

UCSF

UC San Francisco Previously Published Works

Title

Direct Evidence for Prediction Signals in Frontal Cortex Independent of Prediction Error

Permalink

<https://escholarship.org/uc/item/0k39m5p8>

Journal

Cerebral Cortex, 29(11)

ISSN

1047-3211

Authors

Dürschmid, Stefan
Reichert, Christoph
Hinrichs, Hermann
et al.

Publication Date

2019-12-17

DOI

10.1093/cercor/bhy331

Peer reviewed

ORIGINAL ARTICLE

Direct Evidence for Prediction Signals in Frontal Cortex Independent of Prediction Error

Stefan Dürschmid^{1,2}, Christoph Reichert^{1,6}, Hermann Hinrichs^{1,2,3,4,5,6}, Hans-Jochen Heinze^{1,2,3,4,5,6}, Heidi E. Kirsch⁷, Robert T. Knight⁸ and Leon Y. Deouell⁹

¹Department of Behavioral Neurology, Leibniz Institute for Neurobiology, Brennekestr. 6, 39120 Magdeburg, Germany, ²Department of Neurology, Otto-von-Guericke University, Leipziger Str. 44, 39120 Magdeburg, Germany, ³Stereotactic Neurosurgery, Otto-von-Guericke University, Leipziger Str. 44, 39120 Magdeburg, Germany, ⁴German Center for Neurodegenerative Diseases (DZNE), Leipziger Str. 44, 39120 Magdeburg, Germany, ⁵Forschungscampus STIMULATE, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany, ⁶CBBS—Center of Behavioral Brain Sciences, Otto-von-Guericke University, Universitätsplatz 2, 39106 Magdeburg, Germany, ⁷Department of Neurology, University of California, 400 Parnassus Avenue, San Francisco, CA 94122, USA, ⁸Helen Wills Neuroscience Institute and Department of Psychology, University of California, Berkeley, CA 94720, USA and ⁹Edmond and Lily Safra Center for Brain Sciences and Department of Psychology, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91904, Israel

Address correspondence to Stefan Dürschmid, Department of Behavioral Neurology, Leibniz Institute for Neurobiology, Brennekestr. 6, 39120 Magdeburg, Germany. Email: stefan.duerschmid@googlemail.com and Leon Deouell, Edmond and Lily Safra Center for Brain Research and Department of Psychology, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. Email: leon.deouell@mail.huji.ac.il

Abstract

Predictive coding (PC) has been suggested as one of the main mechanisms used by brains to interact with complex environments. PC theories posit top-down prediction signals, which are compared with actual outcomes, yielding in turn prediction error (PE) signals, which are used, bottom-up, to modify the ensuing predictions. However, disentangling prediction from PE signals has been challenging. Critically, while many studies found indirect evidence for PC in the form of PE signals, direct evidence for the prediction signal is mostly lacking. Here, we provide clear evidence, obtained from intracranial cortical recordings in human surgical patients, that the human lateral prefrontal cortex evinces prediction signals while anticipating an event. Patients listened to task-irrelevant sequences of repetitive tones including infrequent predictable or unpredictable pitch deviants. The broadband high-frequency amplitude (HFA) was decreased prior to the onset of expected relative to unexpected deviants in the frontal cortex only, and its amplitude was sensitive to the increasing likelihood of deviants following longer trains of standards in the unpredictable condition. Single-trial HFA predicted deviations and correlated with poststimulus response to deviations. These results provide direct evidence for frontal cortex prediction signals independent of PE signals.

Key words: frontal cortex, high gamma activity, predictive coding, prestimulus activity, temporal cortex

“Prediction is very difficult, especially if it’s about the future”

Niels Bohr

Introduction

Making predictions about upcoming events is a crucial brain function. Predictive coding (PC) theories postulate that the brain iteratively optimizes an internal model of the environment based on sensory inputs (Rao and Ballard 1999; Bastos et al. 2012; Lee and Noppeney 2014; Heilbron and Chait 2018) and generates prediction error (PE) signals if predictions are violated (Winkler and Schröger 2015), to improve the future interaction with the environment. Most PC schemes suggest separate prediction and PE signals/neurons, but separating the 2 in practice has proved challenging (for a comprehensive introduction and review, see Heilbron and Chait (2018)). One important reason is that most evidence for an anticipatory pre-event prediction comes from the ultimate PE signals, elicited after the (un)predicted event has occurred. Hence, recording predictive signals prior to the onset of the stimuli would be strong evidence for prospective, active predictions.

A critical question is also whether predictions are formed automatically (by default) even when the stimuli are not attended. Here, we utilized the high-temporal and -spectral resolution of direct cortical recordings from subdural ECoG electrodes to compare frontal and temporal prediction signals in 5 patients exposed with trains of task-irrelevant and meaningless auditory stimuli in 2 conditions, while attending a visual slide show. The conditions differed in the predictability of deviation from repetitive background stimuli. In “regular” sequences, every deviant followed exactly 4 standards, whereas in “irregular” sequences, deviants were randomly embedded in trains of standard stimuli.

In a previous report, we concentrated on poststimulus activity variations as a response to fully predictable and unpredictable deviants (i.e., on the PE) using the same data set (Dürschmid et al. 2016). Here, we show that in frontal cortex modulation of prestimulus broadband high-frequency amplitude (HFA) heralds ensuing deviants and correlates with the poststimulus PE signal. In contrast, and commensurate with poststimulus activity, prestimulus activity in temporal cortex is insensitive to sequence statistics but reflects only the immediate history.

Methods

Patients

Five epilepsy patients (mean age 33, SD = 9.23) undergoing pre-surgical monitoring with subdural electrodes participated in the experiment after providing their written informed consent. Experimental and clinical recordings were taken in parallel. Recordings took place at the University of California, San Francisco (UCSF) and were approved by the local ethics committees (“Committee for the Protection of Human Subjects at UC Berkeley”). The analysis of the poststimulus effects from these patients with the same data set was previously reported by Dürschmid et al. (2016).

Stimuli

Participants listened to stimuli consisting of 180 ms long (10 ms rise and fall time) harmonic sounds with a fundamental frequency of 500 or 550 Hz and the 3 first harmonics with descending

amplitudes (−6, −9, −12 dB relative to the fundamental). The stimuli were generated using Cool Edit 2000 software (Syntrillium). The stimuli were presented from loudspeakers positioned at the foot of the subject’s bed at a comfortable loudness.

Procedure

While reclined in their hospital bed, participants watched an engaging slide show while sound trains were played in the background. Sound trains included high-probability standards ($P = 0.8$; $f_0 = 500$ Hz) mixed with low-probability deviants ($P = 0.2$; $f_0 = 550$ Hz) in blocks of 400 sounds, with a stimulus onset asynchrony (SOA) of 600 ms. In different blocks, the order of the sounds was either pseudorandom, with a minimum of 3 standard tones before a deviant (irregular condition), or regular, such that exactly every fifth sound was a deviant (Fig. 1A). Thus, under the regular condition, standards and deviants were fully predictable, whereas under the irregular condition, exact prediction was not possible.

Data Recording

The electrocorticogram (ECoG) was recorded at UCSF using 64 platinum-iridium electrode grids arranged in an 8×8 array with 10 mm center-to-center spacing (Ad-Tech Medical Instrument Corporation; see Fig. 2 for grid location). Grids were positioned based solely on clinical needs. Exposed electrode diameter was 2.3 mm. The data were recorded continuously throughout the task at a sampling rate of 2003 Hz.

Preprocessing

We used Matlab 2013b (The Mathworks) for all offline data processing. All filtering were done using zero phase-shift IIR filters. We excluded channels exhibiting ictal activity or excessive noise from further analysis. In the remaining “good” channels, we then excluded time intervals containing artifactual signal distortions such as signal steps of pulses by visual inspection. Finally, we rereferenced the remaining electrode time series by subtracting the common average reference

$$x_{\text{CAR}}(t) = \frac{1}{n} \sum_{c=1}^n x_c(t)$$

calculated over the n good channels c from each channel time series x_c . The resulting time series were used to characterize brain dynamics over the time course of auditory stimulus prediction. For each trial (−1 to 2 s around stimulus onset—sufficiently long to prevent any edge effects during filtering) we band-pass filtered each electrode’s time series in the broadband high frequency range (80–150 Hz; see Supplementary Material). We obtained the analytic amplitude $A_f(t)$ of this band by Hilbert-transforming the filtered time series (HFA). We smoothed the HFA time series such that amplitude value at each time point t is the mean of 10 ms around each time point t . We then baseline-corrected by subtracting from each data point the mean activity of the −700 to −600 ms preceding the stimulus onset (i.e., 100 ms prior to trial $N - 1$) in each trial and each channel.

Prestimulus time series of HFA were used for the following analysis steps (explained in detail in the following). We first parameterized the prediction of upcoming stimuli as the interaction of Stimulus type (standard, deviant) and Block type (regular, irregular) using a time-resolved ANOVA (“I—Estimation of Prediction”). Next, we assessed the involvement of frontal or temporal cortices in this prediction effect (“II—Comparison

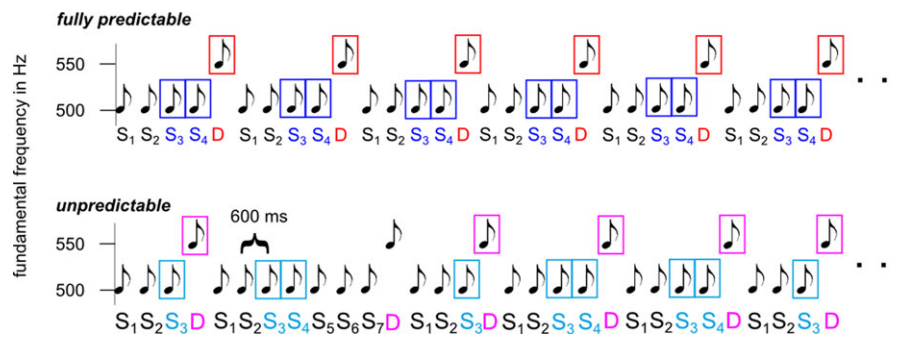


Figure 1. Paradigm. Participants watched a slide show while hearing passively sequences of sounds. High-probability standards mixed with low-probability deviants were presented either unpredictably or were fully predictable (exactly every fifth sound was a deviant). Standards (S_{1-n}) are numbered based on their position relative to the previous deviant. Only standards following at least 2 standards were used for analysis (marked by rectangles)

Between Temporal and Frontal Cortices”). Finally, we tested for an increasing predictability of deviants under the irregular condition following longer trains of standards (“III—Increase in Predictability as a Function of Train Length”).

I—Estimation of Prediction

Given the fixed repetition of 4 standards followed by a deviant under the regular condition, the occurrence of both standards and deviants should be predictable. We assumed that in areas with predictive activity, the activity P prior to (expected) deviants should be different from the brain activity prior to frequent standards:

$$P_{\text{standard}} \neq P_{\text{deviant}}$$

Conversely, since under the irregular condition the system does not know a priori which stimulus will be heard the most frequent class (standard tone) is predicted, and, as a result, the activity P prior to the standards and deviants is equal:

$$P_{\text{standard}} \approx P_{\text{deviant}}$$

Statistically, the difference between conditions can be expressed as an effect of interaction using a 2-way ANOVA with the factors stimulus type (upcoming standard vs. upcoming deviant) and block type (regular vs. irregular), with the stimulus type effect expected to be larger in the regular than irregular condition. We ran this 2-way ANOVA for each electrode (with trials as random variable), at every time point, with HFA as the dependent variable. This leads to 3F-value time series (2 main effects and one interaction: $F_{\text{stimulus type}}$, $F_{\text{block type}}$, $F_{\text{interaction}}$) for each channel with the $F_{\text{interaction}}$ capturing the prediction effect. The level of significance was corrected for multiple comparisons as described below.

Only deviants following the third and the fourth standard in a row (S_3 and S_4 , respectively; see Supplementary Material for a full list of trials subjected to analysis) under the irregular condition were included in the analysis. All deviants following S_5, \dots, S_N were excluded (see Fig. 1 and Supplementary Material). This results in a pool of deviant trials which consist of regular deviants which always occurred after S_4 and irregular deviants following S_3 and S_4 . Note that due to the design of the quasi-random sequence under the irregular condition, with the constraint of at least 3 standards before a deviant, the probability of deviants occurring after S_3 and S_4 was nearly identical (0.17 and 0.2 respectively; see discussion of the hazard function in the following).

The pool of standard trials included only S_3 and S_4 trials under both the regular and irregular conditions. We did not

include the first and second standards after a deviant, since during the prestimulus interval of S_1 a deviant is presented and the prestimulus interval of S_2 might still be influenced by the preceding deviant due to the short ISI. We excluded S_5, \dots, S_N trials under the irregular condition since we hypothesized that the occurrence of deviants would be increasingly expected due to the “hazard function.” That is, we hypothesized that while longer trains of standards under the irregular condition increase the local probability of the standard, the occurrence of deviants also becomes more likely: since a deviant has not occurred for an extended sequence of events, its likelihood increases. By not including irregular deviant following S_5, \dots, S_N we also made the conditions more comparable for analysis, as under the regular conditions deviants never appeared after 5 or more standards. We focused on high-frequency broadband HFA, which in our previous study showed earlier poststimulus deviation signals than low-frequency ERPs (Dürschmid et al., 2016) and differentiated between fully predictable and unpredictable deviation in frontal and temporal cortices (for prediction signal in other bands of the time-frequency spectrum see Supplementary Material).

II—Comparison Between Temporal and Frontal Cortices

Principal component analysis

As noted in step I, the $F_{\text{interaction}}$ captures the prediction effect. We tested whether the $F_{\text{interaction}}$ effect is localized to the temporal or the frontal cortex in the following way. The $F_{\text{interaction}}$ time series were calculated in all channels separately over frontal and temporal regions of interest (ROIs). A principal component analysis (PCA) was used to find the course of a common $F_{\text{interaction}}$ across time, accounting for the highest variance, separately within the set of frontal and temporal channels. Channels loading highly on the first principal component are those that exhibit the strongest variation in terms of interaction amplitude across time.

Data reduction

We chose the channels for which the Pearson correlation r with the principal component exceeded the 75th percentile of all positive r -values. We set this level as a trade-off between a higher statistical power of a smaller number of channels and a stronger generalization across the cortex with a higher number of channels. We averaged the $F_{\text{interaction}}$ -values in these channels and checked whether the averaged $F_{\text{interaction}}$ -values in each region exceeded the empirically determined threshold derived from a surrogate distribution. This surrogate distribution

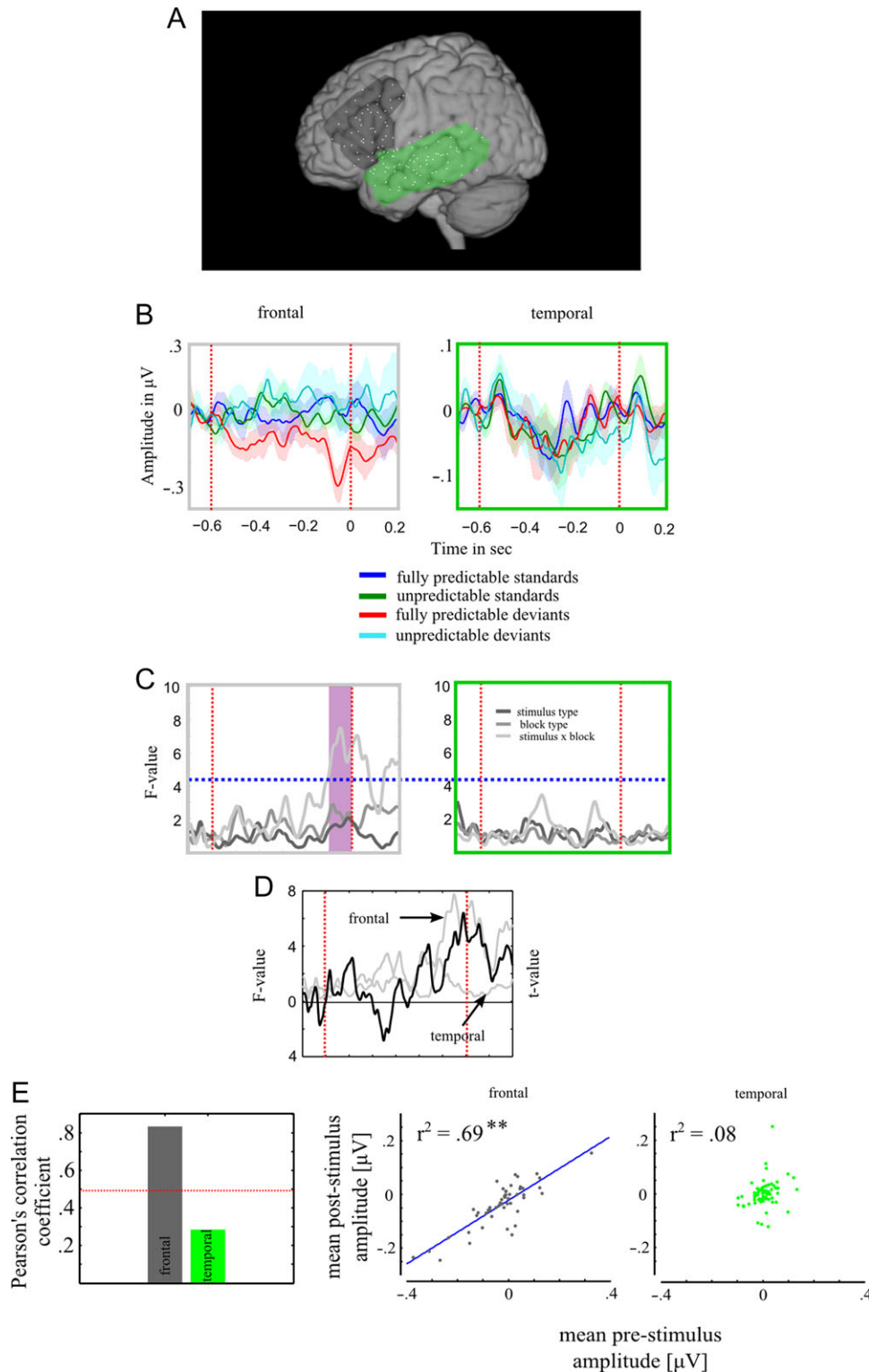


Figure 2. Time-resolved analysis of variance. (A) Frontal (gray) and temporal (green) regions of interest (ROIs). (B) Baseline-corrected HFA modulation prior to both stimulus types under both conditions (shaded areas denote the standard error across channels). (C) Mean F (ME—main effects, IE—Interaction effects) time series of channels loading highly on the frontal (gray frame) and temporal (green frame) $F_{\text{interaction}}$ first principal components. The horizontal dashed blue line indicates the critical $F_{\text{interaction}}$ value based on permutation. The shaded area in the left panel indicates the temporal interval of significant interaction. (D) $F_{\text{interaction}}$ time series for frontal and temporal electrodes (indicated by arrows) together with the t-values (black line) of the difference between the 2 ROIs in the degree of Type X Block interaction. Frontal cortex shows stronger interaction before stimulus onset. (E) Correlation between prestimulus and poststimulus HFA across channels over the frontal and temporal cortices. Left: Pearson's correlation values for each of the 2 ROIs. The dashed line gives the 99% confidence interval of the surrogate distribution. Middle: covariation of prestimulus and poststimulus amplitude of electrodes over the frontal ROI. Each dot represents one electrode. The blue line shows the linear fit to the data. Right: covariation of prestimulus and poststimulus amplitude of electrodes over the temporal ROI. Only in the frontal ROI pre and poststimulus amplitude are significantly correlated.

of the interaction effect was constructed by randomly reassigning the labels (standard, deviant, regular, irregular) to the single trials in 1000 permutations for each channel. This leads to 1000 surrogate $F_{\text{interaction}}$ time series. Significance criterion was an $F_{\text{interaction}}$ -value with $P < 0.01$ within the surrogate distribution of all $F_{\text{interaction}}$ values. We next compared $F_{\text{interaction}}$ effects between frontal and temporal electrodes with an unpaired t-test at each time point between the 2 groups of electrodes (frontal vs. temporal). To determine significance, in 1000 runs we randomly reassigned the labels (temporal vs. frontal) and applied the unpaired t-test.

Group (within-subject) analysis

In the first analysis, we have chosen channels loading highly on the first principal component which are those that exhibit the strongest variation in terms of interaction amplitude across time, regardless of which subject they were taken from. As statistical significance in the analysis across electrodes might be driven by single subjects, we verify that the results presented are valid at a group level. To that end, we repeated the above stages at the single-subject level, using a 2-step procedure (common in fMRI studies). At the first level, we ran the above ANOVA in each subject separately, across trials. Then, as done in the previous section, we ran a PCA on the $F_{\text{interaction}}$ time series for each subject within region, maintained the channels with the highest loading on the first PC, and averaged their $F_{\text{interaction}}$ time series. This led to 2 time series for each subject, one for the temporal channels and one for the frontal. Then, we ran a second level analysis to determine, at each time point and for each region, whether the $F_{\text{interaction}}$ exceeded the significance level, at the group level (i.e., with subjects as random variable). Significance was determined relative to a permutation-derived surrogate distribution of the interaction effect. The distribution was constructed by randomly reassigning the labels (standard, deviant, regular, irregular) to the single trials in 1000 permutations for each channel. This leads to 1000 surrogate $F_{\text{interaction}}$ time series. Significance criterion was an $F_{\text{interaction}}$ -value with $P < 0.01$ within the surrogate distribution of all $F_{\text{interaction}}$ values.

III—Increase in Predictability as a Function of Train Length

Throughout the experiment, we pseudorandomly varied the train length of standards under the irregular condition. This resulted in standard trains of 3–8 standards before deviants. We directly tested whether predictability varies as a function of train length under the irregular condition, congruent with a hazard function (the probability of a deviant increases from 0 after 1 and 2 standards, to 0.17 (1/6) after 3 consecutive standards, and gradually increases to 1 after 8 consecutive standards have occurred in a row). We hypothesized that if HFA modulation correlates with predictability of the next stimulus, then longer standard trains would result in stronger modulation of HFA before the occurrence of deviants. Specifically, we correlated the HFA preceding deviants with the length of the standard train before deviant under the irregular condition. While in the previous analysis, we only used deviants following S_3 and S_4 , here all deviants entered the analysis. To assess significance, Pearson's correlation coefficient of each channel was compared against a surrogate distribution. This surrogate distribution was constructed by randomly reassigning the actual train lengths of single-trial predeviant HFA values in 1000 runs.

For each channel, the confidence intervals (CI; 99.5%) of a normal distribution were determined.

Results

Comparison Between Temporal and Frontal Cortices

We studied 287 channels across all subjects, of which 120 were centered over frontal and temporal cortices. HFA was subject to a Stimulus Type (predeviant, prestandard) \times Block Type (regular, irregular) ANOVA at every time point from -700 to $+200$ ms and we evaluated the interaction term ($F_{\text{interaction}}$) as a signature of predictive activity, separately for all frontal ($N_{\text{frontal}} = 54$) and all temporal channels ($N_{\text{temporal}} = 66$; Fig. 2A). Within each region, we kept the channels loading highly on the first temporal principle component of the $F_{\text{interaction}}$ time series and compared their mean with the empirical surrogate distribution (Step I of data analysis in methods; Fig. 2C). Frontal HFA ($N_{\text{elec}} = 7$) activity showed significant $F_{\text{interaction}}$ values (maximal $F_{\text{interaction}} = 7.76$ $P < 0.00001$, at -51.4 ms) with neither a significant effect of stimulus type (maximal $F_{\text{stimulus type}} = 2.44$) nor of block type (maximal $F_{\text{block type}} = 3.36$) (left panel in Fig. 2C). Temporal activity ($N_{\text{elec}} = 10$) did not show significant F-values for any of the 3 effects (maximal $F_{\text{stimulus type}} = 3.37$; maximal $F_{\text{block type}} = 2.02$; maximal $F_{\text{interaction}} = 3.47$) (right panel in Fig. 2C). The high $F_{\text{interaction}}$ values in frontal cortex correspond in time with a decrease in HFA from -100 ms before and until the onset of deviants, compared with the onset of standards, in the regular blocks (where deviants and standards were predictable) but not in the irregular blocks (Fig. 2B, see Supplementary Material for parallel results at a single-trial level). $F_{\text{interaction}}$ effects were significantly larger in frontal than temporal sites ($t_{15} = 6.49$, permutation based $P < 0.00001$ at -11 ms; Fig. 2D). These results were confirmed at the group level (Supplementary Fig. 2): $F_{\text{interaction}}$ -values in the frontal lobe exceeded the empirical significance threshold ($F_{\text{crit}} = 4.2$) between -0.099 and 0.02 s ($F_{\text{max}} = 6.8$) prior to the onset of the deviants. $F_{\text{interaction}}$ averaged across this interval were significantly different between frontal and temporal cortices ($P < 0.05$; signed-rank test (for paired samples)).

Correlation Between Prestimulus and Poststimulus Responses

Previously, we found that postdeviant HFA was reduced under the regular condition compared with the irregular condition in frontal electrodes (Dürschmid et al. 2016). Since we now found that predictable deviants under the regular condition are heralded by a prestimulus HFA decrease, we tested if the 2 phenomena are correlated. First, both in the frontal ($N_{\text{electrodes}} = 54$) and the temporal ROIs ($N_{\text{electrodes}} = 66$) we correlated HFA preceding stimulus onset (average across -100 to 0 ms) with the amplitude following stimulus onset (average across 0 – 300 ms) across channels. The 2 resulting Pearson's correlation values were tested against a surrogate distribution. This surrogate distribution was constructed by randomly assigning the prestimulus values of each channel with poststimulus values from another channel in 1000 iterations. Based on the distribution of r -values in this permutation analysis, the critical r -value denoting statistical significance was $r = 0.5$. Prestimulus amplitude correlated with poststimulus amplitude in frontal cortex ($r = 0.83$; $P = 0.000002$) but not in the temporal cortex ($r = 0.28$; see Fig. 2E). Next, we tested whether the prestimulus/poststimulus relation is also true at a single-trial level. Hence, we correlated within each electrode the average amplitude in the prestimulus and poststimulus interval across trials. Each individual Pearson's r was compared against a surrogate distribution and

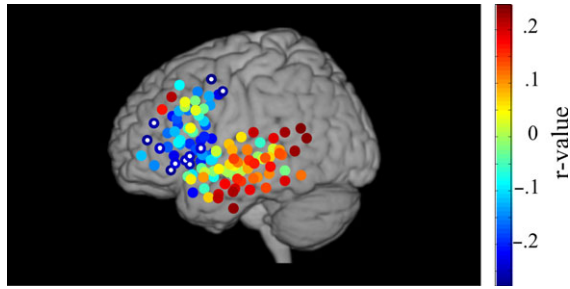


Figure 3. Prefrontal electrodes reflect the hazard function in irregular sequences. Each circle depicts channel positions with the color coding Pearson's correlation coefficient between train length and predeviant HFA. Channels with a white dot show a statistically significant correlation. HFA significantly decreased after longer trains of standards in frontal cortex, while HFA tended to increase with longer trains of standards in temporal cortex.

excluded if smaller than the critical value ($r_{\text{crit}} = 0.1$). This surrogate distribution was constructed by randomly reassigning the prestimulus value of one trial to poststimulus value of another trial by randomly permuting the prestimulus values in 1000 iterations. On average, electrodes in the frontal cortex showed higher r -values than temporal ones (frontal: 0.39; temporal: 0.28; $t_{102} = 3.9$, $P = 0.0002$).

Increase in Predictability as a Function of Train Length

The train length of standards under the irregular condition varied pseudorandomly, allowing us to test whether prestimulus predictive activity varies gradually as a function of train length. We surmised that 2 effects could be operative. Temporally local effects suggest that the probability of a standard tone increases the more standard tones which are played in a row. In contrast, using a more global strategy, the so-called “hazard function” suggests that, given that deviations will happen eventually, expectation of a deviant increases the longer it is since the last deviation. To test whether and where such effects prevail, we correlated predeviant HFA with train length of standards before deviants. Figure 3 shows that the direction of correlation between HFA and standard train length was different between temporal electrodes, showing mostly positive correlations, and frontal electrodes, showing mostly negative correlations. Individually, only the negative correlations in frontal channels reached the permutation critical r -values of $r_{\text{crit}} = \pm 0.19$ (white dots in Fig. 3). Considering that the analysis of the regular versus irregular condition indicated that a decrease in HFA indicates proactive prediction of a deviant, these results suggest that frontal electrodes “apply” predictions even under the irregular condition based on the more global hazard function strategy.

Discussion

PC theories suggest that the brain continuously uses available information to predict forthcoming events and reduce sensory uncertainty (Arnal and Giraud 2012). However, the evidence supporting this notion comes mainly from postevent PE findings (Summerfield et al. 2008; Fogelson et al. 2009; Alink et al. 2010; den Ouden et al. 2010; Todorovic et al. 2011; Winkler and Czigler 2012; Sanmiguel et al. 2013; Bendixen et al. 2014, 2015; Dürschmid et al. 2016), providing only indirect evidence for prediction, since prediction-based neural activity should precede a predicted event. Here, we provide direct evidence for the

prediction of rare deviant events manifested by prestimulus HFA modulation, suggesting an automatic anticipation of the upcoming deviant.

Regular, and thus predictable, deviations were preceded by HFA decrease exclusively in the lateral frontal cortex, observed at both the group and single-trial levels. This complements our previous results, showing that lateral frontal (but not temporal) sites show reduced postevent PE signals to predictable compared with unpredictable stimuli (Dürschmid et al. 2016). Moreover, the predictive prestimulus power reduction correlated with the postdeviant HFA reduction, across both channels and trials, indicating a link between prestimulus HFA decrease and reduced response to predictable deviants (i.e., better prediction leading to less PE). Finally, we found evidence that the frontal but not the temporal cortex followed the statistics of the irregular sequence as well (the “hazard rate”). In sum, these results provide evidence for automatic generation of proactive, anticipatory processes in frontal cortex, which may provide the basis for reduced orienting response to predictable events in an unattended stream. More generally, the results corroborate a hierarchy of prediction in the human brain (Dürschmid et al. 2016). This hierarchy is in line with the notion that early stages of information processing is represented based on bottom-up signals, whereas in higher levels of cortical processing deviations from expectation are registered while predictable components are “filtered out” (Heilbron and Chait 2018).

The Frontal Cortex Follows Complex Statistics of the Input

The comparison between predictable versus irregular deviants pointed to HFA reduction as a signature for predicting a deviation. This observation allowed us to investigate whether anticipatory predictions are generated during irregular, random sequences as well. We found that in frontal cortex, prestimulus HFA decreased as the train of uninterrupted standards became longer. Considering our first conclusion that HFA reduction reflects increasing likelihood of a deviant, this pattern matches well the so-called “hazard function,” in which an imminent event becomes more likely to occur the longer it has not occurred. This suggests that the frontal cortex predictive capacity is not limited to highly structured sequences, but rather, that it generates complex predictions based on sequence probabilities, even in a task-irrelevant irregular stream of events. This progressive increase in deviant prediction resembles the progressive increase in the contingent negative variation (CNV) as a function of distance from the last deviant reported by Chennu et al. (2013), although the CNV effect in Chennu et al. (2013) was only seen when subjects attended the stimuli (especially deviants), whereas in our case stimuli were task irrelevant. The temporal cortex in our study showed a trend toward an opposite effect with an increased prestimulus HFA activity the longer the standard train was. This is consistent with the notion that temporal cortex is based on recent history, such that with longer standard trains, “more of the same” (i.e., another standard) is expected.

Previous Attempts to Corroborate Proactive Prediction

Several studies approached the question of proactive prediction by investigating stimulus omissions (see Heilbron and Chait (2018) for an up-to-date review and discussion). Most omission-locked responses can be considered as violations of a general prediction for the occurrence of a stimulus at a given time

(a temporal prediction). [Sanmiguel et al. \(2013\)](#) had subjects generate environmental sounds by pressing a button. EEG responses to occasional sound omissions were found only when the same sound was repeatedly elicited by the button presses and was thus predictable due to the subjects' intention which does not speak for nonintentional automatic prediction. In a passive task with visual distraction, [Bendixen et al. \(2015\)](#) presented sequential tone pairs in rapid succession. The intra-pair frequencies were identical, whereas the frequencies altered between pairs. Omission-locked responses were found when the identity of the omitted stimulus could be predicted (because it was the second sound in the pair), but not when only its timing could be predicted (because it was the first in the pair). However, subjects may have perceived each pair as an auditory object, and the omission of the second sound in the pair, which elicited the critical omission response, might be a post hoc response to a duration change rather than an anticipatory response.

Rather than looking at poststimulus or postomission responses, our results address the prestimulus time, a time window at which activity modulation has to be ascribed to prediction per se since no error could have been computed. Similarly, [Kok et al. \(2017\)](#) decoded from MEG recordings the orientation of visual grating stimuli, which could be predicted by a preceding auditory stimulus (valid visual stimulus) or not (invalid visual stimulus). Subtracting the signal of valid from invalidly cued gratings revealed differences before stimulus presentation, suggesting the pre-activation of an anticipated sensory template. [Grisoni et al. \(2017\)](#) found EEG evidence for prestimulus anticipatory motor preparation to specific action-verbs predicted by meaningful sentences, but the automatic nature of this prediction is not clear as subjects likely listened to the meaningful sentences. While these studies provide converging evidence for proactive prediction, using MEG or EEG data, the source and type of signal of this predictive activity remain unclear. Taking the advantage of the high signal-to-noise ratio, and the improved spatial resolution of the ECoG data, our findings show that predictable deviants are preceded by frontal cortex HFA decrease not seen in sensory cortex.

Implications for Models of the Poststimulus Mismatch Response

How is prestimulus modulation of the HFA signals related to accounts of the mismatch response elicited by the deviant? Two mechanisms differing with respect to the degree of memory involvement have been proposed by [Fishman and Steinschneider \(2012\)](#). Poststimulus effects like the mismatch negativity may involve different states of neural adaptation (stimulus-specific adaptation ([Ulanovsky et al. 2003](#); [Farley et al. 2010](#))) due to repeated presentation. This creates a model of the recent history, and under an assumption of stationarity, provides a reasonable prediction of future events ([May and Tiitinen 2010](#)). Other models ([Näätänen et al. 2005](#)) suggest that beyond adaptation, stimulus repetition increases the absolute excitability of neurons tuned to values not included in the repeated stimulus. By both accounts, new stimuli elicit a stronger response if not congruent with the current model, which generates a PE signal. However, our observation of predictive predeviant modulation of activity cannot be explained by either mechanism. First, we compared the response with deviants following a similar number of standards in the random and predictable conditions, and overall deviants and standards had the same probability under both conditions. Thus, either adaptation

or lateral excitation should have been similar across conditions. Second, since the effect occurred before the deviant, it cannot be due to activation of nonadapted/excited neurons sensitive to the pitch of the deviant or by a process of comparison. Instead, the results provide evidence of high-level prediction, modifying the poststimulus comparison between the actual input and the ongoing prediction.

Implications for Models of PC

Dynamic causal modeling (DCM) of EEG or MEG studies suggested a hierarchical feedforward-feedback cascade in which the inferior frontal cortex sits at the top, providing top-down predictions to (and receiving PE signals from) the superior temporal gyrus, which in turn provides top-down predictions to (and receives PE signals from) the early auditory cortex ([Garrido et al. 2009](#)). Recently, [Phillips et al. \(2015\)](#) and [Phillips et al. \(2016\)](#) validated the models, originally tested on EEG/MEG data, with ECoG data from 2 patients. However, Phillips et al.'s models suggested that the prediction signal affecting the IFG is limited to temporal deviations (duration deviations and gaps in their study), but not pitch, intensity, or location deviations, whereas our findings showed clear effects of predictability in the ventral frontal cortex when the deviation was in pitch.

Our prestimulus predictive effects were not limited to temporal predictions. In fact, suppression of HFA indexed both the identity (standard or deviant) of the next stimulus in addition to its timing. Moreover, this was observed even though all stimuli were task-irrelevant, meaningless, did not require a response, and had no reward value. Previous findings of anticipatory response typically involved active preparation for an upcoming imperative stimulus, reflected in the CNV recorded on the scalp ([Trillenberg et al. 2000](#); [Janssen and Shadlen 2005](#)), listening to meaningful verbal material ([Grisoni et al. 2017](#)) or reward-prediction signals of different types ([Fiorillo et al. 2003](#)). The current finding provides evidence for ongoing, task-independent, anticipatory predictive signals, operative even before the stimulus occurred.

Previous studies argued that predictions and PE signals are compartmentalized across cortical layers and segregated by spectral content. They suggested that predictions are generated and fed-back by deep (infragranular) layers of the cortex at relatively lower frequencies of alpha/beta, whereas PE are fed forward from superficial (supragranular) layers at high (gamma) frequencies ([Bastos et al. 2012, 2015](#)). The fact that our proactive prediction signal was found in the HFA modulation may seem at odds with this model. However, for several reasons we remain agnostic about how the HFA modulation relates to the more detailed, laminar models of PC. First, the HFA signal should not be mistaken for any narrowband power modulation. Multiple studies using intracranial signals, as well as computational modeling, suggested that the high-frequency broadband signal is a good correlate of population neural firing rate ([Mukamel et al. 2005](#); [Liu and Newsome 2006](#); [Manning et al. 2009](#); [Miller, Sorensen et al. 2009](#); [Ray and Maunsell 2011](#)), making HFA modulation the preferred proxy for asynchronous (nonperiodic) areal activation in ECoG studies ([Miller, Sorensen et al. 2009](#); [Privman et al. 2013](#); [Miller et al. 2014](#); [Coon and Schalk 2016](#); [Kupers et al. 2018](#)). That is, although we parameterize this signal using frequency decomposition, no oscillation (i.e., narrowband periodic activity) is implied. In fact, as argued by Miller and colleagues, the measured HFA may reflect a frequency nonspecific power increase across the spectrum, while changes in the lower frequencies are masked by stronger

oscillatory activity in the lower ranges (Miller et al. 2007; Miller, Zanos et al. 2009). Second, our knowledge about the relationship between activity at specific laminae and how they are reflected in the mesoscopic measurement of the surface electrode is highly limited. Third, whereas the columnar model of PC suggested by Bastos et al. (2012, 2015) specifies some of the components (feedback predictions, feedforward PEs) in frequency content terms, it does not provide that detail about the dynamics of the interlaminar connections (e.g., projection of “expectation neurons” in supragranular layers to deep layers forming the predictions). In fact, the columnar organization vis a vis components of the PC model is still debated (Spratling 2010; Heilbron and Chait 2018). Fourth, it is not clear whether the prestimulus HFA modulation reflects the same prediction signal specified in PC models, or the outcome of this predictive signal (e.g., inhibition of firing rate in anticipation of a deviant). Specifically, current PC models do not account for long-term prospective predictions across hundreds of milliseconds as we see here. Thus, we believe that any speculations from our data to these models would be premature.

Maintaining Parallel and Inconsistent Predictions

Under the PC framework, prediction signals should be transmitted to lower nodes of the network, and PE signals should be carried forward to higher nodes in the network, to allow modification of the current model and influence the next prediction. However, our findings challenge this simple information flow, which must address multiple levels of possibly conflicting predictions (Pieszek et al. 2013). For instance, just prior to a deviant in the regular condition, and also after a long train of standards under the irregular condition, processes based on local effects predict another standard, whereas predictions based on the global statistics predict a deviant. In this situation, it seems efficient to prevent PE signals elicited at the temporal (auditory) cortex from propagating up the hierarchy and modifying a veridical model of the environment. Similarly, it seems that the prediction of an upcoming deviant based on global statistics, present at the frontal cortex, does not propagate down the network to mitigate the PE signal invoked by the expected deviant in the temporal cortex (Schröger et al. 2015). Our results therefore suggest that the flow of information up and down the hierarchy of the network is not as simple as gleaned from typical DCM diagrams (Garrido et al. 2009; Phillips et al. 2015, 2016). We speculate on the functional advantage of maintaining segregated predictions. Specifically, maintaining predictions that account for global regularities allows the prefrontal cortex to efficiently direct attention only to unexpected events (Sussman et al. 2003), whereas for the auditory cortex, detecting all local changes is advantageous for parsing the auditory input into meaningful chunks (e.g., in speech perception).

Relationship Between the Predictive Prestimulus Activity and Attention

Previous selective attention studies have shown prestimulus activation (increased firing rate or BOLD response) prior to task relevant stimuli (Colby et al. 1996; Beck and Kastner 2009) and deactivation prior to task-irrelevant stimuli (Langner et al. 2011; Rodgers and DeWeese 2014). Our study did not use a classic selective attention task but could be considered as involving a competition between the primary task of viewing a slide show, and the potential distraction caused by the auditory stream, especially by deviant events. Thus, the HFA decrease

observed prior to an expected deviant could reflect the same filtering mechanism previously observed during selective attention. Under this premise, the current results suggest that this inhibitory anticipation can be generated selectively, and in predictive manner, in an unattended stream.

In sum, pre and poststimulus HFA responses reveal a unique role for prefrontal cortex in utilizing global regularity to control responses to deviant stimuli. Frontal HFA selectively signals upcoming regular deviants with a decreased amplitude prior to deviant onset. Subsequently, only unpredictable deviants elicit a strong HFA response, putatively related to triggering an orienting response to an environmental perturbation. At the same time, the sensory cortex continues to veridically respond to any change in the stream. Our results highlight a selective role of frontal structures in actively computing predictions to better navigate the environment.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Notes

L.Y.D and R.T.K. conceived and designed the experiment. H.E.K. provided clinical information and helped to eliminate seizure epochs in the ECoG data, S.D. and C.R. analyzed the data, S.D., C.R., H.H., H.J.H, L.Y.D., and R.T.K. interpreted the data S.D., L.Y.D., and R.T.K. wrote the manuscript. *Conflict of Interest:* None declared.

References

- Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. 2010. Stimulus predictability reduces responses in primary visual cortex. *J Neurosci.* 30:2960–2966.
- Arnal LH, Giraud AL. 2012. Cortical oscillations and sensory predictions. *Trends Cogn Sci.* 16:390–398.
- Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012. Canonical microcircuits for predictive coding. *Neuron.* 76:695–711.
- Bastos AM, Vezoli J, Bosman CA, Schoffelen JM, Oostenveld R, Dowdall JR, De Weerd P, Kennedy H, Fries P. 2015. Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron.* 85:390–401.
- Beck DM, Kastner S. 2009. Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Res.* 49:1154–1165.
- Bendixen A, Scharinger M, Strauß A, Obleser J. 2014. Prediction in the service of comprehension: modulated early brain responses to omitted speech segments. *Cortex.* 53:9–26.
- Bendixen A, Schwartz M, Kotz SA. 2015. Temporal dynamics of contingency extraction from tonal and verbal auditory sequences. *Brain Lang.* 148:64–73.
- Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibáñez A, Owen AM, Bekinschtein TA. 2013. Expectation and attention in hierarchical auditory prediction. *J Neurosci.* 33: 11194–11205.
- Colby CL, Duhamel JR, Goldberg ME. 1996. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J Neurophysiol.* 76:2841–2852.
- Coon WG, Schalk G. 2016. A method to establish the spatiotemporal evolution of task-related cortical activity from electrocorticographic signals in single trials. *J Neurosci Methods.* 271:76–85.

- den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE. 2010. Striatal prediction error modulates cortical coupling. *J Neurosci.* 30:3210–3219.
- Dürschmid S, Edwards E, Reichert C, Dewar C, Hinrichs H, Heinze HJ, Kirsch HE, Dalal SS, Deouell LY, Knight RT. 2016. Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proc Natl Acad Sci USA.* 113:6755–6760.
- Farley BJ, Quirk MC, Doherty JJ, Christian EP. 2010. Stimulus-specific adaptation in auditory cortex is an nmda-independent process distinct from the sensory novelty encoded by the mismatch negativity. *J Neurosci.* 30:16475–16484.
- Fiorillo CD, Tobler PN, Schultz W. 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science.* 299:1898–1902.
- Fishman YI, Steinschneider M. 2012. Searching for the mismatch negativity in primary auditory cortex of the awake monkey: deviance detection or stimulus specific adaptation? *J Neurosci.* 32:15747–15758.
- Fogelson N, Wang X, Lewis JB, Kishiyama MM, Ding M, Knight RT. 2009. Multimodal effects of local context on target detection: evidence from p3b. *J Cogn Neurosci.* 21:1680–1692.
- Garrido MI, Kilner JM, Kiebel SJ, Friston KJ. 2009. Dynamic causal modeling of the response to frequency deviants. *J Neurophysiol.* 101:2620–2631.
- Grisoni L, Miller TM, Pulvermüller F. 2017. Neural correlates of semantic prediction and resolution in sentence processing. *J Neurosci.* 37(18):4848–4858.
- Heilbron M, Chait M. 2018. Great expectations: is there evidence for predictive coding in auditory cortex? *Neuroscience.* 389:54–73.
- Janssen P, Shadlen MN. 2005. A representation of the hazard rate of elapsed time in macaque area lip. *Nat Neurosci.* 8:234–241.
- Kok P, Mostert P, de Lange FP. 2017. Prior expectations induce prestimulus sensory templates? *Proc Natl Acad Sci USA* 114:10473–10478.
- Kupers ER, Wang HX, Amano K, Kay KN, Heeger DJ, Winawer J. 2018. A non-invasive, quantitative study of broadband spectral responses in human visual cortex. *PLoS One.* 13:e0193107.
- Langner R, Kellermann T, Boers F, Sturm W, Willmes K, Eickhoff SB. 2011. Modality-specific perceptual expectations selectively modulate baseline activity in auditory, somatosensory, and visual cortices. *Cerebral cortex.* 21:2850–2862.
- Lee H, Noppeney U. 2014. Temporal prediction errors in visual and auditory cortices. *Curr Biol.* 24:R309–R310.
- Liu J, Newsome WT. 2006. Local field potential in cortical area mt: stimulus tuning and behavioral correlations. *J Neurosci.* 26:7779–7790.
- Manning JR, Jacobs J, Fried I, Kahana MJ. 2009. Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J Neurosci.* 29:13613–13620.
- May PJC, Tiitinen H. 2010. Mismatch negativity (mmn), the deviance-elicited auditory deflection, explained. *Psychophysiology.* 47:66–122.
- Miller KJ, Honey CJ, Hermes D, Rao RPN, denNijs M, Ojemann JG. 2014. Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *Neuroimage.* 85(Pt 2):711–720.
- Miller KJ, Leuthardt EC, Schalk G, Rao RPN, Anderson NR, Moran DW, Miller JW, Ojemann JG. 2007. Spectral changes in cortical surface potentials during motor movement. *J Neurosci.* 27:2424–2432.
- Miller KJ, Sorensen LB, Ojemann JG, den Nijs M. 2009. Power-law scaling in the brain surface electric potential. *PLoS Comput Biol.* 5:e1000609.
- Miller KJ, Zanos S, Fetz EE, den Nijs M, Ojemann JG. 2009. Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans. *J Neurosci.* 29:3132–3137.
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. 2005. Coupling between neuronal firing, field potentials, and fmri in human auditory cortex. *Science.* 309:951–954.
- Näätänen R, Jacobsen T, Winkler I. 2005. Memory-based or afferent processes in mismatch negativity (mmn): a review of the evidence. *Psychophysiology.* 42:25–32.
- Phillips HN, Blenkmann A, Hughes LE, Bekinschtein TA, Rowe JB. 2015. Hierarchical organization of frontotemporal networks for the prediction of stimuli across multiple dimensions. *J Neurosci.* 35:9255–9264.
- Phillips HN, Blenkmann A, Kochen S, Bekinschtein TA, Cam-Can, Rowe JB. 2016. Convergent evidence for hierarchical prediction networks from human electrocorticography and magnetoencephalography. *Cortex.* 82:192–205.
- Pieszek M, Widmann A, Gruber T, Schröger E. 2013. The human brain maintains contradictory and redundant auditory sensory predictions. *PLoS One.* 8(1):e53634.
- Privman E, Malach R, Yeshurun Y. 2013. Modeling the electrical field created by mass neural activity. *Neural Netw.* 40:44–51.
- 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.
- Ray S, Maunsell JHR. 2011. Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol.* 9:e1000610.
- Rodgers CC, DeWeese MR. 2014. Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. *Neuron.* 82:1157–1170.
- Sanmiguel I, Saupe K, Schröger E. 2013. I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when”. *Front Hum Neurosci.* 7:407.
- Schröger E, Marzecová A, SanMiguel I. 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. *Eur J Neurosci.* 41:641–664.
- Spratling MW. 2010. Predictive coding as a model of response properties in cortical area v1. *J Neurosci.* 30:3531–3543.
- Summerfield C, Trittschuh EH, Monti JM, Mesulam MM, Egner T. 2008. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci.* 11:1004–1006.
- Sussman E, Winkler I, Schröger E. 2003. Top-down control over involuntary attention switching in the auditory modality. *Psychon Bull Rev.* 10(3):630–637.
- Todorovic A, van Ede F, Maris E, de Lange FP. 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an meg study. *J Neurosci.* 31:9118–9123.
- Trillenberg P, Verleger R, Wascher E, Wauschkuhn B, Wessel K. 2000. Cnv and temporal uncertainty with ‘ageing’ and ‘non-ageing’ s1-s2 intervals. *Clin Neurophysiol.* 111:1216–1226.
- Ulanovsky N, Las L, Nelken I. 2003. Processing of low-probability sounds by cortical neurons. *Nat Neurosci.* 6:391–398.
- Winkler I, Czigler I. 2012. Evidence from auditory and visual event-related potential (erp) studies of deviance detection (mmn and vmmn) linking predictive coding theories and perceptual object representations. *Intern J Psychophysiol.* 83:132–143.
- Winkler I, Schröger E. 2015. Auditory perceptual objects as generative models: setting the stage for communication by sound. *Brain Lang.* 148:1–22.