Title: **Multiple sensitivity profiles to diversity and transition structure in non-stationary input**

Michael J. Tobia[a], Vittorio Iacovella[a], and Uri Hasson[a, b, §]

[a] *Center for Mind/Brain Sciences, The University of Trento, Italy*
[b] *Faculty of Cognitive Sciences, The University of Trento, Italy.*

Manuscript *in press* in NeuroImage (2011)

[§] Corresponding author:
Uri Hasson
Center for Mind/Brain Sciences (CIMeC)
Via delle Regole, 101 I-38060 Mattarello (TN), Italy
Email: uri.hasson@unitn.it
Tel: +39 0461 883082
Fax: +39 0464 80 8654

This research has received funding from the European Research Council under the 7th framework starting grant program (ERC-STG #263318) to U.H. We thank Anthony Dick and Nathan Cashdollar for their suggestions and Mark Makers for general support.

Abstract

Recent formalizations suggest that the human brain codes for the degree of order in the environment and utilizes this knowledge to optimize perception and performance in the immediate future. However, the neural bases of how the brain spontaneously codes for order are poorly understood. It has been shown that activity in lateral temporal cortex and the hippocampus is linearly correlated with the order of short visual series under tasks requiring attention to the input and when series order is invariant over time. Here, we examined if sensitivity to order is manifested in both linear and non-linear BOLD response profiles, quantified the degree to which order-sensitive regions operate as a functional network, and evaluated these questions using a paradigm in which performance of the ongoing task could be completed without any attention to the stimulus whose order was manipulated. Participants listened to a 10-minute sequence of tones characterized by non-stationary order, and fMRI identified cortical regions sensitive to time-varying statistical features of this input. Activity in perisylvian regions was negatively correlated with input diversity, quantified via Shannon's Entropy. Activity in ventral premotor, lateral temporal, and insular regions was correlated linearly, parabolically, or via a step-function with the strength of transition constraints in the series, quantified via Markov Entropy. Granger-Causality analysis revealed that order-sensitive regions form a functional network, with regions showing non-linear responses to order associated with more afferent connectivity than those showing linear responses. These findings identify networks that spontaneously code and respond to diverse aspects of order via multiple response profiles, and that play a central role in generating and gating predictive neural activity.

Keywords: Entropy, disorder, uncertainty, non-monotonic, fMRI

1. Introduction

Predictability and uncertainty as statistical features of the environment are central concepts in contemporary theories of perception, cognition, and behavior, as well as their neural underpinnings (Fiser et al., 2010; Friston, 2009). Individuals learn statistical features of their environment and use this information to formulate rules of order (Reber, 1967), parse streams of input (Saffran et al., 1996) and form categories and concepts (Knowlton et al., 1994). The probability of specific events modulates activity in early sensory processing regions and facilitates perceptual recognition (Summerfield et al., 2008, 2009). Auditory input streams with greater uncertainty among events are associated with more rapid electrophysiological oscillations than streams consisting of stronger statistical relations (Buiatti et al., 2009). In addition, the overall degree of uncertainty in an input stream has been linked to activity levels in a variety of networks that may partially overlap (Bischoff-Grethe et al., 2000; Harrison et al., 2006; Overath et al., 2007; Schubotz et al., 2002; Schubotz, 2007; Strange et al., 2005). Together, these behavioral and neurophysiological findings show that the brain responds to statistical information embedded in sensory signals.

While there is a growing body of scientific literature concerning neural codes for the statistical features of discrete events including their cue validity (Turk-Browne et al., 2010) or their surprise value (Strange et al., 2005), research concerning the neural response to statistical features of temporally extended event sequences is less prominent. In general, work examining the neural coding of uncertainty within extended sequences has assumed there is a linear relationship between the degree of order in a sequence and activity in brain regions sensitive to uncertainty. However, it has been suggested (Hebb, 1958; Loewenstein, 1994) that certain cognitive processes and types of perceptual inference, such as the tendency to seek and

incorporate new information from failed predictions, are most prevalent in contexts of moderate levels of uncertainty, rather than very high or low levels. If organisms are particularly sensitive to moderate levels of uncertainty in their interactions with the environment, certain brain regions may respond to input uncertainty in a non-linear manner with maximal levels of activity occurring for inputs whose uncertainty is neither random nor completely ordered. This suggests that beyond coding for the degree of environmental uncertainty, certain computations are most prominent when there exists an 'optimal' degree of uncertainty.

Another reason for considering a non-linear relation between entropy and brain activity emerges from the theoretical distinction between entropy and complexity. While some theorists hold that more complex systems are simply more random ones, more recent theoretical approaches draw a distinction between randomness and complexity (see Huberman & Hogg, 1986 for a review). In these models, complexity refers to the sophistication of a system, and indicates the degree of difficulty associated with describing the system on an abstract level, or performing the operations necessary to reconstruct a sequence from a given sampling distribution. On this latter approach, the relation between entropy and complexity is not linear, with the most complex states residing somewhat mid-way between randomness and complete determinism. To further illustrate, either a random or fully deterministic sequence consisting of four events (A, B, C, D) can be described easily (e.g., random: "each event is equally likely at any point"; or, deterministic: "repeat ABCD consecutively"). In contrast, describing the rules of a moderately ordered/disordered series constructed from such four events necessitates greater specification of the details relating consecutive events, resulting in longer, more complex descriptions. From a cognitive or neurobiological perspective, coding environmental complexity in this manner, in addition to uncertainty *per se*, may result in multiple levels of representation of

input uncertainty that have potentially different functions. For instance, coding for the complexity of the environment can allow rapid comparisons between sensory inputs to determine whether they share common features at an abstract level. Furthermore, understanding that an environment is complex can allow the brain to organize particular modes of operation necessary for the coding of predictions and the processing of errors from failed predictions. Thus, accounts grounded in the psychology of information seeking and those separating complexity from randomness both suggest that inputs with mid-levels of entropy could engender different sorts of responses than either highly ordered or highly disordered states. Furthermore, processes whose operation manifests a non-linear response profile could depend on a more basic (linear) coding for levels of uncertainty in temporally extended sequences, suggesting a hierarchical organization of linear and non-linear responses to input uncertainty.

The two main goals of this study were therefore to: (i) identify whether a single network codes for different sorts of uncertainty, quantified via Shannon's and Markov Entropy, (ii) determine whether the relation between uncertainty and neural activity is mediated by linear or non-linear transfer functions, and if so, use effective connectivity to determine the nature of information transfer between regions showing linear and non-linear response profiles. Specifically, we examine whether the connectivity between regions showing a non-linear response profile and those showing a linear response profile is marked by more afferent connectivity for the former. This would suggest that regions demonstrating non-linear profiles mediate processes that are dependent on regions coding for uncertainty in a linear manner.

1.1 Defining uncertainty as entropy

Statistical features of temporally extended event sequences, such as *order/disorder*, *predictability*, *randomness* and *complexity* are all related to the notion of uncertainty. Some prior neuroimaging work has defined uncertainty in terms of the degree of physical or metric change in energy between successive tokens, where smaller changes in energy result in more ordered series due to higher autocorrelation (e.g., *Sample Entropy;* Overath et al., 2007). Other work has defined uncertainty in terms of purely symbolic features of an input series such as the relative frequency of tokens or their degree of transition constraints (e.g., Buiatti et al., 2009, McNealy et al, 2006; Strange et al., 2005). Here, we adopt the latter approach and define uncertainty as the information-theoretic quantity *entropy*, which indicates the amount of order/disorder in the signal generator after symbolizing the different tokens in a way that is independent of their physical properties.

Shannon's Entropy (Shannon, 1948) is the statistical uncertainty originating from the form of the sampling distribution giving rise to a series of events. It is maximal when event types are sampled from a uniform distribution (a necessary condition for a random series) for which all events are equally likely, and diminishes as the sampling distribution departs from uniformity, thus conveying information about which events are more likely to appear. Shannon's Entropy is null when the signal is repetitive and there is no uncertainty. The formula for calculating Shannon's Entropy is:

$$\text{Shannon H} = -\sum_x p(x) \log p(x), \qquad \text{eq. 1}$$

for which p is the probability of a given event x.

Markov (conditional) Entropy quantifies the transition constraints between events in a series, and captures relational information not coded by Shannon's Entropy. It is maximal when no transition constraints exist, i.e., when each event is as likely to follow or precede any other

event, and it diminishes to the extent that transition constraints are stronger, vanishing completely in the deterministic case where each event indicates with 100% certainty what the next event would be. The formula for calculating Markov Entropy for a given series of events or given a specific transition matrix is:

$$\text{Markov H} = -\sum_{i=1}^{n} p(i) \sum_{j=1}^{n} p(j|i) \log \, \mathrm{p}(j|i), \qquad \text{eq. 2}$$

in which, *p(i),* is the probability of an event, *i*, and, *p(j|i),* is the probability of event, *j,* given the current event, *i*, and, *n,* indicates the number of different event types.

Quantifying system uncertainty via entropy is commonplace in physical and biological sciences where it is taken to indicate the degree of order/disorder in a system, and is often used synonymously with the notions of randomness and complexity. In our work, we use entropy to refer to randomness, order/disorder and uncertainty (terms which we will use interchangeably).

1.2 Prior work on the neurophysiological basis of tracking uncertainty

Few investigations have parametrically examined the relation between neural activity and entropy for a temporally extended input series (i.e., examined responses to more than two levels of order). Strange and colleagues (2005) assessed the relationship between Shannon Entropy and the BOLD signal during presentation of a series of visual stimuli. They reported a positive linear correlation between Shannon Entropy and BOLD signal in the left anterior hippocampus, and a similar albeit less robust correlation in the right anterior hippocampus. Bischoff-Grethe et al. (2000) reported a negative linear correlation between Markov Entropy and BOLD signal in the posterior lateral temporal cortex bilaterally, also for visual stimuli. Harrison et al. (2006) reported a significant positive linear correlation of mutual information and activity in the left hippocampus, for a sequence of visual stimuli. While these studies aimed at identifying regions

sensitive to environment uncertainty, they did not converge on a single region, but instead suggest that the brain represents diverse aspects of uncertainty, perhaps concurrently.

The ventral premotor cortex has also been linked to the prediction of sequence order for diverse types of stimuli. For instance, Schubotz and von Cramon (2004) studied the neural correlates of sequence prediction, across a range of tasks including sequences of biological motion or sequences of non-biological motionless visual stimuli. They reported activity in the left ventral premotor cortex during predictions for both types of stimuli, which was taken to show that a demand for prediction suffices to activate the left ventral premotor cortex (see Schubotz, 2007 for a review of related findings). However, this latter finding does not address whether the left ventral premotor cortex would mediate spontaneous predictions, to the extent that these occur, in absence of a task demand.

It is important to note that all prior work had assumed that 'entropy sensitivity' is defined as a linear relation between BOLD activity and the degree of entropy for an input sequence, and for this reason, prior work had not examined whether any region demonstrates a relation between entropy and neural activity that is mediated by a *non*-linear transfer function.

Identifying the form of the transfer function mediating the relation between uncertainty and neural activity is important for both cognitive and neurobiological perspectives because it constrains the possible theoretical interpretations concerning the functional role of a particular part of the brain. For example, a linear correlation between uncertainty and neural activity is a necessary condition to interpret the activity of that region as a code for the degree of uncertainty in the input (see Strange et al., 2005). On the other hand, a region showing a U-shaped, inverse-U-shaped, step-up or step-down relation to uncertainty cannot be said to code or track the degree of uncertainty *per se* because it responds to two very different levels of entropy with the same

level of BOLD response. Instead, a region demonstrating such a relation to input entropy would be said to show a second-order representation of uncertainty, perhaps processing the complexity of the input series, or generating predictions in a context where prediction errors are both frequent and informative. Furthermore, as indicated by prior work (Dumontheil et al., 2010; Zarahn et al., 2005) identifying non-monotonic responses to parametric changes is important as it can identify regions associated with either failures of processing for certain parameter levels, or regions involved in endogenously triggered ruminations for very simple stimuli. For this reason, we probed for both linear and non-linear relations between input entropy and BOLD levels.

1.3 Stationarity and non-stationarity of input order

When examining the brain's responses to order of temporally extended input sequences, it is important to consider the temporal window over which statistical information is integrated. This issue pertains to whether the neural coding of order takes place under the premise that observations arise from either a *stationary* or *non-stationary* system. In a stationary system, in which a single sampling distribution is assumed to generate observed events, increased exposure yields a more accurate estimate of a distribution that approaches the true distribution. A non-stationary system, however, lacks a single sampling distribution, as it shifts among generators, making it impossible to discern general statistical features holding for the entire input. Instead, one has to compute "local" features of the recent past and update those continuously. Variations in these local statistics would indicate a non-stationary environment.

Prior work on the brain's coding of uncertainty used stationary sampling distributions to generate experimental stimuli (Bischoff-Grethe et al., 2000; Harrison et al., 2006; Huettel et al., 2002; Strange et al., 2005), and so in those cases, from the viewpoint of a participant, an optimal

strategy for coding entropy would be to update parameter estimates of the environmental statistics by considering all events perceived in a given trial (an updating process corresponding to a 'Bayesian Observer' model). More recent work indicates that the brain assigns greater weight to more recently encountered inputs (Harrison et al., 2011), suggesting that it is indeed sensitive to changes statistical structure at short time scales. However, even this latter finding was identified within the context of a stationary sampling distribution.

No prior work has examined whether the brain is sensitive to entropy in a non-stationary input. In the current work, we addressed this issue by creating stimuli whose statistical features were non-stationary and varied over time. Determining whether the brain tracks entropy in an environment where entropy is constantly changing is important since it will show that rather than assuming a single distribution of events, the brain updates such representations on the basis of local statistics.

1.4 Analytical framework and hypotheses

To determine if tracking input entropy is performed spontaneously in absence of task demands, we presented participants with a series of rapidly presented tones (3Hz) while they were performing an incidental visual monitoring task that did not necessitate any attention to the auditory input (see *Supplementary Materials* for complete audio file). The experimental manipulation was the degree of order in sub-sequences of this continuously-presented auditory stimulus, which allowed varying both Shannon and Markov Entropy in a non-stationary manner, while blood oxygenation level dependent (BOLD) functional magnetic resonance imaging (fMRI) data were collected. The magnitude of these continuously changing entropy values served as models for the BOLD fMRI signal changes in order to determine if neural activity is

related to input uncertainty via linear and/or non-linear link functions. Specifically, two measures of uncertainty (i.e., Shannon and Markov Entropy) were computed from the tone sequence over consecutive 4.5 second sliding windows. Figure 1 presents the procedure by which the regressors were constructed. The model for Shannon Entropy reflected fluctuations in levels of this parameter over the aforementioned 4.5 sec windows. The model for Markov Entropy reflected analogous information, but in addition, from this model we constructed two additional regressors by transforming its values with non-linear transfer functions (U-shaped, step function). From all these models of non-stationary entropy levels, BOLD regressors were created by convolving the entropy models with a hemodynamic response function to create an expected BOLD signal (see Figure 1B, 1C). We did not have any *a-priori* expectation that Shannon's Entropy would relate to neural activity in a non-linear manner because of its analytic relation to the proportion of repeating events in a series. Repetition of stimulus events is known to either suppress or increase (Turk-Browne et al., 2007) neural activity. However, there is no model of repetition suppression suggesting a non-linear relation between the amount of repetition and the amount of neural activity. As such, our hypotheses concerning the possible coding of complexity via the non-linear relation to uncertainty stem from Markov Entropy exclusively.

We had two main hypotheses. First, we wanted to identify brain regions spontaneously sensitive to entropy of a non-stationary input in the absence of any demand to pay attention to that input. On the basis of previous work documenting linear relations between entropy and the BOLD signal in the hippocampus (Harrison et al., 2006; Strange et al., 2005), lateral temporal cortex (Bischoff-Grethe et al., 2000; Overath et al., 2007; Schubotz et al., 2004), and ventral premotor cortex (Schubotz et al., 2002; 2004), we expected to find linear relations in these regions. We also expected to identify data patterns indicating non-linear transfer functions

between entropy and the BOLD response. Finally, we expected that effective connectivity analyses would show that regions demonstrating a linear relation profile between BOLD and input entropy would show processing precedence (i.e., project efferent causal connections) over those regions mediated by a non-linear relation between BOLD and input entropy.

2. Method and Materials

2.1 Participants

Sixteen right-handed volunteers (mean age = 25.44, SD=3.34 years; 8 male) participated in the study. They underwent a medical interview prior to scanning to evaluate exclusion criteria. None reported a history of psychiatric illness, history of substance abuse, nor hearing impairments. Data for one participant could not be analyzed due to poor signal quality, leaving 15 good-quality data sets. The study was approved by the ethical committee of the University of Trento.

2.2 Procedure

Participants passively listened to the tone sequence while engaged in a simple visual change-detection task to maintain alertness. The task was to monitor a fixation cross at the center of the display, and press a response key whenever it changed orientation from "+" to "x" and vice versa. The fixation stimulus changed orientation 32 times during the ten minutes of scanning, at random intervals set separately for each participant. It was presented in white over a black background. Participants were instructed to observe the fixation cross, not try to anticipate its changes since their timing was random, and only press after a change occurred. Prior to the

study, the volume level was adjusted for each participant so that the tones could be clearly and comfortably heard over the scanner noise.

2.3 Stimuli: Auditory parameters, presentation rate and sequence construction

Sensory stimulation consisted of a 10-minute stream of 4 non-randomly alternating auditory tones (frequencies = 262, 294, 330, 394 Hz) presented at a rate of 3.3Hz. Each tone was presented for 250ms with a 50ms inter-stimulus interval, yielding 2000 tone presentations in all. The presentation of each tone was slightly jittered (+/- 5ms) relative to its scheduled presentation timing to avoid temporal synchronization between the presentation rate of the tone sequence and scanner noise generated during functional acquisition. A similar procedure of continuous presentation of visual stimuli has been previously used to study frontal-cortex responses to events departing from a prior pattern (Huettel et al., 2002).

The 10-minute stream was generated so 1st-order Markov Entropy was non-stationary; i.e., order varied across time. This was achieved by setting an initial transition matrix for stimulus generation, and shifting the transition probabilities slightly every 5 cycles – i.e., after generation of 5 tones – according to a predetermined profile that assured strong changes in order over time.

The tone sequence was selected after a series of 100,000 simulations where in each simulation, (*1*) a non-stationary Markov process generated a series of tones, (*2*) Markov and Shannon Entropy (MH, SH) regressors were derived using the sliding-window procedure described below, and (*3*) the correlation between the MH and SH regressors was evaluated using Pearson's correlation coefficient. From these 100,000 simulations, we chose the series in which

the correlation between the MH and SH regressors was minimal (Pearson's correlation coefficient [r] = .03).

2.4 Generation of regressors

The non-stationary MH regressor was constructed by calculating the 1st-order Markov Entropy of the tone sub-sequence presented within 4.5 seconds before each volume acquisition, a time frame that corresponded to 15 tones (pilot studies indicated high sensitivity to order in this time frame). By capturing the transition probability constraints in the sub-sequence encountered immediately prior to volume acquisition, this procedure assigned a Markov Entropy value to each functional volume in the scan. More specifically, for tones presented within this 4.5 sec temporal window, the 1st-order Markov Entropy was calculated from the actual distribution of transition probabilities in the sub-sequence, according to the formula in equation 2. That is, the frequency of each transition type was computed, and these were represented in a transition matrix and assigned a Markov Entropy value for that specific matrix according to equation 2. Figure 1 (panels A and B) illustrates this procedure. We refer to these Markov Entropy values calculated from the most recently encountered 15 items as *instantaneous* MH values ($MH_{instant}$) since they are a parameter reflecting statistical features of a relatively short string rather than the degree of mean surprise of a random variable (following Strange et al., 2005; Harrison et al., 2006; see Garner 1962 for history of using entropy as a statistic in psychological research). We refer to the regressor created from these values as the MHL regressor since it models a linear relation between $MH_{instant}$ values and the neural response.

The MHL regressor was constructed to be non-stationary and frequently fluctuating (see Figure 1A). This allowed frequent changes within the degree of order for the 4.5sec windows. As a result, the autocorrelation of the regressor, i.e., the degree to which a regressor value was

related to recent values, was low and fell to 0.1 at lag-3. Furthermore, the mean difference between the regressor and a lag-2 version of itself was 0.2, confirming that entropy values fluctuated rapidly. The range of MHL values themselves was large, varying from 0.5 to 1.70 (Mean=1.14, SD = 0.21). The regressor consisted of 91 different values after rounding to the second decimal. Note that the maximal MHL value (2 bit) can only be reached for series whose distribution approximates randomness, which typically consist of numerous observations, whereas ours consisted of only 15 observations (tones) each. Two features of the MHL regressor differed from ones that would have been found if sampling from a random stationary distribution. First, the variance of the $MH_{instant}$ values in the regressor was greater than expected from a random process ($p < .001$), as verified against a sampling distribution derived from simulations where the sliding window procedure was applied to permutations of the original MHL regressor. Second, the permutation procedure also showed that the autocorrelation features of the MHL regressor at lags 1 and 2 departed from chance. Autocorrelation values for the MHL regressor were 0.46 at lag-1 and 0.24 at lag-2, both of which exceeded the chance criteria as determined by permutations of MHL ($p < .02$ for lag-1, $p < .001$ for lag-2). To conclude, controlling the construction of the MHL regressor as described above produced a regressor where $MH_{instant}$ had a higher variance, and smoother transitions than would be expected by chance. We similarly constructed a regressor measuring Shannon's Entropy ($SH_{instant}$) for the subsequence within each sliding window. As with the procedure for creating the MHL regressor, this resulted in a regressor that assigned a numeric $SH_{instant}$ value for each volume acquired during the scan, using the formula in equation 1.

[Figure 1 around here: Generation of Markov Entropy regressors]

To identify brain regions showing a non-linear sensitivity profile to $MH_{instant}$, we applied an inverse U-shaped parabolic transformation to the MHL regressor as indicated in equation 3 and

$$MHU = -(MHL^2 - 2MHL), \quad \text{eq. 3}$$

shown in Figure 1C. This transformation re-assigns low BOLD values to higher levels of uncertainty and order, and high BOLD values to intermediate levels of uncertainty, with the highest MHU value corresponding to when MHL=1. This transformation was chosen because it results in a peak value for entropy states between complete randomness and complete order, which are states considered highly complex by several accounts, as reviewed in the Introduction. We refer to this regressor as the MHU regressor since it was used to model a U-shaped, non-linear relation between MH and BOLD responses.

To identify brain regions sensitive to an above-threshold level of order in a step-wise non-linear manner, a stepwise transform (MHS) of the MHL regressor was implemented according to equation 4, with the threshold set at 1.

$$\text{for } MHL \leq 1, MHS=0; \text{ for } MHL > 1, MHS = 1, \quad \text{eq. 4}$$

This transformation re-assigns a "0" to entropy values below the threshold (1 bit), and "1" to entropy values above the threshold and so can separate processes whose function differs between the more and less ordered parts of the entropy continuum. A brain region tracking entropy in this manner shows similar BOLD responses for any input whose order is above the threshold, and a different level of uniform response for any input whose order is below the threshold. For the distribution of Markov Entropy values in our regressor, this step-function captured 25% of the values within the lower part of the step (i.e., MHL < 1), and 75% of the values with the upper part of the step.

To identify brain regions sensitive to points in the auditory series in which a pattern of increasing order over time switches to a pattern of decreasing order or vice versa – i.e., a higher moment of order transition -- we derived a regressor modeling deflection points in the MHL regressor as determined by its first derivative (MHLD regressor henceforth). Deflection profiles were also similarly created for the MHU (MHUD) and SH regressors (SHD) (no deflection regressor was constructed for the MHS regressor since it consisted of only 0,1 values and did not contain trends). Finally, all seven experimental regressors -- MHL, MHU, MHS MHLD, MHUD, SH and SHD -- were convolved with a hemodynamic response function (see Figure 1B, 1C) to create an expected BOLD time course.

2.5 Imaging protocol

High-resolution anatomical and functional data for each participant were acquired with a 4T Bruker/Siemens system. Structural scans were acquired with a 3D T1-weighted MPRAGE sequence (TR/TE=2700/4 ms, flip angle = 7°, isotropic voxel size = 1mm, matrix=256x224; 176 sagittal slices). Two structural volumes were obtained for each participant and averaged to allow accurate image processing within our surface-based analysis pipeline. Four hundred and fifteen single-shot EPI BOLD functional images were acquired using the point-spread-function distortion correction method (Zaitsev et al., 2004) for each participant (TR/TE=1500/34 ms, 25 interleaved slices parallel to AC/PC, voxel size= 4x4x4, slice skip factor = 0.2, Ernst flip angle = 67^{o}; matrix = 64x64; 622.5s overall scan time). The first fifteen volumes were acquired to allow for T1 stabilization effects and were discarded from the analysis.

2.6 Data processing: Physiological noise cleaning, volume registration and surface projection

During the FMRI scan cardiac and respiration data were acquired using a photoplethysmograph and a respiration belt (Bruker), respectively. The impact of physiological noise on the BOLD time series was removed using a combination of AFNI (Cox, 1996) and MATLAB. The fMRI time series were de-spiked, and physiological effects were then modeled using the RETROICOR (Glover et al., 2000) procedure with 13 regressors: 4 for the cardiac series and its harmonics, 4 for the respiratory series and its harmonics, and 5 for respiration variation over time (Birn et al., 2006) and its harmonics. The physiologically-cleaned data set was retained and initiated all subsequent preprocessing and analyses of the functional data in the volume and/or surface domains.

For each participant, functional acquisitions were spatially aligned to a reference acquisition collected during the run, and data were corrected for slice-timing differences. Prior to projection of time series to the surface domain, we removed several sources of variance from the time series data via linear regression including (i) 6 motion parameters estimated during the head motion correction, (ii) linear, second-order and third-order polynomial trends, and (iii) a regressor modeling activity associated with participants' key presses during performance of the incidental task. These latter key-press motor responses were modeled using a Finite Impulse Response function corresponding to the timing of the participant's key press. No temporal or spatial smoothing was conducted as part of preprocessing of the volume data ("volume-domain"). Instead, inter-subject registration and spatial smoothing of the time series was carried out in the surface domain as detailed below.

Inter-subject registration to common space was performed using the procedures implemented in the FreeSurfer software package (v. 4.5, Fischl et al., 1999). First, the two anatomical images of each participant were co-registered and averaged to increase image signal-

to-noise ratio (SNR). The left and right hemispheres of each participant's volume were inflated to a surface representation and aligned to a common template using the warping procedures implemented in FreeSurfer. The automatic parcellation functionality of FreeSurfer was used to derive anatomical parcellations of the cortical surface for each participant for purposes of hippocampus region-of-interest (ROI) analysis in the volume domain.

Projection of functional time series to the surface domain was implemented as follows: The anatomical images were aligned to the functional data using semiautomatic procedures implemented in AFNI (Saad et al., 2009). Alignment was checked and manually adjusted if needed. The functional time series of each voxel in 3D volumetric space were then projected from the 3D EPI volumes onto the 2D cortical surface (imported into SUMA). For each participant the resulting dataset consisted of one matrix [400 time points x 196,000 vertices] per hemisphere, and it was in this 2D surface domain that all subsequent time series analyses were conducted. Spatial smoothing using a 6-mm smoothing kernel (Chung et al., 2005) was applied to the surface time series to increase their temporal SNR. A group-level mean cortical surface was created by averaging participants' individual cortical surfaces. All results were projected to this average cortical representation for display purposes.

2.7 Statistical analysis: Whole-brain regression analysis

Participants' time-series were inspected for head movement, and images acquired during periods with motion in excess of 1 mm were not included in any of the regression analyses (< 1.75% of volumes on average). Unless otherwise mentioned, the first 7 functional volumes during which auditory stimulation was presented were excluded from all analyses to eliminate the impact of the initial 'transient' of BOLD onset on the regression results. We conducted three

whole-brain regression analyses testing for three relations between $MH_{instant}$ and the BOLD response; all included the same Shannon Entropy regressor, but included different sets of Markov Entropy regressors to model the linear, U-shaped, and stepwise-relation between BOLD and Markov Entropy. The first regression model contained the SH, SHD, MHL, and MHLD regressors; the second contained the SH, SHD, MHU, and MHUD regressors; the third contained the SH, SHD, and MHS regressor.

For each participant, Beta weights estimated by the regression at each voxel were scaled by the voxel's mean (i.e., normalized) and propagated to a second-level group analysis. Family-wise Error (FWE) was controlled using cluster-based thresholding following the Monte-Carlo simulation procedures described in Forman et al. (Forman et al., 1995), which were implemented in the cortical surface domain. FWE was corrected using cluster extent thresholding ($p < .05$) with a single voxel extent threshold at $p < .01$.

2.8 Statistical analysis: Inter-subject correlation via eigenvalue synchrony

This analysis was based on the inter-subject correlation method developed by Hanson et al. (2009). In this voxel-wise eigenvalue synchrony analysis, a singular value decomposition (SVD) is applied to a matrix [M] representing the time series for all participants for a given voxel (the [15] participants x [400] acquisition matrix), where M= *U*x*d*x*V*. Hanson et al. defined a reliable inter-subject correlation in a given voxel when the largest singular value – the first element of the diagonal *d* -- was statistically significant (as assed by Roy's Largest Root), and when *all* participants showed the same loading ("+" or "-") on the coefficient corresponding to their entries in the first column of the unitary matrix *U*. This method is highly conservative, and becomes exponentially more so as the number of participants increases, as it requires complete agreement (e.g., 15/15 participants agreeing on a sign corresponds to an overly conservative

false-positive rate of $p < 0.00003$). For this reason, we relaxed the unanimity agreement criterion and required that at least 12 of the 15 participants demonstrate the same sign in the SVD solution ($p < .018$ on a binomial test). We then used cluster-based thresholding to identify clusters where all voxels satisfied this constraint.

Following Hanson et al., the pre-processing for the inter-subject correlation included a number of additional steps because of this analysis' sensitivity to outliers and global mean fluctuations. To this end, for each participant, a global-mean time series of all intracranial voxels was constructed and subtracted from each voxel's time series prior to SVD. This mean-normalized time series was then *Z*-normalized to attenuate the impact of arbitrarily large outliers. The initial 7 volume acquisitions were discarded to attenuate the impact of the initial transition from silence to auditory stimulation on the BOLD response, and to analyze only activity occurring once the BOLD activity had reached an asymptote in auditory cortex. Including the initial 7 acquisitions in the analysis revealed qualitatively different results showing extreme inter-subject correlations only in the vicinity of the primary auditory cortices (see Results).

2.9 Statistical analysis: Granger causality

Granger causality analysis (GCA) examined directional connectivity between two functionally defined networks: (a) those regions showing reliable correlation with the MHL regressor (linear relation profile, LRP) and (b) those regions reliable correlations with the MHU or MHS regressors (non-linear relations profile, NLRP). The analysis tested our hypothesis that there is directional efferent connectivity from regions showing linear relations between BOLD and uncertainty to those networks that track uncertainty in more complex manners. In addition, regions showing sensitivity to deflections in Markov Entropy, as well as those identified via the

inter-subject correlation (ISC) analysis, including the bilateral auditory cortices, formed a third set of regions that was included in the general GCA solution, but was not included in the specific hypotheses tested, as described below.

We used partial GCA (Guo et al., 2008), which is a multivariate technique that evaluates the unique impact of each regressor given the impact of all other regressors. The data submitted to the partial GCA were extracted from the center of mass for each of the 19 functional regions-of-interest. Pre-processing of the data and connectivity analysis were implemented using the Granger Causal Connectivity Toolbox (Seth, 2010) for Matlab. Each time series was truncated to 390 time points, examined for covariance stationarity, detrended, and the temporal mean was removed. Model order of 2 (lag of 2) was determined as the optimal order by the Akaike information criterion. Within GCA, this criterion is used to determine the model order by selecting the maximum lag at which the loss of information during autocorrelation is stabilized and the amount of variance accounted for by the model is optimized. For each connection the directionality was determined by contrasting the difference of influence terms. A connection was defined as reliable on the single participant level if it passed a threshold of $p < .01$. Results of the individual partial GCA solution for each participant were propagated to the group level by identifying connections that were present in more than 3 participants (3/15 participants is significant on a binomial test; $p < .0004$).

To assess the statistical significance of the mean node degree (MND) parameters on the group level, a Monte Carlo simulation with 1000 iterations was conducted. To perform this simulation, in each permutation, for each "surrogate' participant a simulated network was instantiated with the same number of connections as the original matched participant, but with randomly assigned connections. From these 15 surrogates, the MND was calculated at the group

level, and another simulation was initiated. MND values that were either greater than the 95% or lower than the 5% quantiles of this permutation-derived sampling distribution indicated statistically reliable MND values that are not likely to be found by chance given the number of connections found for each of the participants.

To test our main hypothesis concerning the connectivity between the networks demonstrating a linear response profile (LRP) and a non-linear response profile (NLRP), we conducted two paired-samples t-tests contrasting afferent and efferent connectivity within and between these networks, and a Chi-square frequency test to determine if an above-chance proportion of the participants demonstrated the expected relation.

The first analysis focused on the NLRP network and examined whether there is a difference between afferent and efferent connectivity within that network. To this end, for each participant, we counted the number of afferent and efferent connections in the NLRP network and a group level t-test was the conducted to see if there exists a reliable difference between these efferent and afferent counts. The second analysis tested the hypothesis that the NLRP network would be associated with greater afferent connectivity than the LRP network since it tracks input order in a more complex manner. To evaluate this issue, the number of afferent connections for the NLRP and LRP networks was established for each participant, and a group level t-test was conducted.

The last analysis quantified for each participant the number of connections originating within the NLRP network and terminating at the LRP network (NLRP → LRP) and vice versa (LRP → NLRP), and determined whether there is an imbalance between the number of these connections. For each participant, the number of such connections was counted, and participants were coded with 0, 1, or 2 depending on the type of pattern they showed. Specifically,

participants were coded as 0 if they showed no difference between these values, 1 if they showed more LRP→ NLRP connections, and 2 if they showed more NLRP → LRP connections. The significance level was determined using SPSS's function that uses Monte Carlo simulations to establish significance thresholds. Greater information flow from the LRP network into NLRP network would support our hypothesis that uncertainty information is shared with and operated on by other networks performing higher-order computations as a necessary part of deriving the rules of order.

2.10 Statistical analysis: Hippocampus region-of-interest

For each participant, a mask of the left and right hippocampus was established by the FreeSurfer software. These parcellations have been shown to be comparable in accuracy to that of manual parcellation (Fischl et al., 2004). For this ROI, the beta weight corresponding to the regressor of interest, as well as the percentage of voxels for which the regressor showed statistically reliable correlation were propagated from the single participant level to the group analysis.

3. Results

3.1 Performance on incidental task

Performance data for two participants were unavailable due to equipment malfunction. A third participant responded only when the fixation-cross switched from "+" to "x" but not vice versa, and was completely consistent in this pattern. The remaining participants (N = 12) demonstrated high accuracy on the visual target detection task, with a group mean of 94.5% (+/-

4.69 SD) correct responses, indicating that participants were alert during the data acquisition. Average reaction time for the group of participants was 735ms (+/- 346 SD).

[Table 1 around here: Brain regions showing a correlation with aspects of order]

3.2 Coding of order: sensitivity to diversity (Shannon's Entropy)

Shannon's Entropy ($SH_{instant}$) fluctuations were negatively correlated with the BOLD signal time course in cortical regions including the supratemporal plane bilaterally, extending into the parietal operculum and insula, and also the right posterior cingulate (see Figure 2; Table 1). The negative correlation between BOLD and instantaneous Shannon's Entropy indicates that higher entropy (greater randomness) was associated with reduced BOLD activity. These regions identified by the SH regressor were highly similar independent of whether it was inserted into a regression model with the MHL, MHU regressors, or MHS regressors.

[Figure 2 around here: Sensitivity to Shannon's Entropy]

3.3 Coding of order: linear, U-shaped, and step-function sensitivity profiles to transition dependency (Markov Entropy).

BOLD activity was linearly and positively correlated with the MHL regressor in the ventral premotor cortex (vPM) bilaterally (Figure 3; yellow clusters), indicating that increasing randomness was related to increasing activity levels in this region. Brain regions showing a U-shaped relation with $MH_{instant}$ – i.e., activity correlated with the MHU regressor -- consisted mainly of the lateral temporal cortex bilaterally (Figure 3; green clusters) but also the parietal operculum. The sign of the correlation indicated higher activity for extreme levels of order/disorder than intermediate levels. Finally, a region encompassing aspects of the right supratemporal plane extending into the right insula (Figure 3; blue cluster) demonstrated a step-

down response profile, indicating higher activity for low → mid entropy levels than for mid → high entropy levels.

[Figure 3 around here: Sensitivity to Markov Entropy]

3.4 Coding of order: sensitivity to changes in Markov and Shannon Entropy

We also probed for regions sensitive to changes in trends of entropy over time. These changes were defined as deflection points in the MHL, MHU and SH regressors. These deflections have different significance when quantified in the MHL and MHU regressors. For the MHL regressor, they correspond to switch points in which a move towards increasing randomness reverses towards order and vice versa. For the MHU regressor, they correspond to switch points where a move towards mid-scale entropy levels reverses towards either increased randomness or increased order, and vice versa.

Figure 4A shows regions sensitive to trend-changes in the MHL regressor, and Figure 4C shows regions sensitive to trend-changes in the MHU regressor. Deflections in the MHL regressor were associated with reduced activity bilaterally. For MHU, deflections were associated with increased activity. A right inferior frontal region (Figure 4B) showed greater activity for $SH_{instant}$ switches indicating sensitivity to trend-changes in the uniformity of the input.

[Figure 4 around here: Sensitivity to deflections in non-stationary order]

3.5 Coding of order: Model free inter-subject correlation.

To this point, our analyses assessed specific BOLD:$MH_{instant}$ relations, operationalized in the form of specific linear, U-shaped and stepwise transfer functions between instantaneous Markov Entropy and BOLD activity. There are, however, numerous possible relations, i.e., link

functions, that could be assumed beyond those we had explicitly probed for. To identify entropy-sensitive regions without *a-priori* assumptions about the possible linking functions we conducted a model-free analysis based on inter-subject synchronization. The analysis revealed four clusters that were significantly synchronized across participants, shown as clusters of different colors in Figure 5A (see also Table 2).

In the right hemisphere, one cluster was found in the PCC/Precuneus region, a second bordering the lingual gyrus and calcarine sulcus, and a third in fusiform cortex. A fourth cluster was found in the left lingual gyrus. Note that the changes in the shape of the fixation-cross had occurred at different timings for each participant and so cannot account for synchronized activity patterns, but only counteract the possibility to identify them.

In each cluster, we extracted the time series of the eigen vector accounting for maximal inter-subject correlation and evaluated its relation to order fluctuations to attempt to quantify how input order was related to the shared activity pattern in each cluster. We first note that the three eigenvectors extracted from the lingual gyri bilaterally and the right fusiform (clusters #1, #2, #3 in Figure 5A) showed strong pair-wise correlations amongst themselves (Pearson's Rs > 0.33), but none were reliably correlated with the eigenvector extracted from the left PCC/Precuneus. Second, none of these eigen vectors showed correlation with SH. For the time series of the eigenvector extracted from the PCC/Precuneus (Figure 5A cluster #4), a step-down function between BOLD and MH explained a significant proportion of the variance ($r = 0.22$, $p < .00001$) indicating that lower entropy was associated with greater activity in the region. A similar pattern was found for the eigenvector extracted from the left visual cortex, (Figure 5A cluster #1), where a step-down function offered the best fit for the negative relation between entropy and activity ($r = 0.16$, $p = .001$).

The impact of including data acquired during the first 10 seconds of the scan (7 acquisitions that were discarded in other analyses) on the ISC result is illustrated in Figure 5B. The bilateral auditory cortices are clearly synchronized across participants due to the onset of stimulation and the resulting transient in the BOLD signal. This transient can be seen in the eigenvector extracted from the synchrony of these two regions in that it was highly similar to a BOLD impulse response function. The importance of this result is in its internal validation of this analysis method for the current dataset.

[Table 2 around here: Clusters showing inter-subject synchronization]

[Figure 5 around here: regions showing inter-subject correlations during auditory presentation]

3.6 Sensitivity to order in the hippocampus

For each participant, activity in the hippocampus was quantified via two measures: the mean Beta value in the region for each of the theoretical regressors of interest, and the percentage of voxels in the region showing reliable negative or positive correlation with each regressor. Group analysis of the Beta values revealed no reliable sensitivity to any of the theoretical regressors of interest in the hippocampus; i.e., the mean Beta on the group level did not depart significantly from zero. The analysis of percentage of active voxels demonstrated an above-chance percentage of voxels in the right hippocampus that were sensitive to Shannon's Entropy as well as being linearly related to Markov Entropy (M = 18% and 12% of voxels, ps < .05). This suggests that hippocampal sensitivity to order, if existent, is manifested in both above- and below-baseline activity in different subregions of the hippocampus.

3.7 Network connectivity among regions sensitive to transition constraints.

In this analysis we examined the effective connectivity amongst regions showing a linear response profile to $MH_{instant}$ (LRP network) and those showing a non-linear response profile to $MH_{instant}$ (NLRP network). Within the NLRP network, finding greater afferent than efferent connectivity would support the hypothesis that it mediates up-stream processes dependent on the level of order in the environment. Stronger afferent connectivity for the NLRP network compared to the LRP network would support the hypothesis that the latters plays a greater role in mediating lower-level processes.

Figure 6 presents the regions included in the analysis, and for purpose of clarity we classify them according to their response profiles. The LRP network included three clusters (2, 12, and 16). The NLRP network included four clusters comprised of three regions correlated with the MHU regressor and a region correlated with the MHS regressor (4, 9, 11, and 7, respectively). A third set of regions consisted of eight clusters demonstrating sensitivity to deflections in $MH_{instant}$, as well as the left and right primary auditory cortices identified with the ISC analysis. Figure 6 presents group-level afferent and efferent connectivity for all nodes. As the figure shows, there were many connections between the LRP and NLRP networks, indicating that rather than being segregated there was significant information transfer between them.

The first hypothesis we evaluated concerned directional connectivity within the NLRP network. The results indicated a greater number of afferent than efferent connections for this network, $t(14) = -1.987$, $p = .0335$ (one-tailed). This finding confirms our expectation that regions showing non-linear response profiles have an important role in information integration. Our second hypothesis concerned the relative prominence of afferent connections in the NLRP and LRP networks. The data confirmed our prediction that the NLRP network showed more afferent connectivity than the LRP network $t(14) = -3.594$, $p = .0015$ (one tailed), further

supporting the NLRP network's importance for higher-level information integration. We conducted a non-parametric chi-square analysis to determine if on the single participant level, there was an imbalance between the number of LRP → NLRP connections and NLRP → LRP connections. Results of the analysis indicate that a significantly greater number of participants than expected by chance showed the hypothesized relation of stronger efference from the LRP network $\chi^2(2) = 7.6$, $p = .029$. This finding suggests that the NLRP network receives more information from the LRP network than vice versa.

[Figure 6 around here: Lag 2 Granger causal connectivity between order-sensitive regions]

Finally, we wanted to know whether any of the regions identified examined in the GCA was associated with a degree of connectivity greater than would be expected by chance. Monte Carlo simulations established a random sampling distribution against which we could estimate the Node Degree of each of these regions. The results of the simulations indicate that the system under investigation was not the product of a random process as 11 of the 19 nodes were associated with node degrees outside the 5%-95% chance interval ($4 > MND > 5.67$). Eight nodes showed significantly less connectivity than expected by chance including two in the left hemisphere (nodes 4, 9) and six in the right hemisphere (nodes 10, 13, 14, 17, 18 and 19). Three nodes showed significantly greater connectivity than expected by chance including two clusters in the left lingual gyrus (nodes 7, 8) and one cluster in the right middle temporal gyrus (node 11). Two of the three regions showing greater connectivity than chance included a region correlated with the MHU regressor (node 11) and one correlated with the MHS regressor (node 7). The fact that nodes showing a non-linear response profile demonstrated high connectivity suggests their importance in responding to environmental uncertainty.

4. Discussion

In the current work we examined the relation between the BOLD signal and two determinants of order during auditory stimulation that was incidental to the ongoing task. Participants monitored and responded to a randomly changing visual fixation cross while they listened to a train of tones. This tone sequence was generated by a non-stationary first order Markov process that produced a continuous sequence of 4 alternating tones that had fluctuating values of entropy when calculated over 4.5sec sliding windows. A set of analyses consisting of regression models, ISC and GCA revealed that tracking statistics is mediated by interacting cortical regions that respond to different aspects of uncertainty (Shannon's Entropy; SH and Markov Entropy, MH) and that exhibited linear (BOLD:MHL), U-shaped (BOLD:MHU), and step-down (BOLD:MHS) response profiles. An effective connectivity analysis documented asymmetry in connectivity between regions showing linear and non-linear response profiles to MH, indicating stronger communication from the former to the latter, and that regions showing non-linear responses are associated with more afferent than efferent connectivity. All this suggests that regions showing non-linear responses are related to higher-level secondary processes related to the degree of order in the environment. Finally, we identified several cortical regions that were responsive to deflections in the entropy trends themselves, indicating sensitivity to the non-stationary, fluctuating nature of the transition matrices from which the tone sequence was generated.

These findings confirm prior work showing that BOLD signal levels vary with sequence entropy (Bischoff-Grethe et al., 2000; Harrison et al., 2006; Overath et al., 2007; Schubotz et al., 2004; Strange et al., 2005), but extend these in several ways. First, they indicate that order is

related to both linear and non-linear response profiles, and that responses to order are not mediated by a single region, but by distributed networks. Second, they show there exists systematic connectivity patterns between regions showing these different response profiles. Third, they indicate that the tracking of order is performed spontaneously in contexts lacking any decision making component and devoid of any necessity to pay attention to the stimuli whose order is manipulated. Fourth, they show that in a context where order is non-stationary, people track patterns of changes in order levels and are sensitive to deviations in these patterns. Finally, they raise some important qualifications about the potential role of the hippocampus in the coding of order. In what follows we discuss these main findings and their implications for theory development in the field.

4.1 Linear response profiles and their functional significance

The regression analysis identified BOLD signal changes tracking both $SH_{instant}$ and $MH_{instant}$. Shannon's Entropy indicates a degree of predictability of the current observation given prior information which is derived solely from the relative frequency of different tokens, independent of relational information. This information about diversity/frequency is a basis for prediction since in absence of other information, at any given point it is rational to expect the more frequent tokens. A brain region generating predictions in this manner would therefore show greater activity for low-entropy states, i.e., a negative BOLD:$SH_{instant}$ relation. We indeed found such a relation in perisylvian regions of the lateral temporal cortex bilaterally, as well as the insula/claustrum, inferior parietal cortex, and cingulate gyrus. Markov Entropy indicates the predictability of the forthcoming observation taking into account relational information (transition constraints), and was linearly related to the activity level of the ventral premotor

cortex bilaterally, indicating the region shows stronger BOLD response for weaker transition constraints.

Linear relations between BOLD responses and order can be interpreted both in terms of the representation of the degree of entropy, and in terms of processes afforded by entropy. On a neurobiological account inspired by the Bayesian Observer model for the coding of input statistics (e.g., Strange et al., 2005), hippocampal activity has been interpreted as coding for entropy since different levels of activity in the region are associated with different levels of input entropy. As pointed out by Strange et al., this type of internal representation of the statistical features of the environment is fundamental for affording myriad functions including balancing top-down and bottom-up information during perception, generating predictions of the forthcoming sensorium, detecting error signals/mismatches and updating statistical priors accordingly.

However, while a linear relation is a necessary condition for a region to be considered as creating an index for entropy, it is not a sufficient one, and such relations can be interpreted as indicating particular processes linked to the degree of input entropy rather than coding for entropy *per se*. For instance, a negative BOLD:H relation, i.e., greater activity for more ordered conditions, could indicate the involvement of a region in the construction of predictions regarding future inputs, since such predictions are only afforded by low-entropy states. Conversely, a positive BOLD:H relation, i.e., greater activity for less ordered conditions, could indicate the involvement of a region in processing error signals, which would occur more frequently the less ordered an input is. It is in light of these considerations that we interpret the linear relations identified for BOLD:$SH_{instant}$ and BOLD:$MH_{instant}$.

For the BOLD:$SH_{instant}$ relation we identified a negative correlation between uncertainty and BOLD signal change, largely in regions known to code for both simple sounds such as pure tones, but also more complex broadband and speech sounds not typically activated by pure tones (Chevillet et al., 2011). The regions we identify are associated with well-known effects indicating evaluation of current against prior input (e.g., Mismatch-Negativity effects, Schönwiesner et al., 2006), or pattern completion in deterministic sequences (Bekinschtein et al., 2009). Thus, increased activity for more ordered states would be consistent with the possibility of the generation and evaluation of predictions when allowed by statistics, or concomitant error processing when prediction errors occur.

The negative BOLD:$SH_{instant}$ relation is also important in that it rules out several alternative mechanisms that could account for relations between entropy and BOLD. First, it demonstrates that the effect was not driven by neural repetition suppression. Greater sequence order quantified via SH is isomorphic to a less uniform sampling distribution and therefore analytically related to an increasingly frequent presentation of some tokens over others, which leads to a greater proportion of consecutive self-repetitions. To illustrate, a random sequence drawn from a uniform distribution of 4 tokens has a 25% self-repetition rate, i.e., cases in which an item is repeated in succession (e.g., *1-1*-4-3-1-*2-2*-3-1). More ordered sequences, as quantified by SH are less uniform and therefore have more self-repetitions. Therefore, a positive BOLD:$SH_{instant}$ correlation (lower activity with more order) could derive from neural repetition suppression, rather than from coding for the degree of order. In contrast, our finding of a negative BOLD:$SH_{instant}$ profile suggests these regions respond to SH in a manner that counteracts whatever impact neural repetition suppression may induce.

Second, from a computational perspective, greater SH also enables a relatively simple coding scheme whereby more frequent elements are more efficiently coded than less frequent ones by using fewer bits for their representation (Huffman, 1952). A coding scheme of this sort would result in a positive BOLD:SH profile, since more ordered contexts would be associated with less activity as a result of more efficient representation. While such a coding scheme may be neurobiologically plausible, our data provide no support for this mechanism as well as they show higher activity for more ordered contexts. To conclude, the negative linear BOLD:$SH_{instant}$ profile we document is incongruent with a neural repetition suppression account, or with one on which redundancy is associated with more efficient representation, It is however consistent with the possibility that these regions either code for order using an index where greater activity is associated with lower entropy, or that these regions show greater activity for low-order conditions since they are involved in the coding of auditory predictions. Our data cannot distinguish between these two explanations.

For the BOLD:$MH_{instant}$ profile we identified a positive correlation between uncertainty associated with transition constraints and BOLD signal in the ventral premotor cortex bilaterally. This finding converges on the same region identified by Schubotz et al. in numerous experiments examining the neural coding of predictions from ordered sequences (see Bubic et al., 2010; Schubotz, 2007 for reviews). However, whereas those studies typically identify greater activity for ordered vs. random sequences, we identify the opposite pattern, with greater activity for more random sequences. A possible explanation for the divergence is that in contrast to prior work, the current study used a fast presentation rate and was absent of any specific demand to monitor the series being presented or to make explicit predictions. We suggest that in absence of task demands, the region is not involved in generating predictions, and that the negative linear

BOLD:$MH_{instant}$ profile indicates that the region either codes for the degree of MH in the input, or alternatively, is involved in the calculation of error terms for failed predictions, which would occur more frequently in higher entropy states. It is definitely possible that when one is explicitly asked to monitor inputs for the purpose of constructing predictions, this region can additionally code for predictions as suggested by Schubotz et al., and this is indeed an important question for future work.

As discussed to this point, it is likely that some brain regions code for the degree of order in the environment, whereas others are associated with specific processes related to this entropy-index. Linear BOLD:H relations of the sort discussed to this point are compatible with both possibilities. These regions may be part of an intricate system in which the coding of order engages down-stream "second-order" processes in a different network of brain regions, i.e., processes whose operation relates to the degree of order in the environment, but whose activity does not track order in a linear manner. We considered that such regions may show non-linear BOLD:H relations, and indeed we identified several regions showing such non-linear profiles, indicating they neither represent nor compute entropy, but likely utilize or integrate information about environment entropy for diverse functions.

4.2 Non-linear response profiles and their functional significance

We examined two sorts of non-linear responses to input entropy, a step-down response function and a curved, U-shaped response function (see *Methods* for formal definitions). Both of these response profiles cannot be taken as an index of entropy because they manifest a many-to-one relation between entropy levels and the BOLD response so that different levels of input entropy generate similar BOLD responses. In the case of the U-shaped function, such responses are

orthogonal to the degree of entropy itself. The anterior right insula demonstrated a step-down response profile, which associated lower-to-mid levels of entropy with a higher BOLD response than for mid-to-high levels of entropy, suggesting this region is involved in the generation and evaluation of predictions, as it shows stronger activity in low-entropy contexts conducive to such processes.

In addition, we identified lateral temporal regions bilaterally as demonstrating a U-shaped relation to input entropy so that BOLD responses to both highly random and highly ordered conditions were stronger than to moderate entropy levels. U-shaped and inverse U-shaped responses to parametrically varying inputs have been identified in several prior works. For instance, it has been shown that certain prefrontal regions track task difficulty in a U-shaped manner with higher activity for both simple and complex tasks relative to moderately difficult ones (Dumontheil et al., 2010). This pattern has been interpreted in terms of two different processes mediated by the region: when a task is easy, mind wandering may be responsible for the higher activity level, and when the task is difficult, high activity may index thoughts triggered by the task. In this view, maximal attention to the external input *per se*, in absence of internally generated elaboration, is associated with the least activity. An inverse U-shaped response between BOLD activity levels and working memory load had been identified in the hippocampus and bilateral STS regions (Zaharan et al., 2005), and one possible interpretation put forward by the authors was that different types of cognitive processes mediate low-load and high-load maintenance.

Our account for the U-shaped relation between entropy and activity in bilateral temporal regions identified here follows the principle underlying such explanations, and is based on the premise that the profile indicates two different processes mediated by the region. On the one

hand, the higher activity for low-entropy than mid-entropy states could indicate it is implicated in the generation of predictions and their evaluation. On the other hand, the fact that high-entropy states were are also associated with higher activity than mid-entropy states could be related to the fact that high-entropy states are particularly non-engaging and known to be associated with a greater propensity of unrelated thoughts (see Antrobus, 1966 for data linking entropy to stimulus independent thoughts).

4.3 Other entropy-sensitive profiles identified via inter-subject correlations

In addition to the linear and non-linear BOLD:MH profiles identified by the GLM, the ISC analysis also revealed four posterior medial cortical clusters whose time courses were not only synchronized across participants but also demonstrated linear and non-linear BOLD: $MH_{instant}$ profiles.

Because all participants heard the exact same stimulus, it could theoretically be the case that physical features of the stimulus such as the particular shifts in frequencies between the 4 tones used could induce inter-subject synchronization of the BOLD signal across participants, but this interpretation is not strongly supported by our data, as the regions identified do not include ones associated with processing of low-level auditory features such as tonality or pitch. Instead, these findings suggest that the synchronization was driven by higher level features of the stimulus; namely, its statistical structure. Indeed, synchronized time courses extracted via eigenvectors from these regions were related to $MH_{instant}$ either linearly or via a step-down function. Furthermore, these regions of synchronized BOLD activity demonstrated significant causal connectivity in a network consisting of nodes identified by the regression models. In particular, consistent with prior work (Fransson and Marrelec, 2008), the precuneus showed strong connectivity with many other regions.

These findings are consistent with prior work showing synchronization of posterior midline regions to auditory stimulation. For example, reliable ISC during narrative comprehension was reported by Wilson and colleagues (2008), and the sensitivity of the precuneus to auditory input has also been recently demonstrated (Hertz and Amedi, in press). Furthermore, the findings are consistent with prior work showing that posterior midline regions play a role in the representation of sequences; specifically, they show stronger connectivity with frontal regions during inferences based on representing sequence order (Rogers et al., 2010) and respond more strongly to auditory stimuli deviating from expectation (Bekinschtein et al., 2009). Thus, it is likely that these regions demonstrating synchronization across participants are an important part of the brain system that codes for order and subserves associated processes.

4.4 Asymmetric information transfer between regions coding linearly and non-linearly for Markov Entropy

To evaluate our hypothesis regarding information transfer between regions showing a linear response profile (LRP) and those showing a non-linear responses profile (NLRP) to Markov Entropy, we conducted an effective connectivity analysis using Granger Causality. The first finding, that the NLRP network has more afferent than efferent connectivity confirms the notion that these regions code for a higher level representation of entropy and play an important role in integration of statistical information. We suggest that the relatively few efferent connections for NLRP regions, at least within the set we analyzed, indicates that the computational output of NLRP regions may be transmitted to a relatively limited set of network hubs to serve as top-down information. Identifying these targets is an interesting topic for future work. The second finding, that the NLRP network has greater afferent connectivity than the LRP network lends further support to the interpretation of NLRP activity as a higher-order representation of entropy

that integrates multiple sources of information. Finally, the finding that participants showed greater efferent connectivity from LRP to NLRP than vice versa suggests that the linear coding of entropy is an important source of information which is then integrated at the sites showing NLRP activity patterns. To conclude, these findings suggest that not only are there regions that code for entropy linearly and non-linearly, but that these latter regions operate on information received from regions coding linearly for entropy.

4.5 Sensitivity to deflections

Due to the non-stationary progression of order fluctuation in the current study we could identify deflection points marking a temporal point where a trend of either increasing or decreasing entropy is interrupted. Formally, these were points where there was a reversal in the sign of the first derivative of entropy values of the subseries in the stimulus. Several clusters showed significant responses to deflections, including one for SH, eight for MHL, and one in the lateral prefrontal cortex for MHU. Some of these regions, including the pre-SMA, calcarine and lingual gyri, have been documented as tracking auditory stimuli that depart from a complex pattern that can only be established by considering a large set of prior input (Bekinschtein et al., 2009).

It is important to note that these regions are not involved in computing entropy of subsequences within the recent past, but are instead responsive to the history of computed entropy values and patterns of changes in order. Because prior work has examined series where order was stationary, no previous reports of deflections are available in the literature to compare our results. We suggest that responses to such deflections indicate a particular sort of prediction error signal, but one which is associated not with the prediction of a particular stimulus or event, but which refers to the degree of expected order in the near future. These responses to deflections in the progression of entropy levels themselves indicate the complex nature of

processes that are triggered by the degree of input entropy but are not related to the coding of sequence entropy per se.

4.6 Relation to Bayesian Observer accounts

Prior work has advocated a Bayesian Observer model of perceptual brain function (Strange et al., 2005; Harrison et al., 2006), which interpreted BOLD responses to input entropy in terms of a neural code indexing the degree of input uncertainty. By this account, a neural mechanism postulated to be mediated by the hippocampus, regularly updates its representation of environmental statistics on the basis of new information, and generates a different response for different levels of order. A linear BOLD:H relationship, particularly in the hippocampus, would therefore support a Bayesian Observer model. We found such linear relations for $SH_{instant}$ and $MH_{instant}$, in perisylvian and ventral-premotor cortices, respectively, but not in the hippocampus. However, as a whole our data only partially support the generality of a Bayesian observer account for entropy responses. First, as we noted above, while a linear BOLD:H relationship could be interpreted as coding (indexing) the degree of entropy, it is also consistent with the possibility of the generation of more predictions when entropy is low, the evaluation of more errors when entropy is high, or even lower-level neural effects not related to any computation such as particular repetition suppression patterns that are linked to the degree of Shannon's Entropy. Furthermore, the extensive non-linear relations identified here suggest that BOLD responses to order do not reflect just the degree of order detected in the input, but also other entropy-associated processes that are stimulated by changes in order or uncertainty, including but not limited to (*i*) the generation of predictions, (*ii*) the evaluation of predictions against input, and/or (*iii*) the updating of statistical knowledge based on error terms for failed predictions. This

demonstrates that beyond coding for entropy, multiple processes are co-activated while processing the order of temporally extended event sequences.

4.7 Functional heterogeneity of the hippocampus

The current study examined whether the hippocampus codes for transition constraints when stimuli are presented rapidly and in a modality irrelevant to sensory-motor processing. Strange et al. (2005) and Harrison and colleagues (2006) reported that the BOLD signal in the left hippocampus was correlated with fluctuations in Shannon Entropy and the mutual information, respectively. While we indeed identified several networks coding for both $SH_{instant}$ and $MH_{instant}$, indicating sufficient experimental power, our findings did not indicate a clear involvement of the hippocampus in representing either type of entropy exclusively. Specifically, while the number of voxels showing a reliable correlation with $SH_{instant}$ and $MH_{instant}$ exceeded chance, these were divided between positive and negative correlations for any given regressor and consequently, on the group level there was no above-chance correlation, either positive or negative, with any of the regressors. Due to the heterogeneity of correlations and because voxel counts are not a reliable indicator of activation (Cohen and Dubois, 1999), we could not reliably associate hippocampal activity with the coding of order.

Prior work indicates that the hippocampus signals a mismatch between expected and encountered stimuli due to its ability to complete sequences based on partial cues, which in turn affords noticing discordances with these expectations (Kumaran and Maguire, 2006). Other work (Turk-Browne et al., 2010) has shown that the right hippocampus codes for the cue validity of single items. In the cue validity paradigm, highly diagnostic items are informative with respect to what might follow in a manner similar to that of the transition constraints of first order Markov Entropy. A possible explanation for the divergence between our findings and these prior

studies is that they used slow presentation rates (e.g., 1Hz or less) and tasks in which participants were required to direct attention to the stimuli whose order was manipulated. Our more rapid presentation rate could have made it difficult to continuously recall extended sequences based on partial cues. Moreover, even when ordered states existed within sub-sequences of the tone series, these were never completely deterministic in that a cue never indicated what was to follow in an unambiguous manner. This is a crucial difference with respect to prior work in which the beginning of a series was uniquely associated with its ending. The difference between these two types of "prediction" has been noted previously (Grossberg, 1999) and addressing this difference is a topic for future work.

4.8 Limitations and future work

Our work quantified order-related responses to recently encountered stimuli during passive perception. Because five stimuli were presented within each functional volume, our fMRI paradigm could not dissociate systems coding for input order from those tracking the degree of surprise of each individual stimulus. While entropy can be considered as the average surprise of all possible outcomes, series entropy is partially dissociable from stimulus-specific surprise, and prior work using visual stimuli (Strange et al., 2005) documented stimulus-specific surprise effects in premotor, inferior-parietal and occipital regions. However others (Huettel et al., 2002) have documented surprise-related effects in frontal regions and the anterior cingulate for items departing from prior streaks, as well as a correlation between the magnitude of these effects and the length of the prior streak being violated. Work by Bubic et al (2011) indicates that pattern-violating events are registered in different neural systems depending on whether the pattern violated is based on transitions between temporal intervals, spatial configurations, or

specific objects. Finally, numerous auditory studies link surprise effects to the superior temporal plane (e.g., Mustovic et al., 2003). Whether stimulus-specific surprise-related responses are dissociable from sequence-scale entropy responses in contexts such as the one we examine, which are devoid of strategic decisions making, and the relation of such responses to input modality are questions for future work.

Another topic for future research is establishing the parameters determining information decay rate, because it is likely that less recent information has a weaker impact than more recent one in determining the statistical representation of the past. Recent work (Harrison et al., 2011) suggests that when people make decisions that are sensitive to the statistical features of the recent past, the four most recently presented items are weighted particularly strongly in guiding people's behavior. That work suggests that the anterior cingulate is sensitive to entropy fluctuations on this time scale, but posterior regions are sensitive to fluctuations on shorter time scales. The aforementioned work presented stimuli using a stimulus onset asynchrony (SOA) of 2.2sec, and an important topic for future work would be to evaluate such issues as whether this decay rate is a function of temporal interval (e.g., ~8.8sec), the SOA interval, the entropy of the process generating the events, or the number of expectation-violating events in the recent past. Our current work was not designed to address this issue since the stimuli were optimized so that $MH_{instant}$ and $SH_{instant}$ values fluctuated strongly within short time intervals, and increasing the window size in that case results in a much more limited range of values, and consequently, in a design that is not optimal for identifying fluctuations on longer time scales.

5. Conclusions

Our findings indicate that the degree of entropy in temporally extended event sequences may be associated with either positive or negative linear relations between BOLD and entropy, as well as non-linear response profiles. We outlined several mechanisms that may underlie these results, including the construction of an index coding for the degree of input entropy, but also a range of entropy-related processes. Our work extends prior findings in three central aspects. First, they systematically search and identify non-linear entropy related responses indicative of entropy-related processes that cannot be considered as coding for entropy *per se*. Second, a network analysis shows that regions that code for order in linear and non-linear manners maintain an asymmetric information transfer between them, with stronger efference from the former to the latter. Third, they show that the human brain is sensitive not only to the statistical features of series, but also to time-variant (non-stationary) features of order such as changes in the trends of order progression. Finally, the fact that these findings were identified in the context of a study where participants did not need to pay attention to the stimuli whose order was manipulated, and in a stimuli with a high presentation rate suggests that all these processes take place spontaneously under challenging conditions.

The issues we raise play an important role within recent theoretical frameworks focusing on predictive brain functions, which have outlined the computational mechanisms that may underlie the processing of regularity in the auditory domain (Winkler et al., 2009) and mismatch detection/resolution during sensory-perceptual inference (Friston and Kiebel, 2009; Summerfield and Egner, 2009). These frameworks share the premise that predictive codes play an essential role in brain function, and two of these (Friston and Kiebel, 2009; Winkler et al., 2009) attribute a particular weight to how regularities may drive prediction. In addition, prior theoretical work (e.g, Grossberg, 1999) has noted that beyond prediction, order/structure may be important in

"post-diction" of the sort that enables recovery of degraded information that has been recently encountered and still stored in short-term memory.

The neurobiological framework emerging from our study suggests the brain: 1) is continuously sensitive to the degree of order in the environment regardless of goal-directed sensory-motor processing, 2) extracts multiple aspects of statistical information and responds to them via multiple response profiles, 3) is sensitive to order on both short time scales (~4sec) and longer ones, as evident by deactivation elicited by deflections, and 4) consists of brain regions with a high degree of functional connectivity that code both sensory input and integrate diverse information from other nodes in the network. Future work may seek to dissociate systems that code for the degree of environmental order from those that react to special classes of stimuli occurring within ordered sequences, such as diagnostic stimuli that contain information about the future, or surprising stimuli that refute such expectations.

Figure Legends

Figure 1: Generation of Markov Entropy regressors corresponding to non-stationary order within a continuous series of presented tones. Panel A: Construction of Markov Entropy (MH) regressor from extended tone series. For each volume, MH in the recent 15-tone subsequence was calculated, generating a continuous series of MH values (e.g., 1.3., 1.4…, etc). The initial tone series was constructed in a way that assured continuous transitions of MH. This series of MH values was convolved with a hemodynamic response function and used as a regressor in a regression model (panel B). Panel C: An inverse-U transform of the MH series (MHU) was computed to identify regions with similar responses to high and low order, and differential responses to mid-levels of order. The resulting series was convolved with an HRF and used in a separate regression model. The same procedure was applied to generate a step-up transform function of MH and a regressor for Shannon Entropy (these two regressors not shown in figure).

Figure 2: Regions demonstrating sensitivity to Shannon's Entropy in the recent past. All clusters showed a negative correlation between Shannon's Entropy and BOLD activity indicating higher activity for more ordered states.

Figure 3: Regions demonstrating sensitivity to Markov Entropy in the recent past. The figure shows regions demonstrating a positive linear relation with MH (yellow), a U-shaped relation (green), or a step down-relation (blue). The black outline marks regions showing sensitivity to SH.

Figure 4: Regions demonstrating a significant response to deflections in entropy progression as defined by the first derivative of entropy changes. A: Eight functional regions showed a decreased BOLD signal in response to deflections in MH. B: The right inferior frontal gyrus demonstrated a significant BOLD response to deflections in SH. C: A left lateral prefrontal region demonstrated a significant positive BOLD response to deflections in MHU. Cluster colors in the figure delineate different cluster and do not indicate a quantity.

Figure 5. Regions showing inter-subject correlations during auditory presentation. Regions were identified via a Singular Value Decomposition analysis applied to each voxel. The analysis identified voxels and clusters whose activity was driven by a common component. The time

series of these components (plotted below) were extracted from voxels showing the strongest inter-subject correlation. Panel A: Inter-subject synchronization identified in visual cortex and

PCC after clipping the first seven acquisitions associated with the initial activity transient. Panel B: Bilateral primary auditory cortices demonstrate inter-subject synchronization driven by an initial transient corresponding to the initiation of the tone sequence.

Figure 6. Granger-Causal Connectivity for clusters sensitive to Markov Entropy.

Panel A: Regions evaluated in the GCA were those in which activity correlated with Markov Entropy, or with deflections in Markov Entropy, as well as regions in vicinity of primary auditory cortices. The specific relation to Markov Entropy is color coded as indicated in the lower portion of the panel. Node numbers in Panel A correspond to those mapped in Panel B.

Panel B: The panel shows which connections were reliable for each node at the group level. Blue arrows depict efferent connections and green arrows depict afferent connections. In the group network, a red line indicates a bi-directional connection. Nodes from the left hemisphere are labeled 1-9, with node 5 representing the left primary auditory cortex detected by ISC. Nodes from the right hemisphere are labeled 10-19, with node 19 representing the right primary auditory cortex detected by ISC. Nodes marked with a '+' indicate greater connectivity (Node Degree) than expected by chance. Nodes marked with a '–' indicate fewer connections than expected by chance. Note the exclusive efferent connectivity of nodes 15-17 representing the posterior midline cortices.

References

Antrobus J (1968) Information theory and stimulus-independent thought. British Journal of Psychology 59:423-430.

Bekinschtein T, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L (2009) Neural signature of the conscious processing of auditory regularities. Proceedings of the National Academy of Sciences 106:1672-1677.

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

Bubic A, von Cramon DY, Schubotz RI (2010) Prediction, cognition and the brain. Frontiers in Human Neuroscience 4:1-5.

Bischoff-Grethe A, Proper SM, Mao H, Daniels KA, Berns GS (2000) Conscious and unconscious processing of nonverbal predictability in Wernicke's area. Journal of Neuroscience 20:1975-1981.

Bubic, A, von Cramon DY, Schubotz RI. (2011) Exploring the detection of associatively novel events using fMRI. Human Brain Mapping 32:370-81.

Buiatti M, Peña M, Dehaene-Lambertz G (2009) Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. Neuroimage 44:509-519.

Chevillet M, Riesenhuber M, Rauschecker J (2011) Functional correlates of the anterolateral processing hierarchy in human auditory cortex. Journal of Neuroscience 31:9345-9352.

Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC (2005) Cortical thickness analysis in autism with heat kernel smoothing. NeuroImage 25:1256-1265.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research 29:162-173.

Dumontheil I, Gilbert S, Frith C, Burgess P (2010) Recruitment of lateral rostral prefrontal cortex in spontaneous and task-related thoughts. Q J Exp Psychol 63:1740-1756.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis II: inflation, flattening, and a surface-based coordinate system. NeuroImage 9:195-207.

Fischl B, Kouwe Avd, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically parcellating the human cerebral cortex. Cerebral Cortex 14:11-22.

Fiser J, Berkes P, Orban G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends in Cognitive Sciences 14:119-130.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magnetic Resonance in Medicine 33:636-647.

Fransson P, Marrelec G (2008) The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. NeuroImage 42:1178-1184.

Friston K, Kiebel St (2009) Predictive coding under the free-energy principle. Philosophical transactions of the Royal Society of London Series B, Biological sciences 364:1211-1221.

Garner (1962) Uncertainty and Structure as Psychological Concepts. John Wiley and Sons, New York.

Glover GH, Li TQ, Ress D (2000) Image based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magnetic Resonance in Medicine 44:162-167.

Grossberg St (1999) The link between brain learning, attention, and consciousness. Consciousness and Cognition 8:1-44.

Guo St, Seth AK, Kendrick KM, Zhou C, Feng J (2008) Partial Granger causality--Eliminating exogenous inputs and latent variables. Journal of Neuroscience Methods 172:79-93.

Hanson SJ, Gagliardi AD, Hanson C (2009) Solving the brain synchrony eigenvalue problem: conservation of temporal dynamics (fMRI) over subjects doing the same task. Journal of Computational Neuroscience 27:103-114.

Harrison LM, Duggins A, Friston KJ (2006) Encoding uncertainty in the hippocampus. Neural Networks 19:535-546.

Harrison L, Bestmann S, Rosa M, Penny W, Green G (2011) Time scales of representation in the human brain: weighing past information to predict future events. Front Hum Neurosci 5.

Hebb D (1958) The motivating effects of exteroceptive stimulation. American Psychologist 13.

Hertz U, Amedi A (in press) Disentangling unisensory and multisensory components in audiovisual integration using a novel multi-frequency fMRI spectral analysis. NeuroImage.

Huettel, S, Mack PB, McCarthy G. (2002) Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. Nature Neuroscience 5:485-490.

Huffman DA (1952) A method for the construction of minimum-redundancy codes. Proceedings of the IRE 40:1098-1101.

Knowlton B, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. Learning & Memory 1.

Kumaran D, Maguire EA (2006) An unexpected sequence of events: mismatch detection in the human hippocampus. PLoS Biol 4:e424.

Li W (1991) On the relation between complexity and entropy for Markov chains and regular languages. Complex Systems 5:381-399.

Loewenstein G (1994) The psychology of curiosity: A review and reinterpretation. Psychological Bulletin 116:75-98.

McNealy K, Mazziotta JC, Dapretto M (2006) Cracking the language code: Neural mechanisms underlying speech parsing. Journal of Neuroscience 26:7629-7639.

Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., et al. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. NeuroImage*,* 20*:*429-434.

Overath T, Cusack R, Kumar S, von Kriegstein K, Warren JD, Grube M, Carlyon RP, Griffiths TD (2007) An information theoretic characterization of auditory encoding. PLoS Biology 5:e288.

Reber AS (1967) Implicit learning of artificial grammars. Journal of Verbal Learning & Verbal Behavior 6.

Rogers BP, Avery SN, Heckers St (2010) Internal representation of hierarchical sequences involves the default network. BMC neuroscience 11:54.

Saad Z, Glen D, Chen G, Beauchamp M, Desai R, Cox R (2009) A new method for improving functional-to-structural alignment using local Pearson correlation. NeuroImage 44:839-848.

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274:1926–1928.

Schönwiesner M, Novitski N, Pakarinen S, Carlson S, Tervaniemi M, Näätänen R (2006) Heschl's Gyrus, Posterior Superior Temporal Gyrus, and Mid-Ventrolateral Prefrontal Cortex have different roles in the detection of acoustic changes. Journal of Neurophysiology 97:2075-2082.

Schubotz R, von Cramon DY (2002) Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. NeuroImage 15.

Schubotz RI, Von Cramon DY (2004) Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. Journal of Neuroscience 24:5467-5474.

Schubotz RI (2007) Prediction of external events with our motor system: towards a new framework. Trend in Cognitive Sciences 11:211-218.

Seth AK (2010) A MATLAB toolbox for Granger causal connectivity analysis. Journal of Neuroscience Methods 186:262-273.

Shannon CE (1948) A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review 5:3-55.

Strange BA, Duggins A, Penny W, Dolan RJ, Friston KJ (2005) Information theory, novelty and hippocampal responses: unpredicted or unpredictable? Neural Networks 18:225-230.

Summerfield C, Koechlin E (2008) A neural representation of prior information during perceptual inference. Neuron 59:336-347.

Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. Trends in Cognitive Sciences 13:403-409.

Turk-Browne NB, Yi D-J, Leber AB, Chun MM (2007) Visual quality determines the direction of neural repetition effects. Cerebral Cortex 17.

Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM (2010) Implicit perceptual anticipation triggered by statistical learning. Journal of Neuroscience 30:11177-11187.

Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. Cerebral Cortex 18:230-242.

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends in Cognitive Sciences 13:532-540.

Zaitsev M, Hennig J, Speck O (2004) Point spread function mapping with parallel imaging techniques and high acceleration factors: Fast, robust, and flexible method for echo planar imaging distortion correction. Magnetic resonance in medicine 52:1156-1166.

Zarahn E, Rakitin B, Abela D, Flynn J, Stern Y (2005) Positive evidence against human hippocampal involvement in working memory maintenance of familiar stimuli. Cerebral Cortex 15:303-316.

Table 1: Brain Regions with Entropy-Correlated Activity

ST Cluster #	**Talairach**			**Area (mm^2)**	**Max T**
	x	y	z		
Markov Regressor (Linear)					
1. L. Precentral G.	-44	-11	29	644	3.94
2. R. Precentral G.	43	-14	29	760	3.99
Markov Regressor (U-shaped)					
1. L. Insula	-37	-22	20	658	4.13
2. L. Middle Temporal G.	-54	-31	9	423	5.11
3. R. Middle Temporal G.	52	-20	-10	596	3.89
Markov Regressor (Step-Function)					
Right Calustrum	30	5	11	1015	3.16
Right Precentral G.	47	7	15	462	4.15
Shannon Regressor					
1. L. Claustrum	-27	14	9	536	3.43
2. L. STG	-46	-23	6	518	2.81
3. L. Inferior Parietal	-51	-30	28	507	3.84
4. R. Insula	32	-16	16	3359	4.55
5. R. Cingulate G.	6	-25	27	438	2.88
Markov Deflection (Linear)					
1. L. Precuneus	-17	-62	19	950	4.59
2. L. Precentral G.	-49	-3	21	936	5.07
3. L. Lingual G.	-6	-69	-1	704	4.20
4. L. Claustrum / Insula	-30	4	9	562	3.10
5. R. Posterior Cingulate G.	21	-53	6	2355	4.91
6. R. Cingulate G.	9	-29	35	702	4.48
7. R. Cingulate G.	11	7	41	700	2.98
8. R. Postcentral G.	39	-21	41	664	4.14
Markov Deflection (Non-linear)					
1. L. Superior Frontal G.	-21	36	33	457	4.89
Shannon Deflection					
1. R. Middle Frontal G.	39	15	25	434	4.43

Note. Areas sensitive to Shannon's Entropy were extremely similar when modeled with the linear, U-Shaped and Step-Function Markov Regressor. "Max T" refers to the T value found in the voxel showing the maximal effect in each cluster.

Table 2: Brain Regions Showing Inter-Subject Synchronization

ST Cluster #	**Talairach**			**Area (mm^2)**
	x	y	z	
Left Hemisphere				
1.Culmen / Lingual Gyrus	-14	-47	-3	211
Right Hemisphere				
2. Parahippocampal G.	20	-50	5	294
3. Fusiform G.	28	-57	-11	281
4. Posterior Cingulate.	8	-50	19	216

Note. Coordinates for bilateral auditory cortices not shown

Figure 1

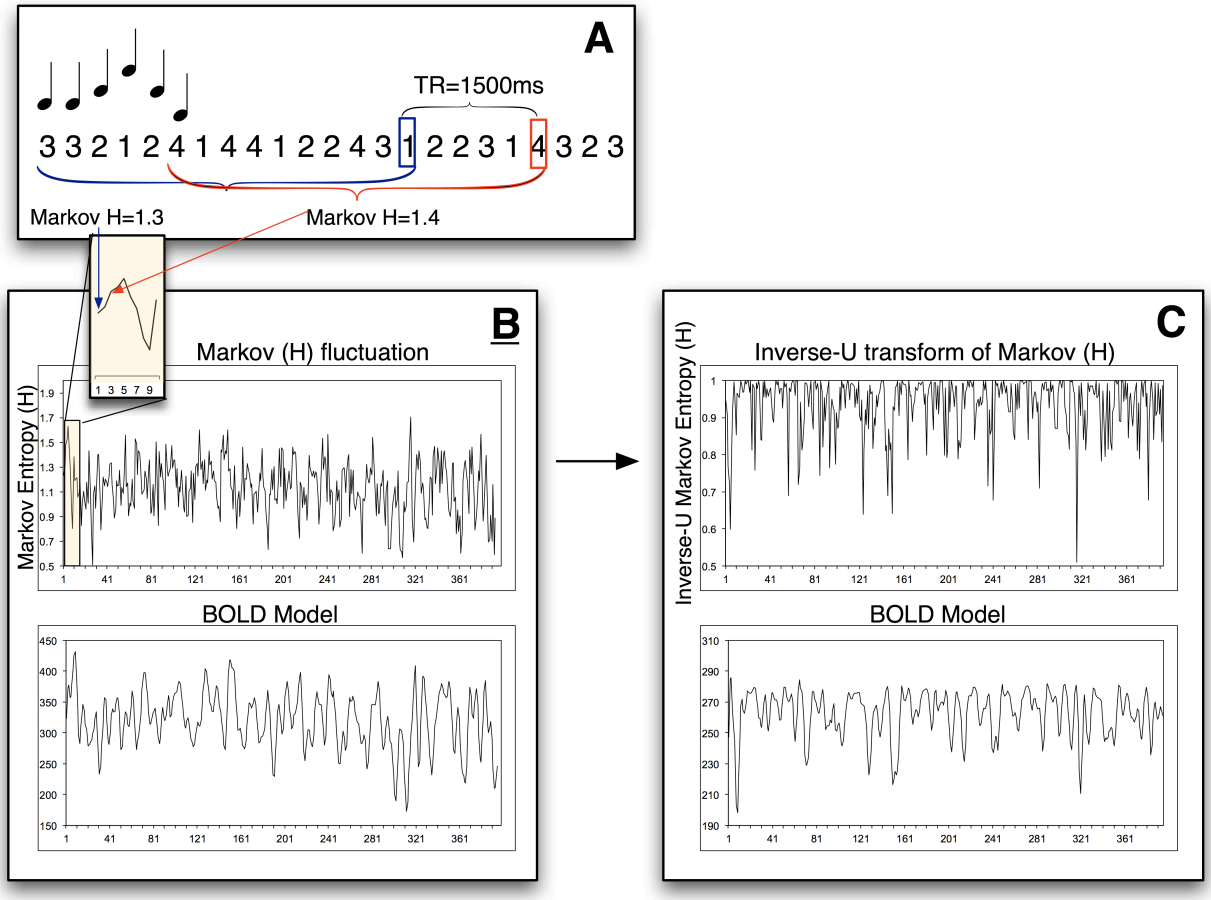
A
TR=1500ms
3 3 2 1 2 4 1 4 4 1 2 2 4 3 1 2 2 3 1 4 3 2 3
Markov H=1.3
Markov H=1.4
B
Markov (H) fluctuation
Markov Entropy (H)
1 3 5 7 9
BOLD Model
C
Inverse-U transform of Markov (H)
Inverse-U Markov Entropy (H)
BOLD Model

Figure 2

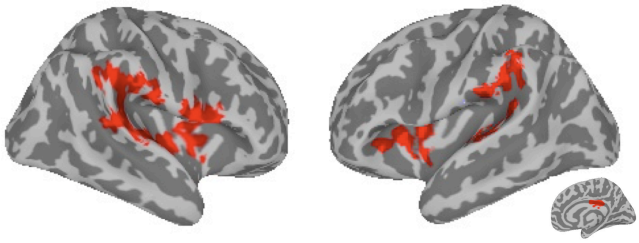

Figure 3

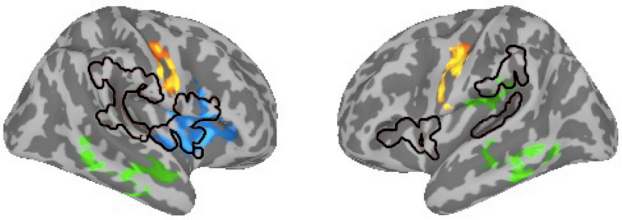

Figure 4

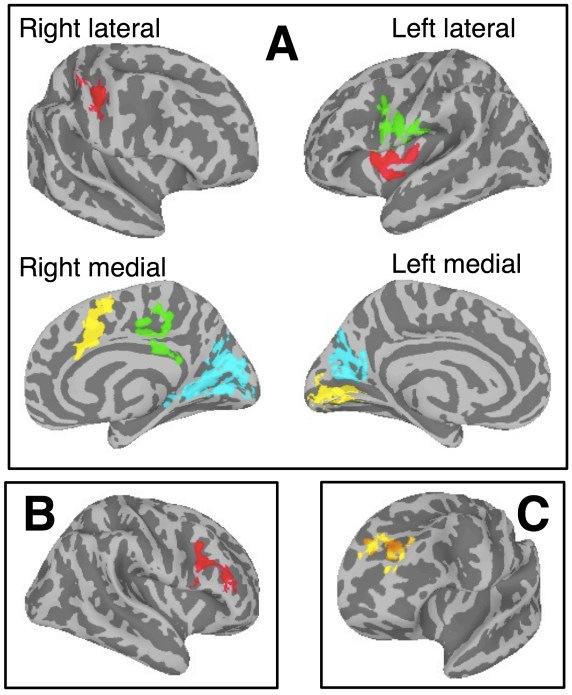

Figure 5

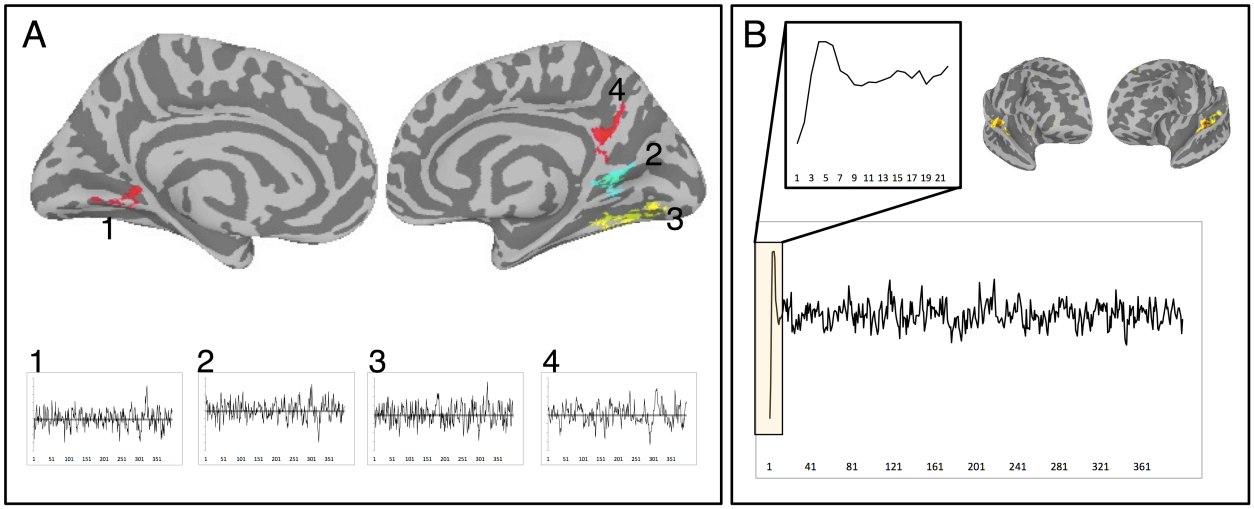

Figure 6