

# Neural Integration of Stimulus History Underlies Prediction for Naturalistically Evolving Sequences

 Brian Maniscalco,<sup>1,2</sup>  Jennifer L. Lee,<sup>2</sup> Patrice Abry,<sup>3</sup> Amy Lin,<sup>1,4</sup> Tom Holroyd,<sup>5</sup> and  Biyu J. He<sup>1,2,6</sup>

<sup>1</sup>National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, <sup>2</sup>Neuroscience Institute, New York University Langone Medical Center, New York, New York 10016, <sup>3</sup>University of Lyon, Ens de Lyon, University Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France, <sup>4</sup>Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, California 90095, <sup>5</sup>National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, and <sup>6</sup>Departments of Neurology, Neuroscience, and Physiology, and Radiology, New York University Langone Medical Center, New York 10016

Forming valid predictions about the environment is crucial to survival. However, whether humans are able to form valid predictions about natural stimuli based on their temporal statistical regularities remains unknown. Here, we presented subjects with tone sequences with pitch fluctuations that, over time, capture long-range temporal dependence structures prevalent in natural stimuli. We found that subjects were able to exploit such naturalistic statistical regularities to make valid predictions about upcoming items in a sequence. Magnetoencephalography (MEG) recordings revealed that slow, arrhythmic cortical dynamics tracked the evolving pitch sequence over time such that neural activity at a given moment was influenced by the pitch of up to seven previous tones. Importantly, such history integration contained in neural activity predicted the expected pitch of the upcoming tone, providing a concrete computational mechanism for prediction. These results establish humans' ability to make valid predictions based on temporal regularities inherent in naturalistic stimuli and further reveal the neural mechanisms underlying such predictive computation.

**Key words:**  $1/f$  activity; magnetoencephalography; naturalistic stimuli; predictive processing; scale-free dynamics

## Significance Statement

A fundamental question in neuroscience is how the brain predicts upcoming events in the environment. To date, this question has primarily been addressed in experiments using relatively simple stimulus sequences. Here, we studied predictive processing in the human brain using auditory tone sequences that exhibit temporal statistical regularities similar to those found in natural stimuli. We observed that humans are able to form valid predictions based on such complex temporal statistical regularities. We further show that neural response to a given tone in the sequence reflects integration over the preceding tone sequence and that this history dependence forms the foundation for prediction. These findings deepen our understanding of how humans form predictions in an ecologically valid environment.

## Introduction

In real-life environments, prior sensory information across multiple time scales continually influences information processing in the present. To date, the relationship between neural integration

of information over time and predictive computations remains mysterious (Hasson et al., 2015). The importance of this question is underscored by the fact that forming valid predictions about the features (Summerfield and de Lange, 2014) and timing (Cravo et al., 2011; Jepma et al., 2012) of upcoming stimuli in the environment is crucial for survival. Predictions about stimulus features can be generated through a variety of mechanisms, including associative learning (Schlack and Albright, 2007), contextual influences (Bar, 2004), imagination (Schacter et al., 2007), task cues (Wyart et al., 2012; de Lange et al., 2013), and extrapolation from statistical regularities in sensory input (Saffran et al., 1996; Alink et al., 2010; Chalk et al., 2010). Because statistical regularities are ubiquitous in natural stimuli (Dong and Atick, 1995; Summerfield and de Lange, 2014), extrapolating from statistical regularities in sensory input should provide a fundamental strategy for forming predictions in natural environments.

Received June 26, 2017; revised Dec. 19, 2017; accepted Dec. 21, 2017.

Author contributions: B.M. and B.J.H. designed research; B.M., J.L.L., and A.L. performed research; P.A. and T.H. contributed unpublished reagents/analytic tools; B.M., J.L.L., and B.J.H. analyzed data; B.M., J.L.L., and B.J.H. wrote the paper.

This work was supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke—National Institutes of Health and the New York University Langone Medical Center. B.J.H. further acknowledges support by the Leon Levy Foundation and a Klingenstein-Simons Neuroscience Fellowship. We thank Ella Podvalny, Adeen Flinker, and Jean-Rémi King for comments on a previous draft of the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Biyu J. He, Neuroscience Institute, NYU Langone Medical Center, 550 1st Avenue, MSB 460, New York, NY 10016. E-mail: biyu.jade.he@gmail.com.

DOI:10.1523/JNEUROSCI.1779-17.2017

Copyright © 2018 the authors 0270-6474/18/381541-17\$15.00/0

Previous studies investigating predictions based on statistical regularities have largely used relatively simple stimuli such as the oddball paradigm, in which a novel stimulus is embedded within a sequence of repeated stimuli (Yaron et al., 2012); the local-global paradigm, in which a sequence including two values contains both local and global regularities (Bekinschtein et al., 2009); and repeated presentations of a fixed sequence of stimuli (Erickson and Desimone, 1999; Meyer and Olson, 2011; Gavornik and Bear, 2014). Although such stimuli allow for tight experimental control, they induce the formation of predictions by repeated presentations of items or sequences of items. In contrast, temporally varying stimuli encountered in the natural environment have rich statistical structures that allow for more complex forms of predictions. At present, whether humans can form valid predictions based on temporal statistical regularities inherent in natural stimuli remains unknown.

In particular, one pervasive statistical feature of natural stimuli is that they exhibit power spectra following a  $P \propto 1/f^\beta$  pattern in spatial (Tolhurst et al., 1992) and temporal frequency domains, where  $\beta$  is an exponent commonly ranging between 0 and 2. In the temporal domain, a  $1/f^\beta$  pattern is observed in the loudness and pitch fluctuations of music, speech, and ambient noise in urban and rural settings (Voss and Clarke, 1975; De Coensel et al., 2003) and in the rhythms of music (Levitin et al., 2012). Previous work has shown that human perception and neural activity are sensitive to the scaling parameter  $\beta$  of dynamic auditory stimuli such as tone sequences featuring  $1/f^\beta$  fluctuations in pitch (Schmuckler and Gilden, 1993; Patel and Balaban, 2000; Garcia-Lazaro et al., 2006; Lin et al., 2016). Therefore,  $1/f^\beta$  statistics are prevalent in natural stimuli, and human perception is sensitive to such statistics. Because a  $1/f^\beta$  type temporal power spectrum indicates temporal dependence and predictability (Eke et al., 2002; He, 2014), these observations raise the intriguing question of whether humans are able to make valid predictions about upcoming items in such sequences.

Here, we investigate these open questions using a novel experimental paradigm that allows us to probe, for the first time, the ability of human subjects to exploit the temporal statistical regularities found in natural stimuli to predict upcoming stimulus features. We created tone sequences with pitch that fluctuated over time following a  $1/f^\beta$  pattern where  $\beta$  ranged from 0 (no temporal dependence, “white noise”) to 2 (strong temporal dependence, “random walk”). The expected pitch of a given tone in the sequence can be expressed as a function of the preceding tone pitches. These stimuli are thus well controlled while also capturing the complex statistical regularities exhibited by auditory stimuli in natural environments. We therefore refer to these tone sequences as “naturalistic” rather than “natural” to highlight that, although the elements of the sequences are man-made pure tones, the dynamics of tone pitch fluctuation reflect the statistics of  $1/f^\beta$  temporal dependence found in natural stimuli. Using a behavioral task based on these stimuli and MEG recordings in humans, we established human subjects’ ability to make valid predictions about upcoming stimulus features based on naturalistic temporal statistical regularities and further uncovered the underlying neural mechanisms as implemented in history-dependent neural processing of incoming information.

## Materials and Methods

**Stimuli creation.** We created auditory tone sequences with pitch fluctuations that had five levels of temporal dependence strength  $\beta$ , ranging from  $\beta = 0$  (no temporal dependence) to  $\beta = 2$  (strong temporal dependence). Sequences with  $\beta < 1$  ( $\beta = 0$  and 0.5) were fractional Gaussian

noise (fGn), whereas sequences with  $\beta > 1$  ( $\beta = 1.01, 1.5, \text{ and } 2$ ) were fractional Brownian motion (fBm). fBm sequences are nonstationary, self-similar sequences with increments that are stationary fGn sequences and the transition from fGn to fBm sequences occurs at  $\beta = 1$  (Mandelbrot and Van Ness, 1968; Eke et al., 2002). We used a circulant embedding algorithm (Helgason et al., 2011) to create six unique 33-element long series for each level of  $\beta$ , as follows:

$$\mathbf{x}_{\beta,i} = [x_1, x_2, \dots, x_{33}], \beta \in \{0, 0.5, 1.01, 1.5, 2\}, 1 \leq i \leq 6$$

where each element  $x_j$  of  $\mathbf{x}_{\beta,i}$  is taken to represent the pitch of the  $j^{\text{th}}$  tone in the sequence. Each  $\mathbf{x}_{\beta,i}$  series was translated and scaled such that its elements ranged from  $\log(220)$  to  $\log(880)$ . This range of pitch values was chosen to span the iso-loudness region of human hearing (Robinson and Dadson, 1956). To investigate the potential effect of overall pitch variability on judgments of sequence statistics, we scaled half of the sequences such that the SD of its corresponding fGn was 70% of its original value (each fBm series is associated with a fGn series of the same Hurst exponent with a power-law exponent  $\beta$  that differs by 2; Eke et al., 2002). This resulted in sequences that spanned a smaller range (green traces in Figure 1B) than the nonscaled sequences (magenta traces). Next, the series were discretized such that each element took on one of 25 values evenly spaced on the log scale with semitone distance. We refer to these discretized series as  $\mathbf{p}_{\beta,i}$ . Each  $\mathbf{p}_{\beta,i}$  thus represents a series of tone pitches with pitch fluctuation that exhibits temporal dependence prescribed by  $\beta$ . To disentangle predictive processing from instantaneous stimulus processing, all sequences converged onto a pitch value of 440 Hz for the 33<sup>rd</sup> and penultimate tone. The 34<sup>th</sup> and final tone of each sequence varied across trials and was selected from one of six possible values of pitch that were 4, 8, or 12 semitone steps above or below 440 Hz (black dots in Fig. 1B). Each unique sequence was presented for a total of 12 times and each of the six final tone pitch values was used twice for each unique sequence in randomized order.

The sound wave for each sequence was constructed according to the following:

$$\mathbf{y}(t)_{\beta,i,j} = A \cos(2\pi \exp(\mathbf{p}_{\beta,i,j}) t / \text{SR} + \varphi_j),$$

where  $t$  denotes sample number,  $i$  denotes sequence number,  $j$  denotes tone number within sequence  $i$ ,  $A$  (amplitude) = 1, and SR (sampling rate) = 44,100 Hz. The 300 ms duration of each tone was chosen for ease of listening, in accordance with prior studies (Patel and Balaban, 2000; Lin et al., 2016). Because each sequence contained 34 tones, it had a total duration of 10.2 s. Cosine waves for each tone  $j$  were concatenated such that there was no silence period between consecutive tones. To prevent clicking noises from occurring between tones,  $\varphi_j$  is computed to ensure that  $\mathbf{y}_{\beta,i,j}$  is a smooth and differentiable function across adjacent  $j$  values. Audio files of sequence examples can be downloaded from [https://med.nyu.edu/helab/sites/default/files/helab/Maniscalco\\_et\\_al\\_stim\\_wav\\_files\\_and\\_figs.zip](https://med.nyu.edu/helab/sites/default/files/helab/Maniscalco_et_al_stim_wav_files_and_figs.zip).

These auditory tone sequences were presented using the PsychPortAudio function of the Psychophysics Toolbox (Brainard, 1997) in MATLAB (The MathWorks). The audio was delivered through specialized ear tubes that were MEG compatible. The Etymotic ER-3 Insert Headphones, which has a frequency response that is flat to 5 kHz, was used. The plastic tubing from the transducer to the earpiece had a speed-of-sound delay of  $\sim 10$  ms, which was corrected in the MEG data analyses.

**Calculation of  $p_{3,4}$ .** Let  $\{x_n, n = 1, \dots, N\}$  be a series of data samples drawn from a stationary zero-mean random process, with covariance function  $r_x$ . Linear prediction of order  $K$ ,  $\hat{x}_n$ , for sample  $n$ , from past samples  $\{x_{n-1}, x_{n-2}, \dots, x_{n-K}\}$  is written as follows:

$$\hat{x}_n = \sum_{k=1}^K a_k x_{n-k},$$

where the vector  $\mathbf{a} = \{a_1, a_2, \dots, a_K\}$  is to be chosen (or estimated) so as to minimize the average (squared) prediction error as follows:

$$\hat{\mathbf{a}} = \min_{\mathbf{a}} \mathbb{E}(x_n - \hat{x}_n)^2$$

Linear algebra leads to an explicit theoretical solution for  $\hat{a}$  (Scharf, 1991):

$$\hat{a}^{theory} = \underline{\underline{R}}_K^{-1} \underline{\underline{r}}_x$$

where  $\underline{\underline{r}}_x$  is the covariance sequence of process  $x$  and  $\underline{\underline{R}}_K$  denotes the  $K \times K$  square matrix, with entry  $(\underline{\underline{R}}_K)_{p,p'} = \underline{\underline{r}}_x(|p - p'|)$  for  $p, p' \in \{1, \dots, K\}^2$ .

If  $\underline{\underline{x}}$  is an fGn sequence ( $-1 < \beta < 1$ ), then covariance  $\underline{\underline{r}}_x$  can be computed as a function of its Hurst exponent  $H$  and variance  $\sigma^2$  (Eke et al., 2002) as follows:

$$\underline{\underline{r}}_x(k) = \frac{\sigma^2}{2} (|k + 1|^{2H} - 2|k|^{2H} + |k - 1|^{2H})$$

where  $H = (\beta + 1)/2$ . We set  $K = 33$  to compute the expected value of the final tone pitch based on the pitch values of all 33 preceding tones.

In summary, the calculation of  $p_{34}^*$  for the (zero-meaned)  $\beta = 0$  and  $\beta = 0.5$  sequences can be written as follows:

$$p_{34|\beta < 1}^* = \sum_{k=0}^{32} a_{k+1, \beta} p_{33-k} \quad (1)$$

where  $a_{k+1, \beta}$  are the optimal prediction weights for an fGn sequence with scaling parameter  $\beta$  and  $p_{33-k}$  are the pitch of the preceding tones. Expected value was calculated for each sequence using zero-meaned pitch series. Because the mean pitch value was  $\log(440)$ , we added  $\log(440)$  to the calculated expected value (Eq. 1) to yield the final value for  $p_{34}^*$  used in analyses. For  $\beta = 0$  sequences, coefficients  $a_{k+1, 0} = 0$ , so the expected value of the upcoming item is equal to the sequence mean, resulting in  $p_{34}^* = \log(440)$  for all  $\beta = 0$  sequences.

fBm sequences ( $1 < \beta < 3$ ) are cumulative sums of fGn sequences with the same Hurst exponent  $H$  and are nonstationary. The preceding methodology for computing expected value only holds for fGn sequences and so cannot be applied to fBm sequences. However, the increments of an fBm sequence with scaling parameter  $\beta$  constitute an fGn sequence with scaling parameter  $\beta - 2$  (Eke et al., 2002). Therefore, to calculate the expected value for fBm sequences, we first computed the difference between successive elements in the fBm sequence, which yielded an fGn sequence (with expected mean = 0). We then computed the expected value for the upcoming item in this fGn sequence, i.e., the expected increment in pitch for the original fBm series, as follows:

$$inc_{34|\beta > 1}^* = \sum_{k=0}^{31} a_{k+1, \beta-2} (p_{33-k} - p_{33-k-1})$$

where  $a_{k+1, \beta-2}$  are the optimal prediction weights for an fGn sequence with scaling parameter  $\beta - 2$  and  $p_{33-k}$  are the pitch of the preceding tones. Finally, we calculated the expected pitch of the 34<sup>th</sup> tone in the fBm sequences ( $\beta = 1.01, 1.5, \text{ and } 2$ ) as the pitch of the 33<sup>rd</sup> tone (which is always 440 Hz) plus the expected increment in pitch, as follows:

$$p_{34|\beta > 1}^* = p_{33} + inc_{34|\beta > 1}^*$$

To facilitate comparison with equation (1),  $p_{34}^*$  for fBm sequences can be written as a weighted sum of preceding pitch values as follows:

$$p_{34|\beta > 1}^* = p_{33} + \left[ a_{1, \beta-2} p_{33} + \left( \sum_{k=1}^{31} (a_{k+1, \beta-2} - a_{k, \beta-2}) p_{33-k} \right) - a_{32, \beta-2} p_1 \right] \quad (2)$$

For a  $\beta = 2$  sequence, the increments form an fGn sequence with  $\beta = 0$ , so the expected increment  $inc_{34}^* = 0$ , entailing that  $p_{34}^* = p_{33} = \log(440)$ .

Herein, we use the term “optimal prediction weights” (sometimes shortened to “prediction weights”) to refer to the weights placed on the

pitch sequence history  $p_{33-k}$  (where  $k \in [0 \ 32]$ ) to compute  $p_{34}^*$  for fGn (Eq. 1) and fBm (Eq. 2) sequences.

**Experimental design and statistical analysis.** Each trial began with presentation of a fixation point in the middle of the screen (Fig. 1A). Subjects were instructed to fixate on the fixation point during presentation of the auditory sequence to minimize eye movements. Likelihood ratings and trend strength ratings were entered using two separate button boxes for the left and right hand, respectively. Subjects were instructed to disregard the final tone in their judgment of trend strength because the final tone did not follow the temporal dependence statistics of preceding tones. Higher trend strength was explained to the subjects as the tendency for a trend of high- or low-pitch values to persist over longer periods of time, which captured the concept of temporal dependence in the time series. Performance feedback about the trend strength rating was presented visually for 2 s after entry of both behavioral responses. The feedback indicated what trend strength rating had been entered by the subject, what the true trend strength of the sequence was, and whether the subject’s trend strength rating was correct, close to correct (off by one level), or incorrect (off by two or more levels).

Trials were split into 12 blocks of 30 trials each and subjects received a self-terminated period of rest after each block. Between blocks, the head position of the subject was measured with respect to the MEG sensor array using coils placed on the left and right preauricular points and the nasion and the subject self-corrected their head position to the same position recorded at the start of the first block using a custom visual-feedback program written by T.H. inspired by previous work (Stolk et al., 2013) to minimize head displacement across the experiment. Details on statistical analyses can be found in the following sections and in the Results.

**Subjects.** The experiment was approved by the Institutional Review Board of the National Institute of Neurological Disorders and Stroke (protocol #14-N-0002). All subjects were right-handed and neurologically healthy with normal hearing and provided written informed consent. Twenty subjects between 21 and 34 years old (mean age 24.1; 15 females) participated in an initial  $\sim 2$  h long behavioral session to familiarize themselves with the task. Overall, the initial group of 20 subjects was able to discriminate sequence  $\beta$  ( $t$  test on  $z$ -transform of Spearman correlation coefficients between sequence  $\beta$  and trend strength rating,  $t_{(19)} = 7.8, p < 0.001$ ) and to make valid predictions ( $p_{34} \times p_{34}^*$  interaction as in the analysis of Fig. 2B,  $F_{(10,190)} = 17.2, p < 0.001$ ; for details, see Results, subjects can make valid predictions about upcoming items based on stimulus sequence history).

We invited subjects exhibiting adequate behavioral performance (Spearman correlation between sequence  $\beta$  and trend strength rating  $> 0.2$  and full usage of the rating scales for both behavioral responses) to perform the task in the MEG scanner. Fourteen subjects were recruited for the MEG portion of the study (lasting  $\sim 3$  h including setup time). Three subjects were excluded from analyses due to excessive MEG artifacts from dental implants or eye or head movements (age range for the remaining subjects: 21–34; mean age: 24.1; 7 females). Behavioral results reported in Figure 2 are from the MEG session and included the same subjects who contributed MEG data.

**Data acquisition and preprocessing.** Experiments were conducted in a whole-head 275-channel CTF MEG scanner. MEG data were collected with a sampling rate of 600 Hz and an anti-aliasing filter at  $< 150$  Hz. Analyses were performed on 271 sensors after excluding four malfunctioning sensors. The analyses were conducted using the Fieldtrip package (Oostenveld et al., 2011) and code custom written in MATLAB. MEG data from each block were demeaned and detrended. To best investigate low-frequency MEG activity, we did not apply a high-pass filter; an earlier study from our laboratory on the same MEG scanner using similar stimuli and much longer trial-length (180 s), as well as resting-state recordings, had demonstrated excellent data quality down to  $\sim 0.005$  Hz (Lin et al., 2016). Powerline noise was removed offline by applying a fourth-order Butterworth band-stop filter in frequency ranges 58–62 Hz and 118–122 Hz. Independent component analysis (ICA) was used to remove artifacts related to eye blinks, eye movements, heartbeat, breathing, and slow movement drift.

**Phase dissimilarity.** We conducted this analysis closely following the methods used in (Luo and Poeppel, 2007). For every trial, sensor, and

subject, we performed a fast Fourier transform on the activity during the 9.9 s presentation of the first 33 tones of the tone sequence using a 6 s length sliding Hanning window that incremented in 100 ms steps. For each time point and frequency bin at each sensor, across-trial phase coherence was computed as follows:

$$C_{phase_{f,t}} = \left( \frac{\sum_{n=1}^N \cos \theta_{f,t,n}}{N} \right)^2 + \left( \frac{\sum_{n=1}^N \sin \theta_{f,t,n}}{N} \right)^2$$

where  $\theta_{f,t,n}$  denotes phase at frequency  $f$ , time  $t$ , and trial  $n$ . Phase coherence ranges from 0 to 1, where 0 indicates random phase across trials and 1 indicates identical phase across trials.

We next defined “identical” stimulus sets as the 12 trials in which the same initial 33-tone sequence was presented and “similar” stimulus sets as trials in which sequences with the same  $\beta$  and  $\sigma$  were presented. To match the number of trials in the two conditions, for “similar” stimulus sets, we randomly selected 12 trials from a total of 36 trials for each level of  $\beta$  and  $\sigma$ . We then computed phase dissimilarity at each frequency bin as the difference in mean phase coherence over time between identical and similar stimulus sets as follows:

$$phasedissimilarity_f = \frac{\sum_{t=1}^T C_{phase_{f,t,identical}}}{T} - \frac{\sum_{t=1}^T C_{phase_{f,t,similar}}}{T}.$$

The following analyses involved performing regressions on MEG activity time locked to the onset of tones. For these analyses, we extracted MEG activity using 50-ms-length sliding windows, which resulted in six windows during the presentation of each tone. For analyses using baseline-corrected MEG activity, because there was no period of silence before each tone (aside from the first tone in the sequence), we performed baseline correction by computing the mean MEG activity in the first 20 ms after tone onset and subtracting it from all time points during the presentation of that tone.

**History tracking analysis.** For every tone  $i$  in the second half of the tone sequence ( $16 \leq i \leq 32$ ) of every trial  $n$ , we extracted a list of the  $k'$  preceding tone pitch values:  $p_{i,n}, p_{i-1,n}, \dots, p_{i-k',n}$ . (Tone 33 was omitted from the analysis because it always had the same pitch of 440 Hz and tone 34 was omitted because its value was selected independently of the prior sequence.) For each sensor  $s$  and trial  $n$ , we extracted MEG activity after the onset of tone  $i$  using the sliding window approach described earlier. We then regressed this value (call it  $M_{s,w,i,n}$ ) onto the current tone as well as the  $k'$  preceding tones as follows:

$$M_{s,w,i,n} = \beta_{0,s,w} + \sum_{k=0}^{k'} \beta_{k+1,s,w} p_{i-k,n} + \varepsilon_{s,w,i,n} \quad (3)$$

where the  $\beta_{k+1,s,w}$  terms describe how MEG activity at sensor  $s$  and time window  $w$  depends upon the pitch of the tone presented  $k$  tones before the current tone  $i$ .

Our aim was to determine what value of  $k'$  (ranging from 0 to 15) provided the best fit to the MEG activity at each sensor and time window. To assess this, we used a cross-validation procedure. For each subject, we partitioned trials into a training set (odd-numbered trials) and a test set (even-numbered trials). We performed 16 regression analyses on the training set (one for each  $k'$  value ranging from 0 to 15). We then used the regression coefficients from the training set to predict MEG activity in the test set. The best value of  $k'$  for a given sensor and time window was defined to be the one that minimized the sum of squared errors of the predicted MEG activity in the test set. This procedure was repeated using even-numbered trials as the training set and odd-numbered trials as the test set. The final optimal  $k'$  assigned to each subject/sensor/time window was the average  $k'$  derived from the two folds. See Figure 4A for a schematic of the analysis. Last,  $k'$  values were averaged across subjects at each

sensor and time window. This analysis was performed both with and without baseline correction.

To assess the statistical significance of the (across-subject) mean  $k'$  values against the null hypothesis that mean  $k' = 0$ , we repeated the cross-validation regression analyses described above using random permutations of the tone sequences in the training set while leaving tone sequences in the test set unshuffled. We randomly permuted the order of the first 32 tones in the training set, with the exception that the value of the current tone pitch  $i$  in each regression was kept the same as in the original sequence. In other words, information about current tone pitch was retained, whereas information about tone sequence history was destroyed. We repeated this procedure 100 times each using odd and even trials as the training set and found the value of  $k'$  yielding minimum error on the test set, yielding 200 permutation samples of  $k'_{perm}$  for each subject, sensor, and time window. Next, we repeatedly sampled (with replacement) from the  $k'_{perm}$  values of each subject and calculated the across-subject average 1000 times, yielding a distribution of mean  $k'$  values at each sensor and time window under the null hypothesis that mean  $k' = 0$ . We defined uncorrected  $p$ -values for each sensor and time window as the proportion of samples in this null distribution that were greater than or equal to the empirically observed mean  $k'$ . These  $p$ -values were subsequently used to define clusters for the cluster correction analysis, in which the sum of  $k'$  across sensors in a cluster,  $k'_{sum}$ , was used as the cluster statistic (see “Cluster correction” section below).

**Prediction effect.** We investigated how MEG activity during the penultimate (33<sup>rd</sup>) tone was modulated by the expected value of the final (34<sup>th</sup>) tone by performing the following regression model:

$$M_{s,w,33,n} = \beta_{0,s,w}^* + \beta_{1,s,w}^* p_{34,n}^* + \varepsilon_{s,w,33,n} \quad (4)$$

Because the penultimate tone always had a pitch of 440 Hz, it was not necessary to include its pitch in the regression analysis. The  $\beta_{1,s,w}^*$  coefficient describes how MEG activity at sensor  $s$  and time window  $w$  after the penultimate tone onset depends on the expected pitch of the final tone,  $p_{34}^*$ . We submitted  $\beta_{1,s,w}^*$  to a group-level analysis (one-sample  $t$  test against zero), which was corrected for multiple comparisons across sensors using a cluster-based permutation test described below.

**Relating history-dependent MEG response with  $p_{34}^*$ .** For each subject, sensor  $s$ , and time window  $w$ , we obtained coefficients  $\beta_{k+1,s,w}$  ( $0 \leq k \leq k'$ ) from the previously conducted history tracking regression analysis on non-baseline-corrected data using the optimal  $k'$  value for that subject, sensor, and time window. Specifically, both the optimal  $k'$  and the coefficients  $\beta_{k+1,s,w}$  were averaged across the two folds for training and testing and noninteger values of  $k'$  were rounded up to the nearest integer (in the rare cases of optimal  $k' = 0$ , it was set to 1 to carry out the regression analysis described below). Next, we partitioned MEG activity in response to the penultimate tone at each sensor  $s$ , time window  $w$ , and trial  $n$  into two components (see Fig. 6A): that which was predicted by the preceding tone sequence and the residual. The former was computed as follows:

$$M_{s,w,33,n}^H = \beta_{0,s,w} + \sum_{k=0}^{k'} \beta_{k+1,s,w} p_{33-k,n} \quad (5)$$

and the residual was computed as follows:

$$M_{s,w,33,n}^{res} = M_{s,w,33,n}^{orig} - M_{s,w,33,n}^H$$

where  $M_{s,w,33,n}^{orig}$  denotes the originally recorded MEG activity. Next, we repeated the prediction effect regression analysis described above, but now using  $M_{s,w,33,n}^H$  and  $M_{s,w,33,n}^{res}$  as the dependent variable as follows:

$$M_{s,w,33,n}^H = \beta_{0,s,w}^{H*} + \beta_{1,s,w}^{H*} p_{34,n}^* + \varepsilon_{s,w,33,n} \quad (6)$$

$$M_{s,w,33,n}^{res} = \beta_{0,s,w}^{res*} + \beta_{1,s,w}^{res*} p_{34,n}^* + \varepsilon_{s,w,33,n}$$

Last, we quantified these observations by comparing the group-averaged regression weights for  $p_{34}^*$  obtained for  $M_{s,w,33,n}^{orig}$ ,  $M_{s,w,33,n}^H$ , and  $M_{s,w,33,n}^{res}$ .

**Relationship between history tracking weights and optimal prediction weights.** Both  $M_{s,w,33,n}^H$  (Eq. 5) and  $p_{34}^*$  (Eqs. 1 and 2) are defined as weighted

sums of the preceding pitch values. We can therefore gain further insight into the relationship between history tracking and prediction by investigating the relationship between the weights on the preceding pitch values that determine  $M^H$  and  $p_{34}^*$ . Recall from Equation 6 that the regression of  $M^H$  onto  $p_{34}^*$  is defined as follows:  $M_{s,w,33,n}^H = \beta_{0,s,w}^{H^*} + \beta_{1,s,w}^{H^*} p_{34,n}^* + \varepsilon_{s,w,33,n}$ . By substituting the definitions of  $M^H$  and  $p_{34}^*$  into the regression equation and solving for the error term  $\varepsilon$ , we can see how the regression error depends on the relationship between the history tracking weights and optimal prediction weights. For fGn sequences:

$$\varepsilon_{s,w,33,n|\beta < 1} = \sum_{k=0}^{k=k'} (\beta_{k+1,s,w} - \beta_{1,s,w}^{H^*} a_{k+1,\beta}) p_{33-k,n} + R_{\beta < 1} \quad (7)$$

where

$$R_{\beta < 1} = (\beta_{0,s,w} - \beta_{0,s,w}^{H^*}) - \beta_{1,s,w}^{H^*} \sum_{k=k'+1}^{k=32} a_{k+1,\beta} p_{33-k}$$

and for fBm sequences:

$$\varepsilon_{s,w,33,n|\beta > 1} = (\beta_{1,s,w} - \beta_{1,s,w}^{H^*} a_{1,\beta-2}) p_{33,n} + \sum_{k=1}^{k=k'} (\beta_{k+1,s,w} - \beta_{1,s,w}^{H^*} (a_{k+1,\beta-2} - a_{k,\beta-2})) p_{33-k,n} + R_{\beta > 1} \quad (8)$$

where

$$R_{\beta > 1} = (\beta_{0,s,w} - \beta_{0,s,w}^{H^*}) - \beta_{1,s,w}^{H^*} (p_{33,n} - a_{k'+1,\beta-2} p_{33-k'-1,n}) + \sum_{k=k'+1}^{k=31} a_{k+1,\beta-2} (p_{33-k,n} - p_{33-k-1,n}).$$

Therefore, for fGn sequences, the regression error for  $M^H$  onto  $p_{34}^*$  partially depends on the  $(\beta_{k+1,s,w} - \beta_{1,s,w}^{H^*} a_{k+1,\beta})$  terms; that is, the difference between the history tracking weights  $\beta_{k+1}$  and the scaled optimal prediction weights  $a_{k+1,\beta}$  (Eq. 1), where the scaling factor  $\beta_{1,s,w}^{H^*}$  is the regression coefficient on  $p_{34}^*$ . Similarly, for fBm sequences, regression error partially depends on  $(\beta_{1,s,w} - \beta_{1,s,w}^{H^*} a_{1,\beta-2})$  and the  $(\beta_{k+1,s,w} - \beta_{1,s,w}^{H^*} (a_{k+1,\beta-2} - a_{k,\beta-2}))$  terms; that is, the difference between the history tracking weight and the scaled optimal prediction weight for the fBm sequence (Eq. 2), where the scaling factor is  $\beta_{1,s,w}^{H^*}$ , the regression coefficient on  $p_{34}^*$  (Eq. 6).

In Figure 7A, we plot optimal prediction weights for different sequence  $\beta$  values (see definition in Eqs. 1 and 2). From the above, we expect  $M^H$  to better predict  $p_{34}^*$  the more closely history tracking weights resemble the scaled optimal prediction weights.

To compare history dependence weights and scaled optimal prediction weights, we focused on the significant sensor cluster showing a prediction effect in the 200–250 ms window (see Fig. 5B, reproduced in Fig. 7B, inset). Because optimal  $k'$  differs across subjects and sensors, first, we averaged optimal  $k'$  across sensors in the significant cluster, training folds, and subjects and rounded the result to the nearest integer for the 200–250 ms and 250–300 ms time windows separately (where significant prediction effects are found). Call this rounded average value  $k'_{\text{avg},w}$ . We found that  $k'_{\text{avg},w} = 5$  and 4 for the 200–250 ms and 250–300 ms time windows, respectively. For these time windows, we then analyzed history dependence weights from the regressions where  $k' = k'_{\text{avg},w}$  for all subjects and sensors to allow for across-subject and across-sensor averaging. History dependence weights from both training folds were averaged across sensors in the significant cluster and then averaged across subjects. Next, we computed the scaled optimal prediction weights; that is, the  $\beta_{1,s,w}^{H^*} a_{k+1,\beta}$  terms in Equation 7 and the  $\beta_{1,s,w}^{H^*} a_{1,\beta-2}$  and  $\beta_{1,s,w}^{H^*} (a_{k+1,\beta-2} - a_{k,\beta-2})$  terms in Equation 8 to compare them to the history dependence weights. To determine the  $\beta_{1,s,w}^{H^*}$  values, we computed the across-subject, across-sensor average regression coefficient  $\beta_{1,s,w}^{H^*}$  for the 200–250 ms and 250–300 ms time windows, respectively. We aver-

aged the scaled prediction weights computed for each level of sequence  $\beta$  to arrive at a single set of scaled prediction weights for each time window (dashed lines in Fig. 7B).

Inspection of the history dependence weights suggests that they feature a pattern in which the pitch of the current and previous tone have large but opposing effects on MEG activity during the current tone (Fig. 7B). We investigated this effect further by calculating the across-sensor correlations between the history dependence weights for the current and previous pitches. In particular, if the current and previous pitch have opposing effects on MEG activity, then we might expect that the weights of the current and previous pitch on MEG activity are inversely correlated across sensors. Therefore, using history dependence weights defined by  $k'_{\text{avg},w}$  as described above, we investigated the across-sensor lagged correlation between the first history dependence weight ( $k = 0$ ; i.e., regression weight for the current tone pitch) and subsequent history dependence weights ( $k \geq 1$ ; i.e., regression weights for earlier tones in the sequence) for sensors in the 200–250 ms prediction cluster. We computed the correlation for all possible lags for each subject. Statistical significance was assessed by computing the Fisher's  $z$  transform of the correlation coefficients and submitting the  $z$ -values to an across-subject, one-sample  $t$  test against zero.

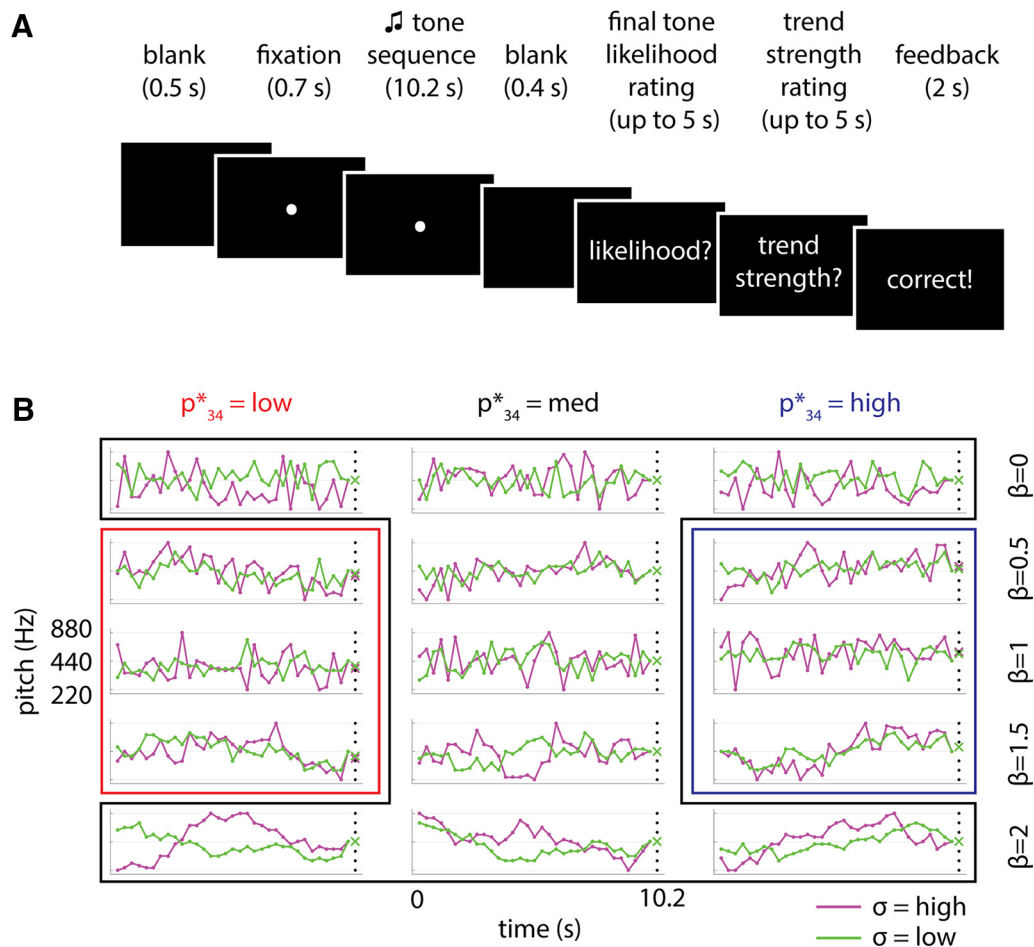
**Cluster correction.** To correct for multiple comparisons across sensors in the analyses of MEG data, we used cluster-based permutation tests (Maris and Oostenveld, 2007). For a given statistical test performed at each sensor, clusters were defined as neighboring sensors exhibiting test statistics of the same sign (e.g., positive  $t$ -values) and  $p < 0.05$ . The neighbors for each sensor were defined using the CTF275\_neighb.mat template in Fieldtrip. For each such cluster, a summary cluster statistic was defined as the sum of the absolute values of the test statistics across all sensors in the cluster. We generated a null distribution of cluster statistics by repeating the following procedure 1000 times: we randomly permuted the data independently for each subject (described in detail below) and used the same permutation across all sensors. Using the permuted data, we again performed the statistical test for each sensor and extracted the maximum cluster statistic across all clusters for the current permutation iteration. This yielded a null distribution of 1000 cluster statistics. We then used this null distribution to assign a  $p$ -value to clusters occurring in the original, unshuffled data. We defined a measure of effect size for clusters in the original data,  $d_{\text{cluster}}$ , as follows:

$$d_{\text{cluster}} = \frac{(CS_{\text{cluster}} - \text{mean}(CS_{\text{null}}))}{\text{std}(CS_{\text{null}})},$$

where  $CS_{\text{cluster}}$  is the cluster statistic for the cluster in the original data and  $CS_{\text{null}}$  is the null distribution of cluster statistics derived from the permutation procedure.  $d_{\text{cluster}}$  thus measures the effect size of a cluster statistic as the number of SDs by which it exceeds the mean of the null distribution.

For ANOVA analyses (Fig. 3B), we permuted the data by randomly shuffling the assignment of data cells to levels of the ANOVA factor under investigation for each subject. We considered clusters in the original data to be significant if their  $p$ -value was  $< 0.05$ . For regression analyses (Fig. 5B), we used the following procedure. For a given regression coefficient  $j$  at time window  $w$ ,  $\beta_{j,w}$ , we performed group-level analysis at every sensor using a one-sample  $t$  test against 0 on  $\beta_{j,w}$  across subjects. For each subject, we permuted the data by randomly shuffling the across-trial correspondence between the dependent variable (MEG activity) and independent variable (the term weighted by  $\beta_{j,w}$ ). We considered clusters in the original data to be significant if their  $p$ -value was  $< 0.025$  (corresponding to a two-tailed test at  $p < 0.05$ ). For cluster correction of  $k'$  (Figs. 4B, 5A), we defined clusters as spatially neighboring sensors exhibiting  $k' > 0$  at  $p < 0.05$  (uncorrected) and took the cluster statistic to be the sum of the  $k'$  values of all sensors in a cluster. The permutation procedure and the construction of the null distribution is described in the section “History tracking analysis.” We considered clusters in the original data to be significant if their  $p$ -value was  $< 0.05$ .

Permutation-based cluster correction has the advantage of being a nonparametric test and therefore not dependent upon distributional assumptions about the data. Inspection of the marginal distributions of



**Figure 1.** Task and stimuli. **A**, Trial structure. At the end of each trial, subjects judged the likelihood of the final tone, given the preceding tone sequence, on a scale of 1–5. Subjects then rated the “trend strength” of the sequence on a scale of 0–4. Feedback on the accuracy of the trend strength rating was provided at the end of every trial. **B**, Full stimulus set. For each level of  $\beta$  (rows), we generated 6 sequences, 3 of which ranged from 220 to 880 Hz (magenta traces, high  $\sigma$ ) and 3 of which spanned a slightly smaller pitch range (green traces, low  $\sigma$ ). For all sequences, the penultimate (33<sup>rd</sup>) tone was 440 Hz and the final (34<sup>th</sup>) tone was drawn from one of 6 possible values 4, 8, or 12 semitone steps above or below 440 Hz (black dots). For  $\beta = 0.5, 1,$  and  $1.5$ , sequences were chosen to have expected pitch for the final tone ( $p_{34}^*$ ) spanning low ( $< 440$  Hz; red outline), medium (440 Hz; black outline), and high ( $> 440$  Hz; blue outline) ranges (indicated by  $x$ ). For  $\beta = 0$  and  $2$ , the expected final tone pitch was always 440 Hz. Each sequence was presented 12 times, with the 6 possible values for final tone pitch distributed evenly and randomly across the 12 presentations, for a total of 360 trials.

MEG recordings within single subjects, sensors, and trials confirmed that MEG data was approximately normally distributed.

**Replication of main findings in an independent sample.** We performed a follow-up experiment the primary purpose of which was to investigate the influence of tone duration on the history tracking effect. Results for this analysis will be reported in a future study. For the purposes of the present study, we used this new dataset to investigate whether the primary behavioral and neural effects of interest would replicate in an independent sample. For the new experiment, we generated a new set of tone sequences using the same procedures described above with the following exceptions. Tone sequences had scaling parameter  $\beta$  values limited to 0.5, 0.99, or 1.5 and the sequence  $\sigma$  manipulation was omitted. Within each sequence, tone duration was 150 ms, 300 ms, or 600 ms. Subjects completed 324 trials in total (108 trials at each tone duration setting). The behavioral task was identical to that of the main experiment.

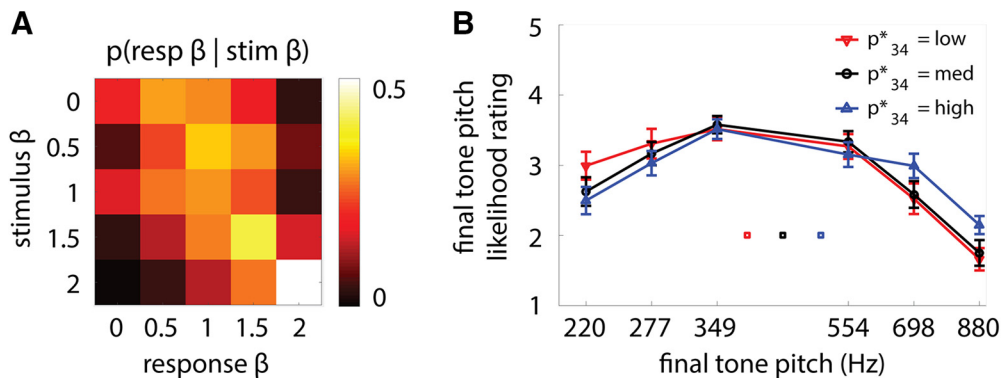
In total, 26 subjects completed the experiment in the MEG (lasting  $\sim 3$  h including setup time), of which 7 were prescreened for behavioral performance (as described for the main experiment), whereas the remaining 19 subjects performed the task for the first time in the MEG scanner without any behavioral prescreening. Six out of 26 subjects were excluded due to poor performance (not using the full range on rating scales) or excessive MEG artifacts caused by head movements, yielding a final group of 20 subjects for analyses (age range: 19–34; mean age 25.0; 11 females).

Behavioral results reported are from the MEG session and included the same 20 subjects who contributed MEG data. To investigate replication of the primary effects of interest from the main experiment, we analyzed data from the follow-up experiment using only trials with a 300 ms tone duration (108 trials per subject). For these trials, we assessed the behavioral prediction effect (cf. Figs. 8A, 2B), the neural prediction effect (cf. Figs. 8B, 5B), and the relationship between the neural prediction and history tracking effects (cf. Figs. 8C, 6D). All MEG preprocessing and data analysis procedures were conducted as described above for the main experiment.

## Results

### Task paradigm

On every trial, human subjects listened to a  $\sim 10$ -s-long sequence of 34 tones, each 300 ms in duration. The pitch (i.e., log of sound-wave frequency) fluctuation of the first 33 tones in the sequence followed different levels of temporal dependence as prescribed by the parameter  $\beta$ , which measures the slope of the temporal power spectrum of the pitch time series (following a  $P \propto 1/f^\beta$  form). Higher  $\beta$  (i.e., steeper power spectrum) corresponds to stronger temporal dependence (Fig. 1B). Sequence  $\beta$  spanned five levels: 0, 0.5, 1, 1.5, and 2, with  $\beta = 0$  corresponding to no temporal dependence and  $\beta = 2$  corresponding to strong temporal depen-



**Figure 2.** Behavioral performance. **A**, Trend strength rating. Heat map of the group-average distribution of trend strength ratings (transformed into the  $\beta$  scaling as “response  $\beta$ ”) as a function of true sequence  $\beta$  (“stimulus  $\beta$ ”). Responses concentrated along the diagonal, suggesting that subjects could meaningfully perform the trend strength rating task. Spearman’s  $\rho$  for stimulus  $\beta$  and response  $\beta$  was significantly above chance for every subject (all  $p < 0.001$ ,  $n = 360$  trials for each subject), with an average rho of 0.45. **B**, Final tone pitch likelihood rating. Final tone likelihood rating is plotted as a function of presented final tone pitch ( $p_{34}$ ; abscissa) and expected final tone pitch ( $p_{34}^*$ ; red, low; black, medium; blue, high). A repeated-measures ANOVA revealed a significant interaction between  $p_{34}$  and  $p_{34}^*$  ( $F_{(10,100)} = 5.7$ ,  $p < 0.001$ ,  $n = 11$ ), which can be observed by noting the crossover effect of the red and blue curves. Red, black, and blue squares indicate mean expected final tone pitch values for the respective conditions. Error bars denote SEM across subjects.

dence. Thirty unique sequences were used in the experiment. For each level of  $\beta$ , three sequences spanned the full range of 220–880 Hz (Fig. 1B, magenta traces) and three sequences spanned a slightly smaller range of pitch (average range = 276–701 Hz; Fig. 1B, green traces, for details, see Materials and Methods). All sequences converged on a pitch of 440 Hz for the penultimate (33<sup>rd</sup>) tone to disentangle neural activity related to instantaneous stimulus processing from that related to predictive processing (see below).

Given knowledge about a sequence of  $N - 1$  values and the sequence’s autocorrelation structure, it is possible to mathematically compute the expected value of the  $N^{\text{th}}$  item in the sequence (see Materials and Methods section “Calculation of  $p_{34}^*$ ”). We refer to the expected value of the final tone pitch as  $p_{34}^*$ . For  $\beta = 0.5, 1$ , and 1.5, sequences varied in  $p_{34}^*$ , exhibiting low (<440 Hz; Fig. 1B, red outline), medium (440 Hz; black outline), or high (>440 Hz; blue outline) expected values (indicated by the  $\times$  at the end of each sequence in Fig. 1B). Average values of  $p_{34}^*$  for the low and high categories were 388.5 Hz (range: [370, 415] Hz) and 504.2 Hz (range: [466, 554] Hz), respectively. For  $\beta = 0$  sequences, because there is no temporal dependence,  $p_{34}^*$  is simply the mean value of the sequence and is 440 Hz. For  $\beta = 2$  sequences, because they are random walks with no drift,  $p_{34}^*$  is the same as the pitch of the 33<sup>rd</sup> tone and is also 440 Hz.

Each of the 30 unique sequences was presented 12 times (in randomized order), for a total of 360 trials. For each unique sequence, the actually presented 34<sup>th</sup> (and final) tone pitch varied across the 12 presentations and was randomly selected from one of six values evenly spaced around the penultimate tone pitch of 440 Hz (black dots in Fig. 1B). After listening to the tone sequence on a given trial, subjects made two responses (Fig. 1A). First, they rated how probable the final tone pitch was given the preceding tone sequence on a scale of 1–5. Second, they indicated the “trend strength” of the first 33 tones of the sequence on a scale of 0–4, where each trend strength rating corresponded to a level of  $\beta$ .

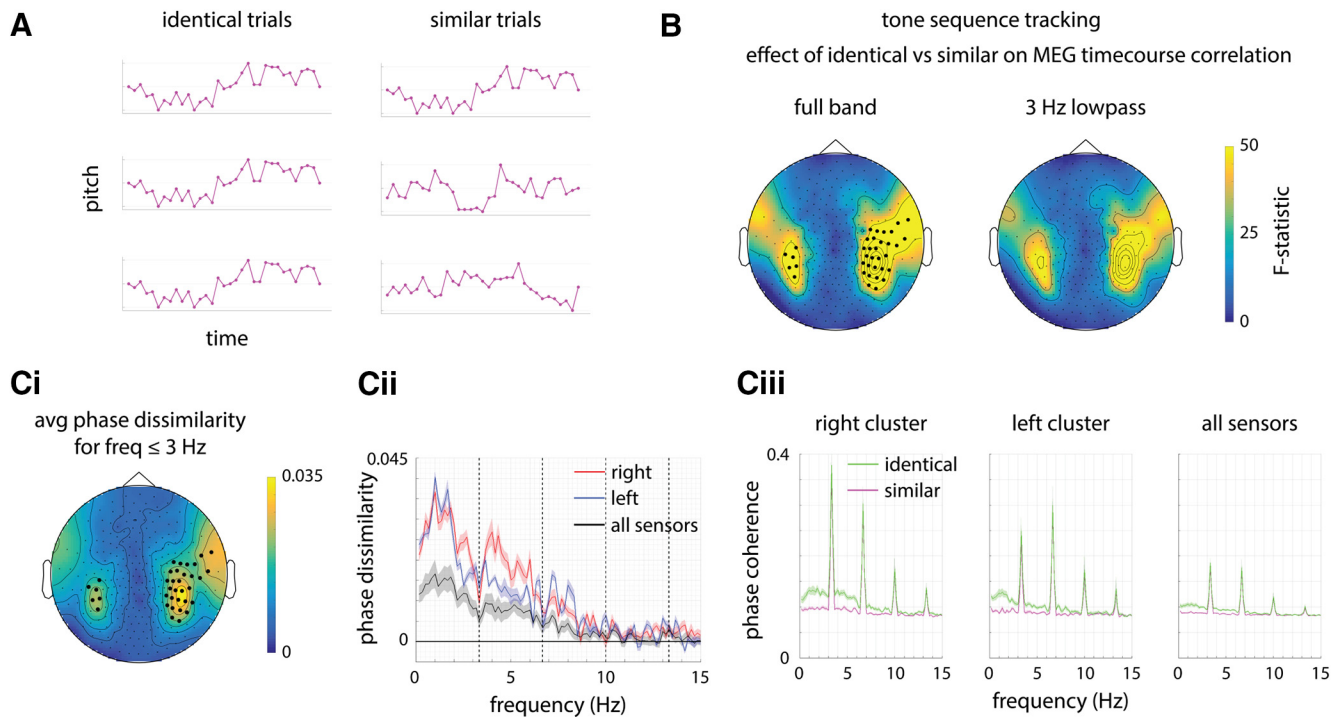
### Subjects can make valid predictions about upcoming items based on stimulus sequence history

We first replicated previous findings that subjects can form valid judgment about sequence temporal dependence (Schmuckler and Gilden, 1993; Lin et al., 2016). The across-trial correlation

between sequence  $\beta$  and trend strength rating was significant for all subjects (all  $p < 0.0001$ ,  $n = 360$  trials for each subject; mean  $\rho = 0.45$ ). Figure 2A plots the group average of stimulus–response matrix, with the distribution of responses concentrated along the diagonal, indicating that subjects’ responses about “trend strength” closely tracked sequence  $\beta$ .

Can subjects capitalize on the temporal dependence in a sequence to make valid predictions about upcoming items in the sequence? To investigate this question, we performed the following analysis. For each subject, we calculated the average likelihood rating for the final tone pitch as a function of its expected value ( $p_{34}^*$ ; low, medium, high), its actually presented value ( $p_{34}$ ) and overall variability in the pitch sequence ( $\sigma$ ; magenta vs green traces in Fig. 1B). We then submitted the likelihood ratings to a 6 ( $p_{34}$ )  $\times$  3 ( $p_{34}^*$ )  $\times$  2 ( $\sigma$ ) repeated-measures ANOVA. There was a main effect of  $p_{34}$  ( $F_{(5,50)} = 20.2$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.67$ ), reflecting the overall inverse U shape of the function (Fig. 2B); as expected, subjects rated final tone pitch to be less likely the more it deviated from the penultimate tone pitch of 440 Hz. Crucially, we also observed a  $p_{34} \times p_{34}^*$  interaction ( $F_{(10,100)} = 5.7$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.36$ ). The interaction can be observed by noting that, when the presented final tone pitch was low (220 Hz), subjects rated the final tone to be more likely when its expected pitch was also low (paired  $t$  test for  $p_{34}^* = \text{low}$  vs  $p_{34}^* = \text{high}$ ,  $t_{(10)} = 3.2$ ,  $p = 0.009$ , difference = 0.50); in contrast, when the presented final tone pitch was high (880 Hz), subjects rated the final tone to be more likely when its expected pitch was also high ( $t_{(10)} = -3.4$ ,  $p = 0.007$ , difference =  $-0.49$ ). Therefore, subjects’ likelihood ratings for the final tone pitch were sensitive to its expected value, which depends on the previous sequence history. This result suggests that subjects were indeed able to capitalize on sequence temporal dependence to make valid predictions. No other main effect or interaction in the ANOVA was significant (all  $p > 0.15$ ). Therefore, likelihood ratings did not appear to be sensitive to the pitch sequence variability ( $\sigma$ ) manipulation in our stimulus set.

See Materials and Methods “Subjects” section for results from an additional behavioral dataset, which replicated the above findings. We now proceed to investigate the neural mechanisms underlying the predictive computation: What aspects of neural activity encode a particular sequence and what aspects of neural activity contribute to prediction? Finally, how is such prediction constructed?



**Figure 3.** Slow cortical dynamics track tone sequences. **A**, Analysis schematic. We defined “identical” trials as trials in which an identical tone sequence was presented repeatedly and “similar” trials as trials in which tone sequences were statistically similar (same  $\beta$  and  $\sigma$ ) but nonidentical. We identified neural activity encoding tone sequences by comparing MEG activity across-trial correlation/coherence for identical versus similar trials. **B**, Correlation analysis. Across-subject mean  $F$ -values corresponding to the effect of “identical” versus “similar” trial pairs from within-subject ANOVAs using across-trial correlation of the MEG activity time course as the dependent variable. A repeated-measures ANOVA across subjects shows a single significant cluster including all sensors after correction by cluster-based permutation test ( $p < 0.001$ ,  $n = 11$ ). Left, Result obtained using full-band (0.05–150 Hz) MEG data. Right, Result obtained using 0.05–3 Hz data. Black dots indicate sensors where tone sequence tracking effect in the full band data was strongest ( $F > 50$ ). **C**, Phase dissimilarity analysis. **Ci**, Group-average phase dissimilarity (difference in phase coherence between identical and similar set of trials) in the  $\leq 3$  Hz band is plotted across sensors. Black dots are the same as in **B** and were used to define left and right sensor clusters for subsequent analysis. **Cii**, Phase dissimilarity by frequency for the sensor clusters shown in **Ci**, as well as for all sensors. Dashed vertical lines indicate tone presentation frequency (3.3 Hz) and its harmonics. Shaded areas denote SEM across subjects. **Ciii**, Phase coherence across identical (green) and similar (magenta) set of trials, respectively, for the sensor clusters shown in **Ci**, as well as for all sensors. Peaks in the plots correspond to tone presentation frequency and its harmonics.

### Slow cortical dynamics track naturalistic tone sequences

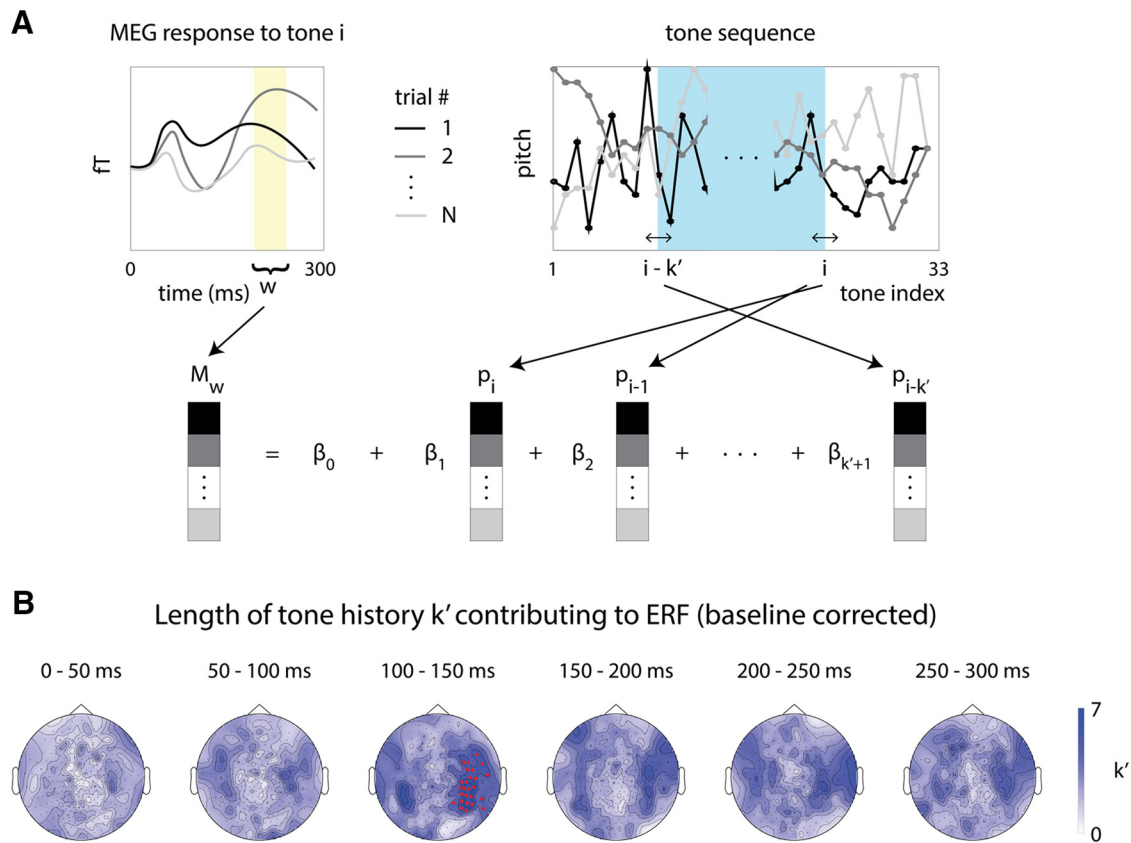
To identify neural activity that encodes tone sequences, we searched for components of the MEG activity that were reliably modulated by repeated presentations of an identical tone sequence above and beyond sequences that were only statistically similar. For each subject, we defined “identical” trial pairs as pairs of trials in which the same initial 33-tone sequence was presented and “similar” trial pairs as pairs of trials in which the initial 33-tone sequences were different but exhibited the same sequence  $\beta$  and variability  $\sigma$  (Fig. 3A). For each sensor, we computed the Pearson correlation of MEG activity time courses during the presentation of the first 33 tones in a sequence for all “identical” and “similar” trial pairs. These correlation values were Fisher- $z$  transformed and averaged across trial pairs for the two groups of trials separately and then submitted to a repeated-measures ANOVA across subjects. There was a strong effect of “identical” versus “similar” trial pair including all sensors (Fig. 3B left; all sensor-level  $p$ -value  $< 0.047$ ), which formed a single significant cluster comprising all 271 sensors [ $p < 0.001$ , corrected by cluster-based permutation test (henceforth “cluster-corrected”); cluster size = 271,  $d_{\text{cluster}} = 14.9$ ]. The magnitude of the effect was particularly high in sensors overlying bilateral auditory cortices, with a right lateralization. There was no significant interaction effect with the temporal dependence ( $\beta$ ) or variability ( $\sigma$ ) of stimulus sequences. Strikingly, these findings were almost identical when MEG activity was low-pass filtered at  $< 5$ , 3, or 1 Hz before computing trial-pair correlations (the  $< 3$  Hz result is shown in Fig.

3B, right). Therefore, repetitions of an identical tone sequence consistently modulated MEG activity in the low-frequency range.

To evaluate the frequency content of this activity more thoroughly, we computed a phase dissimilarity measure (for details, see Materials and Methods) for each sensor at each frequency, which quantifies the degree of phase locking to repetitions of an identical stimulus sequence above and beyond that elicited by statistically similar sequences and was computed as the difference in phase coherence between “identical” and “similar” set of trials (Luo and Poeppel, 2007). Reassuringly, mean phase dissimilarity in the  $\leq 3$  Hz range (Fig. 3Ci) has a similar spatial pattern to the results from the correlational analysis described above (Fig. 3B). Phase dissimilarity as a function of frequency averaged across all sensors is shown in Figure 3Cii (black trace), which revealed that phase dissimilarity did not peak at a particular frequency, but rather was spread across all frequencies  $< 10$  Hz, with the highest values in the lowest frequencies. This pattern suggests phase locking in arrhythmic rather than oscillatory brain activity (He, 2014). The phase dissimilarity plots averaged across sensors in the left and right clusters defined from full-band correlational analysis (black dots in Fig. 3B, left) exhibit similar patterns (Fig. 3Cii, red and blue traces).

The phase dissimilarity curves exhibited sharp dips at the frequency of tone presentation (3.3 Hz) and its harmonics (Fig. 3Cii, dashed vertical lines). To gain additional insight into this pattern, we plotted across-trial phase coherence separately for “identical” and “similar” trials (Fig. 3Ciii). Interestingly, whereas





**Figure 4.** History-dependent neural responses. **A**, Analysis schematic. MEG activity in response to the  $i^{\text{th}}$  tone in the sequence is regressed onto the pitch of tone  $i$ , as well as the pitch values of the preceding  $k'$  tones, for all values of  $i$  occurring in the second half of the tone sequence. The regression is performed separately for each MEG sensor, in a sliding window  $w$  (length 50 ms) and a cross-validation procedure is used to determine the value of  $k'$  that provides the best fit to the data. **B**, Length of history dependence in baseline-corrected MEG signals (i.e., evoked responses). The optimal  $k'$  value determined for each sensor is plotted for each time window after tone onset. A right-lateralized sensor cluster in the 100–150 ms time window survived correction by cluster-based permutation test (red dots;  $p = 0.01$ ,  $n = 11$ ), where MEG response to a given tone is influenced by up to 7 prior tones.

both “identical” and “similar” trials display strong phase locking at the frequency of tone presentation and its harmonics, only identical trial pairs exhibit additional arrhythmic, low-frequency phase locking. Therefore, phase locking at 3.3 Hz and its harmonics reflects consistent modulation of MEG activity locked to the onset of each tone (likely contributed by the event-related field, ERF, due to tone presentation) regardless of the evolving pitch sequence. In contrast, phase locking in the low-frequency arrhythmic activity tracks the specific sequence of pitch values. Together, these results suggest that tone sequences with naturalistic temporal statistical regularities consistently modulate slow, arrhythmic brain activity.

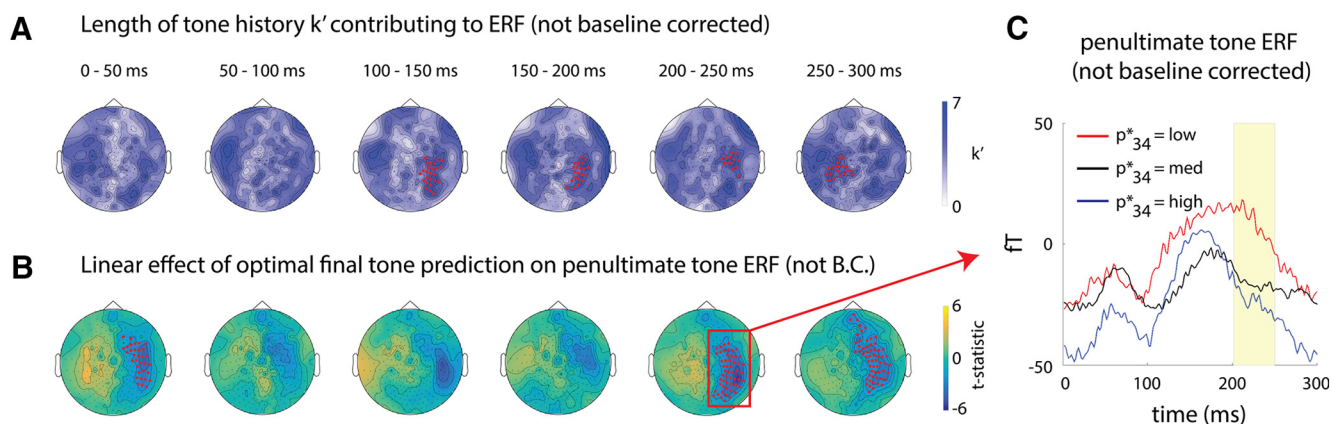
**Sequence integration in neural activity**

The above results may reflect two possible scenarios: the MEG activity might track the fluctuating tone pitches instantaneously or integrate information in the sequence over a period of time. Next, we tested whether MEG activity at a given moment depended not only on the current tone pitch but also on the history of preceding tone pitches.

To this end, we regressed MEG activity after the onset of a given tone (using a 50-ms-long sliding window applied to full-band data) onto the pitch values of the previous  $k'$  tones, for each value of  $k'$  ranging from 0 to 15. We determined which value of  $k'$  provided the best fit to the data at each sensor and time window using a cross-validation approach (see Materials and Methods, “History tracking analysis” section). The cross-validated  $k'$  value

thus describes the length of sequence history that affects neural activity at a given moment. A schematic of this analysis is shown in Figure 4A. We performed this analysis using both conventional baseline-corrected tone ERFs and MEG activity without baseline correction. The former reveals whether the evoked response to a given tone is modulated by preceding tones. However, applying baseline correction also removes any potential effect in neural activity that accumulates over the course of sequence processing. Such ongoing (i.e., non-baseline-corrected) activity is a good candidate for carrying a continuously updated prediction for the upcoming tone pitch (investigated in the next section). Therefore, we also applied the regression analysis on non-baseline-corrected data.

The result using baseline-corrected MEG responses is shown in Figure 4B. As expected,  $k'$  values for all sensors are near 0 within the first 50 ms of tone onset. However, over time, groups of sensors exhibiting sensitivity to the preceding tone sequence begin to emerge. This effect achieves statistical significance in the time window of 100–150 ms ( $p = 0.01$ , cluster-corrected; cluster size = 26,  $d_{\text{cluster}} = 4.2$ , mean  $k'$  in cluster = 5.4, max  $k'$  in cluster = 7.3; red dots in Fig. 4B), when a right-lateralized group of sensors exhibit tone-evoked responses that are sensitive to the pitch of up to 7 previous tones. Therefore, the evoked response to a given tone depends on preceding tones in the sequence. This result suggests history-dependent stimulus processing and is consistent with the idea that, during neural processing of naturalistic temporally extended stimuli, past information continu-



**Figure 5.** Neural correlates of history integration and prediction. **A**, Length of history dependence in non-baseline-corrected MEG signals (i.e., ongoing/accumulated activity). Same as Figure 4B, but for non-baseline-corrected data. Red dots indicate sensors belonging to significant clusters (all  $p < 0.05$ , cluster-based permutation test,  $n = 11$ ). **B**, Neural correlate of prediction. Using a 50 ms sliding window, non-baseline-corrected MEG activity during the penultimate tone (which always had a pitch of 440 Hz) was regressed onto the expected pitch of the final tone to determine which sensors showed a significant prediction effect. Maps plot the  $t$ -values corresponding to a group-level one-sample  $t$  test on regression coefficients at each sensor and time location. Red dots indicate sensors belonging to significant clusters (all  $p < 0.05$ , cluster-based permutation test,  $n = 11$ ). **C**, MEG activity during penultimate tone as a function of expected final tone pitch. The non-baseline-corrected MEG signal for sensors in the significant cluster from the 200–250 ms window (red dots in **B**) as a function of the expected pitch of the final tone. A monotonic relationship can be seen in the significant time window. Note that, although expected pitch is binned into three levels here for illustration purposes, the regression analysis depicted in **B** used continuously varying expected values for pitch.

ously influences information processing in the present (Hasson et al., 2015).

The result using non-baseline-corrected MEG data is shown in Figure 5A. This analysis also revealed a right-lateralized sensor cluster in the 100–150 ms window exhibiting significant sensitivity to up to 7 previous tones ( $p = 0.001$ , cluster-corrected; cluster size = 25,  $d_{\text{cluster}} = 5.2$ , mean  $k'$  in cluster = 5.3, max  $k'$  in cluster = 6.8), which persisted until 250 ms after tone onset (150–200 ms:  $p = 0.01$ , cluster corrected, cluster size = 21,  $d_{\text{cluster}} = 3.5$ , mean  $k'$  in cluster = 5.1, max  $k'$  in cluster = 6.8; 200–250 ms:  $p = 0.045$ , cluster corrected, cluster size = 13,  $d_{\text{cluster}} = 2.1$ , mean  $k'$  in cluster = 5.7, max  $k'$  in cluster = 7.0). An additional left-lateralized sensor cluster emerged in the 250–300 ms window ( $p = 0.017$ , cluster-corrected; cluster size = 19,  $d_{\text{cluster}} = 3.3$ , mean  $k'$  in cluster = 5.3, max  $k'$  in cluster = 6.4). Therefore, ongoing (non-baseline-corrected) MEG activity exhibits stronger history dependence than the baseline-corrected, tone-evoked responses, suggesting a build-up effect over time that contains information integrated over sequence history.

### Neural correlate of prediction

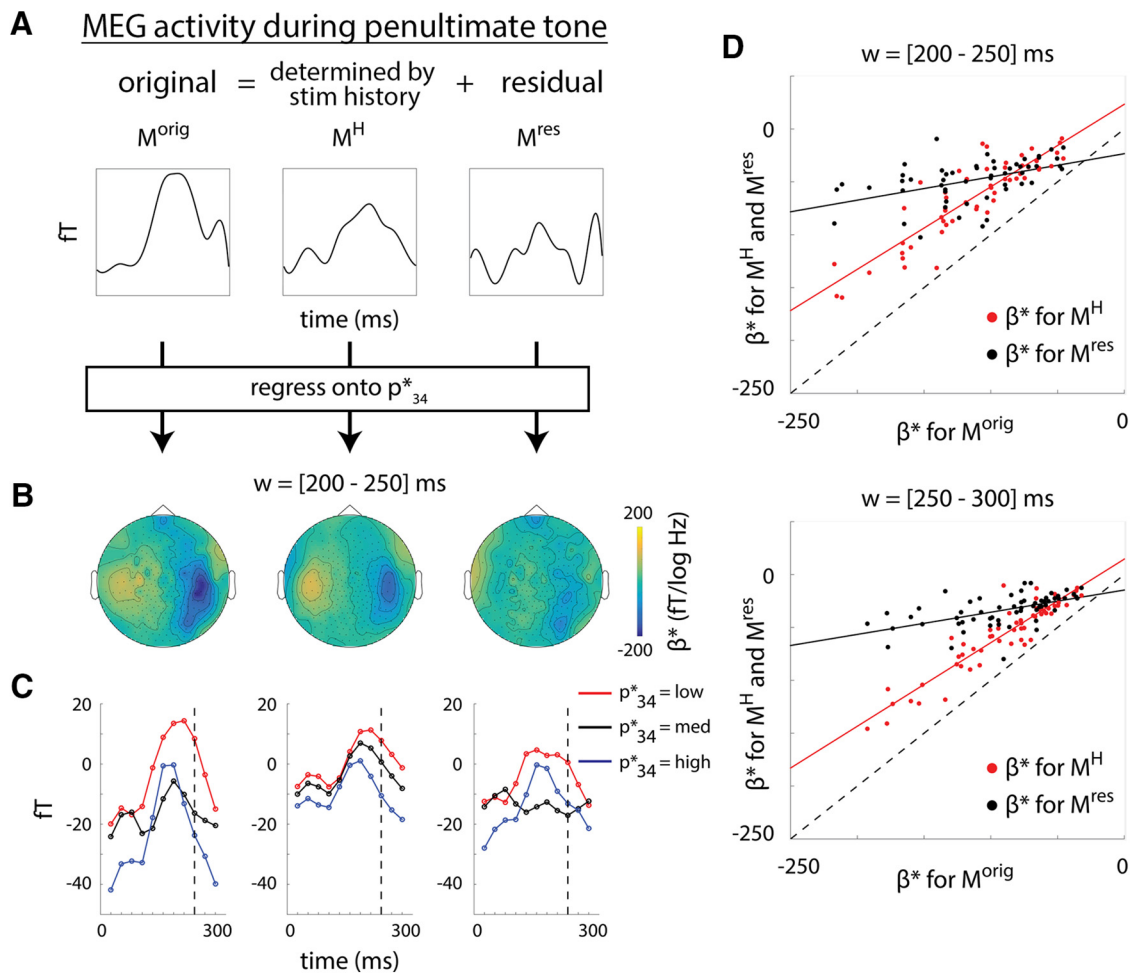
To identify neural mechanisms underlying sequence prediction, we investigated whether brain activity during the penultimate (33<sup>rd</sup>) tone (which always had a pitch of 440 Hz) contains information about the expected pitch of the upcoming final tone,  $p_{34}^*$  (see Materials and Methods, “Prediction effect” section). Because sensory input during the penultimate tone is identical across all trials, this analysis allowed a clean assessment of neural activity underlying the predictive computation without the confounds of concurrent sensory processing. Similar to the previous analysis, we investigated MEG activity both with and without baseline correction to focus on tone-evoked responses and slow fluctuations that build up over multiple tones, respectively. Using ongoing, non-baseline-corrected activity, the activity in a right lateralized sensor cluster predicted the expected final tone pitch in early (0–50 ms) and late (200–300 ms) time windows (Fig. 5B, red dots, cluster-corrected; 0–50 ms:  $p = 0.02$ , cluster size = 33,  $d_{\text{cluster}} = 3.1$ ; 200–250 ms:  $p = 0.001$ , cluster size = 53,  $d_{\text{cluster}} = 6.5$ ; 250–300 ms:  $p < 0.001$ , cluster size = 59,  $d_{\text{cluster}} = 6.1$ ). The

average activity time courses in the significant sensor cluster from the 200–250 ms time window are shown in Figure 5C for trials with low, medium, and high expected final tone pitch, respectively, which exhibit a monotonic relationship with the expected pitch of the upcoming tone. No significant effect was found using baseline-corrected data. Therefore, the neural signature of prediction is carried by slow cortical dynamics that accumulates information over the tone sequence.

### Sequence integration accounts for prediction

The above result reveals a neural correlate of subjects’ predictions about upcoming, but yet unrepresented, stimulus. A remaining important question is the computational mechanisms leading to the construction of such predictive neural activity. Given that predicting upcoming items in a sequence requires integrating information over sequence history and extrapolating based on its temporal statistical regularity, could the sequence integration effect observed in our earlier analysis directly form the basis for predictive computation? We reasoned that, if this were the case, then the MEG activity during the penultimate tone that is predicted by sequence history integration (Fig. 5A) should also be predictive of the expected final tone pitch (as assessed in Fig. 5B). This is a substantive empirical question because the results thus far are consistent with the possibility that the neural integration of tone sequence history revealed in our analysis serves no predictive function.

We investigated this question using the following analysis. For each subject, we computed the component of (non-baseline-corrected) MEG activity during penultimate tone presentation that is accounted for by stimulus history (including the present tone) based on the history dependence analysis described earlier (Fig. 5A; for details, see Materials and Methods, “Relating history-dependent MEG response with  $p_{34}^*$ ” section). Call this activity  $M^{\text{H}}$ , where the “H” superscript denotes sequence history (Fig. 6A, middle). We then computed the residual activity not accounted for by sequence history,  $M^{\text{res}}$  (Fig. 6A, right). Therefore,  $M^{\text{H}}$  captures the portion of neural activity that linearly depends on tone pitch history, whereas  $M^{\text{res}}$  captures any additional nonlinear dependencies as well as other influences. We next repeated the above



**Figure 6.** Neural integration of stimulus history accounts for prediction. **A**, Analysis schematic. The (non-baseline-corrected) MEG activity during the penultimate tone (left column) can be expressed as the component accounted for by stimulus history integration (middle column) plus residual (right column). Prediction effect for each signal is then assessed by regressing the activity onto the expected final tone pitch, as in Figure 5B. **B**, Regression coefficients of the expected final tone pitch for the original MEG activity during the penultimate tone (left, from the 200–250 ms window), its history-dependent component (middle), and the residual (right). Maps show mean coefficients across subjects. **C**, MEG activity time courses during penultimate tone as a function of expected final tone pitch. Left, Reproduced from Figure 5C, but with 50-ms-length, half-overlapping sliding windows. Middle and right, Component constructed from history integration effect and the residual, respectively. **D**, Comparison of regression coefficients for  $M^{\text{orig}}$ ,  $M^{\text{H}}$ , and  $M^{\text{res}}$  across sensors. For sensor clusters showing significant prediction effects in the time windows 200–250 ms (top) and 250–300 ms (bottom) (red dots in Fig. 5B), the original and history-dependent MEG activity exhibit similar sensitivity to  $p_{34}^*$  (red dots and lines; 200–250 ms:  $r = 0.89$ ,  $p < 10^{-15}$ ,  $n = 53$ ; 250–300 ms:  $r = 0.93$ ,  $p < 10^{-15}$ ,  $n = 59$ ). In comparison, the residual activity is much less sensitive to  $p_{34}^*$  (shown by the black dots being closer to the ordinate value of 0) and less similar to the original response (200–250 ms:  $r = 0.48$ ,  $p = 0.0003$ ,  $n = 53$ ; 250–300 ms,  $r = 0.56$ ,  $p = 4 \times 10^{-6}$ ,  $n = 59$ ). Dashed lines indicate unity.

tone pitch prediction analysis by regressing MEG activity onto  $p_{34}^*$  (as in Fig. 5B), but using  $M^{\text{H}}$  and  $M^{\text{res}}$  as the dependent variables.

The scalp topography of regression weights in the window of 200–250 ms are shown in Figure 6B. Consistent with our hypothesis that sequence history integration reflected in MEG activity contributes to its predictive sensitivity to the upcoming tone’s expected pitch, the originally observed relationship between MEG activity and expected pitch was closely mirrored by  $M^{\text{H}}$ , but not  $M^{\text{res}}$ . Using the significant sensor cluster showing a prediction effect in the 200–250 ms window (Fig. 5C, reproduced in Fig. 6C, left, using a 50 ms sliding window), we further found that the component of its activity that is attributable to sequence history ( $M^{\text{H}}$ ; Fig. 6C, middle), but not the residual activity ( $M^{\text{res}}$ ; Fig. 6C, right), shows a linear relationship with  $p_{34}^*$ .

We quantified these observations by comparing the regression weights for  $p_{34}^*$  obtained using the original MEG activity during the penultimate tone ( $M^{\text{orig}}$ ) with those obtained using  $M^{\text{H}}$  and  $M^{\text{res}}$  for the sensor clusters exhibiting a significant prediction

effect in the 200–250 ms and 250–300 ms time windows (Fig. 5B, red dots). The regression weights are plotted across sensors for  $M^{\text{orig}}$  against  $M^{\text{H}}$  (Fig. 6D, red) and for  $M^{\text{orig}}$  against  $M^{\text{res}}$  (Fig. 6D, black). There was a strong correlation between the regression weights for  $M^{\text{orig}}$  and  $M^{\text{H}}$  (200–250 ms:  $r = 0.89$ ,  $p < 10^{-15}$ , regression coefficient = 0.78,  $n = 53$ ; 250–300 ms:  $r = 0.93$ ,  $p < 10^{-15}$ , regression coefficient = 0.79,  $n = 59$ ). Therefore, the sensitivity of the MEG activity to the expected upcoming tone pitch is largely explained by its linear sensitivity to the preceding tone sequence. In contrast, the relationship between the regression weights for  $M^{\text{orig}}$  and  $M^{\text{res}}$  was considerably weaker though non-negligible (200–250 ms:  $r = 0.48$ ,  $p = 0.0003$ , regression coefficient = 0.22,  $n = 53$ ; 250–300 ms:  $r = 0.56$ ,  $p = 4 \times 10^{-6}$  and regression coefficient = 0.21,  $n = 59$ ). This suggests that the linear history dependence of the MEG activity cannot entirely explain the predictive effect, and nonlinear history dependence may carry additional prediction effect.

Overall, these results suggest that linear integration of sequence history contained in slow, arrhythmic neural activity con-

stitutes a prominent component of predictive processing for upcoming items in the sequence.

### Sequence integration weights resemble scaled optimal prediction weights

Further insight into the relationship between history-dependent processing and prediction follows from the fact that both the history-dependent MEG activity  $M^H$  and the optimal predicted pitch  $p_{34}^*$  are defined as weighted sums of the preceding pitch values. Therefore, we can investigate the relationship between history-dependent processing and prediction by analyzing the relationship between the weights on the pitch sequence history that determine  $M^H$  and  $p_{34}^*$ .

As shown in the Materials and Methods section “Relationship between history tracking weights and optimal prediction weights,” error in the regression of pitch-history-dependent MEG activity  $M^H$  onto the expected pitch  $p_{34}^*$  depends on the discrepancy between the weights characterizing history dependence in MEG activity and a scaled version of the weights determining optimal pitch prediction (Eqs. 7 and 8), where the scaling factor is the  $\beta_1^{H^*}$  coefficient determined by the regression of  $M^H$  onto  $p_{34}^*$  (Eq. 6). Optimal prediction weights for  $\beta = 0$  and  $\beta = 0.5$  sequences (as defined in Eq. 1) are plotted in Figure 7A, top row. Optimal prediction weights for  $\beta = 1.01$  and  $\beta = 1.5$  sequences (as defined in Eq. 2) are plotted as solid lines in Figure 7A, bottom row, along with the prediction weights for the corresponding fGn sequences (dashed lines), which were used to derive the optimal prediction weights for  $\beta > 1$  sequences (see Materials and Methods for details).

We investigated how history tracking weights compared with scaled optimal prediction weights in the cluster of sensors showing significant prediction effect in the 200–250 ms time window (Fig. 7B, inset, displaying the result from Fig. 5B). We computed the average  $k'$  value in this cluster for the 200–250 ms and 250–300 ms time windows (time period with significant prediction effects; Fig. 5B),  $k'_{\text{avg}, w}$ , and then compared averaged history tracking and scaled prediction weights across sensors in the cluster extending  $k'_{\text{avg}, w}$  tones back. As expected, history dependence weights for MEG activity exhibited a pattern qualitatively similar to that of scaled optimal prediction weights (Fig. 7B). This suggests that modulation of neural activity due to pitch sequence history follows a pattern similar to the manner in which the expected pitch for the upcoming tone depends on pitch sequence history and that this similarity between neural history dependence and the calculation of expected pitch is what underlies the neural prediction effect.

The results in Figure 7B suggest that the pitch of the current ( $k = 0$ ) and previous ( $k \geq 1$ ) tones have opposite effects on MEG activity. To quantify this effect, we computed the across-sensor lagged correlation between the weight on the current pitch,  $\beta_1$  (corresponding to  $k = 0$ ), and the weight on pitch  $k$  tones back ( $k > 0$ ),  $\beta_{k+1}$ , for sensors in the prediction cluster (Fig. 7B, inset). An example correlation between  $\beta_1$  and  $\beta_2$  (corresponding to weights on the current and immediately preceding tones) for a single subject is shown in Figure 7C. For each level of lag ( $k$ ), we performed a one-sample  $t$  test on the Fisher's  $z$ -transformed correlation coefficients to assess significance at the group level. We found that the lag-1 correlation was significant at both the 200–250 ms ( $t_{(10)} = -3.0$ ,  $p = 0.01$ , Cohen's  $d = -0.92$ ) and 250–300 ms ( $t_{(10)} = -2.3$ ,  $p = 0.04$ , Cohen's  $d = -0.69$ ) time windows (Fig. 7D), indicating that, for sensors in the significant prediction cluster, the current and previous tone pitch had opposing effects on MEG activity. The lag-4 correlation was also significant for the

200–250 ms time window ( $t_{(10)} = -2.3$ ,  $p = 0.04$ , Cohen's  $d = -0.69$ ), consistent with the observation that both history tracking weight and scaled prediction weight have negative values for  $k = 4$  (Fig. 7B).

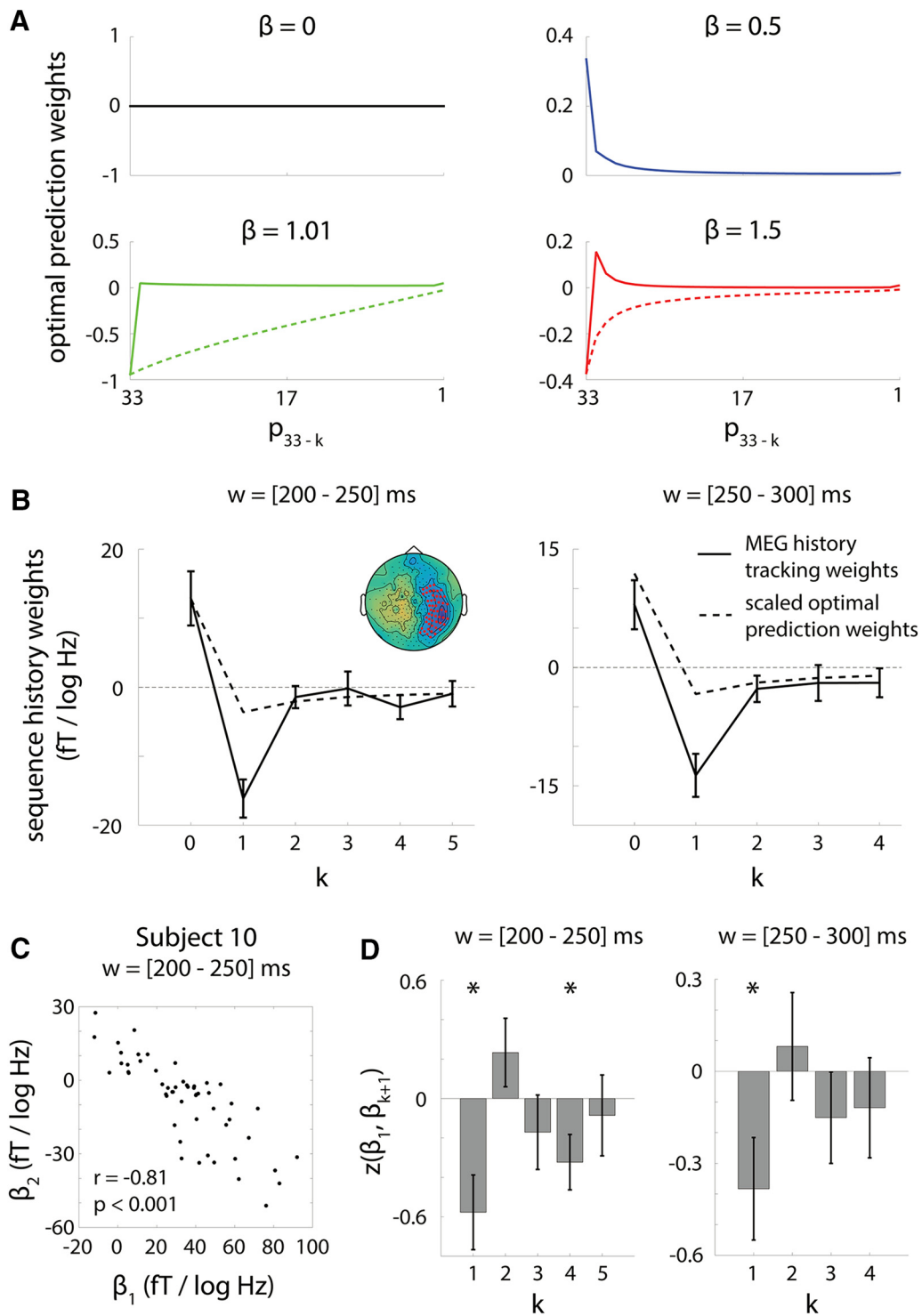
### Replication of main findings in an independent sample

To assess the robustness of our main findings, we analyzed an independent dataset from 20 additional subjects who performed a slightly modified version of the experiment (for details, see Materials and Methods, “Replication of main results in an independent sample” section).

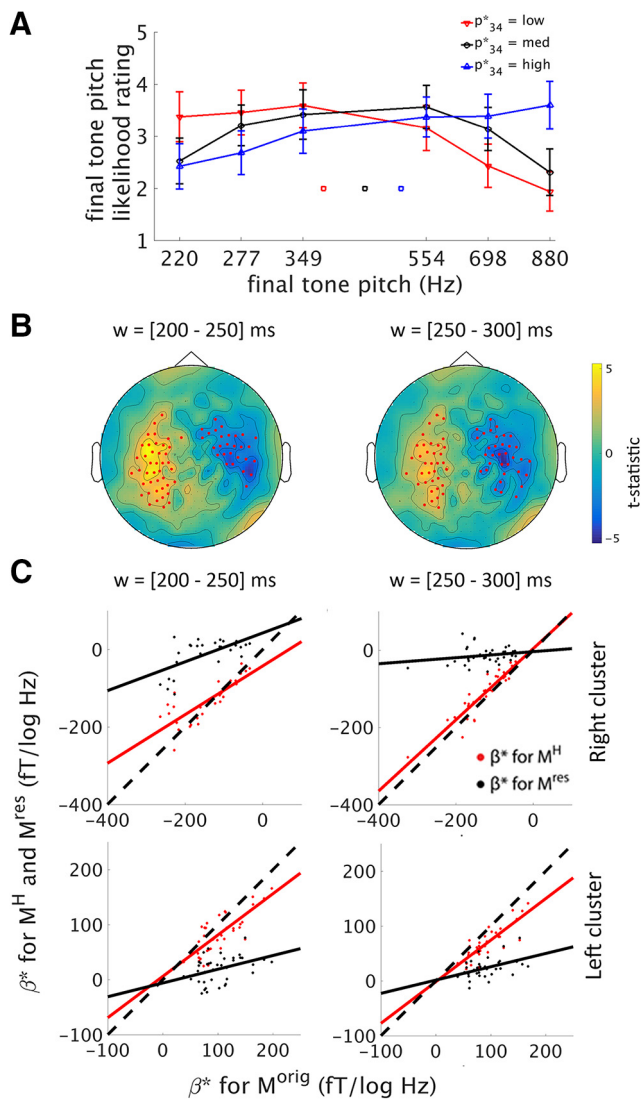
Consistent with earlier results (Fig. 2A and Lin et al., 2016), subjects were able to discriminate sequence  $\beta$  ( $t$  test on  $z$ -transform of Spearman correlation coefficients between sequence  $\beta$  and trend strength rating,  $t_{(19)} = 5.5$ ,  $p = 3 \times 10^{-5}$ , Cohen's  $d = 1.23$ ). Following the analysis of Figure 2B, we first established that subjects exhibited a behavioral prediction effect by performing a  $6 (p_{34}) \times 3 (p_{34}^*)$  repeated-measures ANOVA on likelihood ratings for the final tone pitch. The  $p_{34} \times p_{34}^*$  interaction was highly significant ( $F_{(10,190)} = 20.4$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.52$ ), reflecting the fact that subjects' likelihood ratings for the final tone pitch were sensitive to the expected pitch (Fig. 8A). The interaction can be observed by noting that, when the presented final tone pitch was low (220 Hz), subjects rated the final tone to be more likely when its expected pitch was also low (paired  $t$  test for  $p_{34}^* = \text{low}$  vs  $p_{34}^* = \text{high}$ ,  $t_{(19)} = 3.7$ ,  $p < 0.002$ , difference = 0.95); in contrast, when the presented final tone pitch was high (880 Hz), subjects rated the final tone to be more likely when its expected pitch was also high ( $t_{(19)} = -7.9$ ,  $p = 2 \times 10^{-7}$ , difference =  $-1.66$ ). There was also a main effect of  $p_{34}$  ( $F_{(5,95)} = 10.1$ ,  $p = 9 \times 10^{-8}$ ,  $\eta_p^2 = 0.35$ ), reflecting the inverse U shape of the function.

Second, following the analysis of Figure 5B, we established that subjects exhibited a neural prediction effect by regressing non-baseline-corrected MEG activity during the penultimate tone onto expected final tone pitch. Replicating the key finding in the main experiment (Fig. 5B), we found a right-lateralized cluster of sensors exhibiting the prediction effect in the 200–250 ms and 250–300 ms time windows (Fig. 8B, red dots, cluster-corrected; 200–250 ms:  $p = 0.009$ , cluster size = 28,  $d_{\text{cluster}} = 4.5$ ; 250–300 ms:  $p = 0.004$ , cluster size = 34,  $d_{\text{cluster}} = 5.3$ ). In these time windows, there was also a significant left-lateralized prediction cluster (Fig. 8B, red dots, cluster-corrected; 200–250 ms:  $p < 0.001$ , cluster size = 42,  $d_{\text{cluster}} = 8.6$ ; 250–300 ms:  $p = 0.003$ , cluster size = 36,  $d_{\text{cluster}} = 5.5$ ). We found a similar left-lateralized group of sensors exhibiting a prediction effect in the same time windows in the main experiment (Fig. 5B), although this group of sensors did not pass the significance threshold established by cluster correction in the main experiment.

Third, following the analysis of Figure 6D, we established that the prediction effect is related to the linear history tracking effect. As in Figure 6, we partitioned the original MEG data,  $M^{\text{orig}}$ , into the component accounted for by history tracking,  $M^H$ , and the residual,  $M^{\text{res}}$ . We then compared the regression weights obtained by regressing  $M^{\text{orig}}$ ,  $M^H$ , and  $M^{\text{res}}$  onto  $p_{34}^*$  for the 200–250 ms and 250–300 ms time windows. The regression weights are plotted for  $M^{\text{orig}}$  against  $M^H$  (Fig. 8C, red) and for  $M^{\text{orig}}$  against  $M^{\text{res}}$  (Fig. 8C, black), separately for the right- and left-lateralized significant prediction clusters (top and bottom rows of Fig. 8C, respectively). For both clusters, there was a strong correlation between the regression weights for  $M^{\text{orig}}$  and  $M^H$  (right cluster, 200–250 ms:  $r = 0.81$ ,  $p < 3 \times 10^{-7}$ , regression coefficient = 0.63,  $n = 28$ ; right cluster, 250–300 ms:  $r = 0.93$ ,  $p < 3 \times 10^{-15}$ , regression coefficient = 0.92,  $n = 34$ ; left cluster, 200–250 ms:



**Figure 7.** Relationship between history integration weights and optimal prediction weights. **A**, Optimal prediction weights as a function of sequence  $\beta$ . For  $1/f^\beta$ -type sequences, the expected value of the upcoming item is a weighted sum of the previous items; these weights (“optimal prediction weights”) are plotted as solid lines for each  $\beta$  level. For  $\beta < 1$ , expected value depends on the  $a_{k+1}$  weights (Eq. 1). For  $\beta > 1$ , expected value depends on the weights  $[a_1, (a_2 - a_1), \dots, (a_{k+1} - a_k), -a_{k+1}]$  (Eq. 2). For  $\beta > 1$  sequences, dashed lines show optimal prediction weights for the corresponding fGn sequence with scaling parameter  $\beta - 2$ , which were used to derive the weights in Equation 2 and shown in solid lines. **B**, History integration weights resemble scaled optimal prediction weights. In the time window of 200–300 ms, the dependence of MEG activity on pitch history (solid lines) in the significant prediction cluster (inset left column, red sensors; reproduced from Fig. 5B, 200–250 ms window) reflected the pattern exhibited by the scaled optimal prediction weights averaged across sequence  $\beta$  (dashed lines). This resemblance accounts for why history-dependent MEG activity carries the prediction effect (Fig. 6; Eqs. 7 and 8). **C**, Correlation between history integration weights on current and previous pitches. To further explore the apparent opposing effects of the current and previous pitch on MEG activity illustrated in **B**, we computed across-sensor correlations in the prediction cluster between the weight on the current pitch,  $\beta_1$ , and weights on previous pitches,  $\beta_{k+1}$ ,  $k \geq 1$ . Here, we show an example scatterplot of the across-sensor relationship between  $\beta_1$  and  $\beta_2$  for one subject. **D**, Inverse correlation between current and previous pitch weight. Across subjects, the influence of the current and previous tone pitch on MEG activity was significantly inversely correlated for the 200–250 ms ( $t_{(10)} = -3.0, p = 0.01$ ) and 250–300 ms ( $t_{(10)} = -2.3, p = 0.04$ ) time windows, as assessed by a one-sample  $t$  test on the Fisher’s  $z$ -transform of the correlation coefficients. Asterisks denote  $p < 0.05$ . Error bars denote SEM across subjects.



**Figure 8.** Reproduction of main findings in an independent dataset. We collected MEG data from 20 additional subjects on a modified version of the task paradigm (for details, see Materials and Methods). **A**, Behavioral prediction effect (cf. Fig. 2B). A  $6(p_{34}) \times 3(p_{34}^*)$  repeated-measures ANOVA on likelihood ratings for the final tone pitch yielded a significant  $p_{34} \times p_{34}^*$  interaction ( $F_{(10,190)} = 20.4, p < 0.001$ ), indicating that subjects' likelihood ratings for the final tone were sensitive to the expected pitch value. **B**, Neural correlate of prediction (cf. Fig. 5B). As in the main experiment, a right-lateralized cluster of sensors (red dots) exhibited a significant linear relationship with expected pitch in the 200–250 ms and 250–300 ms time windows. A left-lateralized cluster also emerged as a significant neural correlate of prediction in these data (a similar pattern was observed in the main experiment, which did not pass statistical significance). **C**, Linking history tracking and prediction (cf. Fig. 6D). Across sensors in the right-lateralized prediction cluster (top row) and left-lateralized prediction cluster (bottom row), the original and history-dependent MEG activity exhibit similar sensitivity to  $p_{34}^*$  (red dots and lines). In comparison, the residual activity is much less sensitive to  $p_{34}^*$  (shown by the black dots being closer to the ordinate value of 0) and less similar to the original response. Dashed lines indicate unity. For statistics, see Results.

$r = 0.80, p < 3 \times 10^{-10}$ , regression coefficient = 0.75,  $n = 42$ ; left cluster, 250–300 ms:  $r = 0.85, p < 9 \times 10^{-11}$ , regression coefficient = 0.76,  $n = 36$ ). Therefore, this result reproduces the original finding in the main experiment (Fig. 6D) that the sensitivity of the MEG activity to the expected upcoming tone pitch is largely explained by its linear sensitivity to the preceding tone sequence. In addition, as in the main experiment, the relationship between the regression weights for  $M^{\text{orig}}$  and  $M^{\text{res}}$  was consider-

ably weaker though non-negligible (right cluster, 200–250 ms:  $r = 0.63, p < 4 \times 10^{-4}$ , regression coefficient = 0.37,  $n = 28$ ; right cluster, 250–300 ms:  $r = 0.21, p = 0.24$ , regression coefficient = 0.08,  $n = 34$ ; left cluster, 200–250 ms:  $r = 0.40, p = 0.009$ , regression coefficient = 0.25,  $n = 42$ ; left cluster, 250–300 ms:  $r = 0.45, p = 0.006$ , regression coefficient = 0.24,  $n = 36$ ).

In summary, the primary behavioral and neural findings pertaining to pitch prediction and its relationship to pitch history tracking (Figs. 2B, 5B, 6D) replicated in a second independent dataset (Fig. 8), demonstrating the robustness of these findings.

## Discussion

In summary, we used precisely constructed tone sequences with temporal statistical regularities representative of temporally varying stimuli in the natural environment to investigate predictive processing in the human brain. We first established that humans can form valid predictions about upcoming items in such sequences by extrapolating from sequence history. We then reveal the underlying neural mechanisms for such predictive processing by showing that: (1) these naturalistic stimuli consistently modulate slow, arrhythmic neural activity; (2) ongoing neural activity predicts the expected pitch of the upcoming tone; and (3) such predictive activity is constructed by neural integration of sequence history. These findings further our understanding about predictive computations in the human brain significantly and reveal an important role for slow, arrhythmic brain activity in the processing and prediction of naturalistic stimuli.

We observed that the phase of arrhythmic neural activity in the  $<10$  Hz range tracked pitch fluctuation in a tone sequence. This observation generalizes a previous finding showing that, when tone sequences with  $1/f^\beta$  pitch fluctuation are amplitude modulated at 40 Hz, the phase of a 40 Hz entrained neural oscillation tracks the evolving pitch sequence (Patel and Balaban, 2000). Here, we show that, in the absence of such artificial amplitude modulation, the phase of low-frequency arrhythmic neural activity tracks the evolving pitch sequence. In addition, we demonstrate that slow, ongoing neural activity that accumulates information over multiple tones plays a crucial role in integrating sequence history and predicting upcoming stimulus input. These findings complement previous research showing phase locking of low-frequency neuronal oscillations to rhythmic stimuli, which may align high-excitability phases to events within a stimulus stream and thereby enhance perceptual sensitivity (Schroeder and Lakatos, 2009). Such phase locking to periodic stimuli (a tone every 300 ms) was also evident in our data (Fig. 3Ciii). However, our results show that the processing of an evolving sequence of pitches over multiple tones was carried by slow, arrhythmic neural activity (Fig. 3Cii). This slow, arrhythmic neural activity was previously shown to have behavioral, developmental, and clinical significance (Monto et al., 2008; He et al., 2010; He, 2014; Voytek et al., 2015; Lin et al., 2016). Our results significantly extend this prior literature by demonstrating that it serves an important role in processing and predicting natural stimuli, consistent with an earlier hypothesis (He et al., 2010).

The current study advances our understanding of the neural mechanisms of prediction in two important directions. First, our results establish that human subjects can make valid predictions of upcoming stimulus features for naturalistic stimulus sequences by extrapolating from their temporal statistical regularities and reveal a neural correlate for such predictions. Second, we describe a mechanism for how the content of prediction is constructed based on neural processing of sequence history. Below we elaborate on these findings.

Previous studies on stimulus feature prediction based on statistical regularities have often used repeated presentations of sequences, items, or pairs of items. A notable exception is Garrido et al. (2013), who presented subjects with sequences of tones with pitches randomly drawn from Gaussian distributions and found that the width of the Gaussian distribution influenced the magnitude of mismatch negativity elicited by tones at the tail of the distribution. In the present study, we take this idea further by using tone sequences with temporal statistical regularities similar to those encountered in the natural environment. For such sequences, prediction is derived by extrapolating statistical regularities in the stimulus, not by repetition. Indeed, our task design ensured that prediction of the final tone pitch could not be derived from associative learning over repeated sequence presentations because the final tone pitch on each trial was randomly selected from the same set of six possible pitch values for every sequence. Instead, the statistical regularity inherent in sequence history  $p_1, \dots, p_{N-1}$  established an expectation for what values  $p_N$  would likely have, given that it should be a natural continuation of the overall trend of the preceding tones. To our knowledge, our findings are the first demonstration that human subjects can make valid predictions based on naturalistic temporal statistical regularities.

Previous studies using instructive cues and associative learning have revealed correlates of prediction in neural activity before and during the presentation of a predicted stimulus that is congruent with a categorical prediction (Erickson and Desimone, 1999; Schlack and Albright, 2007; Puri et al., 2009; Esterman and Yantis, 2010; Hanks et al., 2011; Kok et al., 2013; de Lange et al., 2013). Using a novel experimental design to probe predictions formed by extrapolating temporal statistical regularities, we discovered a novel neural correlate of prediction during continuous sensory processing that is formed based on evolving stimulus history and carries a continuous-valued predictive content. The neural correlate of prediction was found in non-baseline-corrected data, but not in baseline-corrected data. This suggests that prediction was carried by slow, ongoing neural activity that accumulates over the evolving sequence, rather than by fast, evoked responses to each tone. This result is consistent with a previous study showing that the magnitude of a slow EEG potential increases across consecutive trials in which a frequent “global standard” tone sequence was played as if in anticipation of the eventual presentation of the rare “global deviant” sequence (Chennu et al., 2013). Nonetheless, this previous finding may also reflect a general process of attentional readiness in anticipation of the eventual presentation of a task-relevant stimulus (Tecce, 1972). In the current experiment, the subject’s attentional state was equated across different expected values of final tone pitch, given the random order of sequence presentation and identical task demands, allowing us to identify a neural correlate of the content of prediction without the confound of attention.

An important further question is how the content of prediction is actually formed. For paradigms in which prediction is established by associative learning, the presentation of a stimulus may activate the neural representation of its associate via lateral projections in sensory regions formed during learning (Albright, 2012). For paradigms based on repeated presentations of items or sequences, adaptation and plasticity are likely at play (Yaron et al., 2012; Gavornik and Bear, 2014). For predictions established by an explicit instructive cue, the cue may activate the neural representation of its predictive content via a combination of top-down and bottom-up factors (Albright, 2012; Summerfield and

de Lange, 2014). However, prediction derived from extrapolation of sequence history, such as that investigated herein, likely requires a different set of mechanisms given its reliance on statistical trend instead of fixed repetitions or explicit instruction.

Here, we describe a mechanism for constructing the predicted feature of an upcoming stimulus in an evolving stimulus sequence. We show that neural activity processing a given tone in a sequence contains linear dependence on up to seven prior tones. We further demonstrate that this linear integration of sequence history in ongoing (i.e., accumulated, not-baseline-corrected) neural activity largely accounted for the neural substrate of prediction about upcoming stimulus. Although future investigation is needed to determine the extent to which nonlinear dependence on sequence history may facilitate predictive processing, our finding points to a potential general mechanism for sequence prediction that is remarkably simple: neural integration of stimulus history forms the prediction for the upcoming stimulus feature directly. Indeed, an interesting question for future investigation is the extent to which the mechanism uncovered herein generalizes to stimulus sequences with non-1/f-type temporal dependences.

To our knowledge, neither linear integration of pitch sequence history nor the linear effects of pitch prediction in neural activity were previously reported. Using MEG source analysis, an earlier study demonstrated a monotonic relationship between pitch and the amplitude and latency of tone-evoked responses in auditory cortex (Krumbholz et al., 2003), providing precedent for linear effects of pitch processing in MEG activity.

The opposing effects of the current and previous tone pitch on neural activity (Fig. 7B–D) may seem to be consistent with the predictive coding framework in which the integration of pitch sequence history (representing the expected pitch) opposes the neural response to the current pitch, thus generating a prediction error signal. However, it is important to note that, in our analyses linking history integration with prediction (Figs. 6, 7), the influence of the current and past pitches on neural activity do not represent a prediction error for the current pitch, but rather constitute a prediction for the upcoming pitch. In addition, the opposing effects of the current and previous tone pitch are best understood as arising from the mathematics of prediction in sequences exhibiting naturalistic temporal dependence (Fig. 7A), which produces the opposing pattern seen in neural history dependence (Fig. 7B).

A line of previous work revealed a hierarchy of temporal integration windows in neural activity (Gauthier et al., 2012; Honey et al., 2012; Chaudhuri et al., 2015), but it was unknown whether such neural integration of stimulus history plays a role in predictive computations (Hasson et al., 2015). Our results fill in this missing link by demonstrating that neural integration of stimulus history contains predictive information about upcoming stimulus features. Currently, it remains unclear whether the window of history integration uncovered herein is determined by absolute time (in seconds) or by the amount of information (in bits) (Hasson et al., 2015); future studies manipulating stimulus duration upon the current design could tease apart these intriguing possibilities.

The integration of stimulus sequence history in neural activity uncovered herein bears some resemblance to the recently described phenomenon of serial dependence in vision (Fischer and Whitney, 2014; Liberman et al., 2014; but see Fritsche et al., 2017) in which a misperception of the currently presented stimulus occurs due to a bias toward recent stimulus history. Studies on serial dependence in vision have so far used randomly varying

stimuli and it was postulated that misperception in such a paradigm may in fact be predictive/adaptive when faced with natural stimuli that contain temporal dependence (Burr and Cicchini, 2014; Liberman et al., 2014). Here, using stimulus sequences that capture the temporal dependence structure of natural stimuli, we show that history integration encoded in neural activity indeed forms the basis for predictive computation.

In conclusion, we demonstrate that human subjects can exploit temporal dependence in naturalistic stimuli to predict upcoming stimulus features. We further identify a neural mechanism underlying such prediction whereby the neural activity processing the current item in a sequence encodes the predicted feature of the upcoming item. This predictive component of neural activity is in turn constructed based on a largely linear integration of sequence history.

## Notes

Supplemental Material for this article is available at [https://med.nyu.edu/helab/Maniscalco\\_et\\_al\\_2018\\_SI](https://med.nyu.edu/helab/Maniscalco_et_al_2018_SI). In this analysis we address an aliasing effect found in the  $B = 1.01$  sequences discovered after data collection. We discuss the consequences of this effect and demonstrate that the main results reported in the manuscript remain if  $B = 1.01$  sequences are excluded from analyses. This material has not been peer reviewed.

## References

- Albright TD (2012) On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron* 74:227–245. [CrossRef Medline](#)
- Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. *J Neurosci* 30:2960–2966. [CrossRef Medline](#)
- Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617–629. [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](#)
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Burr D, Cicchini GM (2014) Vision: efficient adaptive coding. *Curr Biol* 24:R1096–R1098. [CrossRef Medline](#)
- Chalk M, Seitz AR, Seriès P (2010) Rapidly learned stimulus expectations alter perception of motion. *J Vis* 10:2. [CrossRef Medline](#)
- Chaudhuri R, Knoblauch K, Gariel MA, Kennedy H, Wang XJ (2015) A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* 88:419–431. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2011) Endogenous modulation of low frequency oscillations by temporal expectations. *J Neurophysiol* 106:2964–2972. [CrossRef Medline](#)
- De Coensel B, Botterdooren D, De Muer T (2003) 1/f noise in rural and urban soundscapes. *Acta Acustica* 89:287–295.
- de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci* 33:1400–1410. [CrossRef Medline](#)
- Dong DW, Atick JJ (1995) Statistics of natural time-varying images. *Network: Computation in Neural Systems* 6:345–358. [CrossRef](#)
- Eke A, Herman P, Kocsis L, Kozak LR (2002) Fractal characterization of complexity in temporal physiological signals. *Physiol Meas* 23:R1–R38. [CrossRef Medline](#)
- Erickson CA, Desimone R (1999) Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci* 19:10404–10416. [Medline](#)
- Esterman M, Yantis S (2010) Perceptual expectation evokes category-selective cortical activity. *Cereb Cortex* 20:1245–1253. [CrossRef Medline](#)
- Fischer J, Whitney D (2014) Serial dependence in visual perception. *Nat Neurosci* 17:738–743. [CrossRef Medline](#)
- Fritsche M, Mostert P, de Lange FP (2017) Opposite effects of recent history on perception and decision. *Curr Biol* 27:590–595. [CrossRef Medline](#)
- Garcia-Lazaro JA, Ahmed B, Schnupp JW (2006) Tuning to natural stimulus dynamics in primary auditory cortex. *Curr Biol* 16:264–271. [CrossRef Medline](#)
- Garrido MI, Sahani M, Dolan RJ (2013) Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS Comput Biol* 9:e1002999. [CrossRef Medline](#)
- Gauthier B, Eger E, Hesselmann G, Giraud AL, Kleinschmidt A (2012) Temporal tuning properties along the human ventral visual stream. *J Neurosci* 32:14433–14441. [CrossRef Medline](#)
- Gavornik JP, Bear MF (2014) Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci* 17:732–737. [CrossRef Medline](#)
- Hanks TD, Mazurek ME, Kiani R, Hopp E, Shadlen MN (2011) Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J Neurosci* 31:6339–6352. [CrossRef Medline](#)
- Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn Sci* 19:304–313. [CrossRef Medline](#)
- He BJ (2014) Scale-free brain activity: past, present, and future. *Trends Cogn Sci* 18:480–487. [CrossRef Medline](#)
- He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. *Neuron* 66:353–369. [CrossRef Medline](#)
- Helgason H, V P, Abry P (2011) Fast and exact synthesis of stationary multivariate gaussian time series using circulant embedding. *Signal Processing* 91:1123–1133. [CrossRef](#)
- Honey CJ, Theisen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76:423–434. [CrossRef Medline](#)
- Jepma M, Wagenmakers EJ, Nieuwenhuis S (2012) Temporal expectation and information processing: a model-based analysis. *Cognition* 122:426–441. [CrossRef Medline](#)
- Kok P, Brouwer GJ, van Gerven MA, de Lange FP (2013) Prior expectations bias sensory representations in visual cortex. *J Neurosci* 33:16275–16284. [CrossRef Medline](#)
- Krumholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B (2003) Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cereb Cortex* 13:765–772. [CrossRef Medline](#)
- Levitin DJ, Chordia P, Menon V (2012) Musical rhythm spectra from bach to joplin obey a 1/f power law. *Proc Natl Acad Sci U S A* 109:3716–3720. [CrossRef Medline](#)
- Liberman A, Fischer J, Whitney D (2014) Serial dependence in the perception of faces. *Curr Biol* 24:2569–2574. [CrossRef Medline](#)
- Lin A, Maniscalco B, He BJ (2016) Scale-free neural and physiological dynamics in naturalistic stimuli processing. *eNeuro* 3: pii: ENEURO.0191–16.2016. [CrossRef Medline](#)
- Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54:1001–1010. [CrossRef Medline](#)
- Mandelbrot BB, Van Ness JW (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10:422–437. [CrossRef](#)
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190. [CrossRef Medline](#)
- Meyer T, Olson CR (2011) Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci U S A* 108:19401–19406. [CrossRef Medline](#)
- Monto S, Palva S, Voipio J, Palva JM (2008) Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J Neurosci* 28:8268–8272. [CrossRef Medline](#)
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869. [CrossRef Medline](#)
- Patel AD, Balaban E (2000) Temporal patterns of human cortical activity reflect tone sequence structure. *Nature* 404:80–84. [CrossRef Medline](#)
- Puri AM, Wojciulik E, Ranganath C (2009) Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res* 1301:89–99. [CrossRef Medline](#)



- Robinson DW, Dadson RS (1956) A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics* 7:166–181. [CrossRef](#)
- Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928. [CrossRef](#) [Medline](#)
- Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci* 8:657–661. [Medline](#)
- Scharf LL (1991) *Statistical signal processing: detection, estimation, and time series analysis*. Reading, MA: Addison-Wesley.
- Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* 53:881–890. [CrossRef](#) [Medline](#)
- Schmuckler MA, Gilden DL (1993) Auditory perception of fractal contours. *J Exp Psychol Hum Percept Perform* 19:641–660. [CrossRef](#) [Medline](#)
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18. [CrossRef](#) [Medline](#)
- Stolk A, Todorovic A, Schoffelen JM, Oostenveld R (2013) Online and offline tools for head movement compensation in MEG. *Neuroimage* 68:39–48. [CrossRef](#) [Medline](#)
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756. [CrossRef](#) [Medline](#)
- Tecce JJ (1972) Contingent negative variation (CNV) and psychological processes in man. *Psychol Bull* 77:73–108. [CrossRef](#) [Medline](#)
- Tolhurst DJ, Tadmor Y, Chao T (1992) Amplitude spectra of natural images. *Ophthalmic and Physiological Optics* 12:229–232. [Medline](#)
- Voss RF, Clarke J (1975) '1/f noise' in music and speech. *Nature* 258:317–318. [CrossRef](#)
- Voytek B, Kramer MA, Case J, Lepage KQ, Tempesta ZR, Knight RT, Gazzaley A (2015) Age-related changes in 1/f neural electrophysiological noise. *J Neurosci* 35:13257–13265. [CrossRef](#) [Medline](#)
- Wyart V, Nobre AC, Summerfield C (2012) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A* 109:3593–3598. [CrossRef](#) [Medline](#)
- Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76:603–615. [CrossRef](#) [Medline](#)