A place for time: the spatiotemporal structure of neural dynamics during natural audition

Greg J. Stephens,[1,2] Christopher J. Honey,[3] and Uri Hasson[3]

[1]*Department of Physics and Astronomy, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands;* [2]*Okinawa Institute of Science and Technology, Tancha, Onna-son, Okinawa, Japan; and* [3]*Department of Psychology and Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey*

Stephens GJ, Honey CJ, Hasson U. A place for time: the spatiotemporal structure of neural dynamics during natural audition. *J Neurophysiol* 110: 2019–2026, 2013. First published August 7, 2013; doi:10.1152/jn.00268.2013.—We use functional magnetic resonance imaging (fMRI) to analyze neural responses to natural auditory stimuli. We characterize the fMRI time series through the shape of the voxel power spectrum and find that the timescales of neural dynamics vary along a spatial gradient, with faster dynamics in early auditory cortex and slower dynamics in higher order brain regions. The timescale gradient is observed through the unsupervised clustering of the power spectra of individual brains, both in the presence and absence of a stimulus, and is enhanced in the stimulus-locked component that is shared across listeners. Moreover, intrinsically faster dynamics occur in areas that respond preferentially to momentary stimulus features, while the intrinsically slower dynamics occur in areas that integrate stimulus information over longer timescales. These observations connect the timescales of intrinsic neural dynamics to the timescales of information processing, suggesting a temporal organizing principle for neural computation across the cerebral cortex.

fMRI; natural addition; natural cognition; neural dynamics

TO SUCCESSFULLY PERCEIVE AND act within the natural world, we must process sensory information and organize our behavior over a wide range of timescales (Cole 1995; Dong and Atick 1995; Voss and Clarke 1975). In spoken language for example, spectral features are composed into phonemes (tens of milliseconds), phonemes are connected into words (hundreds of milliseconds), words into sentences (few seconds), and sentences into paragraphs (tens of seconds) and paragraphs into increasingly longer constructions. Neural dynamics with timescales across many orders of magnitude are also observed across numerous recording methodologies, input stimuli, and organisms (Baria et al. 2011; He et al. 2010; Honey et al. 2012; Lerner et al. 2011; Linkenkaer-Hansen et al. 2001; Schneidman et al. 2006). However, what is the relationship between the timescales on which information unfolds in the world and the timescales of the dynamics in neural systems?

Prior work using functional magnetic resonance imaging (fMRI) and electocorticography (ECoG) suggests that brain regions process information over different timescales and that this processing is organized hierarchically (Hasson et al. 2008; Honey et al. 2012; Lerner et al. 2011). Regions at lower levels of the hierarchy integrate sensory information (e.g., word acoustics) over short periods of time, while higher order regions integrate information (e.g., the meaning conveyed by a set of sentences or paragraphs) over seconds and minutes. These findings support a distributed and hierarchical model of information integration, in which the temporal integration capacity increases gradually from early to higher order areas. An intriguing hypothesis is that the timescales of neural dynamics of a given brain area reflect the timescales on which that region typically integrates information. If so, then neural dynamics should vary according to a well-defined topographic map as a function of the information processing timescales.

To map the timescale of neural dynamics, prior studies have used fMRI and ECoG to measure the relative proportion of slow (<0.1 Hz) and fast fluctuations of neural activity in different regions. Differences in low-frequency content of the blood oxygen level-dependent (BOLD) signal have been reported across sets of predefined brain networks (He 2011; Wu et al. 2008) as well as in the coupling between networks (Salvador et al. 2008). The timescales of resting BOLD dynamics have also been explored at the voxel level (Baliki et al. 2011; Baria et al. 2011; Yan et al. 2009) with chronic pain sufferers exhibiting unusually fast dynamics within medial prefrontal, posterior cingulate, and lateral parietal cortices (Baliki et al. 2011). However, these studies did not examine the relationship between timescales of neural dynamics and the timescales of information processing.

Using ECoG we recently explored the relationship between timescales of neural dynamics and timescales of stimulus dynamics within lateral portions of the left hemisphere of the human cerebral cortex (Honey et al. 2012). Prior LFP studies had demonstrated the neurophysiological reality of slow timescale fluctuations in neuronal activity (Leopold et al. 2003) and EEG work had shown the relevance of slow fluctuations to behavior (Monto et al. 2008). In our ECoG study we observed, first, that higher order regions of the cerebral cortex contain proportionally more slow fluctuations of neuronal activity than early auditory and visual cortices. Second, we observed that slow (<0.1 Hz) fluctuations of neuronal activity were reliably modulated by an audiovisual movie. Finally, we observed that the reproducibility of slow fluctuations in high-order areas was reduced when the long timescale content of the stimulus was disrupted by temporally scrambling the movie. These ECoG data suggested a connection between slow neuronal population dynamics and temporally coherent information structure over long timescales (e.g., connected series of events within the narrative).

While the high temporal resolution of ECoG allowed us to precisely characterize the neural dynamics near each electrode,

Address for reprint requests and other correspondence: G. J. Stephens, Dept. of Physics, Faculty of Exact Sciences, Vrije Universitiet Amsterdam, 1081 De Boelelaan, 1081 HV Amsterdam, The Netherlands (e-mail: g.j.stephens @vu.nl).

the sparse coverage of the electrode arrays precluded an understanding of the topography of neural dynamics across all cortical areas. Here, we used fMRI to study neural dynamics across the entire cortex and asked: *1*) whether a hierarchical topography of neural dynamical timescales, ranging from fast dynamic in sensory areas to slow dynamic in high-order areas, can be seen using the BOLD signal; *2*) whether this large-scale topography is stable across conditions of rest and auditory narrative processing; and *3*) whether slower (faster) dynamics correspond to an area accumulating information over longer (shorter) timescales.

METHODS

Subjects. Nineteen subjects (ages 20–36 yr) participated in the fMRI study. Eleven subjects were scanned while listening to the story and eight subjects were scanned in the resting state condition. The story data were used by Lerner et al. (2011) to characterize the hierarchy of temporal receptive windows (TRW). In the current study, we characterize the power spectrum of each voxel time series in the same data set, and relate this spectrum to *1*) the spectrum observed in a separate resting state data set and *2*) TRW indices computed in the same dataset. Conditions in which the head motion was >1 mm or in which the BOLD signal was corrupted were discarded from the analysis. Additional subjects were scanned till data from 11 subjects were collected for each condition. All procedures were approved by the Princeton University Committee on Activities Involving Human Subjects, and all subjects had normal hearing and provided written informed consent.

MRI acquisition. Subjects were scanned in a 3T head-only MRI scanner (Allegra; Siemens, Erlangen, Germany). A custom radio frequency coil was used for the structural scans (NM-011 transmit head coil; Nova Medical, Wakefield, MA). For fMRI scans, 300 volumes were acquired using a T2*-weighted echo planar imaging pulse sequence [repetition time (TR) = 1,500 ms; echo time (TE) = 30 ms; flip angle = 75°], each volume comprising 25 slices of 3-mm thickness with 1-mm gap (in-plane resolution = 3×3 mm^2). Slice acquisition order was interleaved. In addition, a set of 160 T1-weighted high-resolution ($1 \times 1 \times 1$ mm^3) anatomical images of the same orientation as the EPI slices were acquired for each subject with a magnetization-prepared rapid-acquisition gradient echo (MP-RAGE) pulse sequence (TR = 2,500 ms, TE = 4 ms, slice thickness = 1 mm, no gap, in-plane resolution = 1×1 mm^2, field of view = 256 mm^2) and used for cortical segmentation and three-dimensional (3D) reconstruction. To minimize head movement, subjects' heads were stabilized with foam padding. Stimuli were presented using Psychophysics toolbox. MRI-compatible headphones (MR Confon; Magdeburg, Germany) were fitted to provide considerable attenuation to the scanner noise and to present the audio stimuli to the subjects.

Stimuli and experimental design. Stimuli for the experiment were generated from a 7-min real life story ("Pie-man" told by Jim O'Grady) recorded at a live storytelling performance ("The Moth" storytelling event, New York City). Subjects listened to the story from beginning to end (intact condition). In addition subjects listened to scrambled versions of the story, which were generated by dividing the original stimulus into segments of different timescales (paragraphs, sentences, and words) and then permuting the order of these segments.

The timescale of neural dynamics is a property of an individual brain region and can be measured in an individual subject on an individual trial. We quantified the dynamical timescale by first calculating the power spectrum of the BOLD responses while subjects listened to the intact story, and then averaging the spectra across subjects.

By contrast, the information processing timescale was defined through changes in the reliability of neural responses to different scrambled stimuli. For example, a region would be said to have a long

information processing timescale if it responded reliably when intact paragraphs were presented but not when the word order was scrambled (see Lerner, et al, 2009).

To generate the scrambled stimuli, the story was segmented manually by identifying the end points of each word, sentence, and paragraph. Two adjacent short words were assigned to a single segment in cases where we could not separate them. Following segmentation, the intact story was scrambled at three timescales: short: "words" (W; 608 words, 0.7 ± 0.5 s each); intermediate: "sentences" (S; 69 sentences, 7.7 ± 3.5 s each); and long: "paragraphs" (P; 11 paragraphs, 38.1 ± 17.6 s each). Laughter and applause were classified as single word events (4.4% of the words). Twelve seconds of neutral music and 3 s of silence preceded and 15 s of silence followed each playback in all conditions. These music and silence periods were discarded from all analyses. A typical session comprised five runs, each consisting of the presentation of one condition. Presentation order was pseudorandomized across subjects. Attentive listening to the story was confirmed using a simple questionnaire at the conclusion of the experiment.

Regions of interest (ROIs) were defined based on the timescale of information processing. This was done entirely using the responses to the various scrambled stimuli. Subsequently, we characterized the dynamical timescales of the neural responses within each ROI. The dynamical timescales were measured based on the response to the intact story (i.e., based on a dataset independent from that used to define the ROIs).

Resting state data. Eight subjects (7 additional subjects and 1 subject from the "Pieman" experiment) also participated in a separate scanning session in which data were collected for 10 min while at rest (mixture of eyes closed and eyes open). To assess the dynamical timescales in the absence of stimulus-evoked responses, the timescales of neural dynamics were also measured in the resting state condition, again via computation of the power spectrum of the BOLD signal.

Data preprocessing. fMRI data were preprocessed with the Brain-Voyager software package (Brain Innovation, version 1.8) and with additional software written with MATLAB (The MathWorks, Natick, MA). Preprocessing of functional scans included linear trend removal and high-pass filtering (3 cycles per experiment). In all experiments, this cutoff is longer than the window used to estimate the power spectra. To correct for head motion, we used a 3D algorithm that adjusts for small head movements by rigid body transformations of all slices to the first reference volume. Detected head motions were <1 mm in size, which is well within the range of typical movements observed in other imaging studies. All functional images were transformed into a shared Talairach coordinate system so that corresponding brain regions were roughly spatially aligned. To account for inaccuracies in registration across subjects, the data were spatially smoothed with a Gaussian filter of 6-mm full width at half-maximum value.

Spectral analysis. BOLD time series from every voxel were normalized to zero mean and unit variance to account for differences in variance of nonneural origin (e.g., distance from head coil). In addition, for each subject and each voxel, we used a best-fit linear model to remove the spatial-mean brain dynamics. After subtraction of the global mean, the voxel time series were renormalized to unit variance. Since the variance is equal across all voxels, all spectra have the same integrated area. The power spectra were estimated using Welch's method with a Tukey window (*r*) of width 100 s and 50% overlap. The spectra were clustered using the *k*-means algorithm employing a Euclidean metric and 20 replicates. Error bars throughout the article denote bootstrap errors in the mean. The proportion of low-frequency power was quantified through $\alpha = \int_0^{0.04} P(f) df$, where $P(f)$ is the power estimated as above. We find similar results when changing the number of clusters (± 1) and the low-frequency cutoff by ± 0.01 Hz. Power-law exponents were estimated by assuming $P(f) = cf^{-\beta}$ in the frequency range $f_{min} = 0.01$ Hz and $f_{max} = 0.33$ Hz. Best-fit parameters were determined using nonlinear least squares

with trust-region optimization. Errors in the exponents reflect 95% confidence intervals.

Coherence analysis. For each voxel time series $x(t)$, we compute the coherence between x and its $n − 1$ subject average $y(t)$ as

$$C_{xy}(f) = \frac{|\rho_{xy}|(f)}{\rho_{xx}(f)\rho_{yy}(f)}$$

where ρ_{xy} is the cross spectral density and ρ_{xx} is the spectral density. We then average C_{xy} across subjects and voxels in each region. The coherence ranges between 0 and 1 and provides a direct link between the frequency content of the dynamics and the correlation across subjects.

TRW indices. To assess the timescale of processing for each brain area, we first performed an intersubject correlation analysis. This analysis provides a measure of the reliability of the responses to a temporally complex stimulus, such as a story, by comparing the BOLD response time courses across different subjects (intersubject correlation; henceforth inter-SC). Inter-SC coefficients were calculated on a voxel-by-voxel basis (in Talairach space) within each condition (backward, words, sentences, and paragraphs) by comparing the responses across all listeners (inter-SC).

At every voxel, the Pearson product-moment correlation coefficient ρ_k of *subject k* was computed as:

$$\rho_k = \rho(r_k, \bar{r}) = \frac{r_k(t) \cdot \bar{r}(t)}{\sqrt{[r_k(t) \cdot r_k(t)][\bar{r}(t) \cdot \bar{r}(t)]}}$$

where $r_k(t)$ is the response time course of a voxel to the stimulus presentation for a *subject k*, $\bar{r}(t) = \sum_{i \neq k} r_i(t)$ is the response time course averaged across all subjects except k, and $r_k(t) \cdot \bar{r}(t)$ is the inner product of $r_k(t)$ and $\bar{r}(t)$. Next, the average correlation $\rho = \sum_i \rho(r_i, \bar{r})$ was computed at every voxel. Statistical significance of each ρ value was computed using a bootstrapping procedure (Lerner et al. 2011).

Finally, to correct for multiple comparisons, we applied the Benjamini-Hochberg-Yekutieli false-discovery procedure that controls the false discovery rate (FDR) under assumptions of dependence (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001).

To determine the timescale of processing in each brain area, we used the inter-SC correlations across conditions (see Fig. 5A). A voxel was assigned the label "short TRW" (red), when it was significantly coupled in all inter-SC maps (backward, words, sentences, and paragraphs levels). A voxel was assigned the label "intermediate TRW" when it was significantly coupled in all inter-SC maps, except backward (yellow) or words and backward (green). A voxel was assigned the label "long TRW" (blue) when it was significantly coupled only in the paragraph scrambled inter-SC maps. Importantly, note that for the calculation of the TRW indices we did not use the data from the intact story; the intact story was separately used to assess the power spectrum at each ROI or voxel. In addition we also calculated a continuous TRW index (see Fig. 5B) as TRW index = $(\rho_{\text{Full Story}} − \rho_{\text{Words}})$. Values close to 0 indicate that the voxel responded equally to both the intact story and fine scrambled story; relatively large positive values indicate that the responses are higher for the intact condition than the word-scrambled condition.

RESULTS

When subjects listen to a 7-min real-life story, the timescales of their neural dynamics exhibit a well-defined gradient along the cortical surface. This gradient ranges from faster dynamics in early sensory areas to the slower dynamics further along the cortical hierarchy (Fig. 1). To quantify these dynamics we compute the power spectrum for each brain area within each listener and then average the power spectra across the ensemble of listeners ($n = 11$). To provide an impression of the
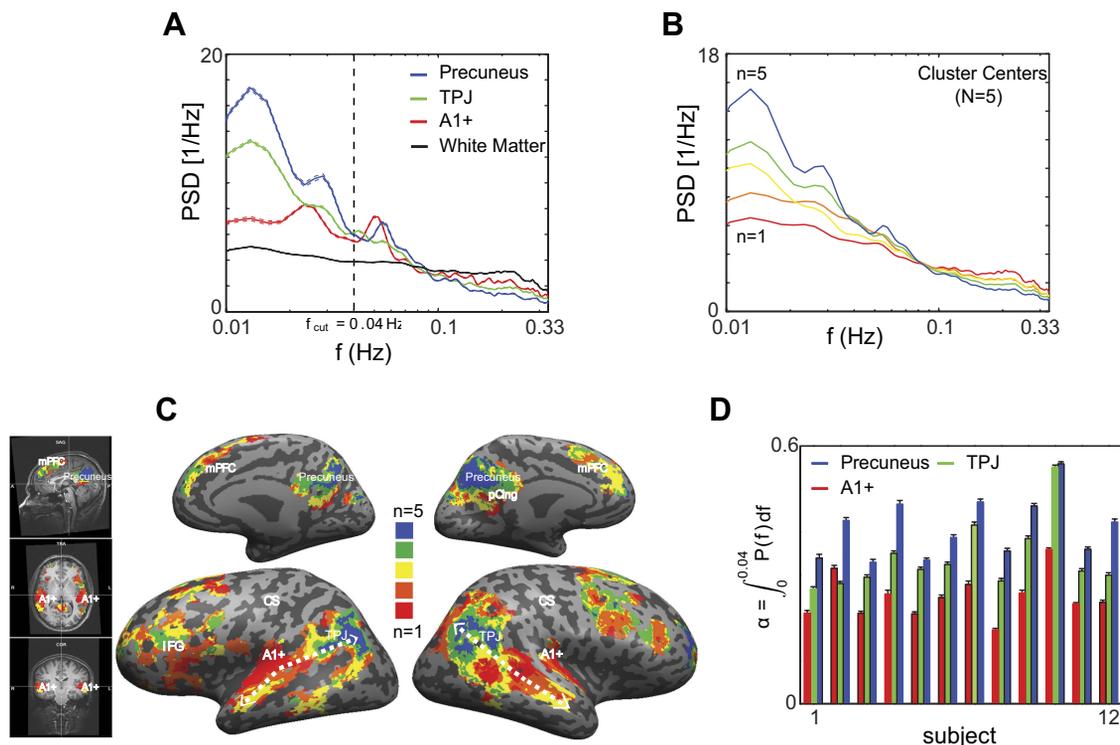
Fig. 1. Spatiotemporal patterning of neural dynamics across the brain. *A*: power spectrum in three illustrative brain regions [early auditory cortex (A1+), temporoparietal junction (TPJ), and precuneus] and white matter, computed from 11 subjects listening to an audio story in the scanner. *B*: *k*-means clustering with 5 clusters reveals a natural separation in the shapes of the power spectra across the brain. *C*: map of cluster identity restricted to significantly coupled brain regions. *D*: gradient in timescales is also clearly visible in individual listeners. We characterize the shape of the power spectrum through the cumulative low-frequency power $\alpha = \int_0^{0.04} P(f) df$ and we plot α for individual listeners in selected brain regions. Error bars in *D* and dotted lines in *A* denote SE in the mean across the voxels in each region. PSD, postsynaptic density; mPFC, medial prefrontal cortex; IFG, inferior frontal gyrus; CS, cortical surface.

power spectrum we first present the results from three illustrative brain regions, with three different timescales of information processing (as revealed in a previous study, Lerner et al. 2011): early auditory cortex (A1+, fast information processing timescale), temporoparietal junction (TPJ; intermediate information processing timescale), and the precuneus (long information processing timescale). The assessment of the timescales of the neural dynamics (as measured by the power spectrum analysis in these ROIs) was then followed by a comprehensive voxel-wise calculation of dynamical timescales across the whole brain.

The power spectra in these three ROIs show a clear shift in the dynamics, with frequencies <0.1 Hz increasingly prevalent from A1+ to the TPJ to the Precuneus (Fig. 1A). We note that our assignment of (relative) neural dynamics timescales is based on the broadband shape of the power spectrum and not on any band-limited peaks in the power spectrum. Next, we measure the dynamics within each brain area (voxel) and cluster the power spectra using k-means with five clusters. We restrict the analysis to regions that respond reliably to the story across subjects (inter-SC, FDR corrected $q < 0.05$, see METHODS). We find that the cluster centers can be naturally ordered by the distribution of their low-frequency content (Fig. 1B), and we show a brain map of cluster identity (Fig. 1C). The timescale gradient presents a neural topography: early auditory areas exhibit faster dynamics (red) and the dynamics become slower (yellow to green to blue) as one moves along the posterior and anterior superior temporal gyrus (see white dotted arrows in Fig. 1C), with the slowest dynamics in the precuneus and medial prefrontal gyrus. Thus the analysis revealed a clear change in the

broadband profile of the power spectrum, indicating a shift across areas from faster to slower dynamics.

To characterize the proportion of low-frequency power in each subject we calculate α, the cumulative power below a fixed threshold of 0.04 Hz within each ROI for each subject. We note that while each voxel has the same cumulative power since the time series are normalized to unit variance, the distribution of power across frequencies can still vary. The same gradual gradient of neural dynamics is apparent for individual listeners (Fig. 1D). Finally, the white matter and ventricles display a nearly flat power spectrum (Fig. 1A, black line) corresponding to unstructured white noise and suggesting that spectral clustering methods can be used to filter such areas at the preprocessing stage.

The varying topography of voxel power spectra is also observed when we focus on those neural dynamics that are locked to the stimulus (Fig. 2). When subjects listen to a story, their BOLD dynamics can be decomposed into a "noise" component (from, e.g., instrumental, physiological and cognitive variability) and a "signal" component (which is induced by the stimulus). Averaging the responses across subjects increases the relative magnitude of the shared signal component. For each ROI we average across the listener ensemble and then compute the power spectrum. Within a noisy brain area in which the responses are uncorrelated across subjects, this procedure would result in a flattened mean-listener spectrum. In Fig. 2A we present the power spectra of the listener-averaged dynamics using the same ROIs as in Fig. 1A. Within these ROIs we find that the average spectrum is not flat, which indicates that the responses contain a reliable component
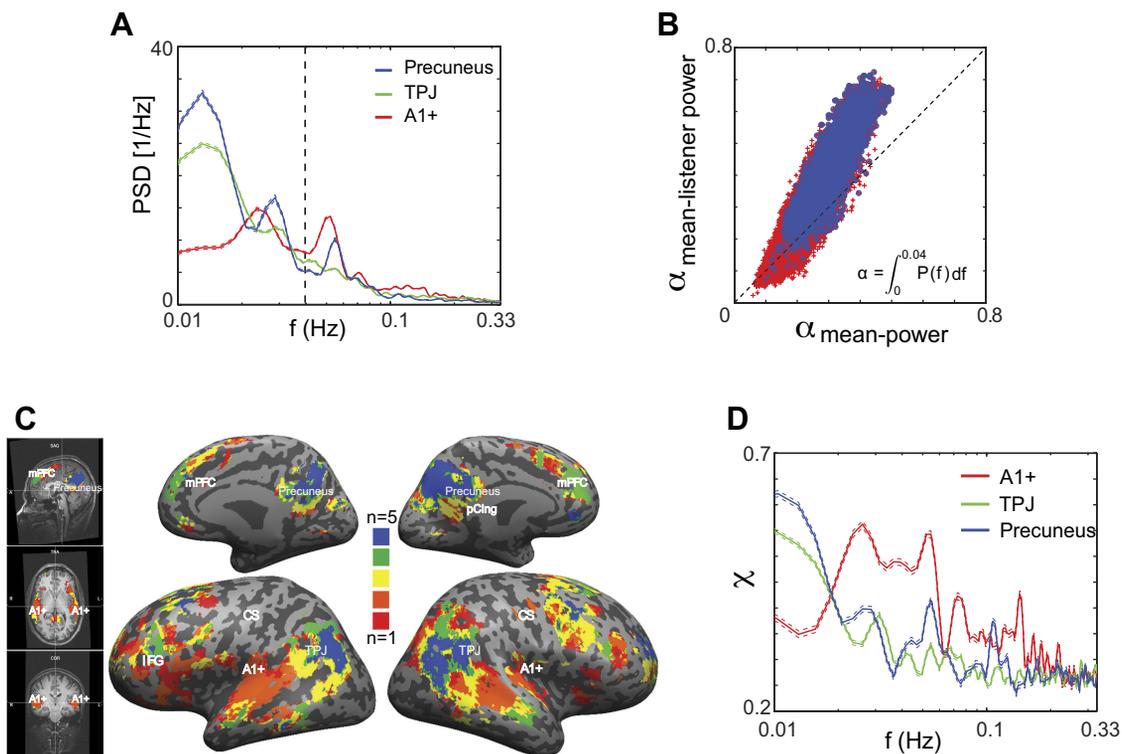
Fig. 2. Common brain dynamics across listeners exhibits accentuated low-frequency power. *A*: power spectrum of the mean-listener activity along the same pathway as in Fig 1A. We average each voxel time series across different listeners, compute the power spectrum of the resulting mean signal, and then average the spectra across voxels within each region. *B*: for each voxel, the cumulative low-frequency power in the mean listener is enhanced relative to the mean power across listeners. Blue denotes significantly coupled voxels identified previously (Lerner et al. 2011). *C*: mean-listener α displayed across the brain. *D*: spectral coherence among listeners exhibits increasing correlation in low-frequency bands.

across subjects. Furthermore, the shift in power from fast to slow dynamics across ROIs indicates that different temporal aspects of the stimulus are driving the responses in each of the ROIs. Note especially the enhancement of slow power in the TPJ and precuneus.

The relative increase in low-frequency power, from early to higher order areas, is more pronounced in the average dynamics than for individual listeners (Fig. 2, *B–D*). This can be seen on a voxel-by-voxel level, in which α preserves its relative gradient, with a steeper (above the diagonal) increase in low-frequency power for the average signal for areas with high alpha relative to the mean power within individuals (Fig. 2*B*). A marked topographic timescale map was also observed when we cluster the neural dynamics of the average signal using the same *k*-means procedure (Fig. 2*C*). The functional relevance of low-frequency dynamics is underscored by a spectral coherence analysis (see METHODS) that can identify the frequency ranges within the power spectrum that are correlated across subjects (Fig. 2*D*). While in early auditory areas, the coherence was stronger at relatively higher frequencies (0.02–0.2 Hz), in the TPJ and the precuneus the coherence was stronger at frequencies below 0.02 Hz.

The timescales of neural dynamics retained their spatial ordering during rest, as well as during three additional conditions in which the story's temporal structure was manipulated (Fig. 3, Fig. 4, *B* and *C*). This suggests that the timescale gradient reveals an intrinsic property of neural circuits, which may be modulated, but not completely controlled, by environmental stimuli. In Fig. 3, we plot the low-frequency content measured while subjects were at rest and while they listened to time-reversed, sentence-scrambled and paragraph-scrambled

versions of the intact story. The relative difference in the proportion of low-frequency power was preserved across all conditions at the voxel-by-voxel level. In addition, we observed a slight increase in slow-frequency dynamic across all voxels during the resting condition and during the reverse speech condition compared with the intact audio story (i.e., the points fall above the diagonal line).

A similar gradient of timescales was revealed when timescales were characterized via the slope of a power law fit to the BOLD spectra. The power law exponent β gradually increases from early auditory areas to high-order areas, and across the *k*-means power clusters in both the story data (Fig. 4*A*) and the resting data (Fig. 4*B*). Moreover, as with the cumulative power matric, the exponent slopes seems to be preserved, on a voxel-by-voxel basis, across the story and the resting data sets (Fig. 4*C*). These findings may support recent suggestions of scale-free neural dynamics in the human brain (He 2011; He et al. 2010), although we note that BOLD data is not optimal for distinguishing between scale-free activity and other dynamical regimes.

The power spectrum gradient coincides with the TRWs gradient found using scrambled auditory stimuli (Fig. 5), and this suggests a functional link between the processing timescale of a region and its spectral dynamics characterization. In previous work we assessed the timescale of processing within each brain area by scrambling the same auditory story at the level of words, sentences, and paragraphs (Lerner et al. 2011). We found that early auditory areas responded reliably for all scrambling levels while the reliability in higher brain areas increased gradually in correspondence with the length of coherent temporal structures in the stimulus. By analogy with the notion of a spatial receptive field, we defined the TRW of a neural circuit as the length of time before a response during which sensory information may affect that response (Lerner et al. 2011). To directly connect the TRW indices to the voxel dynamics we averaged the power spectra within short (red), intermediate (yellow, green), and long (blue) TRW regions (Fig. 5*A*). The results reveal a clear gradient: areas with short TRWs (less sensitive to the scrambling level) have faster dynamics, while areas with long TRWs (which accumulate information over long timescales) have slower dynamics. A similar correspondence between the region temporal receptive window and proportional power was observed on a voxel-by-voxel basis (Fig. 5*B*).

DISCUSSION

We have shown that human brain dynamics elicited by the processing of natural auditory stimuli are remarkably structured, with faster timescales in early sensory areas and slower dynamics in higher order regions. We find a similar temporal gradient under resting state conditions within the scanner, and this pattern is enhanced by the common, synchronized dynamics of complex experience. The neural timescale gradient (defined by the voxel power spectrum) correlates with the temporal receptive window gradient (defined by the reliability of neural response as a function of stimulus structure). These data therefore connect the time-varying structure of the stimulus to intrinsic neural dynamics and suggest a large-scale organizing principle in which cortical areas that accumulate information over longer timescales exhibit a larger fraction of

Fig. 3. Cumulative low-frequency power in the intact story is similar to the resting state and to multiple stimulus conditions. We show scatter plots of the low-frequency power α in the intact story and resting state (*A*), reversed speech (*B*), scrambled sentences (*C*), and scrambled paragraphs (*D*) conditions. We quantify the similarity with the correlation coefficient ρ and the error is computed using bootstrap permutations among the voxels.

Fig. 4. Power-law exponents of the spectral clusters also display the timescale gradient. *A*: normalized power spectrum plotted in a log-log plot for each of the *k*-means power clusters in the intact story. We fit each spectrum to a power-law function between $f_{min} = 0.01$ Hz and $f_{max} = 0.33$ Hz and the values for the exponent β appear below the plot, ordered from lower α values (*left*) to higher α-values (*right*). *B*: similar analysis performed on the resting state data. *C*: for each voxel, the power-law exponent in resting state vs. the intact story.

low-frequency dynamics. Indeed, theoretical work (Friston 2009; Jaaskelainen et al. 2011; Wiskott and Sejnowski 2002) has argued that invariances on increasingly long timescales may support invariant object recognition (Berkes and Wiskott 2005) and that hierarchical coupling of neural dynamics may underlie prediction and perception across multiple timescales (Kiebel et al. 2008).

What is the origin of the differences in dynamics among regions? One possibility is that the differences in timescales reflect differences in intraregional properties such as local recurrent wiring (Brody et al. 2003; Durstewitz et al. 2000), membrane excitability (Marom 1998) and varieties of short-term synaptic plasticity (Zucker and Regehr 2002). A second

Fig. 5. Timescale gradient in the voxel power spectrum parallels the map of temporal receptive windows found using scrambled auditory stimuli (Hasson et al. 2008; Lerner et al. 2011). *A*: voxel power spectra averaged within temporal receptive windows (TRW) categories. TRW categories with longer time windows are associated with an increased proportion of low-frequency power in the voxel dynamics. *B*: scatter plot of α vs. the TRW index (a continuous measure of the TRW category). The correlation between α and the TRW index is $\rho = 0.47 \pm 0.01$, where the error is computed using bootstrap among the population of voxels.

possibility is that the dynamical differences originate at the level of inter-regional connectivity (see, e.g., Salvador et al. 2008; Wu et al. 2008). In particular, if individual regions in the processing hierarchy act as low-pass filters on their inputs, then higher order regions would exhibit slower dynamics because they receive input signals that have already passed through more filtering stages (Baria et al. 2013). A third potential origin of slow cortical dynamics is via slow subcortical modulation. While we cannot settle the issue here, it is likely that both region-specific and large-scale network influences contribute to the variation of timescales across regions.

Our analysis reveals a robust spatial organization of the timescales of regional BOLD dynamics, both at rest and across multiple stimulation conditions (Fig. 3). Moreover, in agreement with other studies (Baria et al. 2011; He et al. 2010), we observed a global increase in low-frequency power relative to the intact story (i.e., substantially more dots lie above the diagonal line) during rest. Interestingly, a similar global increase was observed during the reverse story condition. This is of particular interest, given that the spectrum of the audio envelope is unchanged when the stimulus is played in reverse. Based on that, we conclude that the global increase in low-frequency power can be dissociated from the low-level characteristics of the stimulus power spectrum (which is identical in the intact and reverse conditions). Rather, the BOLD dynamical timescales are more related to the presence or absence of meaningful input across the intact, rest, or/and reverse speech conditions.

Along the auditory pathway our results are consistent with the processing of sequentially longer timescales by networks with slower dynamics. However, we cannot currently generalize our findings to the visual cortex, as it was shown that early visual areas generally display more low-frequency content than high-order ventral (though not dorsal) visual areas (Baria et al. 2011; He et al. 2010). It is unclear whether the audio and visual discrepancies arise from functional differences between the two sensory pathways or from other factors such as hemodynamic-respiratory coupling (Birn et al. 2006). This discrepancy between auditory and visual sensory pathways is an important issue for future investigation.

Consistent with prior studies (Baria et al. 2013), we find that the very slowest neural dynamics are not seen in the frontal cortex but in the vicinity of the angular gyrus and the precuneus. The finding that lateral frontal areas exhibit mostly

intermediate or slow dynamics is consistent with their role in the comprehension and production of sentences (Hickok and Poeppel 2004) and in working memory and executive control functions (Baddeley et al. 1996). The finding that the slowest dynamics are in precuneus and angular gyrus is consistent with their role as cortical network hubs (Baria et al. 2013; Hagmann et al. 2008) and in episodic memory (Andrews-Hanna et al. 2010; Ranganath and Ritchey 2012).

The results reveal for the first time that differences in the neural dynamic across brain areas can be linked to differences in the information processing timescale (Fig. 5). First, the neural dynamic gradient is preserved even in the absence of stimuli (Fig. 3), suggesting that it is an intrinsic property of the neural circuits. Second, the neural response dynamics align with the stimulus dynamics. While areas with fast dynamic (e.g., early auditory areas) are coupled to rapidly varying stimulus dynamic (e.g., single phonemes or words), areas with mid-range dynamic (such as posterior STS and inferior frontal gyrus) are coupled to the intermediate stimulus temporal structure (e.g., the structure of sentences), and areas with the slowest dynamics (e.g., precuneus and medial prefrontal cortex) are coupled to the slowly varying stimulus dynamic (e.g., stimulus narrative; Honey et al. 2012; Kauppi et al. 2010; Lerner et al. 2011). Furthermore, when the temporal structure of the stimulus is disrupted via scrambling, this interferes with the alignment of neural dynamics to the stimulus structure in areas with long processing timescale (Honey et al. 2012). Together, these results suggest that areas with faster neural dynamic accumulate information over short timescales, while areas with slower neural dynamic accumulate information over longer timescales.

Finally, we note that neural dynamics appear heavy tailed in their high frequencies, consistent with claims of critical and scale-free dynamics (Chialvo 2010) based on in vitro field experiments (Friedman et al. 2012) as well as fMRI (Aguirre et al. 1997; Expert et al. 2011; He 2011) and ECoG (Solovey et al. 2012) measurements in humans. Here, we too observe heavy tails in the power spectra, Fig. 4. However, our results also suggest that there exists a systematic diversity in the neural dynamics (and thus the power exponents) among different brain networks and that this diversity tracks the diverse demands of real-world information processing (see also He 2011).

To conclude, the present study revealed a functional link between the multiple timescales of neural dynamics and the multiple timescales of real-world information accumulation. The circuit level mechanisms underlying neural and cognitive timescale diversity should be the focus of future studies.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: G.J.S., C.J.H., and U.H. conception and design of research; G.J.S., C.J.H., and U.H. analyzed data; G.J.S., C.J.H., and U.H. interpreted results of experiments; G.J.S., C.J.H., and U.H. prepared figures; G.J.S., C.J.H., and U.H. drafted manuscript; G.J.S., C.J.H., and U.H. edited and revised manuscript; G.J.S., C.J.H., and U.H. approved final version of manuscript; C.J.H. and U.H. performed experiments.

REFERENCES

Aguirre GK, Zarahn E, D'Esposito M. Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage* 5: 199–212, 1997.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL. Functional-anatomic fractionation of the brain's default network. *Neuron* 65: 550–562, 2010.

Baddeley A, Sala SD, Robbins TW, Baddeley A. Working memory and executive control [and Discussion]. *Phil Trans R Soc B* 351: 1397–1404, 1996.

Baliki MN, Baria AT, Apkarian AV. The cortical rhythms of chronic back pain. *J Neurosci* 31: 13981–13990, 2011.

Baria AT, Baliki MN, Parrish T, Apkarian AV. Anatomical and functional assemblies of brain BOLD oscillations. *J Neurosci* 31: 7910–7919, 2011.

Baria AT, Mansour A, Huang L, Baliki MN, Cecchi GA, Mesulam MM, Apkarian AV. Linking human brain local activity fluctuations to structural and functional network architectures. *Neuroimage* 73: 144–155, 2013.

Benjamini Y, Hochberg Y. Controlling the false discovery rate–a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser B* 57: 289–300, 1995.

Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29: 1165–1188, 2001.

Berkes P, Wiskott L. Slow feature analysis yields a rich repertoire of complex cell properties. *J Vis* 5: 579–602, 2005.

Birn RM, Diamond JB, Smith MA, Bandettini PA. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage* 31: 1536–1548, 2006.

Brody CD, Hernández A, Zainos A, Romo R. Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cereb Cortex* 13: 1196–1207, 2003.

Chialvo DR. Emergent complex neural dynamics. *Nat Phys* 6: 744–750, 2010.

Cole BJ. Fractal time in animal behavior: the movement activity of Drosophila. *Anim Behav* 50: 11, 1995.

Dong DW, Atick JJ. Statistics of natural time-varying images. *Netw Comput* 6: 345–358, 1995.

Durstewitz D, Seamans JK, Sejnowski TJ. Neurocomputational models of working memory. *Nat Neurosci* 3, *Suppl*: 1184–1191, 2000.

Expert P, Lambiotte R, Chialvo DR, Christensen K, Jensen HJ, Sharp DJ, Turkheimer F. Self-similar correlation function in brain resting-state functional magnetic resonance imaging. *J R Soc Interface* 8: 472–479, 2011.

Friedman N, Ito S, Brinkman BA, Shimono M, DeVille RE, Dahmen KA, Beggs JM, Butler TC. Universal critical dynamics in high resolution neuronal avalanche data. *Phys Rev Lett* 108: 208102, 2012.

Friston K. The free-energy principle: a rough guide to the brain? *Trends Cogn Sci* 13: 293–301, 2009.

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the structural core of human cerebral cortex. *PLoS Biol* 6: e159, 2008.

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N. A hierarchy of temporal receptive windows in human cortex. *J Neurosci* 28: 2539–2550, 2008.

He BJ. Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *J Neurosci* 31: 13786–13795, 2011.

He BJ, Zempel JM, Snyder AZ, Raichle ME. The temporal structures and functional significance of scale-free brain activity. *Neuron* 66: 353–369, 2010.

Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92: 67–99, 2004.

Honey CJ, Thesen T, Donner Tobias H, Silbert Lauren J, Carlson Chad E, Devinsky O, Doyle Werner K, Rubin N, Heeger David J, Hasson U. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76: 423–434, 2012.

Jaaskelainen IP, Ahveninen J, Andermann ML, Belliveau JW, Raij T, Sams M. Short-term plasticity as a neural mechanism supporting memory and attentional functions. *Brain Res* 1422: 66–81, 2011.

Kauppi JP, Jääskeläinen IP, Sams M, Tohka J. Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Front Neuroinform* 4: 5, 2010.

Kiebel S, Daunizeau J, Friston K. A hierarchy of time-scales and the brain. *PLoS Comput Biol* 4: e1000209, 2008.

Leopold DA, Murayama Y, Logothetis NK. Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cereb Cortex* 13: 422–433, 2003.

Lerner Y, Honey CJ, Silbert LJ, Hasson U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci* 31: 2906–2915, 2011.

Linkenkaer-Hansen K, Nikouline VV, Palva JM, Ilmoniemi RJ. Long-range temporal correlations and scaling behavior in human brain oscillations. *J Neurosci* 21: 1370–1377, 2001.

Marom S. Slow changes in the availability of voltage-gated ion channels: effects on the dynamics of excitable membranes. *J Membr Biol* 161: 105–113, 1998.

Monto S, Palva S, Voipio J, Palva JM. Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J Neurosci* 28: 8268–8272, 2008.

Ranganath C, Ritchey M. Two cortical systems for memory-guided behaviour. *Nat Rev Neurosci* 13: 713–726, 2012.

Salvador R, Martinez A, Pomarol-Clotet E, Gomar J, Vila F, Sarro S, Capdevila A, Bullmore E. A simple view of the brain through a frequency-specific functional connectivity measure. *Neuroimage* 39: 279–289, 2008.

Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440: 1007–1012, 2006.

Solovey G, Miller KJ, Ojemann JG, Magnasco MO, Cecchi GA. Self-regulated dynamical criticality in human ECoG. *Front Integr Neurosci* 6: 44, 2012.

Voss RF, Clarke J. 1-F-noise in music and speech. *Nature* 258: 317–318, 1975.

Wiskott L, Sejnowski TJ. Slow feature analysis: unsupervised learning of invariances. *Neural Comput* 14: 715–770, 2002.

Wu CW, Gu H, Lu H, Stein EA, Chen JH, Yang Y. Frequency specificity of functional connectivity in brain networks. *Neuroimage* 42: 1047–1055, 2008.

Yan L, Zhuo Y, Ye Y, Xie SX, An J, Aguirre GK, Wang J. Physiological origin of low-frequency drift in blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI). *Magn Reson Med* 61: 819–827, 2009.

Zucker RS, Regehr WG. Short-term synaptic plasticity. *Annu Rev Physiol* 64: 355–405, 2002.