Behavioral/Cognitive

Hierarchical Organization of Frontotemporal Networks for the Prediction of Stimuli across Multiple Dimensions

Holly N. Phillips,[1,2] Alejandro Blenkmann,[3,4] Laura E. Hughes,[1,2] Tristan A. Bekinschtein,[2,5,6] and James B. Rowe[1,2,6]

[1]Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0SZ, United Kingdom, [2]Medical Research Council, Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom, [3]National Scientific and Technical Research Council (CONICET), Buenos Aires C1033AAJ, Argentina, [4]Institute of Cellular Biology and Neuroscience "Prof E. De Robertis" (IBCN), School of Medicine, University of Buenos Aires–CONICET, Buenos Aires C1121ABG, Argentina, [5]Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom, and [6]Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB United Kingdom

Brain function can be conceived as a hierarchy of generative models that optimizes predictions of sensory inputs and minimizes "surprise." Each level of the hierarchy makes predictions of neural events at a lower level in the hierarchy, which returns a prediction error when these expectations are violated. We tested the generalization of this hypothesis to multiple sequential deviations, and we identified the most likely organization of the network that accommodates deviations in temporal structure of stimuli. Magnetoencephalography of healthy human participants during an auditory paradigm identified prediction error responses in bilateral primary auditory cortex, superior temporal gyrus, and lateral prefrontal cortex for deviation by frequency, intensity, location, duration, and silent gap. We examined the connectivity between cortical sources using a set of 21 generative models that embedded alternate hypotheses of frontotemporal network dynamics. Bayesian model selection provided evidence for two new features of functional network organization. First, an expectancy signal provided input to the prefrontal cortex bilaterally, related to the temporal structure of stimuli. Second, there are functionally significant lateral connections between superior temporal and/or prefrontal cortex. The results support a predictive coding hypothesis but go beyond previous work in demonstrating the generalization to multiple concurrent stimulus dimensions and the evidence for a temporal expectancy input at the higher level of the frontotemporal hierarchy. We propose that this framework for studying the brain's response to unexpected events is not limited to simple sensory tasks but may also apply to the neurocognitive mechanisms of higher cognitive functions and their disorders.

Key words: dynamic causal modeling; magnetoencephalography; mismatch negativity; prediction and prediction error

Introduction

Brain function can be conceived as a hierarchy of generative models that optimizes predictions of sensory inputs and prediction errors (Friston and Kiebel, 2009). Under this generalized prediction hypothesis, top-down predictions are compared with bottom-up sensory inputs and return prediction errors when unexpected stimuli occur (Rao and Ballard, 1999; Kiebel et al., 2008; Friston, 2009; Chennu et al., 2013; Lieder et al., 2013b). Prediction errors underlie the event-related potential (ERP) response to sensory stimuli that

violate learned regularities, such as the mismatch negativity (MMN) in oddball tasks (Näätänen et al., 1993; Garrido et al., 2009a; Kimura et al., 2011) and rule sequences (Bekinschtein et al., 2009; Wacongne et al., 2011; El Karoui et al., 2014).

Evidence for hierarchical prediction in human frontotemporal cortex comes from functional brain imaging. Deviant auditory stimuli evoke neural responses in bilateral auditory cortex, superior temporal gyri, and prefrontal cortex (Giard et al., 1990; Doeller et al., 2003; Molholm et al., 2005; Rinne et al., 2005; Cheng et al., 2013; Chennu et al., 2013) from which Garrido et al. (2008) found clear evidence for a frontotemporal hierarchy of prediction and prediction error message passing. The core features of this model have been replicated (Garrido et al., 2007a,b, 2009b; Dietz et al., 2014) and studied in the context of coma (Boly et al., 2011), drug treatment (Schmidt et al., 2013), and aging (Cooray et al., 2014; Moran et al., 2014). However, previous approaches lacked a mechanism to explain neural responses to absent stimuli or subtler differences in their temporal structure. Temporal regularities over multiple events are associated with high-order representations of environment and action in the prefrontal cortex (Zhang and Rowe, 2015), whereas frontostriatal interactions have been associated with learning and prediction of temporal regularities (Grahn and Rowe, 2013). We therefore proposed that inter-

S D$_2$ S D$_1$ S D$_4$ S D$_3$ S D$_5$ S D$_4$ S D$_1$ S D$_5$ S ...

Figure 1.   MMN Optimum-1 paradigm (Näätänen et al., 2004). Standard tones (S) alternate with different deviant tones (D$_\#$).

nally generated inputs to prefrontal cortex are a feature of the generative models for MMN tasks with temporal regularities.

Previous imaging studies of human predictive coding in the context of MMN tasks have focused on a narrow set of deviant stimuli dimensions. However, Näätänen et al. (2004) and Hughes and Rowe (2013) observed differences in MMN amplitudes in EEG sensors and MEG dipoles for tones that deviated by frequency, intensity, location, duration, or silent gap (Näätänen et al., 2004). The latter two deviants differed significantly from the others in their distinct temporal structure. Thus, a hierarchical model should encompass multiple stimulus characteristics.

We therefore explored whether different violations of sensory regularities are associated with different interactions in a fronto-temporal hierarchy. We also tested the hypothesis that there are internally generated temporal predictions, revealed in response to deviations of temporal structure. We used a Bayesian approach to compare alternate hierarchical neural networks for auditory predictive coding with multiple deviant types. We identified the optimal connectivity pattern from a set of 21 generative network models, including a subset of models with top-down expectations that influence the prefrontal cortex to implement high-level predictions of events.

Materials and Methods

Participants. Eleven healthy adults participated in the study (seven males; mean age, 26 years, range, 18–37 years). Normal binaural hearing was confirmed immediately before the main experiment, and stimuli were presented at 60 dB above hearing threshold. Participants gave informed written consent, and the study was approved by the local research ethics committee.

Auditory paradigm. The "Optimum-1" multi-MMN paradigm of Näätänen et al. (2004) was used to investigate MMN responses to multiple deviant types (Fig. 1). This is a time-efficient variant of the classic oddball task, alternating standard tones with one of several different deviant tones, evoking equivalent MMN amplitude responses to each type of deviant equivalent to a classic task with low-frequency oddball events (Näätänen et al., 2004). Standard tones (75 ms in duration, with 7 ms ramp up and ramp down) contained three sinusoidal partials of 500, 100, and 1500 Hz. These alternated with deviant tones that differed in one of five dimensions: shortened duration (25 ms), frequency (550, 1100, and 1650 Hz or 450, 900, and 1350 Hz), intensity (±6 dB), location of sound source (right or left instead of binaural), or the presence of a silent gap of the middle 25 ms.

Tones were presented every 500 ms using E-Prime software (Psychology Software Tools) via plastic tubes and earpieces. Deviant tones were presented in a pseudo-random order such that a deviant type never appeared twice in a row and each deviant type would appear at least once in a sequence of 10 tones. A total of 900 standard and 900 deviant tones were played in three blocks of 5 min. Fifteen standard tones were played at the beginning of each block.

MEG data acquisition and processing. Data were collected with a 306-channel Vectorview system in a magnetically shielded room (Elekta Neuromag), including a magnetometer and two orthogonal planar gradiometers at each of the 102 positions. Paired EOG electrodes recorded vertical and horizontal eye movements, and five head-position indicator coils monitored head position. The three-dimensional locations of the coils and three anatomical fiducials (nasion and left and right preauricular points) were recorded using a 3D digitizer (Fastrak; Polhemus). Movement compensation and downsampling from 1 kHz to 250 Hz was completed using Maxfilter software (Elekta Neuromag). The remaining preprocessing steps were completed using SPM8 software

(Wellcome Trust Centre for Neuroimaging, University College London). This included high-pass filtering at 1 Hz and low-pass filtering at 40 Hz using Butterworth filters in forward and reverse directions and epoching −100 to 400 ms around each tone onset with baseline correction of the −100 to 0 ms period. Automatic artifact rejection used thresholding of EOG electrodes at 200 μV. Trials were averaged using robust averaging (Wager et al., 2005), followed by an additional low-pass filter at 40 Hz to remove high-frequency noise that can be introduced by robust averaging.

Source space analysis. Source reconstruction of the ERPs to standard and deviant tones was completed for the gradiometer data using SPM8. The forward model (leadfield) was estimated from a single shell template cortical mesh of each participant's anatomical T1-weighted MR image (3D MPRAGE sequence; TR, 2250 ms; TE, 2.99 ms; flip angle, 9°; field of view, 240 × 256 × 160; 1 mm slice thickness; collected on a 3T Siemens Tim Trio scanner), coregistered by digitized fiducial markers and >60 scalp points. Source waveforms of standard and deviant tones were extracted for each participant by applying the inverted leadfield matrix to estimate the six equivalent current dipoles (ECDs) of the anatomically defined sources of the MMN. The six sources approximated previously published work with the auditory oddball task (Garrido et al., 2009b; Moran et al., 2014; see Fig. 3): bilateral primary auditory cortex (A1; MNI coordinates: [−42, −22, 7], [46, −14, 8]), superior temporal gyri (STG; [−61, −32, 8], [59, −25, 8]), and inferior frontal gyri (IFG; [±46, 20, 8]). Source locations were fixed, allowing orientation of the dipoles to vary.

Group differences between standard and deviant waveform mean amplitudes were assessed over the characteristic MMN time window of 100–200 ms using paired t tests ($p < 0.05$) at each MMN source location. The false discovery rate (FDR) was used to correct for multiple comparisons. We also present the effect size, r. To support our hypothesis that different deviant tones are associated with different networks in the frontotemporal hierarchy, we tested the differences between individual deviants: here we used source reconstruction for each deviant MMN waveform (the difference between standard tone and each deviant tone). We then used a repeated-measures ANOVA with two factors, deviant type (duration, frequency, gap, intensity, and location) and source location (bilateral A1, STG, and IFG). This used the peak amplitudes and latencies of the MMN waveforms. Mauchly's test was used to indicate which statistics required correction using Greenhouse–Geisser estimates caused by sphericity assumption violations. Finally, each deviant mean MMN amplitude was tested for significance at the six locations using a one-sample t test to be sure significant differences found in the ANOVA were not attributable to a lack of a MMN response in some deviant types ($p < 0.05$ threshold for significance, FDR correction for multiple comparisons).

Network analysis. We used dynamic causal modeling (DCM) of the effective connectivity between the six specified sources. The models included standard and deviant tones, together with their modulation by deviance, and are not restricted to modeling the mismatch response. Sources of standard and individual deviant tones were reconstructed separately using the forward modeling described above and inverted using the SPM8 DCM10 standard algorithm with default settings.

With biophysically constrained neural mass models, DCM makes inferences about the mechanisms behind observed ERPs, the coupling between ECD sources, and how experimental stimuli changes this coupling (David et al., 2006; Kiebel et al., 2006, 2009). Twenty-one generative models (Fig. 2) were used to model alternative hypotheses of the mechanism underlying the MMN, based on the anatomically motivated networks (Giard et al., 1990; Rinne et al., 2000; Optiz et al., 2002; Doeller et al., 2003; Molholm et al., 2005, Garrido et al., 2008). These described the effective connectivity between temporal and frontal sources for the time window of 0–250 ms from the onset of each stimulus, encompassing the MMN interval. All connections between MMN sources were bidirectional and modulated after Garrido et al. (2007b). DCM is agnostic as to the direct or indirect route of connections via monosynaptic or polysynaptic pathways.

The first six models were a conceptual replication of Garrido et al. (2008). They begin with driving inputs into bilateral primary auditory cortex, with or without intrinsic connections within these sources (mod-
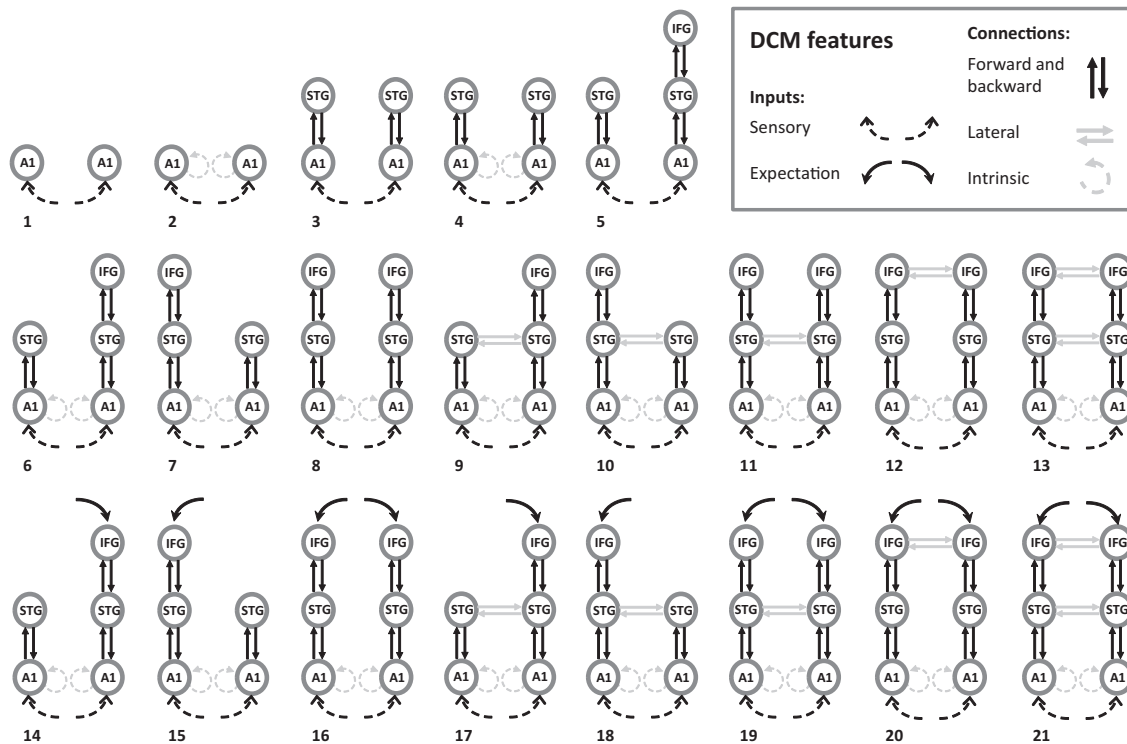
Figure 2. The 21 DCM models used to compare frontotemporal networks associated with predictive coding. The first six models replicate the model space of Garrido et al. (2008). Models 7 and 8 add left IFG nodes, models 9–13 add lateral connections, and models 14–21 add expectation driving inputs for event onsets onto the IFG.

els 1 and 2), followed by the presence of bidirectional connections to bilateral STG (models 3 and 4) and bidirectional connections to right IFG (models 5 and 6). These models have been examined using oddball and roving MMN paradigms (Garrido et al., 2008, 2009b). Model 6 had the highest model evidence using Bayesian model selection (BMS; Penny et al., 2004).

We began by examining these six models to test whether the Optimum-1 MMN paradigm replicates this result for all deviant types together and separately. Then our model space was extended to include further cortical areas, inputs, and connections as follows.

First, we added bidirectional connections between left STG and left IFG and bilateral connections to both IFGs (Garrido et al., 2009b) to create models 7 and 8, respectively. These models were motivated from the debate in the literature to determine whether the left IFG is a source of the MMN (Giard et al., 1990; Alho et al., 1994; Rinne et al., 2000, 2005; Jemel et al., 2002; Opitz et al., 2002; Cheng et al., 2013).

Second, bidirectional modulated lateral connections were added between STG sources and IFG sources (models 9–13). Lateral connections are used in several DCM studies (Boly et al., 2011; Schmidt et al., 2013; Cooray et al., 2014), although only Boly et al. (2011) performed comparisons between models with and without these connections, finding that models with lateral connections had the greatest evidence in healthy adults.

Finally, we added event expectation driving inputs into the IFG sources (models 14–21), motivated by our hypothesis that with regular stimuli, there will be internally generated expectations of when a stimulus would occur (without specifying its properties).

Bayesian model selection. BMS was used to compare the generative models (Penny et al., 2004) and discover which best explains the neural responses. BMS compares the free-energy estimate (F) of the bound on the log of model evidence of each model, $\ln p(y|m)$ (the probability of the data y given each model m). This measure of model evidence adjusts model fit for model complexity to reduce overfitting (Kiebel et al., 2009; Stephan et al., 2010). In the main analysis, we used a fixed-effects (FFX) approach, assuming our population of healthy participants use the same network architecture but have connection strength variation (Stephan et al., 2010; Dietz et al., 2014). In addition, we repeated the model and

family analyses using a random-effects (RFX) approach to consider possible bias on FFX results attributable to participant outliers.

The model with the highest model evidence is often referred to as the "winning" model. A difference in model evidence between the winning and "second place" models (ΔF) of three units or more is comparable with a Bayes factor of 20, and by convention, this is regarded as strong evidence for one model over another (Kass and Raftery, 1995; Stephan et al., 2010). We also calculated the posterior probability of each model to demonstrate the probability of that model given the neural responses within the current model space. BMS was first used to compare model evidence for the first six models to replicate the model of Garrido et al. (2007a). Then we used BMS for the complete model space.

Finally, we performed a *post hoc* comparison of model families to assess the significance of connections required by individual deviant types by removing uncertainty about other model structural aspects (Penny et al., 2010). Specifically looking at lateral connections and prefrontal expectancy inputs, we split the model space into five families: (1) the original models of Garrido et al. (2007a; G; models 1–5); (2) models without lateral connections or prefrontal expectancy inputs (lp; models 6–8); (3) models with lateral connections (Lp; models 9–13); (4) models with prefrontal expectancy inputs (lP; models 14–16); and (5) models with both lateral connections and prefrontal expectancy inputs (LP; models 17–21). DCM and BMS were completed using SPM8.

Results

MMN source waveforms

The source waveforms were reconstructed (Fig. 3) using ECDs for each of the six MMN source locations to be used in the following DCM analysis. From these, the mean waveform amplitudes were calculated over the MMN characteristic time window of 100–200 ms for standard and all deviant tones and were compared using paired sample t tests. Each source location had a significant difference between standard and deviant tones (right A1: $t = 6.28$, $p = 0.0001$; left A1: $t = 4.11$, $p = 0.0021$; right STG: $t = 4.54$, $p = 0.0011$; left STG: $t = 2.58$, $p = 0.0274$; right IFG: $t = 7.75$, $p < 0.0001$; left IFG: $t = 4.22$, $p = 0.0018$) and remained

significant after FDR correction for multiple comparisons. All comparisons achieved $r > 0.5$ indicative of large effect sizes.

Individual deviant tone MMN waveforms are shown in Figure 4. Differences in MMN peak amplitudes and latencies were assessed using separate repeated-measure ANOVAs with two factors, ECD locations and deviant type. The main effects of deviant type and ECD location violated the sphericity assumption for peak amplitudes [$\chi^2(9) = 19.48$ and $\chi^2(14) = 57.85$, respectively] as well as the main effect of location for peak latency [$\chi^2(9) = 19.48$]. Thus the degrees of freedom were corrected by the Greenhouse–Geisser method. For peak amplitudes, there was a significant main effect of ECD location ($F_{(1.85,18.5)} = 9.84$, $p = 0.001$) and of deviant type ($F_{(2.05,20.5)} = 7.79$, $p = 0.003$). For peak latency, there was a significant main effect of deviant type ($F_{(4,40)} = 6.16$, $p = 0.001$), but not of ECD location. For both peak amplitudes and peak latencies, there was not a significant interaction between ECD location and deviant type ($p > 0.05$). To test whether these effects may be attributable to a lack of a MMN response, the mean MMN amplitude of each deviant type at each location was tested for significance using a one-sample t test. Using FDR correction for multiple comparisons, we found all mean MMN amplitudes to be significant ($p < 0.05$).

Hierarchical network models

We initially compared the first six models in Figure 2, which conceptually replicate the models used by Garrido et al. (2008) using Bayesian model selection with a FFX approach. Figure 5A shows very strong evidence in favor of model 6 (the winning model) with the highest relative log-evidence (F). ΔF indicates the difference between the winning and second place model evidence. All deviant dimensions achieve $\Delta F > 3$, which is equivalent to a Bayes factor of ~20 and represents strong evidence in favor of the winning model. The posterior probability for model 6 exceeds 0.99 for each deviant type, demonstrating the high probability of this model given the evidence, within the current model space. Repeating the analysis with a RFX approach to account for possible individual participant bias also reveals model 6 to be the winning model (with the highest model exceedance probability, P) for each deviant type: duration, $p = 0.92$; gap, $p = 0.98$; frequency, $p = 0.84$; intensity, $p = 0.92$; location, $p = 0.61$; all deviants together, $p = 0.96$. Model 6 has bidirectional connections between bilateral A1 and STG and between right STG and right IFG plus intrinsic connections in bilateral A1 as shown in Figure 5B. The winning model 6 for all deviants is equivalent to the winning model of Garrido et al. (2008), though that study used a roving MMN paradigm with frequency deviants only.

Using BMS with a FFX approach, Figure 6 shows the relative log-evidence (F) and posterior probabilities for the full model space (Fig. 2, models 1–21) along with the winning models for
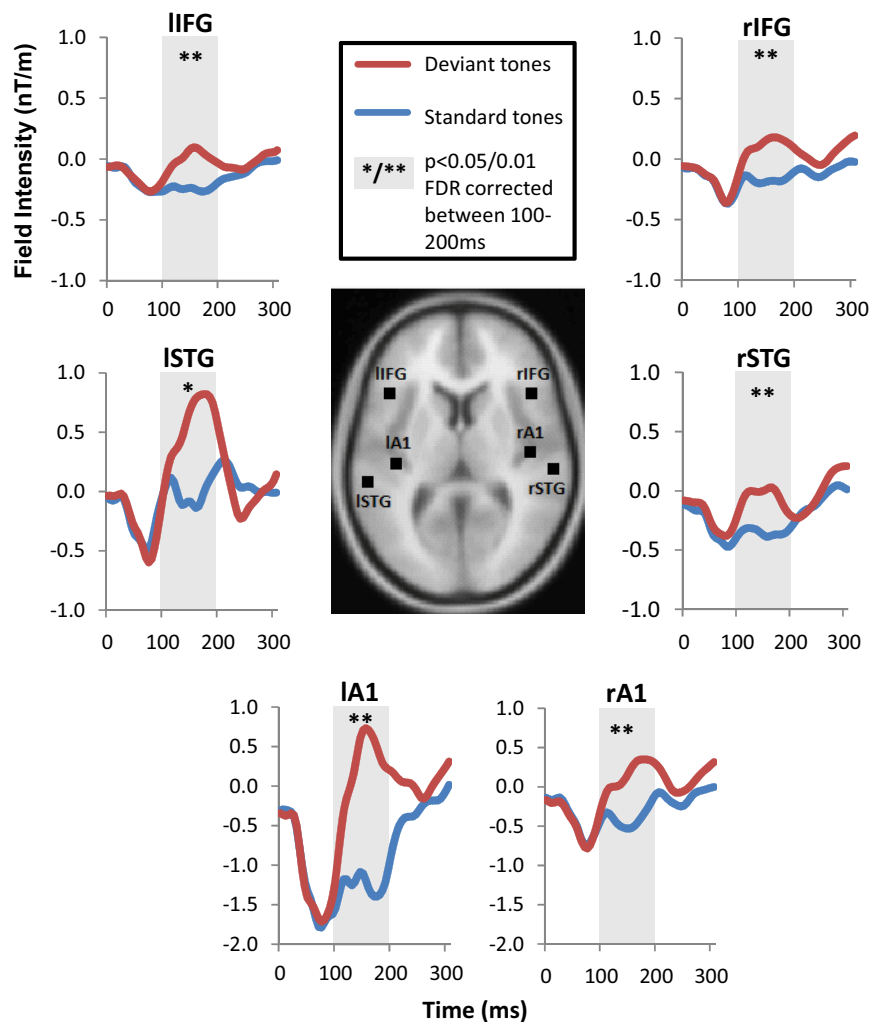

Figure 3. Mean source waveforms to standard (blue) and all deviant (red) events in bilateral A1, STG, and IFG MMN sources, with anatomical locations illustrated in the center. The shaded areas indicate the 100 –200 ms period of the characteristic MMN-evoked response. Significant differences between standard and deviant tones using paired t tests were evident at each location as indicated (*$p < 0.05$ and **$p < 0.01$), with FDR correction for multiple comparisons. rA1, Right A1; lA1, left A1; rSTG, right STG; lSTG, left STG; rIFG, right IFG; lIFG, left IFG.

each of the separate deviant types. For each deviant type, the winning model included all the features from model 6 plus bidirectional connections between bilateral STG and/or IFG sources. For duration and gap deviant tones, the models with the highest log-evidence included expectation inputs into IFG sources and bidirectional lateral connections between IFG sources. The gap deviant also included lateral connections between STG sources (model 21) whereas duration did not (model 20). The model with the highest log-evidence for the frequency deviant included lateral connections between IFG sources (model 12), and both the intensity and location deviants' winning models included lateral connections between STG sources (model 11). All winning models exceeded a posterior probability 0.99. The differences (ΔF) between the winning and second place models are shown in Figure 7 below each winning model diagram. This analysis was repeated with a RFX approach, which showed agreement over winning models for frequency, intensity, and location deviants. For the gap deviant, there was equipoise between model 21 (FFX winning model) and model 20. The winning model for the duration deviant changed to model 16, which had the same architecture as the FFX winning model 20, but without interhemispheric
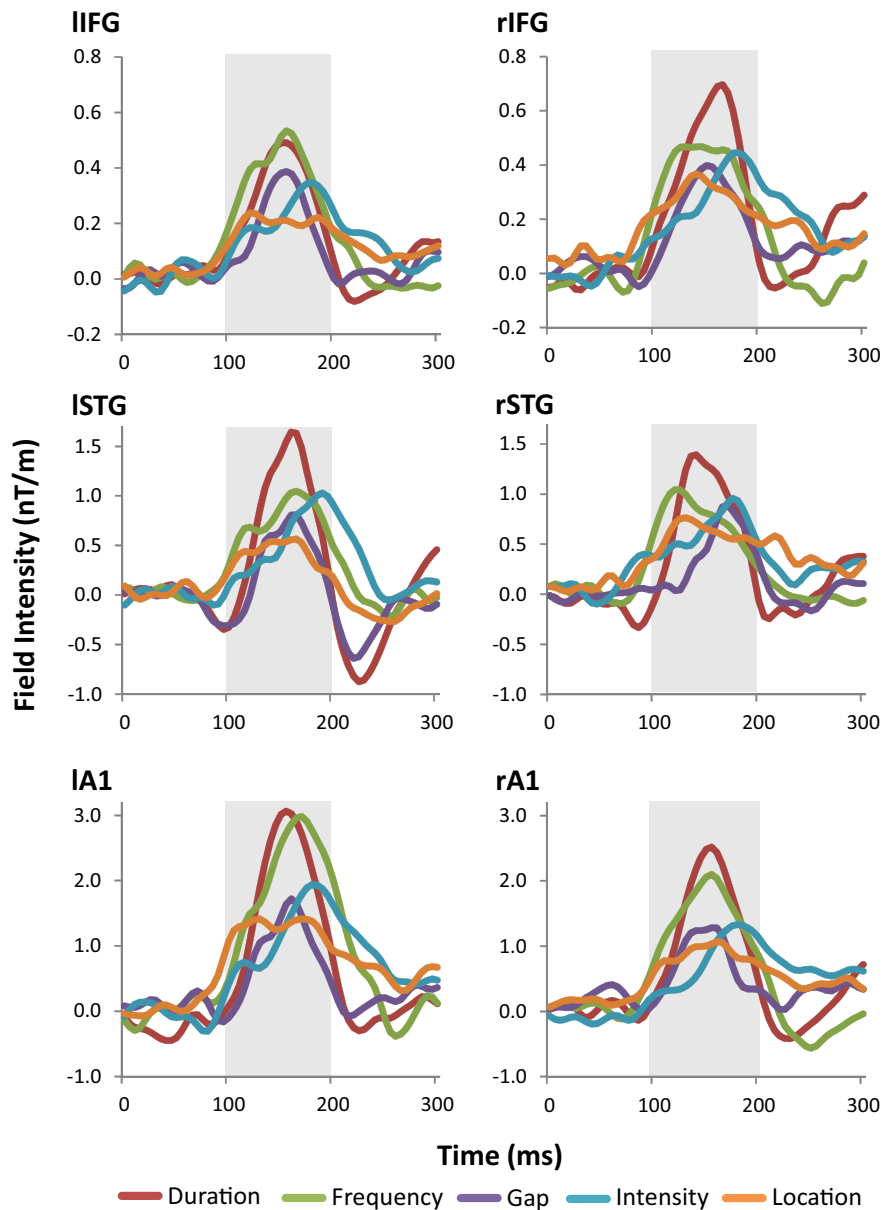
Figure 4. MMN dipole waveforms for each deviant type, for each MMN source location. The gray area highlights the characteristic MMN period (100–200 ms). At every location, for all deviant types, there is a visible MMN waveform during this characteristic period. The mean amplitude of each waveform over 100–200 ms was tested using a one-sample *t* test comparison against zero mean, and all were significant (*p* < 0.05, FDR corrected). Please note that the color used here for each deviant type will be consistent across subsequent figures.

and LP), one the same as the FFX result and the other a nested model within it that lacks an interhemispheric connection. For both models, there is still inclusion of prefrontal expectancy inputs. Thus, there is general agreement between FFX and RFX approaches.

Discussion

This study provides evidence for hierarchical frontotemporal networks supporting the prediction of sensory information and responses to violations of these predictions. There was evidence for the following key features in the most likely network: (1) reciprocal feedforward and feedback connections between auditory cortex and STG, connections between STG and IFG bilaterally, and interhemispheric interactions; and (2) internally generated expectations as driving inputs to prefrontal cortex at the uppermost level of the model's hierarchy present for temporal structure violations.

Whereas previous studies have focused on single dimensions of deviance, we identified differences in the hierarchical frontotemporal networks underlying the response to multiple types of deviants. For all deviants, we replicate previous results from a classic oddball (Garrido et al., 2009b) and a roving (Garrido et al., 2008) paradigm in the context of an equivalent model space (Fig. 5). The results accord with the predictive coding hypothesis in which feedback predictions and feedforward prediction errors pertain to each layer of the hierarchy (Friston and Kiebel, 2009; Carlin et al., 2011). Extending the model space enabled us to test additional hypotheses of the frontotemporal interactions related to sensory prediction and mismatch error signaling. We found that winning models for frequency, intensity, and location differed subtly from each other in terms of the lateral connections between IFG and/or STG sources but were within a family of structurally similar models. Together, these data suggest generalized features in the hierarchical networks for the response to multiple types of sensory deviation.

Some previous studies modeled unilateral prefrontal cortical sources (Garrido et al., 2008; Boly et al., 2011; Schmidt et al., 2013), and others used bilateral sources (Hughes et al., 2013). We formally compared unilateral versus bilateral models and found very strong evidence in favor of bilateral frontal cortical sources, consistent with the bilateral evoked responses (Fig. 6). Moreover, we found evidence for top-down predictions in explaining the data for stimuli that differed in their duration of temporal profile. We term these high-order inputs temporal expectancy predictions. An important corollary of this prefrontal expectancy input is that it enables the network to predict auditory and STG activity even when a stimulus is omitted altogether, as in some forms of

connections. Across FFX and RFX findings, only deviants with temporal structure changes required prefrontal inputs, and all deviants required bilateral prefrontal sources.

Finally, we used *post hoc* family-level inference to assess the importance of lateral connections and prefrontal expectancy inputs for the individual deviants. The FFX results in Figure 7 show that deviants that violate temporal tone structure (duration and gap) require prefrontal expectancy inputs (model family LP), whereas the remaining deviant types do not (model family Lp). For all deviant dimensions, the winning model families include lateral connections. The analysis was repeated with a RFX approach for verification that FFX results were not biased by outliers. Figure 7B shows that frequency, intensity, location, and gap deviants have matched winning models across approaches. For the duration deviant, there is equipoise between two models (lP
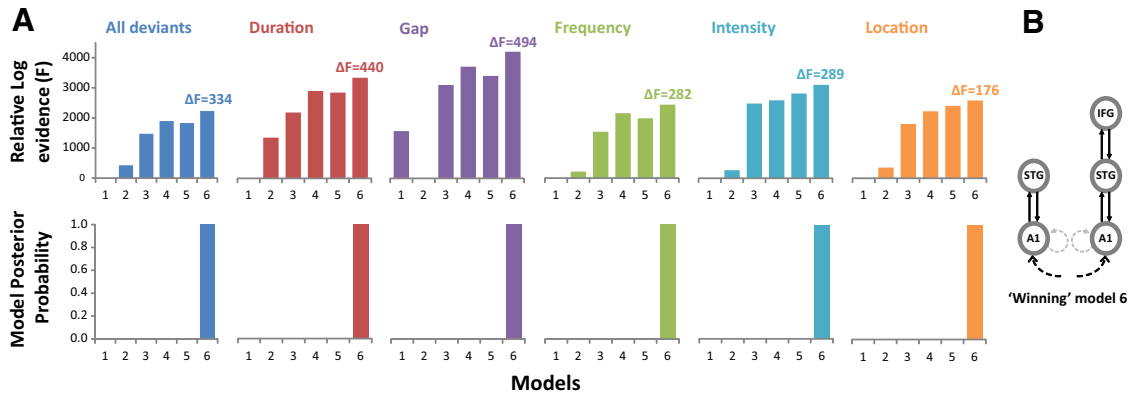
Figure 5. Bayesian model selection results for the first six models. *A*, Top row, Deviant types together and separately consistently show model 6 to win with the highest log-evidence (F) relative to the model with lowest log-evidence, which is set to zero. This is a conceptual replication of the principal result of Garrido et al. (2008), which used a roving MMN paradigm with frequency deviations. The ΔF above each winning model is the difference between the highest model log-evidence and the second place model log-evidence. A $\Delta F > 3$ in favor of the winning model is equivalent to a Bayes factor of 20, and a $\Delta F > 5$ is equivalent to a Bayes factor of 150. For all deviant types, model 6 has a $\Delta F > 5$ over the second place model. Bottom row, Model 6 consistently achieves a posterior probability of >0.99 compared with the other five models. *B*, The winning model with the highest log-evidence (F) for all deviant types together and individually.
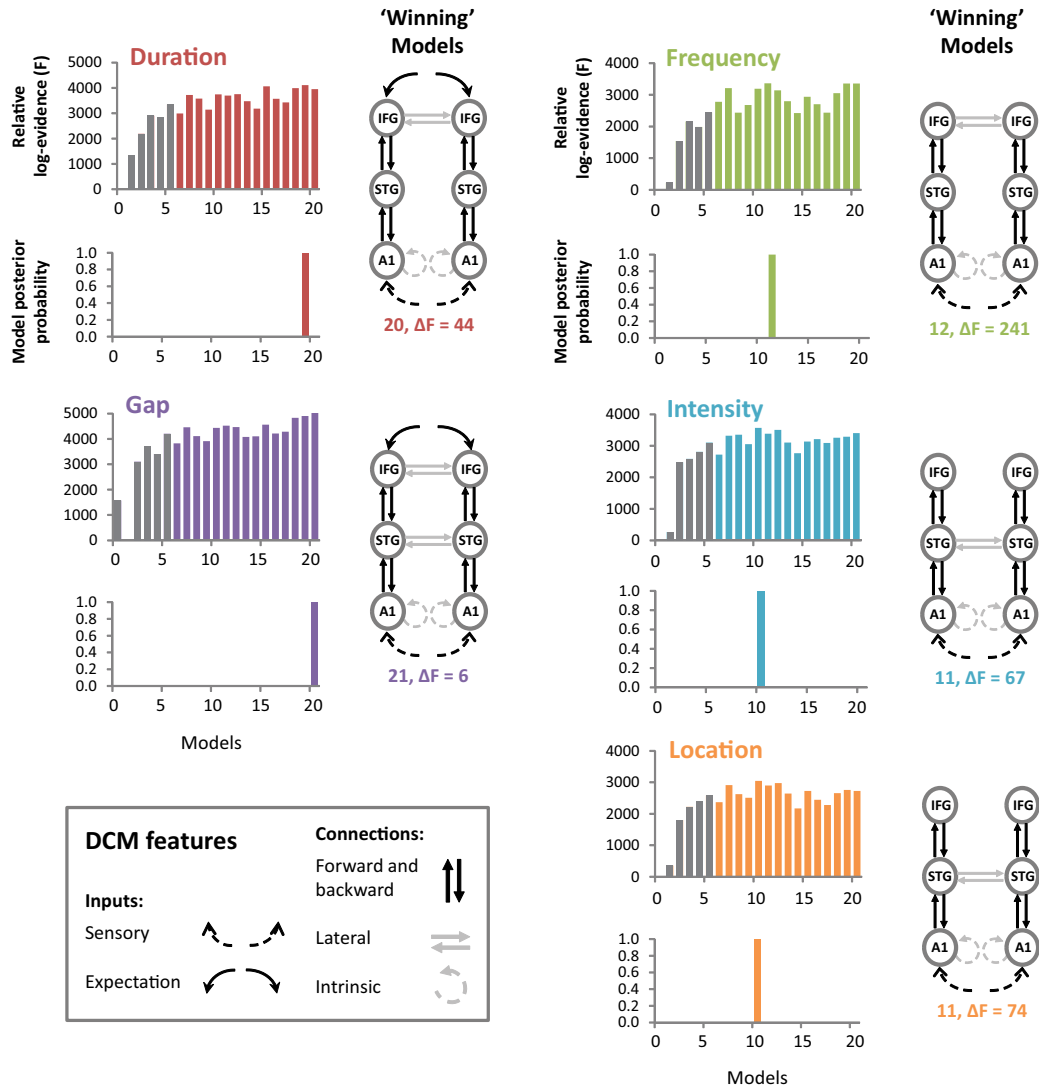

Figure 6. Bayesian model selection for each deviant type with full model space. For each deviant type, we include the relative log-evidence and model posterior probabilities for every model, a diagram of the winning model, and the difference between winning and second place relative log-evidence (ΔF). In each plot of log-evidence, the first six bars in gray indicate the models used by Garrido et al. (2008). Both duration and gap deviants (left) reveal bilateral prefrontal expectancy inputs into the IFG sources, with lateral connections between IFG sources. The gap deviant also has lateral connections between STG sources. Frequency, intensity, and location deviant winning models do not include expectancy inputs (right). They differ within themselves by lateral connections, where frequency deviants include interhemispheric IFG sources whereas intensity and location deviants include STG interhemispheric connections. For all deviant types, the winning models have a $\Delta F > 5$ and, therefore, are considered to have very strong evidence in favor of those models.
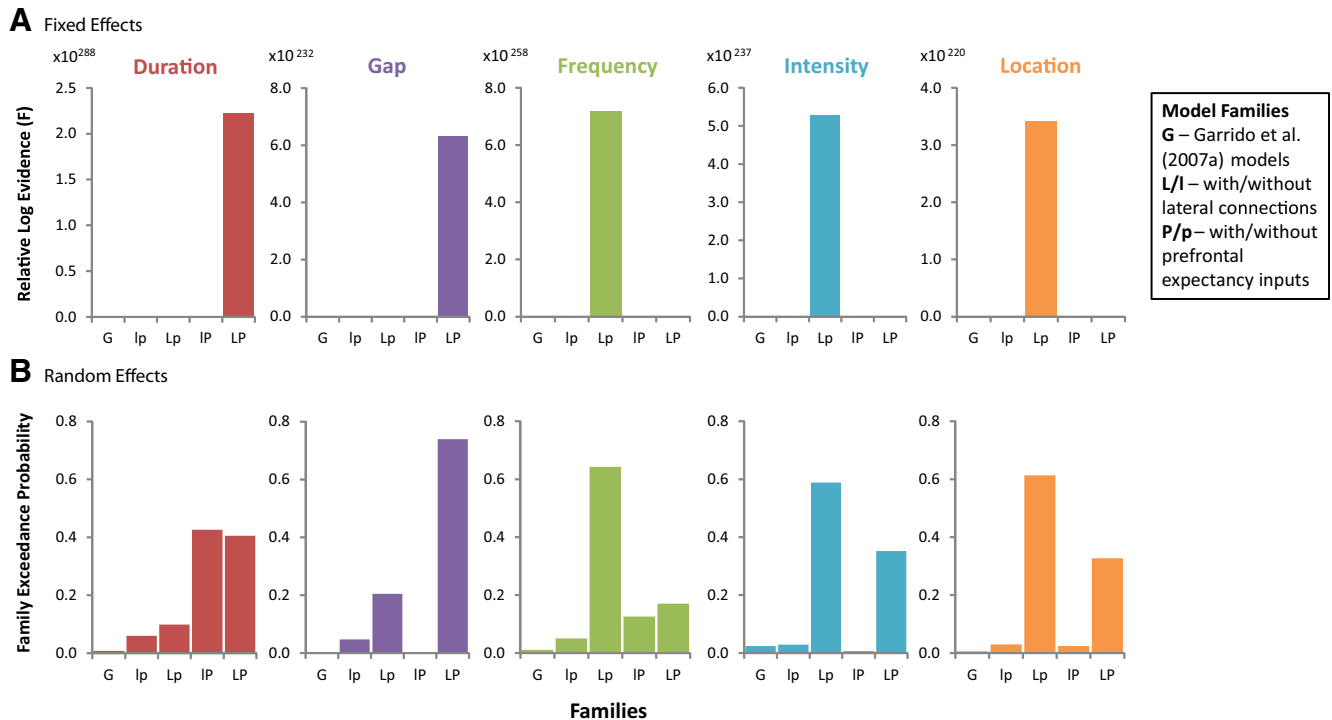
Figure 7. Family-wise model selection for each deviant dimension using both fixed- and random-effects approaches. The following model families were investigated: G, Original model space of Garrido et al. (2007a); lp, models without lateral connections or prefrontal expectancy inputs; Lp, models with lateral connections but not prefrontal expectancy inputs; lP, models with prefrontal expectancy inputs but not lateral connections; LP, models with both lateral connections and prefrontal expectancy inputs. ***A***, FFX analysis. Duration and gap deviants had the highest model evidence (*F*) for the family of models with prefrontal expectancy inputs with lateral connections (LP). The posterior probability for this family was >0.99, showing strong evidence for this family for duration and gap deviant dimensions. Frequency, intensity, and location deviants instead showed strong evidence for the family of models with lateral connections but without prefrontal inputs (Lp). For these three deviant dimensions, the posterior probability for this family was >0.99. ***B***, RFX analysis. The winning families (with the highest family exceedance probability) for frequency, intensity, location, and gap deviants match the FFX winning families. For the duration deviant, the two model families with prefrontal inputs have almost equal exceedance probabilities, one the same as the FFX result and the other without an interhemispheric connection.

the MMN task (Raij et al., 1997; Hughes et al., 2001; Wacongne et al., 2011).

The prefrontal temporal expectancy input was important for explaining the response to duration and gap deviants, which, unlike the other deviants, are defined by a violation in the temporal structure of stimuli. Analogous temporal expectancy has been observed for beat prediction (Zanto et al., 2006; Pecenka and Keller, 2011; Teki et al., 2011a; Fujioka et al., 2012). Teki et al. (2011b) suggesting that a striato–thalamo–cortical circuit, including putamen and prefrontal areas, is involved in relative beat-based timing. Grahn and Rowe (2009, 2013) also found putamenal involvement in the prediction tones in a regular beat pattern, interacting with frontal cortex. Connectivity between the striatum and IFG would be well suited to afferent expectancy projects, and such connections are supported by autoradio-graphic tracer studies in rhesus monkeys (Yeterian and Pandya, 1991) and diffusion-weighted MR imaging in humans (Croxson et al., 2005), from putamen and caudate (Lehéricy et al., 2004; Novak et al., 2015). Furthermore, Postuma and Dagher (2006) identified functional connectivity between IFG, caudate, and rostral putamen.

Additionally, local context-dependent circuits have been proposed, integrated with a core timing circuit (Merchant et al., 2013). These have been observed in human visual, auditory, and parietal areas (Leon and Shadlen, 2003; van Wassenhove and Nagarajan, 2007; Bueti et al., 2008) and in *in vitro* recordings (Johnson et al., 2010), suggesting time-dependent cellular properties can allow local circuits to encode specific stimulus timing

(Karmarkar and Buonomano, 2007). In contrast, the core timing circuit is observed across multiple stimulus modalities (Merchant et al., 2008), in the basal ganglia (Meck et al., 2008) and prefrontal cortex (Coull et al., 2011), which may have a monitoring role over stimulus durations (Rao et al., 2001) and extended sequences with regularities (Zhang and Rowe, 2015). Thus, our expectancies for temporal deviants may originate from prefrontal cortex itself or inputs from the striatum.

Although duration and gap deviants were distinct from other deviant types in the likely network model, there were important similarities between all deviants (Fig. 6), including the presence of bilateral prefrontal sources and interhemispheric connections. MMN responses to multiple deviant dimensions have been observed in EEG (Giard et al., 1995; Jemel et al., 2002; Näätänen et al., 2004; Petermann et al., 2009; Fisher et al., 2011; Chennu et al., 2013) and MEG (Hughes and Rowe, 2013). Both amplitude and latency differences were observed in auditory cortex. Several mechanisms have been proposed to explain these MMN effects. Under the adaptation and change detection hypotheses, MMNs are produced when deviant tones elicit activity from nonsuppressed neurons (May and Tiitinen, 2010) or differ from a memory trace of standard tones (Schröger and Winkler, 1995), respectively. Both imply different generator locations dependent on deviant dimension. Under the model adjustment hypothesis, MMN reflects the updating of a model of standard tones (Winkler, 2007). In contrast, the predictive coding hypothesis proposes that the MMN reflects prediction errors. In a direct

comparison, predictive coding is more likely than the "phenom-enological" hypotheses (Garrido et al., 2008; Lieder et al., 2013a).

The frontotemporal network underlying the detection of un-expected sensory events provides a robust framework to study the impact of disease. The auditory MMN paradigm is advantageous in part because it does not require attention or behavioral re-sponses (for review, see Näätänen et al., 2012). For example, Hughes and Rowe (2013) showed reduced β-band frontotempo-ral coherence in behavioral-variant frontotemporal dementia (bvFTD). Although coherence is not directional (Fries, 2005; Bastos et al., 2012), it has been suggested that β coherence reflects feedback predictions, which in the case of Hughes and Rowe (2013) would likely be from frontal to temporal cortex. In the light of the current findings, we speculate that frontal cortical degeneration in bvFTD could alternatively impair the impact of prefrontal temporal expectancies.

Other disorders have also been investigated. Schizophrenia reduces duration–MMN amplitudes (Michie et al., 2000) in pro-portion to symptoms (Kärgel et al., 2014). In dyslexia, Kujala et al. (2006) observed reduced frequency deviant amplitudes, which correlated with reading skill (Baldeweg et al., 1999). Additionally, Morlet and Fischer (2014) and Daltrozzo et al. (2007) confirmed that duration deviants are robust for predicting coma outcome. These studies suggest that using the MMN for clinical research could benefit from tailoring deviant types to disorders to achieve maximal decoding ability.

Furthermore, using the MMN to study the neural responses to unexpected events is not limited to simple sensory tasks but also applies to the neurocognitive basis of higher cognitive functions (Clark, 2013). The use of a hierarchy of generative models to predict the sensorium is suggested to be the common framework behind learning (Friston and Stephan, 2007; Fletcher and Frith, 2009; Moran et al., 2013, 2014), recognition (Egner et al., 2010; Muckli, 2010), attention (Clark, 2013), and motor control (Fris-ton et al., 2011). Here we show that even for a simple passive task with small differences in auditory stimuli, the hierarchical gener-ative model is flexible to predict specific deviations. This supports the notion that the brain optimizes connectivity to better predict its environment in both low-level perception and higher cogni-tive functions (Moran et al., 2014).

There are limitations to this study. Our winning models are selected from a defined model space based on our hypothesis and prior literature, but it could be argued that other networks might be better still. However, DCM is a hypothesis testing framework, rather than an exploratory model-search technique [see Friston et al. (2013) in response to Lohmann et al. (2012)]. This is not because it is computationally intensive but because there is a greater risk of overfitting with large model sets (Friston et al., 2013). Thus, we kept to the recommendation to choose necessary and sufficient model space with which to test hypotheses. Second, we used the same prior source locations for each MMN source regardless of the deviant type examined. Molholm et al. (2005) suggested modest location differences between frequency and de-viant types within the auditory and prefrontal cortices, using fMRI. However MEG is tolerant of minor (millimeter) deviations of the site of sources, in part because of its inherently lower spatial resolution. More important is the orientation of the dipole, which remained free (Garrido et al., 2007a). Third, one could potentially model the source of temporal expectancies acting on the prefrontal cortex. We did not do so, in part because striatal sources are not well observed in MEG and because potential pre-frontal sources could not be specified a priori. We speculate that the temporal expectancy inputs act as a "pacemaker" prediction

of temporal regularities in stimulus trains. But additional studies would be needed to test the hypothesis that expectancy inputs are important for the response to deviations from temporal iso-chrony. This could be undertaken using deviations from iso-chrony and omission instead of the qualitative differences in regular stimuli that we used.

In conclusion, the auditory multi-mismatch task reveals the presence of hierarchical frontotemporal networks for the predic-tion of sensory events and response to sensory deviants. We show the flexibility of this generative model hierarchy to predict mul-tiple variations in auditory dimensions, including the temporal structure of stimuli. Furthermore, we provide new evidence for internally generated temporal expectations that influence pre-frontal cortex. The role of these higher-level expectations may be particularly relevant in hierarchical networks that support higher cognitive functions and their disorders.

References

Alho K, Woods DL, Algazi A, Knight RT, Näätänen R (1994) Lesions of frontal cortex diminish the auditory mismatch negativity. Electroen-cephalogr Clin Neurophysiol 91:353–362. CrossRef Medline

Baldeweg T, Richardson A, Watkins S, Foale C, Gruzelier J (1999) Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. Ann Neurol 45:495–503. CrossRef Medline

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. Neuron 76:695–711. CrossRef Medline

Bekinschtein TA, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L (2009) Neural signature of the conscious processing of auditory regularities. Proc Natl Acad Sci U S A 106:1672–1677. CrossRef Medline

Boly M, Garrido MI, Gosseries O, Bruno MA, Boveroux P, Schnakers C, Massimini M, Litvak V, Laureys S, Friston K (2011) Preserved feedfor-ward but impaired top-down processes in the vegetative state. Science 332:858–862. CrossRef Medline

Bueti D, Bahrami B, Walsh V (2008) Sensory and association cortex in time perception. J Cogn Neurosci 20:1054–1062. CrossRef Medline

Carlin JD, Calder AJ, Kriegeskorte N, Nili H, Rowe JB (2011) A head view-invariant representation of gaze direction in anterior superior temporal sulcus. Curr Biol 21:1817–1821. CrossRef Medline

Cheng CH, Baillet S, Hsiao FJ, Lin YY (2013) Effects of aging on neuromag-netic mismatch responses to pitch changes. Neurosci Lett 544:20–24. CrossRef Medline

Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibáñez A, Owen AM, Bekinschtein TA (2013) Expectation and attention in hier-archical auditory prediction. J Neurosci 33:11194–11205. CrossRef Medline

Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36:181–204. CrossRef Medline

Cooray G, Garrido MI, Hyllienmark L, Brismar T (2014) A mechanistic model of mismatch negativity in the ageing brain. Clin Neurophysiol 125:1774–1782. CrossRef Medline

Coull JT, Cheng RK, Meck WH (2011) Neuroanatomical and neurochemi-cal substrates of timing. Neuropsychopharmacology 36:3–25. CrossRef Medline

Croxson PL, Johansen-Berg H, Behrens TE, Robson MD, Pinsk MA, Gross CG, Richter W, Richter MC, Kastner S, Rushworth MF (2005) Quanti-tative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. J Neurosci 25:8854–8866. CrossRef Medline

Daltrozzo J, Wioland N, Mutschler V, Kotchoubey B (2007) Predicting coma and other low responsive patients outcome using event-related brain po-tentials: a meta-analysis. Clin Neurophysiol 118:606–614. Medline

David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ (2006) Dynamic causal modelling of evoked responses in EEG and MEG. Neu-roimage 30:1255–1272. CrossRef Medline

Dietz MJ, Friston KJ, Mattingley JB, Roepstorff A, Garrido MI (2014) Effec-tive connectivity reveals right-hemisphere dominance in audiospatial perception: implications for models of spatial neglect. J Neurosci 34:5003–5011. CrossRef Medline

Doeller CF, Opitz B, Mecklinger A, Krick C, Reith W, Schröger E (2003) Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. Neuroimage 20: 1270–1282. CrossRef Medline

Egner T, Monti JM, Summerfield C (2010) Expectation and surprise determine neural population responses in the ventral visual stream. J Neurosci 30:16601–16608. CrossRef Medline

El Karoui I, King JR, Sitt J, Meyniel F, Van Gaal S, Hasboun D, Adam C, Navarro V, Baulac M, Dehaene S, Cohen L, Naccache L (2014) Event-related potential, time-frequency, and functional connectivity facets of local and global auditory novelty processing: an intracranial study in humans. Cereb Cortex pii:bhu143. CrossRef Medline

Fisher DJ, Grant B, Smith DM, Knott VJ (2011) Effects of deviant probability on the "optimal" multi-feature mismatch negativity (MMN) paradigm. Int J Psychophysiol 79:311–315. CrossRef Medline

Fletcher PC, Frith CD (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci 10:48–58. CrossRef Medline

Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–480. CrossRef Medline

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

Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. Philos Trans R Soc Lond B Biol Sci 364:1211–1221. CrossRef Medline

Friston K, Mattout J, Kilner J (2011) Action understanding and active inference. Biol Cybern 104:137–160. CrossRef Medline

Friston K, Daunizeau J, Stephan KE (2013) Model selection and gobbledygook: response to Lohmann et al. Neuroimage 75:275–278, discussion 279–281. CrossRef Medline

Friston KJ, Stephan KE (2007) Free-energy and the brain. Synthese 159: 417–458. CrossRef Medline

Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J Neurosci 32:1791–1802. CrossRef Medline

Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ (2007a) Dynamic causal modelling of evoked potentials: a reproducibility study. Neuroimage 36:571–580. CrossRef Medline

Garrido MI, Kilner JM, Kiebel SJ, Friston KJ (2007b) Evoked brain responses are generated by feedback loops. Proc Natl Acad Sci U S A 104: 20961–20966. CrossRef Medline

Garrido MI, Friston KJ, Kiebel SJ, Stephan KE, Baldeweg T, Kilner JM (2008) The functional anatomy of the MMN: a DCM study of the roving paradigm. Neuroimage 42:936–944. CrossRef Medline

Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009a) The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol 120:453–463. CrossRef Medline

Garrido MI, Kilner JM, Kiebel SJ, Friston KJ (2009b) Dynamic causal modeling of the response to frequency deviants. J Neurophysiol 101:2620–2631. CrossRef Medline

Giard MH, Perrin F, Pernier J, Bouchet P (1990) (1990) Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. Psychophysiology 27:627–640. CrossRef Medline

Giard MH, Lavikahen J, Reinikainen K, Perrin F, Bertrand O, Pernier J, Näätänen R (1995) Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. J Cogn Neurosci 7:133–143. CrossRef Medline

Grahn JA, Rowe JB (2009) Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. J Neurosci 29:7540–7548. CrossRef Medline

Grahn JA, Rowe JB (2013) Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. Cereb Cortex 23:913–921. CrossRef Medline

Hughes HC, Darcey TM, Barkan HI, Williamson PD, Roberts DW, Aslin CH (2001) Responses of human auditory association cortex to the omission of an expected acoustic event. Neuroimage 13:1073–1089. CrossRef Medline

Hughes LE, Rowe JB (2013) The impact of neurodegeneration on network connectivity: a study of change detection in frontotemporal dementia. J Cogn Neurosci 25:802–813. CrossRef Medline

Hughes LE, Ghosh BC, Rowe JB (2013) Reorganisation of brain networks in frontotemporal dementia and progressive supranuclear palsy. Neuroimage 2:459–468. CrossRef Medline

Jemel B, Achenbach C, Müller BW, Röpcke B, Oades RD (2002) Mismatch negativity results from bilateral asymmetric dipole sources in the frontal and temporal lobes. Brain Topogr 15:13–27. CrossRef Medline

Johnson HA, Goel A, Buonomano DV (2010) Neural dynamics of in vitro cortical networks reflects experienced temporal patterns. Nat Neurosci 13:917–919. CrossRef Medline

Kärgel C, Sartory G, Kariofillis D, Wiltfang J, Müller BW (2014) Mismatch negativity latency and cognitive function in schizophrenia. PloS One 9:e84536. CrossRef Medline

Karmarkar UR, Buonomano DV (2007) Timing in the absence of clocks: encoding time in neural network states. Neuron 53:427–438. CrossRef Medline

Kass RE, Raftery AE (1995) Bayes factors. J Am Stat Assoc 90:773–795. CrossRef

Kiebel SJ, David O, Friston KJ (2006) Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. Neuroimage 30:1273–1284. CrossRef Medline

Kiebel SJ, Daunizeau J, Friston KJ (2008) A hierarchy of time-scales and the brain. PLoS Comput Biol 4: e1000209. CrossRef Medline

Kiebel SJ, Garrido MI, Moran R, Chen CC, Friston KJ (2009) Dynamic causal modeling for EEG and MEG. Hum Brain Mapp 30:1866–1876. CrossRef Medline

Kimura M, Schröger E, Czigler I (2011) Visual mismatch negativity and its importance in visual cognitive sciences. Neuroreport 22:669–673. CrossRef Medline

Kujala T, Lovio R, Lepistö T, Laasonen M, Näätänen R (2006) Evaluation of multi-attribute auditory discrimination in dyslexia with the mismatch negativity. Clin Neurophysiol 117:885–893. CrossRef Medline

Lehéricy S, Ducros M, Van de Moortele PF, Francois C, Thivard L, Poupon C, Swindale N, Ugurbil K, Kim DS (2004) Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. Ann Neurol 55:522–529. CrossRef Medline

Leon MI, Shadlen MN (2003) Representation of time by neurons in the posterior parietal cortex of the macaque. Neuron 38:317–327. CrossRef Medline

Lieder F, Daunizeau J, Garrido MI, Friston KJ, Stephan KE (2013a) Modelling trial-by-trial changes in the mismatch negativity. PLoS Comput Biol 9:e1002911. CrossRef Medline

Lieder F, Stephan KE, Daunizeau J, Garrido MI, Friston KJ (2013b) A neurocomputational model of the mismatch negativity. PLoS Comput Biol 9:e1003288. CrossRef Medline

Lohmann G, Erfurth K, Müller K, Turner R (2012) Critical comments on dynamic causal modelling. Neuroimage 59:2322–2329. CrossRef Medline

May PJ, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47:66–122. CrossRef Medline

Meck WH, Penney TB, Pouthas V (2008) Cortico-striatal representation of time in animals and humans. Curr Opin Neurobiol 18:145–152. CrossRef Medline

Merchant H, Zarco W, Prado L (2008) Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. J Neurophysiol 99:939–949. CrossRef Medline

Merchant H, Harrington DL, Meck WH (2013) Neural basis of the perception and estimation of time. Ann Rev Neurosci 36:313–336. CrossRef Medline

Michie PT, Budd TW, Todd J, Rock D, Wichmann H, Box J, Jablensky AV (2000) Duration and frequency mismatch negativity in schizophrenia. Clin Neurophysiol 111:1054–1065. CrossRef Medline

Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ (2005) The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. Cereb Cortex 15:545–551. CrossRef Medline

Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ (2013) Free energy, precision and learning: the role of cholinergic neuromodulation. J Neurosci 33:8227–8236. CrossRef Medline

Moran RJ, Symmonds M, Dolan RJ, Friston KJ (2014) The brain ages optimally to model its environment: evidence from sensory learning over the adult lifespan. PLoS Comput Biol 10: e1003422. CrossRef Medline

Morlet D, Fischer C (2014) MMN and novelty P3 in coma and other altered states of consciousness: a review. Brain Topogr 27:467–479. CrossRef Medline

Muckli L (2010) What are we missing here? Brain imaging evidence for higher cognitive functions in primary visual cortex V1. Int J Imaging Syst Technol 20:131–139. CrossRef

Näätänen R, Paavilainen P, Tiitinen H, Jiang D, Alho K (1993) Attention and mismatch negativity. Psychophysiology 30:436–450. CrossRef Medline

Näätänen R, Pakarinen S, Rinne T, Takegata R (2004) The mismatch negativity (MMN): towards the optimal paradigm. Clin Neurophysiol 115: 140–144. CrossRef Medline

Näätänen R, Kujala T, Escera C, Baldeweg T, Kreegipuu K, Carlson S, Ponton C (2012) The mismatch negativity (MMN)–a unique window to disturbed central auditory processing in ageing and different clinical conditions. Clin Neurophysiol 123:424–458. CrossRef Medline

Novak MJU, Seunarine KK, Gibbard CR, McColgan P, Draganski B, Friston K, Clark CA, Tabrizi SJ (2015) Basal ganglia-cortical structural connectivity in Huntington's disease. Hum Brain Mapp 36:1728–1740. CrossRef Medline

Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E (2002) Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage 15:167–174. CrossRef Medline

Pecenka N, Keller PE (2011) The role of temporal prediction abilities in interpersonal sensorimotor synchronization. Exp Brain Res 211:505–515. CrossRef Medline

Penny WD, Stephan KE, Mechelli A, Friston KJ (2004) Comparing dynamic causal models. Neuroimage 22:1157–1172. CrossRef Medline

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. PLoS Comput Biol 6:e1000709. CrossRef Medline

Petermann M, Kummer P, Burger M, Lohscheller J, Eysholdt U, Döllinger M (2009) Statistical detection and analysis of mismatch negativity derived by a multi-deviant design from normal hearing children. Hear Res 247: 128–136. CrossRef Medline

Postuma RB, Dagher A (2006) Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. Cereb Cortex 16:1508–1521. CrossRef Medline

Raij T, McEvoy L, Mäkelä JP, Hari R (1997) Human auditory cortex is activated by omissions of auditory stimuli. Brain Res 745:134–143. CrossRef Medline

Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87. CrossRef Medline

Rao SM, Mayer AR, Harrington DL (2001) The evolution of brain activation during temporal processing. Nat Neurosci 4:317–323. CrossRef Medline

Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, NäätänenR (2000) Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage 12:14–19. CrossRef Medline

Rinne T, Degerman A, Alho K (2005) Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: an fMRI study. Neuroimage 26:66–72. CrossRef Medline

Schmidt A, Diaconescu AO, Kometer M, Friston KJ, Stephan KE, Vollenweider FX (2013) Modeling ketamine effects on synaptic plasticity during the mismatch negativity. Cereb Cortex 23:2394–2406. CrossRef Medline

Schröger E, Winkler I (1995) Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. Neurosci Lett 193:185–188. CrossRef Medline

Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ (2010) Ten simple rules for dynamic causal modeling. Neuroimage 49: 3099–3109. CrossRef Medline

Teki S, Grube M, Griffiths TD (2011a) A unified model of time perception accounts for duration-based and beat-based timing mechanisms. Front Integr Neurosci 5:90. CrossRef Medline

Teki S, Grube M, Kumar S, Griffiths TD (2011b) Distinct neural substrates of duration-based and beat-based auditory timing. J Neurosci 31:3805–3812. CrossRef Medline

van Wassenhove V, Nagarajan SS (2007) Auditory cortical plasticity in learning to discriminate modulation rate. J Neurosci 27:2663–2672. CrossRef Medline

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A 108:20754–20759. CrossRef Medline

Wager TD, Keller MC, Lacey SC, Jonides J (2005) Increased sensitivity in neuroimaging analyses using robust regression. Neuroimage 26:99–113. CrossRef Medline

Winkler I (2007) Interpreting the mismatch negativity. J Psychophysiol 21: 147–163. CrossRef

Yeterian EH, Pandya DN (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. J Comp Neurol 312:43–67. CrossRef Medline

Zanto TP, Snyder JS, Large EW (2006) Neural correlates of rhythmic expectancy. Adv Cogn Psychol 2:221–231. CrossRef

Zhang J, Rowe JB (2015) The neural signature of information regularity in temporally extended event sequences. Neuroimage 107:266–276. CrossRef Medline