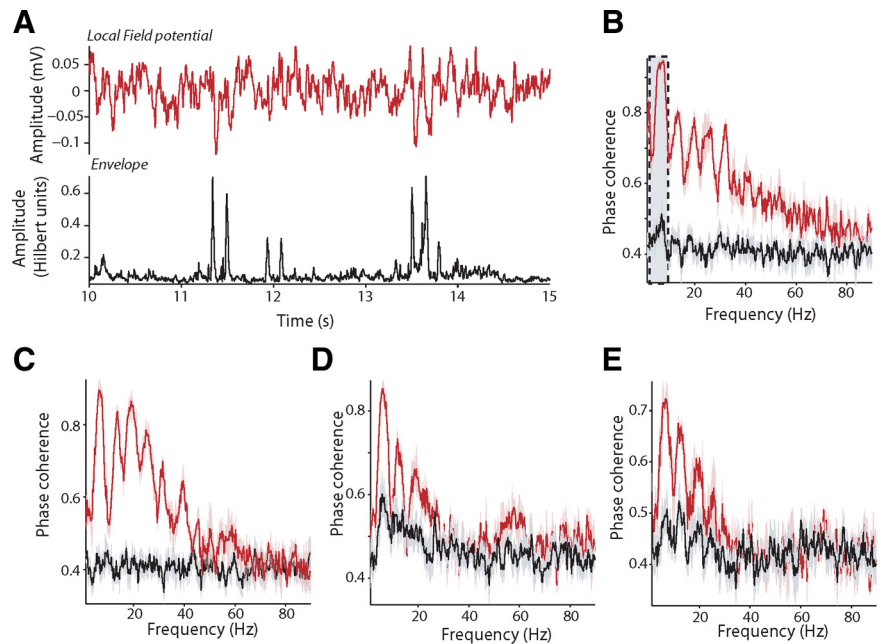
Figure 3. LFP locks to the temporal structure of the scene. *A*, Top, Expanded traces of the LFP response between 53 and 59 s from the cortical site shown in Figure 2*A*. x-axes depicts time in seconds; y-axes depicts amplitude in millivolts. Bottom, Wideband envelope of the savanna acoustic scene. x-axes depicts time in seconds; y-axes depicts Hilbert amplitude. *B*, Phase coherence between the LFP and the savanna scene averaged over all intact pairs (red line) and shuffled pairs (black line) for the cortical site shown in *A*. x-axes depicts frequency in hertz; y-axes depicts phase coherence. Error bars denote SEM over either three intact or six shuffled pairs. *C*–*E*, Phase coherence between the LFP and the savanna scene averaged over all intact pairs (red line) and shuffled pairs (black line) for three other cortical sites. Figure conventions as in *B*. For additional details on intact versus shuffled phase coherence, see supplemental Figure S4 (available at www.jneurosci.org as supplemental material).

Figure 4*B* shows an expanded version of the peristimulus time histogram in the −2 to 5 s region (Fig. 4*A*, gray box) along with the wideband envelope of the sound. Like LFP activity, multiunit spiking responses also track the envelope of the sound. Figure 4*C* shows the intact and shuffled phase coherence between the spiking activity and the modulations of the savanna scene for this cortical site. As in the LFP responses, spiking activity follows the temporal modulations in the 5–7 Hz range. The average intact phase coherence in the 5–7 Hz band (mean ± SD, 0.96 ± 0.01) and the average shuffled phase coherence in the 5–7 Hz band (mean ± SD, 0.45 ± 0.027) was significantly different ($t_{(7)}$ = 31.12, $p < 0.05$) for spiking activity in this cortical site. Figure 4*D*–*F* shows the phase coherence between the spiking activity and the envelope of the savanna acoustic scene for three other cortical sites, all of them displaying robust phase coherence between the savanna scene and the spiking activity at ~5–7 Hz ($p < 0.05$).

We next analyzed the phase coherence pattern for the population of cortical sites (*n* = 60) and all three natural scene types. Figure 5*A* shows the mean intact savanna–LFP phase coherence (averaged over the three exemplars) for each of the 60 cortical sites (thin gray lines) along with the population averages for intact savanna scene–LFP phase coherence (thick red line) and shuffled savanna scene–LFP phase coherence (thick black line). Several cortical sites show robust phase coherence between the LFP and the savanna scenes. The average intact savanna scene–LFP phase coherence for several sites again showed the peak in the 5–7 Hz region observed in the spectrotemporal analysis of the savanna acoustic scene (Fig. 1*D*,*F*). Figure 5*B* shows the average rank-ordered intact savanna scene–LFP phase coherence in the 5–7 Hz band for the population of cortical sites, along with the
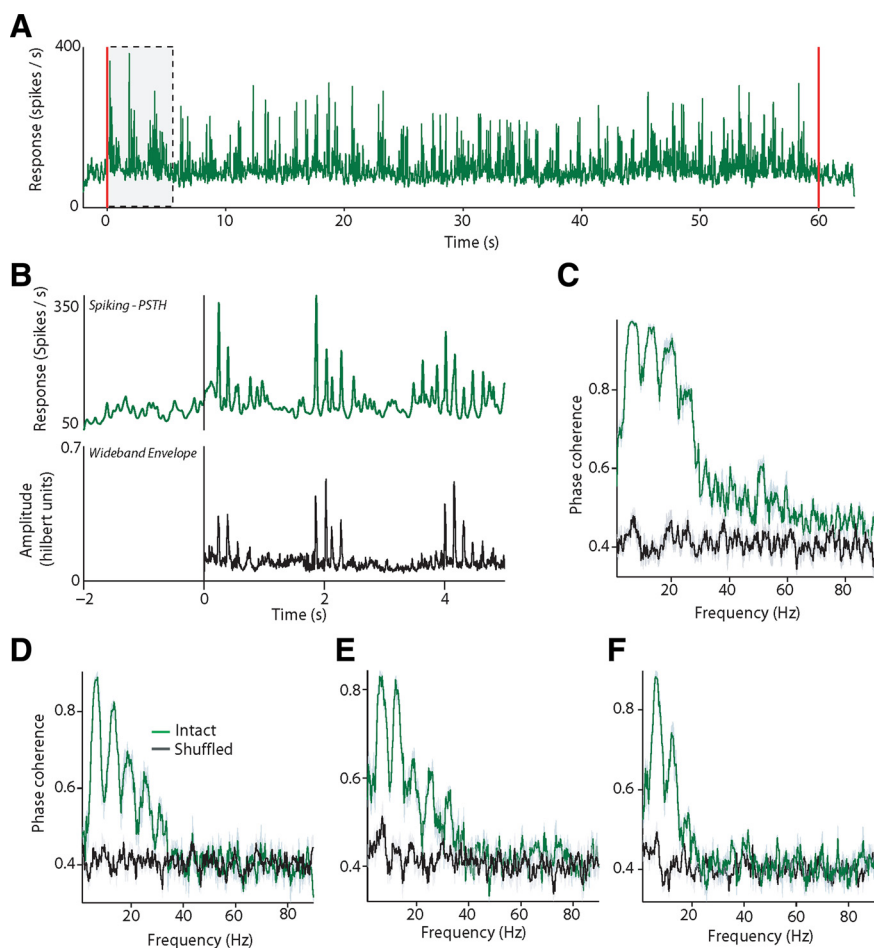
Figure 4. Single exemplars of spiking activity in response to natural savanna scenes. **A**, Peristimulus time histogram to a single savanna exemplar as a function of time for a single cortical site averaged over five repetitions. *x*-axes depicts time in seconds; *y*-axes depicts response in spikes per second. Red lines denote onset and offset of acoustic stimulation. **B**, Top, Expanded traces of the peristimulus time histogram (PSTH) between −2 and 5 s from the cortical site shown in **A**. *x*-axes depicts time in seconds; *y*-axes depicts amplitude in spikes per second. Bottom, Wideband envelope of the savanna acoustic scene during the same time period. *x*-axes depict time in seconds; *y*-axes depicts Hilbert transform amplitude. **C**, Phase coherence between the spiking activity and the savanna scene averaged over all intact pairs (red line) and shuffled pairs (black line) for the cortical site shown in **A** and **B**. *x*-axes depict frequency; *y*-axes depict phase coherence. Error bars denote SEM over either three intact or six shuffled pairs. **D–F**, Phase coherence between the spiking activity and the savanna scene averaged over all intact pairs (red line) and shuffled pairs (black line) for three other cortical sites. Figure conventions as in **C**.

corresponding shuffled coherence in the same frequency band. We used a two-sample *t* test to determine whether there were significant differences between intact and shuffled coherence for each cortical site. According to this criterion, 55% (33 of 60) of cortical sites showed significant differences between the intact savanna scene–LFP phase coherence compared with the shuffled savanna scene–LFP phase coherence. We repeated the same analysis for the riverine forest and rainforest acoustic scenes, and these are shown in supplemental Figure S5, *A,B* and *C,D* (available at www.jneurosci.org as supplemental material). Statistics are provided in Table 1. In particular, for the riverine forest acoustic scene, a clear peak was observed in the phase coherence at 25 Hz, a frequency that matches the dominant temporal modulation of the riverine forest acoustic scene (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). This provides additional support for the claim that the LFP response to a natural scene involves, at least in part, the phase locking of the LFP response to the temporal structure of the scene.

We repeated the population analysis for the multiunit spiking activity in response to the natural scenes. Figure 5*C* shows the

mean intact savanna scene–spike phase coherence as a function of frequency for the population of 60 cortical sites along with the average intact savanna scene–spike and shuffled savanna scene–spike phase coherence. Figure 5*D* shows the average intact and shuffled phase coherence in the 5–7 Hz band on a site-by-site basis. Twenty-nine percent (17 of 60) of cortical sites showed significantly larger intact savanna scene–spike phase coherence in the 5–7 Hz band compared with shuffled savanna scene–spike phase coherence. Corresponding analyses for the riverine forest and rainforest acoustic scenes are shown in supplemental Figure S5, *E,F* and *G,H* (available at www.jneurosci.org as supplemental material) along with statistical details in Table 2.

We note that phase coherence between LFP and the natural scenes did not always imply phase coherence between spiking activity and natural scenes. In fact, fewer sites showed coherence between spiking activity and the natural scenes compared with LFP responses. For example, whereas 55% of cortical sites showed larger intact phase coherence between the savanna scene and the LFP signal, only 29% did so for the spiking activity. Similar patterns were observed for the riverine forest and rainforest acoustic scenes (Tables 1, 2).

In summary, savanna scenes lead to a change in the spectral structure of the LFP responses with power in frequency regions between 2–9 and 16–30 Hz enhanced relative to baseline and with frequencies in the 9–16 Hz region not modulated relative to baseline. For the other two scenes, a very similar pattern of results was observed. Analysis of the phase coherence between the envelope of the scene and LFP and spiking responses suggests that the power modulation in these frequency ranges is at least in part attributable to an entrainment of the LFP response to the rhythmic structure in the scene. This entrainment of LFP and spiking responses was observed even in frequencies as high as 25 Hz and could well be higher.

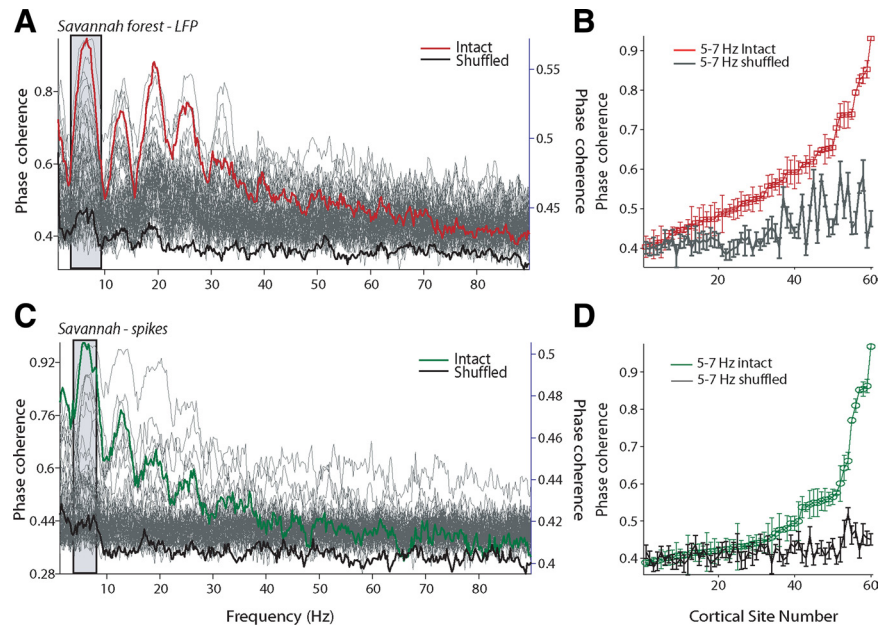## Relationship between neural activity and spectral structure of natural acoustic scenes

Thus far, we have focused on the relationship between temporal modulations of wideband envelopes of natural scenes with modulations in LFP and spiking activity. However, across the multiple exemplars of two of our scene categories (savanna and riverine forest), there was consistent spectral energy concentrated in the 5–6 kHz region, and this dominated the temporal envelope of the sound (Fig. 1) (supplemental Fig. S2*A,B*, available at www.jneurosci.org as supplemental material). To capture the influence of this spectral modulation, we bandpass filtered the stimulus into finer-grained spectral frequency bands and then recomputed the phase coherence on a band-by-band basis. Such an analysis allowed us to identify the precise contribution of spectral

and temporal structure driving both LFP and spiking responses. Figure 6*A* shows the mean savanna scene–LFP phase coherence for a single cortical site as a function of both the spectral and temporal modulation frequency. The LFP from this site locks robustly to the modulations in spectral frequency ranges >5 kHz, at ~5–7 Hz. Figure 6*B* shows another such example of robust locking of the LFP to the spectral modulations in the 5–6 kHz region. Figure 6*C* shows the proportion of cortical sites for the savanna scenes that showed enhanced phase coherence relative to shuffled phase coherence as a function of both spectral and temporal frequency for the savanna scene. Nearly 50% of sites showed robust phase locking in the LFP to savanna sounds in the spectral frequency range of 5–8 kHz and temporal frequency from 5 to 7 Hz. Results for the two other classes of acoustic scenes were similar. Up to 25% of cortical sites show robust following responses to the spectrotemporal modulations for the riverine forest and rainforest acoustic scenes (supplemental Fig. S6*A*,*B*, available at www.jneurosci.org as supplemental material). Similar results were observed for spiking activity (data not shown).

Figure 5. Phase coherence between neural activity and scene for the population of cortical sites. *A*, Average phase coherence between the LFP and savanna scene averaged over the population of cortical sites for intact (thick red lines) versus shuffled (thick black line). Thin gray traces represent individual sites for intact phase coherence measures. *x*-axes depict frequency in hertz; *y*-axes on the left depict phase coherence for the single exemplars, and *y*-axes on the right depict the phase coherence for the average. *B*, Average rank ordered intact and shuffled coherence between savanna scene and the LFP for specific frequency bands. *x*-axes depict cortical site number; *y*-axes depict phase coherence. Error bars denote SEM over either three intact or six shuffled pairs. *C*, Average phase coherence between the spikes and savanna scene averaged over the population of cortical sites for intact (thick green lines) versus shuffled (thick black line). Figure conventions as in *A*. *D*, Average rank-ordered intact and shuffled coherence between savanna scene and the spikes for specific frequency bands. Figure conventions as in *B*. For relationship of spikes and LFPs to other acoustic scenes, see supplemental Figure S5 (available at www.jneurosci.org as supplemental material).

Overlap between spectral selectivity of a cortical site and scene structure predicts phase locking of LFP and spiking activity

The previous analysis suggests that, even when the stimulus was partitioned into different spectral frequency bands, only a proportion of sites locked to these natural scenes. We attempted to identify whether there was any predictor that we could use to characterize cortical sites and thereby decide how interactions between spiking activity and LFPs would be mediated in response to natural scene dynamics. One obvious first-order predictor for how well an auditory cortical site may respond to the spectrotemporal modulations in a natural scene would be its spectral frequency selectivity (i.e., its tuning curve). For example, because both the savanna and riverine forest scenes contained dominant energy in the 5–6 kHz band, it is reasonable to hypothesize that those cortical sites that show spectral selectivity in this region would also show maximal phase coherence between the temporal modulations in natural scenes and LFP activity.

For every cortical site, spectral frequency tuning was estimated by using tone pips spaced logarithmically between 125 Hz and 20 kHz presented at different attenuation levels (for more details, see Materials and Methods). The relationship between spectral frequency tuning and phase coherence was then examined. Figure 7*A* shows robust phase coherence between the temporal modulations and both the LFP and spiking activity for the savanna scene. The phase coherence between the savanna scene and the LFP as a function of both spectral and temporal frequency is depicted in Figure 7*B*, and the spectral selectivity of the multiunit cluster is displayed in Figure 7*C*. The spectral frequency selectivity of this cortical site is in the 5–8 kHz region, which overlaps with the predominant spectral structure for the savanna acoustic scene (Fig. 1*C*). In contrast, Figure 7, *D* and *E*,

shows the same analysis for another such cortical site in response to the savanna acoustic scene but one that has spectral frequency selectivity in the low-frequency range, at ~1 kHz (Fig. 7*F*). This site shows little phase coherence. These exemplars suggest that the closer the spectral frequency selectivity of a cortical site is to the dominant spectral modulation in the natural scene, the stronger will be its phase coherence with the temporal modulations of the natural scene. To quantify this, we computed the mean phase coherence for all cortical sites with spectral selectivity in the 4–8 kHz region (overlapping with the dominant spectral content of the natural scenes) and those with spectral selectivity in the 0–4 kHz region.

If our premise that the overlap with spectral energy in the 5–6 kHz decides the degree of phase locking, then there should be significant differences between phase coherence estimates for these two sets of cortical sites. Indeed, that is what we found. For cortical sites with spectral selectivity in the 4–8 kHz region in response to the savanna scene, the mean ± SD phase coherence between the scene and LFP in the 5–7 Hz band was 0.65 ± 0.15. For the cortical sites with spectral selectivity in the 0–4 kHz region, the mean ± SD phase coherence in the 5–7 Hz band was 0.53 ± 0.09 ($t_{(55)} = 3.39, p < 0.05$). A similar pattern was seen for spiking activity. The mean ± SD phase coherence between the savanna scene and spiking activity in the 5–7 Hz band was 0.66 ± 0.19 for cortical sites with spectral selectivity in the 4–8 kHz region, whereas for cortical sites in the 0–4 kHz region, it was 0.45 ± 0.07 ($t_{(55)} = 5.96, p < 0.05$). We repeated the exact same analysis for the riverine forest acoustic scene. Comparing the cortical sites with spectral selectivity in the 4–8 kHz region versus those with spectral selectivity in the 0–4 kHz region revealed that

Table 1. Statistics on phase coherence between local field potentials and the wideband stimulus envelope for the population of cortical sites

	% sites	*t* statistic		% enhancement		Mean ± SD coherence	
		Minimum	Maximum	Minimum	Maximum	Intact	Shuffled
Savanna (5–7 Hz)	43	4.06	21.98	12.5	91	0.65 ± 0.11	0.46 ± 0.07
Riverine forest (23–27 Hz)	35	4.46	37.74	10	112	0.66 ± 0.11	0.42 ± 0.03
Rainforest (0–20 Hz)	45	4.2	33.86	5	73	0.53 ± 0.06	0.43 ± 0.04

% sites refers to the percentage of sites with significant differences between intact and shuffled phase coherence in a given frequency band. *t* statistic refers to the *t* value obtained by a two-sample *t* test between intact and shuffled phase coherence in relevant frequency bands for each scene. % enhancement refers to the enhancement of intact phase coherence over the shuffled phase coherence. Mean coherence refers to the mean across all cortical sites in the specified frequency band. $p < 0.005$ for all tests.

Table 2. Statistics on phase coherence between spiking activity and the wideband stimulus envelope for the population of cortical sites

	% sites	*t* statistic		% enhancement		Mean ± SD coherence	
		Minimum	Maximum	Minimum	Maximum	Intact	Shuffled
Savanna (5–7 Hz)	29	4.07	31.12	15	133	0.66 ± 0.16	0.44 ± 0.04
Riverine forest (23–27 Hz)	20	4.05	18.61	12	78	0.60 ± 0.12	0.42 ± 0.04
Riverine forest (0–10 Hz)	31	4.46	33.42	5	60	0.58 ± 0.12	0.42 ± 0.04
Rainforest (0–20 Hz)	33	6.58	34.90	9	62	0.55 ± 0.08	0.43 ± 0.02

Conventions as in Table 1.
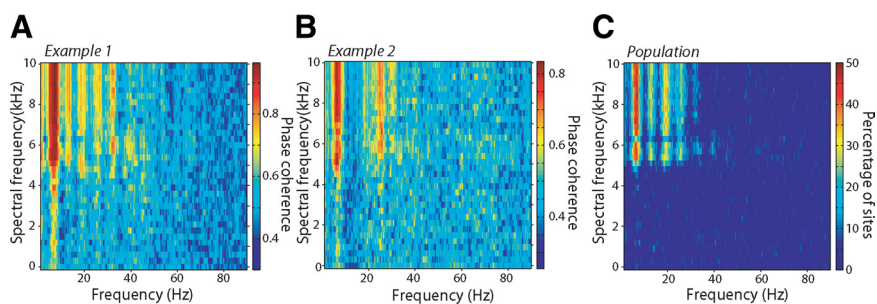
Figure 6. Phase coherence between LFP and natural scenes as a function of both spectral and temporal modulation frequency. *A*, Phase coherence between the LFP and narrowband envelopes of a savanna acoustic scene for a single cortical site as a function of both spectral and temporal frequency bands. *x*-axes depict frequency in hertz; *y*-axes depict spectral frequency in kilohertz. Color bar denotes phase coherence. *B*, Another example of a cortical site showing robust phase coherence in the 5–10 kHz band and for multiple temporal modulation frequencies. Figure conventions as in *A*. *C*, Percentage of cortical sites that showed enhanced phase coherence between savanna scenes and LFP as a function of both spectral and temporal frequency. *x*-axes depict temporal frequency in hertz; *y*-axes depict spectral frequency in kilohertz. Color bar denotes percentage of sites. For the same analysis for other acoustic scenes, see supplemental Figure S6 (available at www.jneurosci.org as supplemental material).

the mean ± SD phase coherence in the 23–27 Hz was 0.64 ± 0.15 for LFPs and 0.57 ± 0.17 for spiking activity in cortical sites with spectral selectivity in the 4–8 kHz region. In contrast, for the cortical sites with spectral selectivity in the 0–4 kHz region, the mean ± SD intact phase coherence in the 23–27 Hz band for LFPs was 0.5 ± 0.08 and 0.42 ± 0.05 for spiking activity. Two-sample *t* tests again revealed significantly enhanced phase coherence in the 23–27 Hz band for cortical sites with spectral selectivity in the 4–8 kHz region versus those in the 0–4 kHz region ($t_{(55)} = 3.86$, $p < 0.05$ for LFPs; $t_{(55)} = 4.83$, $p < 0.05$ for spikes). Thus, this analysis suggests that spectral selectivity of a site is at least a partial predictor of the degree to which it would follow temporal modulations of specific spectral frequencies.

Spike–field phase coherence in a given cortical site is related to its sensitivity to the natural scene

Our analysis thus far suggests that both the phase of LFP and spiking activity are locked to the temporal dynamics of the natural scenes and that overlap between the spectral structure of the scene and the spectral selectivity of the cortical site influences the degree of locking. Here we address the relationship between spiking activity and LFPs as a function of how well the LFP locks to

scene structure. We focused only on the savanna and riverine forest acoustic scenes because they contained consistent spectrotemporal structure that we could use as an index of how well a given cortical site locked to the modulations in the scenes.

Figure 8*A* shows three analyses: the phase coherence between the savanna scene and the LFP, the phase coherence between the savanna scene and spiking activity, and the phase coherence between the spiking activity and the LFP (spike field) in response to a savanna scene. All signals are from the same cortical site. We used the correlation coefficient between the savanna scene–LFP phase coherence and the spike–field coherence to index the degree to which these coherence functions were similar. The mean ± SD correlation coefficient between savanna–LFP phase coherence and spike–field phase coherence averaged over the three exemplars for this cortical site was 0.66 ± 0.1, suggesting that the spike–field phase coherence in the auditory cortical response to the savanna scene was primarily determined by the temporal modulations in the savanna scene. Figure 8*B* shows another example of a close relationship between spike–field phase coherence and savanna scene–LFP phase coherence. The mean ± SD correlation coefficient in this case was 0.81 ± 0.05.

For each cortical site, we computed the correlation coefficient between spike–field phase coherence and the savanna scene–LFP phase coherence for each of the exemplars, allowing us to estimate at the population level how well the savanna scene–LFP phase coherence predicts spike–field phase coherence. Figure 8*C* plots this relationship between the savanna scene–LFP phase coherence in the 0–20 Hz band and the correlation coefficient between spike–field phase coherence and the savanna scene–LFP phase coherence. The data suggest that, as the scene–LFP phase coherence increases, so does the similarity of spike–field phase coherence ($R^2 = 0.61$, $p < 0.05$). This suggests that the stronger the locking of the phase of the LFP to the savanna scene, the greater likelihood that the spiking activity will also lock to this

same phase. Thus, the temporal structure of the savanna scene may exert a common relationship on both the spiking activity and the LFP. In other words, the spike–field relationship is well explained by the savanna scene structure when sites lock strongly to the scene. Similar results were observed for the riverine forest acoustic scene (supplemental Fig. S7*A*, available at www.jneurosci.org as supplemental material). There was again a significant relationship between riverine forest scene–LFP phase coherence and its similarity to the spike–field phase coherence albeit smaller ($R^2 = 0.21$, $p < 0.05$).

The degree of temporal locking by the LFP to the scene was a reasonably good predictor of the shape of the spike–field coherence. However, even in cortical sites in which the phase locking by the LFP and spiking activity to the scene was weak, there was still robust spike–field coherence. Figure 9, *A* and *B*, shows two such cortical sites in which there was only weak locking to the acoustic scene for both spiking activity and LFPs, but there was still robust spike–field phase coherence. Such enhanced spike–field phase coherence could be a result of stimulus-related but not stimulus-locked processing or attributable to intrinsic circuit dynamics. Therefore, to identify the frequency range over which spikes locked to the LFP during natural scene processing, we computed the average spike–field coherence over all 60 cortical sites during the stimulus period and the baseline. Figure 9*C* shows the average spike–field coherence in response to the savanna scenes for the stimulus period and the baseline period. Spike–field coherence seems to be enhanced relative to the baseline for frequencies up to 30 Hz. Figure 9*D* plots the percentage enhancement of the spike–field coherence during the stimulus period relative to the baseline for the population of cortical sites. The most robust enhancement of spike–field phase coherence was in the lower-frequency range between 0 and 20 Hz and some regions between 20 and 30 Hz ($p <$ 0.05). Similar results were observed for the two other scene categories with upper limits of the phase coherence between spikes and the LFP again robust up to 30 Hz (supplemental Fig. S7, available at www.jneurosci. org as supplemental material).

Discussion

Our study examined the representation of natural acoustic scenes by LFPs and spiking activity in auditory cortical circuits. In our approach, we characterized the statistical structure of three types of acoustic scenes recorded in different primate habitats—savanna, riverine forest, and rainforest—and found that they con-
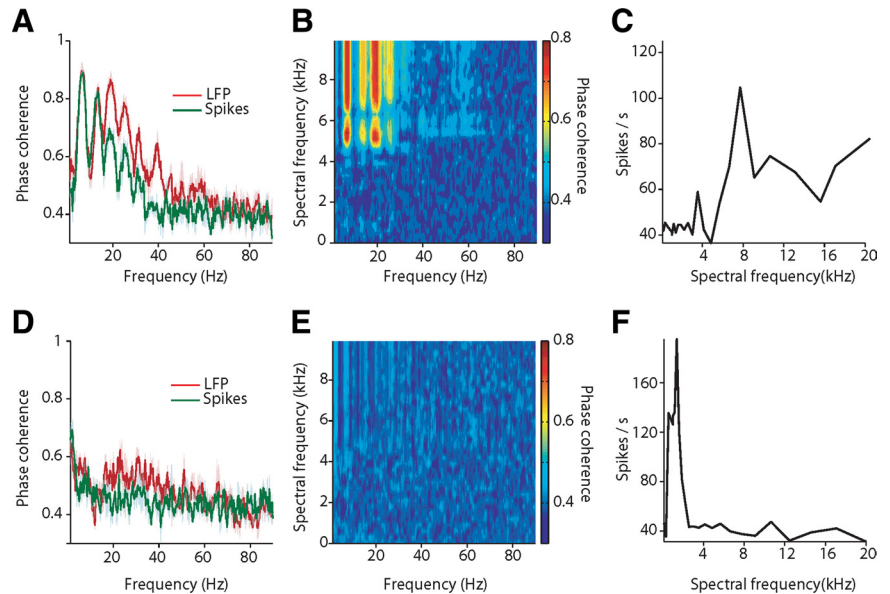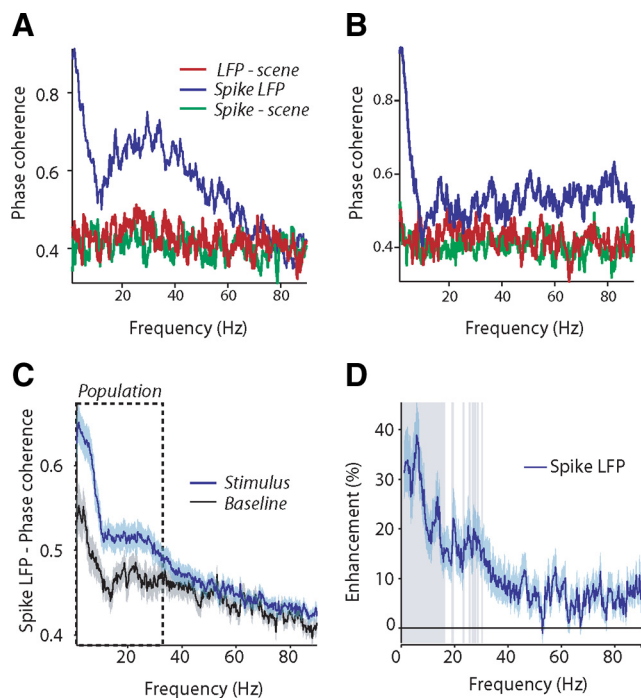

Figure 7. Phase coherence between LFP, spiking activity, and natural scenes as a function of the spectral selectivity of the cortical site. ***A***, Phase coherence between the wideband envelope of the natural scene and both LFP and spiking activity for a single cortical site. *x*-axes depict frequency in hertz; *y*-axes depict phase coherence. Note robust phase coherence between scene and neural activity for this site. ***B***, Phase coherence between the LFP and narrowband envelopes of the natural scenes for the cortical site shown in ***A*** as a function of both spectral and temporal modulation frequencies. *x*-axes depict frequency in hertz; *y*-axes depict spectral frequency in kilohertz. Color bar denotes phase coherence. ***C***, Spectral selectivity (tuning curve) for the cortical site for the cortical site shown in ***A*** and ***B***. *x*-axes depict spectral frequency in kilohertz; *y*-axes depict response in spikes per second. ***D***, Phase coherence between the wideband envelope of the natural scene and both LFP and spiking activity for another cortical site. Note the lack of locking to the savanna scene. ***E***, Phase coherence between the LFP and narrowband envelopes of the natural scenes for the cortical site shown in ***D*** as a function of both spectral and temporal modulation frequencies. Conventions as in ***B***. ***F***, Spectral selectivity (tuning curve) for the cortical site shown in ***D*** and ***E***. Conventions as in ***C***. Note the selectivity of this site is in the frequency range of 1 kHz.
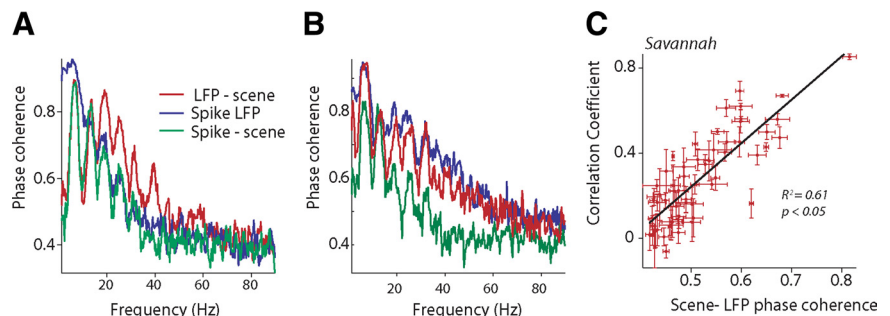
Figure 8. Phase coherence between spiking activity and LFPs show strong relationships to temporal modulations in natural scenes. ***A***, Phase coherence between the spikes and the savanna scene (green line), spikes and LFP (blue line), and LFP and savanna scene (red line) for a single cortical site. *x*-axes depict frequency in hertz; *y*-axes depict phase coherence. ***B***, Phase coherence between spikes and LFP, spikes and scene, and LFP and scene for another cortical site. Figure conventions as in ***A***. ***C***, Left, Plot of correlation coefficient between savanna scene–LFP phase coherence and spike–field phase coherence for the population of cortical sites. Horizontal error bar denotes the SE for phase coherence over the three intact pairs. Vertical error bars denote the SE over correlations performed for the three intact pairs. *x*-axes depict savanna scene–LFP phase coherence; *y*-axes denotes the correlation coefficient. For similar results with riverine forest acoustic scenes, see supplemental Figure S7*A* (available at www. jneurosci.org as supplemental material).

tain considerable temporal structure up to 25 Hz. We then analyzed the impact of scene structure on LFPs and spiking activity in the primary auditory cortex. These natural scenes modulate the power of the auditory LFP in two distinct regions, from 2 to 9 Hz and 16 to 30 Hz. In contrast, ongoing rhythmic dynamics of auditory cortex in the 9–16 Hz band were essentially unaffected by the scenes. Phase coherence showed that at least part of the changes in the power of the LFP during the scene were attrib-

Figure 9. Average spike LFP phase coherence for the population of cortical sites. *A*, Phase coherence between the spiking activity and a savanna scene (green line), spikes and LFP (blue line), and LFP and savanna scene (red line) for a cortical site showing no locking to the scene. *x*-axes depict frequency in hertz; *y*-axes depict phase coherence. *B*, Example of another cortical site that shows no locking to the scene but still displays coherence between spikes and the LFP. Figure conventions as in *A*. *C*, Average phase coherence between spikes and LFP for the population of cortical sites during the presentation of the scene and baseline. *x*-axes depict frequency in hertz; *y*-axes depict phase coherence between the spikes and the LFP. *D*, Percentage enhancement of the phase coherence between spikes and LFP for the population of cortical sites during the presentation of the scene relative to the baseline. *x*-axes depict frequency in hertz; *y*-axes depict percentage enhancement. For similar results with riverine forest and rainforest acoustic scenes, see supplemental Figure S7, *B* and *C* (available at www.jneurosci.org as supplemental material).

utable to phase locking of the LFP and spiking activity to the scene structure, with locking present even in frequencies as high as 25 Hz. A key predictor of phase locking was the overlap between the spectral structure of the scene and spectral selectivity of the cortical site. Finally, spike–field coherence analysis suggested that, in auditory cortex, at least for passive auditory scene processing, much of the locking of spikes to the LFP is restricted to frequencies <30 Hz.

Interaction between stimulus dynamics and ongoing activity in the representation of auditory scenes

We observed that the presentation of natural scenes led to enhanced power in the LFP between 2 and 9 Hz and between 16 and 30 Hz. In contrast, there was no change in the ongoing rhythmic activity in auditory cortex in the frequency ranges from 9 to 16 Hz. These observations that natural scene stimulation leads to enhanced power in low-frequency bands are partially consistent with a recent study of the encoding of natural movies in the visual cortex of anesthetized monkeys (Belitski et al., 2008). In this movie study, natural stimulation led to an enhancement of power in the frequencies from 1 to 8 Hz, with modulations in the 12–24 Hz band essentially uninformative. Perhaps the modulation we observe in this higher-frequency band from 16 to 30 Hz is because our scenes contained energy up to 30 Hz, whereas natural visual

movies usually contain energy only up to 10 Hz (Dong and Atick, 1995).

The 9–16 Hz rhythm, which we report as the ongoing dynamics of the auditory cortical circuit, is probably a result of the animal not being engaged in a task with these natural scenes (Fontanini and Katz, 2005, 2008). Many such rhythms in similar frequency ranges (~10 Hz) located in the motor cortex (mu rhythms), visual cortex (alpha rhythms), auditory cortex (tau rhythms), and gustatory cortex are associated with a "disengaged" state (Hari and Salmelin, 1997; Lehtela et al., 1997; Fontanini and Katz, 2005). The presence of this rhythm in both our baseline and stimulus periods suggests that, despite the presence of the natural scene (a stimulus), our monkey subjects remain disengaged. Thus, the interactions between natural scene structure and cortical activity that we observe may be taking place entirely in this "tau" behavioral state. Investigating how the LFP and spiking representations of these scenes change while the animal is actively engaged with the scene is the next logical step, as demonstrated in rats and monkeys using simple stimuli (Ryan et al., 1984; Otazu et al., 2009).

We observed only a very weak modulation of the gamma-band power relative to the baseline, and the enhancement was not significant. Previous studies of the LFP in response to sensory stimulation have emphasized the gamma band as important for encoding sensory stimuli (Henrie and Shapley, 2005; Liu and Newsome, 2006) and natural visual movies (Belitski et al., 2008). It is unclear whether the differences observed are attributable to our recordings being performed in awake animals with a larger variety of scenes or in a different cortical structure (auditory vs visual cortex). Observations similar to ours were made in a recent study comparing the encoding of natural auditory and visual scenes in auditory and visual cortices (Belitski et al., 2010). In that study, strong modulations were observed in the gamma band for encoding visual movies but weak modulations in the gamma band in auditory cortex for encoding auditory scenes.

Phase relationships between LFPs, spikes, and naturalistic stimuli

In response to natural acoustic scenes, depending on the scene and the cortical site, the temporal dynamics of the scene can modulate the phase of the LFP and spiking activity up to 25 Hz. These results confirm speculations that LFPs would lock to the temporal modulations present in natural stimuli, such as speech (Lakatos et al., 2005; Schroeder et al., 2008), and visual movies (Belitski et al., 2008, 2010). For example, in a recent study of the primary auditory cortex, the phase of ongoing delta oscillations locked to the onset of auditory stimulation, which was delivered at a delta rate (Lakatos et al., 2005). A study of the responses in primary visual cortex in an anesthetized monkey to natural movies found that phase of LFP was reliably modulated up to 12 Hz (Montemurro et al., 2008). In a similarly designed study of auditory cortex, Kayser et al. (2009) found a reliable modulation of the phase of the LFP up to 30 Hz. Neither of these studies however related the stimulus structure to the LFP. Our approach investigating the responses of primary auditory cortex to three different categories of natural acoustic scenes extends and confirms that LFP phase is indeed reliably modulated by the temporal dynamics of the naturalistic stimulus at multiple timescales (i.e., in multiple frequency bands).

Our results suggest that the entrainment of LFPs at several frequencies to the temporal structure of the stimulus is a partial contributor to the change in the power spectrum of the LFP during natural scene stimulation. These results support theoret-

ical predictions positing that a cortical neuronal assembly composed of excitatory and inhibitory neurons would encode the slow dynamic features (<8 Hz) of naturalistic stimuli into LFP fluctuations in slow frequencies (Mazzoni et al., 2008). That is, oscillations in the low-frequency range of the stimulus are reproduced in the LFP, causing a peak in the spectrum at exactly the input frequency but with little effect on the rest of the LFP spectrum. Our data are consistent with this idea but extend it two ways. First, there are reliable power modulations in low-frequency bands even when there is no energy in the stimulus in the same frequency band. Second, entrainment of LFPs can occur even in frequencies as high as 25 Hz, and the true upper limit may well be higher. Indeed, as studies of the auditory steady state response suggest, the upper limit of entrainment in auditory cortex may well be as high as 80 Hz (Ross et al., 2000).

We have emphasized first-order interactions between scene dynamics and the LFP. Other phenomena, beyond entrainment, may influence network activity. These include spectral mixing (Ahrens et al., 2002), the resetting of an ongoing oscillation (Rizzuto et al., 2003; Ghazanfar and Chandrasekaran, 2007), or the emergence of additional synchronized and stimulus-evoked activity (Ross et al., 2000; Shah et al., 2004). Addressing these issues would provide more insight into how cortical networks operate in real natural environments and thereby guide adaptive behavior (Schroeder et al., 2008; Schroeder and Lakatos, 2009).

We show that spiking activity in auditory cortex is related to the phase of the LFP up to 25 Hz. Coherence between spiking activity and the LFP has been demonstrated before, but such demonstrations have been confined primarily to the gamma band (Fries et al., 2001; Womelsdorf et al., 2006), with these studies usually demonstrating suppressed spike–field coherence relative to baseline at low frequencies. What is the significance of such locking of spiking activity at low frequencies seen in our data? One potential use for such locking of spikes to the phase of the LFP, at least in sensory cortices, is to build a temporal code, in which spiking activity could exploit the LFP as an internal reference frame, the so-called phase-of-firing code (Montemurro et al., 2008; Panzeri et al., 2010). For example, in the hippocampus, the timing of spikes relative to the phase of theta band (4–8 Hz) oscillations carries information about the position and heading in space of the animal (Huxter et al., 2008). Similarly, in auditory and visual cortices, the phase of firing with respect to low-frequency (1–8 Hz) LFPs carries a significant amount of information about complex naturalistic visual and auditory stimuli (Montemurro et al., 2008; Kayser et al., 2009). Our results suggest that even LFP frequencies up to 30 Hz might be informative and serve as a reference frame for encoding sensory stimuli. Additional use of reverse correlation methods, natural acoustic scenes, and information theoretic methods would help elucidate the role of the LFP and spike–field relationships in the encoding of natural scenes (Tiesinga et al., 2008; Quian Quiroga and Panzeri, 2009; Panzeri et al., 2010).

Distributed place and stimulus-locked representation of natural acoustic scenes

We observed that spiking activity in auditory cortex can lock to the temporal structure of natural scenes, a result reminiscent of previous research on the responses of primary auditory cortex to simple stimuli (Eggermont and Smith, 1995; Bieser and Müller-Preuss, 1996; Wang et al., 2003; Joris et al., 2004) and vocalizations (Wang et al., 1995; Nagarajan et al., 2002). Our results extend such observations and suggest that such mechanisms operate in response to spectrotemporally complex and temporally

extended natural scenes and do so for both LFPs and spikes. In addition, this stimulus-locked mechanism operates only for certain cortical sites, only those in which the spectral selectivity overlapped with the spectral structure of the stimulus. This suggests that primary auditory cortex uses a dual place and temporal representation for the encoding of natural stimuli, whereby the encoding of complex stimuli is represented by the discharge patterns of distributed neuronal populations (Creutzfeldt et al., 1980; Wang et al., 1995; Gehr et al., 2000; Rotman et al., 2001; Nagarajan et al., 2002). The best support for this "distributed encoding" model comes from a study showing that the spectro-temporal discharge patterns of spatially distributed populations in the anesthetized marmoset primary auditory cortex were correlated with the spectrotemporal components of conspecific vocalizations (Wang et al., 1995). Similar patterns were seen for cat auditory cortex (Gehr et al., 2000; Rotman et al., 2001). Our results suggest that this mechanism operates also for natural acoustic scenes. That is, at the level of primary auditory cortex, neurons are responding to the spectrotemporal contents of the acoustic milieu and could thus provide a representation for additional auditory scene analysis (Elhilali et al., 2009).

References

Ahrens KF, Levine H, Suhl H, Kleinfeld D (2002) Spectral mixing of rhythmic neuronal signals in sensory cortex. Proc Natl Acad Sci U S A 100:15176–15181.

Attias H, Schreiner CE (1997) Temporal low-order statistics of natural sounds. In: Neural information processing systems (Mozer MC, Jordan MI, Petsche T, eds), pp 27–34. Vancouver: Massachusetts Institute of Technology.

Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S (2008) Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. J Neurosci 28:5696–5709.

Belitski A, Panzeri S, Magri C, Logothetis NK, Kayser C (2010) Sensory information in local field potentials and spikes from visual and auditory cortices: time scales and frequency bands. J Comput Neurosci. Advance online publication. Retrieved September 11, 2010. doi:10.1007/s10827-010-0230-y.

Bieser A, Müller-Preuss P (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. Exp Brain Res 108:273–284.

Bracewell R (1999) The Fourier transform and its applications, Ed 3. Boston: McGraw-Hill.

Brosch M, Budinger E, Scheich H (2002) Stimulus-related gamma oscillations in primate auditory cortex. J Neurophysiol 87:2715–2725.

Brown CH (2003) Ecological and physiological constraints for primate vocal communication. In: Primate audition: ethology and neurobiology (Ghazanfar AA, ed), pp 127–150. Boca Raton, FL: CRC.

Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. In: Advances in the study of behavior (Slater PJB, Snowdon CT, Roper TJ, Brockmann JH, Naguib M, eds), pp 151–209. San Diego: Academic.

Buzsáki G (2006) Rhythms of the brain. New York: Oxford UP.

Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. Science 304:1926–1929.

Cohen YE, Theunissen F, Russ BE, Gill P (2007) Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. J Neurophysiol 97:1470–1484.

Creutzfeldt O, Hellweg FC, Schreiner C (1980) Thalamocortical transformation of responses to complex auditory stimuli. Exp Brain Res 39:87–104.

Dong DW, Atick JJ (1995) Statistics of natural time-varying images. Network 6:345–358.

Eggermont JJ, Smith GM (1995) Synchrony between single-unit activity and local field potentials in relation to periodicity coding in primary auditory cortex. J Neurophysiol 73:227–245.

Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA (2009) Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61:317–329.

Fontanini A, Katz DB (2005) 7 to 12 Hz activity in rat gustatory cortex

reflects disengagement from a fluid self-administration task. J Neurophysiol 93:2832–2840.

Fontanini A, Katz DB (2008) Behavioral states, network states, and sensory response variability. J Neurophysiol 100:1160–1168.

Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291:1560–1563.

Gehr DD, Komiya H, Eggermont JJ (2000) Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls. Hear Res 150:27–42.

Ghazanfar AA, Chandrasekaran CF (2007) Paving the way forward: integrating the senses through phase-resetting of cortical oscillations. Neuron 53:162–164.

Ghazanfar AA, Turesson HK, Maier JX, van Dinther R, Patterson RD, Logothetis NK (2007) Vocal-tract resonances as indexical cues in rhesus monkeys. Curr Biol 17:425–430.

Ghazanfar AA, Chandrasekaran C, Logothetis NK (2008) Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. J Neurosci 28:4457–4469.

Hackett TA, Stepniewska I, Kaas JH (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 394:475–495.

Hari R, Salmelin R (1997) Human cortical oscillations: a neuromagnetic view through the skull. Trends Neurosci 20:44–49.

Henrie JA, Shapley R (2005) LFP power spectra in V1 cortex: the graded effect of stimulus contrast. J Neurophysiol 94:479–490.

Huxter JR, Senior TJ, Allen K, Csicsvari J (2008) Theta phase-specific codes for two-dimensional position, trajectory and heading in the hippocampus. Nat Neurosci 11:587–594.

Jarvis MR, Mitra PP (2001) Sampling properties of the spectrum and coherency of sequences of action potentials. Neural Comput 13:717–749.

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577.

Juergens E, Guettler A, Eckhorn R (1999) Visual stimulation elicits locked and induced gamma oscillations in monkey intracortical- and EEG-potentials, but not in human EEG. Exp Brain Res 129:247–259.

Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron 61:597–608.

Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94:1904–1911.

Lehtela L, Salmelin R, Hari R (1997) Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. Neurosci Lett 222:111–114.

Lewicki MS (2002) Efficient coding of natural sounds. Nat Neurosci 5:356–363.

Liu J, Newsome WT (2006) Local field potential in cortical area MT: stimulus tuning and behavioral correlations. J Neurosci 26:7779–7790.

Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 23:3963–3971.

Maier JX, Chandrasekaran C, Ghazanfar AA (2008) Integration of bimodal looming signals through neuronal coherence in the temporal lobe. Curr Biol 18:963–968.

Mazzoni A, Panzeri S, Logothetis NK, Brunel N (2008) Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. PLoS Comput Biol 4:e1000239.

Mitzdorf U (1987) Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. Int J Neurosci 33:33–59.

Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S (2008) Phase-of-firing coding of natural visual stimuli in primary visual cortex. Curr Biol 18:375–380.

Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM (2002) Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. J Neurophysiol 87:1723–1737.

Nauhaus I, Busse L, Carandini M, Ringach DL (2009) Stimulus contrast modulates functional connectivity in visual cortex. Nat Neurosci 12:70–76.

Nelken I, Rotman Y, Bar Yosef O (1999) Responses of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157.

Nelson MJ, Pouget P, Nilsen EA, Patten CD, Schall JD (2008) Review of signal distortion through metal microelectrode recording circuits and filters. J Neurosci Methods 169:141–157.

Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. Nat Neurosci 12:646–654.

Panzeri S, Brunel N, Logothetis NK, Kayser C (2010) Sensory neural codes using multiplexed temporal scales. Trends Neurosci 33:111–120.

Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA (2002) Temporal structure in neuronal activity during working memory in macaque parietal cortex. Nat Neurosci 5:805–811.

Pfingst BE, O'Connor TA (1980) A vertical stereotaxic approach to auditory cortex in the unanesthetized monkey. J Neurosci Methods 2:33–45.

Pollack G (2000) Who, what, where? Recognition and localization of acoustic signals by insects. Curr Opin Neurobiol 10:763–767.

Quian Quiroga R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 10:173–185.

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 16:1661–1687.

Recanzone GH, Guard DC, Phan ML (2000) Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J Neurophysiol 83:2315–2331.

Rizzuto DS, Madsen JR, Bromfield EB, Schulze-Bonhage A, Seelig D, Aschenbrenner-Scheibe R, Kahana MJ (2003) Reset of human neocortical oscillations during a working memory task. Proc Natl Acad Sci U S A 100:7931–7936.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. J Acoust Soc Am 108:679–691.

Rotman Y, Bar-Yosef O, Nelken I (2001) Relating cluster and population responses to natural sounds and tonal stimuli in cat primary auditory cortex. Hear Res 152:110–127.

Ryan AF, Miller JM, Pfingst BE, Martin GK (1984) Effects of reaction time performance on single-unit activity in the central auditory pathway of the rhesus macaque. J Neurosci 4:298–308.

Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci 32:9–18.

Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A (2008) Neuronal oscillations and visual amplification of speech. Trends Cogn Sci 12:106–113.

Shah AS, Bressler SL, Knuth KH, Ding M, Mehta AD, Ulbert I, Schroeder CE (2004) Neural dynamics and the fundamental mechanisms of event-related brain potentials. Cereb Cortex 14:476–483.

Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114:3394–3411.

Slabbekoorn H (2004) Habitat-dependent ambient noise: consistent spectral profiles in two African forest types. J Acoust Soc Am 116:3727–3733.

Tiesinga P, Fellous JM, Sejnowski TJ (2008) Regulation of spike timing in visual cortical circuits. Nat Rev Neurosci 9:97–107.

Wang X, Merzenich MM, Beitel R, Schreiner CE (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol 74:2685–2706.

Wang X, Lu T, Liang L (2003) Cortical processing of temporal modulations. Speech Commun 41:107–121.

Waser PM, Brown CH (1986) Habitat acoustics and primate communication. Am J Primatol 10:135–154.

Womelsdorf T, Fries P, Mitra PP, Desimone R (2006) Gamma-band synchronization in visual cortex predicts speed of change detection. Nature 439:733–736.

Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P (2007) Modulation of neuronal interactions through neuronal synchronization. Science 316:1609–1612.

Zeitler M, Fries P, Gielen S (2006) Assessing neuronal coherence with single-unit, multi-unit, and local field potentials. Neural Comput 18:2256–2281.