Delta–Beta Coupled Oscillations Underlie Temporal Prediction Accuracy

Luc H. Arnal¹, Keith B. Doelling¹ and David Poeppel¹,²

¹Department of Psychology, New York University, New York, NY 10003, USA and ²NYUAD Institute, New York University Abu Dhabi, 129188 Abu Dhabi, UAE

The ability to generate temporal predictions is fundamental for adaptive behavior. Precise timing at the time-scale of seconds is critical, for instance to predict trajectories or to select relevant information. What mechanisms form the basis for such accurate timing? Recent evidence suggests that (1) temporal predictions adjust sensory selection by controlling neural oscillations in time and (2) the motor system plays an active role in inferring "when" events will happen. We hypothesized that oscillations in the delta and beta bands are instrumental in predicting the occurrence of auditory targets. Participants listened to brief rhythmic tone sequences and detected target delays while undergoing magnetoencephalography recording. Prior to target occurrence, we found that coupled delta (1–3 Hz) and beta (18–22 Hz) oscillations temporally align with upcoming targets and bias decisions towards correct responses, suggesting that delta-beta coupled oscillations underpin prediction accuracy. Subsequent to target occurrence, subjects update their decisions using the magnitude of the alpha-band (10–14 Hz) response as internal evidence of target timing. These data support a model in which the orchestration of oscillatory dynamics between sensory and motor systems is exploited to accurately select sensory information in time.

Keywords: auditory, motor, neuronal oscillations, sensorimotor, timing

Introduction

Accurately predicting when future events will occur facilitates sensory processing, speeds up behavior, and optimizes the allocation of attentional resources in time (Correa et al. 2005; Nobre et al. 2012). Predictive timing, by analogy to the notion of predictive coding (Knill and Richards 1996; Friston 2005) in the time domain, requires the construction of an internal model of observed temporal regularities to infer precisely the occurrence of future events. Because these regularities can happen at different time-scales (seconds, hours, days . . .), distinct neural systems and computations may be used to generate temporally adaptive predictions (Ivry and Schlerf 2008; Morillon et al. 2009). The seconds time-scale, in particular, is highly relevant for online human behavior, as a large number of phenomena pertaining to perception and action (i.e., speech, movement etc.) occur at this scale. Here we investigate how the brain extracts temporal regularities that emerge from the perception of isochronous beats at this time-scale.

The inclination to automatically synchronize our movements to external rhythms suggests that temporal regularities are particularly relevant to the motor system (Schubotz et al. 2000; Grahn and Brett 2007; Bengtsson et al. 2009; Teki et al. 2011). This system is arguably at the origin of the contingent negative variation (CNV), an anticipatory electrophysiological component that precedes relevant sensory or motor events (Walter et al. 1964; Pfurtsch et al. 2003; Praamstra et al. 2006; Cravo et al. 2011) and facilitates performance when task timing is predictable (Hillyard 1973; Niemi and Nätänen 1981; Stefanics et al. 2010). It is also involved in duration perception, in particular for durations shorter than 2 s (Ivry and Schlerf 2008; Morillon et al. 2009). Recent considerations have suggested that the motor system internally simulates movement synchronized with future events to anticipate their occurrence and facilitate their processing (Schubotz 2007; Tian and Poeppel 2010; Arnal 2012; Arnal and Giraud 2012). According to this idea, corollary discharges that are contingent on movement simulation propagate to sensory areas to align ongoing activity with predicted events. Here, we hypothesized that temporal predictions are instantiated through sensorimotor oscillatory interactions. Such a mechanism would arguably permit the extraction of temporal regularities in order to infer when an event should occur.

Cortical oscillations are usually seen as a means towards flexibly communicating between distant neuronal populations (Engel et al. 2001; Fries 2005). Recent findings also suggest that oscillations, which reflect fluctuations of neuronal excitability, can be temporally adjusted to optimize sensory selection (Schroeder and Lakatos 2009). Accordingly, expectations align the phase of delta oscillations in time, which in turn accelerates response timing (Lakatos et al. 2008; Stefanics et al. 2010). Furthermore, ongoing delta phase modulates the weighting of sensory events in the decision-making process during sequential information processing (Wyart et al. 2012; Cravo et al. 2013). Beta oscillations arguably play a complementary function during temporal expectations and the accumulation of sensory evidence (Donner et al. 2009; Saleh et al. 2011; Fujioka et al. 2012; de Lange et al. 2013).

In sum, sensory selection (stimulus detection or sensory weighting) can be passively or actively (through prediction) regulated through modulation of the prestimulus oscillatory state. However, whether they govern subjects’ accuracy remains unclear. That is, whether oscillations play an instrumental role in determining whether an event occurs at the expected time has not been elucidated.

This study aimed at determining the neurophysiological mechanisms underpinning (1) the generation of a temporal prediction and (2) the subsequent evaluation of this prediction. We used a delayed-target detection task in which subjects were required to detect whether the last tone of an isochronous sequence was delayed with regard to the beat (Fig. 1A). Our analysis design (Fig. 1B) allows us to distinguish how 2 crucial stages of the processing chain are implemented: (1) the predictive stage, which reflects the neural activity patterns used to predict "when" and (2) the decisional stage, in which subjects' decisions (i.e., subjective reports) are made (see Materials and methods and Fig. 1C).
The factor that was asked to judge whether the last tone (target) was delayed or not with regard to the beat. Participants were instructed to listen to a sequence of 4 or 5 tones and were explicitly asked to stay completely still until providing the response. The target occurred either at the expected time (Fig. 1A) or was delayed by either 75 ms (condition Δ75, 33.3% of the stimuli) or 150 ms (condition Δ150, 33.3% of the stimuli) with regard to the delay. The factor determined whether participant responded correct or not with regard to the delay. The factor Decision relies on participant’s subjective report, that is, whether the target was perceived as “normal” or “delayed”. (C) Sequential model for predictive processing and decision-making. We assumed that contrasting the different levels of the factors Accuracy and Decision permits to reveal 2 distinct processing stages, respectively: (1) the prediction stage and (2) the decision stage (see Materials and methods).

Material and Methods

Participants
Nineteen participants (11 women, age range 18–42) completed the experiment after providing written informed consent and received a compensation for their participation. All participants were right-handed, with normal hearing and no history of neurological disorders. The experimental protocol was approved by the New York University Institutional Review Board.

Procedures and Stimuli
The experiment consisted of a “delayed-target detection task” during which participants listened to an isochronous sequence of 4 or 5 tones (150 ms duration per tone, 5 ms cosine ramp onset and offset, presented at 60 dB SPL). We used 3 different stimulus onset asynchronies (SOA = 800, 1000, or 1200 ms), pseudo-randomized across trials. Participants were asked to detect whether the last tone (target) was delayed with regard to the beat of the sequence (see Fig. 1A). Before each trial, subjects were visually informed whether the sequence included 4 or 5 tones and were explicitly asked to stay completely still until providing the response. The target occurred either at the expected time (Fig. 1B condition Δ0, 33.3% of the stimuli), or was delayed by either 75 ms (condition Δ75, 33.3% of the stimuli) or 150 ms (condition Δ150, 33.3% of the stimuli) with regard to the beat. At the end of each trial, participants used their left hand to indicate whether the target occurred at the predicted time (i.e., on the beat of the sequence; response: “normal”) or not (response: “delayed”). These “listen” trials were followed by a “produce” trial during which participants were instructed to reproduce the tone sequence by pressing a button that elicited a tone, and to detect whether the last tone was delayed with regard to their last button press. However, because participants’ performance was at chance for this secondary task, we focus the analysis here on “listen” trials. Because no feedback was given, the subjects were unaware of their performance, which ensured that their attention was equally engaged in both tasks. Finger and motor response contingencies were randomized across subjects. Delays and SOA parameters were selected on the basis of a series of pilot experiments (run on a different set of participants) in which psychophysical delay detection thresholds were determined as a function of SOAs. This pretesting allowed us to select a range of delays ensuring that on average—and for all SOAs—participants were better than chance at detecting the delays. Behavioral performance was assessed for each delay and SOAs by measuring the proportion of correct responses. We used 3 different SOAs in order to induce temporal uncertainty between trials, which ensured that subjects were actively engaged across the whole experiment. While the SOA slightly affected subjects’ performance at detecting the target (see Results), we systematically tested its effect on each neural measure, which never reached significance. We also performed reaction time analyses on this dataset and decided, in the absence of informative findings related to response speed, to focus on participants’ performance only.

Experimental Design and Sequential Processing Model
We assumed that contrasting the different levels of the factors Accuracy (Correct vs. Incorrect) and Decision (response: “normal” vs. “delayed”) can be used to reveal 2 distinct processing stages, respectively: a predictive stage (1) and a decisional stage (2), (see Fig. 1C). We hypothesized that the associated physiological effects should provide evidence for the oscillatory mechanisms implemented to perform these critical processing steps. (1) Prediction stage: because predictive mechanisms are deployed before the occurrence of the target, we anticipated that contrasting correct and incorrect responses (Accuracy factor) should reflect the success versus failure of these mechanisms, respectively. This contrast—applied to prestimulus activity—should therefore reveal the prediction stage, that is, the neural mechanisms used to accurately predict “when”. (2) Decision stage: since this refers to the processing stage at which the participant makes a decision about the target, we assumed that contrasting the trials corresponding to subject responses (“normal” vs. “delayed”) should reveal the neural mechanisms that determine subjects’ decisions. Because the decision depends on the sensory evidence, we hypothesized that related effects should be visible after the target occurrence.

Our primary questions rest on a complicated experiment, whose data have the dimensions of space, pre- and poststimulus time, as well as frequency, and whose design is multifactorial in nature. We therefore spend some time describing the sequence of analyses and motivating this sequence in relation to our questions.

MEG Recordings and Data Processing
Neuramagnetic signals were measured using a 157-channel whole-head axial gradiometer system (KIT, Kanazawa, Japan). Five electromagnetic coils were attached to a participant’s head to monitor head position during MEG recording. The locations of the coils were determined with respect to 3 anatomical landmarks (nasion, left and right preauricular points) on the scalp using 3D digitizer software (Source Signal Imaging, Inc.) and digitizing hardware (Polhemus, Inc.). The coils were localized to the MEG sensors, at both the beginning and the end of the experiment. The MEG data were acquired with a sampling rate of 1000 Hz, filtered online between 1 and 200 Hz, with a notch filter at 60 Hz.

MEG recordings were noise-reduced off-line using the CALM algorithm (Adachi et al. 2001). Data analysis was performed using the Fieldtrip (http://fieldtrip.fcdonders.nl; Oostenveld et al. 2011) and
EEGlab (Delorme and Makeig 2004) packages, and additional programs developed in MATLAB (The MathWorks, Natick MA). Trials were visually inspected, and those with obvious artifacts were removed. An independent component analysis as implemented in FieldTrip was used to correct for eye blink, eye movement, and heartbeat-related artifacts. The activity of malfunctioning sensors (<2 per subject) was interpolated by computing the average of neighboring sensors.

Auditory and Motor Functional Localizers
In order to extract neural activity in specific anatomical regions, one usually considers the use of source reconstructions using inverse solutions. However, we decided not to pursue this strategy for several reasons: first, source localization of distinct frequency bands might possibly result in differential sensitivity and spatial selectivity between frequency bands, which might introduce unwanted confounds when considering cross-frequency interactions; second, we did not have access to single subjects' anatomical magnetic resonance images, which considerably reduces the accuracy of the method. As we aimed at testing a putative role of cross-frequency interactions, we therefore decided to use functional localizers at the topographical level rather than anatomical constraints in source space.

To detect neural activity from motor and auditory regions while ensuring functional selectivity, we chose to predefine auditory and motor sensor clusters of interest by using the following functional localizers. As we were primarily interested in early auditory cortical responses to target sounds, we selected channels using the magnitude of the M100 response to a 1000 Hz sinusoidal tone as an independent pretest functional localizer. We selected, for each participant, the 5 sensors with the largest M100 response amplitude in each pole of the 2 gradiometer-based contour maps reflecting the underlying dipoles in auditory cortex (20 channels per subject). Analyses of auditory neural responses are computed on this sensor selection, unless otherwise stated. Prior to the experiment, we also ran a self-paced left hand button press experiment, which allowed us to functionally identify a subset of 15 central sensors that maximally captured neural activity over motor areas and did not overlap with the auditory sensor selection. This allowed us to functionally determine that the negative effect revealed on central sensors in Fig. 3C likely reflected activity generated in motor areas.

**Time–Frequency Analysis on Auditory Sensors**
A time–frequency wavelet transform was then applied to each trial (1000 ms pre- to 1000 ms post-target, zero-padded) at each MEG sensor using a wavelet (m = 7) analysis (0.5 Hz resolution from 1 to 10 Hz; 1 Hz resolution from 10 to 50 Hz). This analysis resulted in an estimate of oscillatory power at each time sample and at each frequency between 1 and 50 Hz. (As no significant effect was observed above 25 Hz, we restrain time–frequency rendering in our figures to the 1–25 Hz range for clarity.) In order to isolate the potential contribution of the tone preceding the target and the motor response following the target, we focused our analyses on the 450 ms pre- to 500 ms post-target time-window (see Fig. 2A). Importantly, note that figures represent neural activity aligned to the physical onset of the target rather than to the expected time of target occurrence. While this latter option was systematically tested for each analysis performed, it did not change any of the results. Each trial was then normalized (z-score) at each peri-stimulus time bin using the average and standard deviation (SD) across trials and conditions. Note that because of the potential influence of the previous tone, we preferred this method rather than using the prestimulus baseline. However, results were unchanged when using a −500 to −100 ms prestimulus time-window as a baseline. As a consequence, data were distributed normally, which allowed us to use standard parametric tests (e.g., paired t-test, repeated-measures analysis of variance [ANOVA]) to assess the statistical significance of observed effects on our experimental factors. After correcting for multiple comparisons (see below), these tests allowed us to identify time- (300 to −100 ms) and frequency- (delta, 1–3 Hz; beta 18–22 Hz) windows of interest for further analyses (Fig. 2B).

To ensure that the observed effects were not driven by a single condition (SOA for instance), we systematically controlled for such potential bias by testing the main effects and interaction of SOA and delays, which never reached significance (see below). Note as well that since none of the comparisons computed on phase-locking between trials was significant, we only describe results related to the other measures. We also systematically tested for the effect of Delay and its interaction with Accuracy and Decision. We found significant differences between correct and incorrect responses (Accuracy effect) irrespective of whether the target was delayed or not. We found that the delay only significantly affected the Decision effect presented in Figure 4. Therefore, Figures 2B and 3(A–F) present the contrast between correct and incorrect responses, regardless of whether the target was delayed or not.

**Phase, Power and Cross-Frequency Coupling Analysis on Filtered ERs**
We then aimed to further explore the prestimulus accuracy effect at the topographical level (Fig. 3), by focusing on the phase and power of the frequency bands identified in Figure 2B. To do so, we first reduced the dimensionality of our data by band-pass filtering single-trial event related fields (ERs) in delta (1–3 Hz) and beta (18–22 Hz) frequency bands, using a zero-phase lag FIR filter, as implemented in the EEGLab toolbox. Oscillatory phase and power were calculated using the Hilbert transform of the filtered signal; power was defined as the squared absolute value and phase as the angular component of the Hilbert transform. Single-trial filtered ERs were then Z-scored using the same procedure as described above. Trials were grouped by accuracy outcome (correct vs. incorrect), which were compared using the following measures (and statistical tests): power and phase-locking

![Figure 2. Time–frequency analysis of pre- to poststimulus MEG activity and accuracy effects. (A) Mean oscillatory activity of responses to target tones in auditory sensors (selected individually, see Materials and methods). Time is expressed relative to target onset (black vertical line). (B) Accuracy effect (correct minus incorrect stimulus) on oscillatory activity. Effect is expressed in t-scores; contours indicate significance thresholds of P < 0.05, 0.01, and 0.001 for those clusters that survived the correction for multiple comparisons (see Materials and methods and Maris and Oostenveld 2007). Black dashed boxes highlight prestimulus (−300 to −100 ms) accuracy effects in delta (1–3 Hz) and beta (18–22 Hz) power, corresponding to a stronger increase in these 2 frequency bands for targets that are subsequently correctly perceived.](https://academic.oup.com/cercor/article-abstract/25/9/3077/2926125)
Accuracy effect on delta phase
(1–3 Hz, −300 to −100 ms)

Accuracy effect on beta power
(18–22 Hz, −300 to −100 ms)

Delta-beta coupling
(1–3/18–22 Hz, −300 to −100 ms)

In Figure 3, we demonstrate that the accuracy effects on delta (1–3 Hz) phase, beta (18–22 Hz) power, and delta-beta cross-frequency coupling are significant. We employ the circular-to-linear correlation to assess the coupling between delta and beta oscillations. We also apply a Watson-Williams test to determine whether the phase differences across different conditions are statistically significant.

(continued...)

To perform phase analyses at the sensor group level (Fig. 3B), we computed the circular average of the phase differences across a group of sensors. We then applied the Watson-Williams test to determine if the phase differences across different conditions are statistically significant.

where \( t_0 \) is the center of the time-window of interest, and \( f \) is the mean frequency of the frequency band of interest (delta 1–3 Hz). This correction allows us to group time points within the window of interest and ensures greater reliability of the phase estimate to which the Watson-Williams test is applied. To assess cross-frequency coupling at the topographical level (Fig. 3A), we calculated the circular-to-linear correlation between the phase of delta- and the power of beta-oscillations at each sensor in the time window of interest (−300 to −100 ms). The coupling time-course illustrated in Figure 3F was obtained by calculating the circular-to-linear correlation between delta phase-course (circular average across selected sensors) and the beta power (averaged across the same sensor selection). To compute the interaction between Delay and Accuracy factors (Fig. 4B), we first calculated the correlation between the delay (0, 75, or 150 ms) and the magnitude of neural
Finally, we calculated the corrected t-scores (reported as Fisher-transformed Pearson’s r-values) between correct and incorrect responses. One potential limitation inherent to our recording constraints is that high-pass filtering online at 1 Hz possibly impedes the detection of entrained oscillatory activity at the rate of the stimulus (ranging from 0.83 to 1.25 Hz). According to the notion that neural oscillations are preferentially tuned to a limited range of frequencies—consistent with other studies that use similar filtering procedures (Stefanics et al. 2010; Wyart et al. 2012)—we reasoned that filtering data between 1 and 3 Hz reliably isolates the intrinsic delta oscillation. This assumption is supported by 2 observations (see also Results): first, we observed consistent phase-locked prestimulus oscillatory activity in this delta (1–3 Hz) band for all tested SOA values. This shows that after several isochronous tones, the beat consistently entrains intrinsic delta-band oscillations across trials regardless of the stimulation frequency. Second, prestimulus accuracy effects on the phase of delta oscillations were reliable regardless of the SOA value. Altogether, that suggests that this procedure reliably captures intrinsic delta-band oscillations and allows assessing how their deployment in time impacts subject accuracy.

Correction for Multiple Comparisons
To assess the statistical difference between the experimental conditions while controlling for multiple comparisons, we performed nonparametric cluster analyses (Maris and Oostenveld 2007) for each of the aforementioned statistical tests. All the results presented are corrected for multiple comparisons using this method unless otherwise stated. This method is based on the comparison of the statistic of clusters of adjacent points (significant at P < 0.05) in the time-, frequency- or sensor-dimensional space and the same statistic computed on randomly permuted data. The nonparametric statistics were performed by repeating 1000 times the calculation of a permutation test where the experimental conditions are randomly intermixed within each subject. For each of these permutations, we extracted the maximum (absolute value) cluster-level statistic. Finally, we calculated the corrected P-values (reported as p corr) by comparing the values of the cluster-level statistics of our original data with the cluster-level statistics of all permutations.

Results

Behavioral Results
Subjects’ performance was above chance (76.18 ± 2.32% correct, P < 10^-3), and significantly varied as a function of the Delay (F2,18 = 12.26; P < 10^-3). Performance was higher for Δt75 (80.31 ± 2.56%, P < 10^-3) and Δt150 (85.6 ± 2.77%, P < 10^-3) than Δt3 condition (62.6 ± 4.60%, P = 0.01). Overall, subjects’ performance slightly decreased with increasing SOA (main effect of SOA (F2,18 = 7.94; P < 10^-3); SOA by performance correlation: Pearson’s r = −0.22; P < 10^-3), suggesting that temporal judgments were more accurate for the higher stimulation rates.

Prestimulus Oscillatory Activity Predicts Behavioral Accuracy
We first assessed the time-frequency profile of responses aligned to the target tones in auditory sensors. Figure 2A shows that target tones elicited a typical power increase in theta band (4–8 Hz) followed by a broadband suppression in alpha (9–14 Hz) and beta (15–25 Hz) bands (Fig. 2A). Conversely, the power in alpha and beta bands increased in the time-period preceding the occurrence of the target. In order to reveal the mechanisms underpinning the anticipation of the forthcoming sound, we sorted the trials as a function of subjects’ accuracy (Fig. 2B). By contrasting correct and incorrect trials, we observed that ~200 ms prior to stimulus appearance, the power in delta (1–3 Hz) and beta (18–22 Hz) frequency bands was significantly higher for correct trials. This effect did not differ between delays (prestimulus ~300 to ~100 ms Accuracy-by-Delay interaction: delta band: F2,90 = 0.23; P = 0.78; beta band: F2,90 = 2.18; P = 0.11), suggesting that the power in these frequency bands co-varied with the accuracy of temporal judgments regardless of the Delay. The contrast between correct and incorrect trials also revealed that correctly anticipated targets were associated with smaller theta band power poststimulus responses. Again, no difference was observed between delays (poststimulus 100–150 ms Accuracy-by-Delay interaction for theta (6–8 Hz) band: F2,90 = 0.21; P = 0.80).

As hypothesized, delta and beta oscillations are implicated when anticipating the occurrence of a sound. To clarify the oscillatory mechanisms underlying temporal predictions, we focused the following phase and power analyses on these specific frequency bands. Finally, to investigate the regional specificity of these findings, we broadened these analyses to all MEG channels.

Prestimulus Oscillatory Effects on Accuracy
To better understand the impact of entrained oscillations on the participants’ ability to make accurate temporal judgments, we compared the phase distribution, the power, and the

---

Figure 4. Time–frequency analysis of Delay and Decision effects. (A) Decision effect (delayed minus normal conditions) on mean oscillatory activity. Time is expressed relative to target onset (black vertical line). Effect is expressed in t-scores; contours indicate significance thresholds of P < 0.05, P < 0.01 and P < 0.001 for those time–frequency clusters that survived the correction for multiple comparisons (see Materials and methods). (B) Delay-by-Accuracy interaction effect on mean oscillatory activity. This map reveals the difference of the parametric effect of delays between correct and incorrect conditions (paired t-tests using individual slope values, see Materials and methods). The white dotted square highlights the time- (200–400 ms poststimulus) and frequency- (10–14 Hz) window of interest for subsequent plotting. (C) alpha-band activity extracted from the time–frequency window of interest (right dotted square in panel B) expressed in t-score for each delay (y-axis) separately between correct (blue) and incorrect (red) conditions. Error bars indicate SEM. Gray dotted ellipses show data-points grouped as a function of the Decision (normal vs. delayed).
phase-power coupling between correct and incorrect trials at the topographic level. Because we found a prestimulus Accuracy effect in auditory sensors (~300 to −100 ms relative to target appearance, Fig. 2B), we pragmatically confined the observation of the correct minus incorrect contrast at the topographic level during this time window.

We first tested the Accuracy effect on phase-angle distributions during the prestimulus period of interest. We contrasted the angle distributions of correct and incorrect conditions using the Watson–Williams test on single-trial ERPs filtered to the delta band (1–3 Hz). Figure 3A shows the effect of Accuracy at the topographic level, during the time-window of interest. This analysis revealed several clusters (mostly over anterior sensors overlapping with the functionally localized auditory and motor sensors) showing consistent angle distribution differences between correct and incorrect conditions. We then computed this difference across time on averaged sensors within the cluster that survived the correction for multiple comparisons. By averaging the angle at each time-point during this time-period (while correcting for the theoretical angle difference between these time-points, see Materials and methods), we found that angle distributions for correct and incorrect trials are in opposite phase ($P < 10^{-5}$, Fig. 3B). Therefore, both prestimulus power as well as instantaneous phase in the delta band influenced participants’ ability to make accurate temporal judgments. Again, these differences were consistent regardless of the Delay, suggesting that correct and incorrect responses are associated with opposite mean phase angles. The predictive entrainment of delta oscillations aligns to an optimal phase to improve perceptual judgments. When aligned in the opposite phase, the stimuli were more likely to be incorrectly categorized.

We then applied the Accuracy contrast on prestimulus (~300 to −100 ms) beta power at the topographic level (Fig. 3C). This revealed an interesting spatial pattern, showing opposite power effects between auditory and motor sensors. Prestimulus beta activity over lateral sensors (overlapping with audition functionally localized sensors, see Materials and methods) was positive in correct trials (Fig. 3D), whereas it was negative in incorrect trials. Interestingly, this pattern reversed for central channels, overlapping with functionally defined motor sensors. Again this effect was significant regardless of the Delay, suggesting that the ability to make accurate temporal judgments was also determined by the balance of beta activity between auditory and motor systems during this time window.

**Delta-Phase/Beta-Power Frequency Coupling**

We further tested for coordinated activity between the delta phase and beta power to provide a measure of nonlinear coupling at the topographic level. Figure 3E shows that, during the prestimulus period of interest, cross-frequency coupling between delta phase and beta power was maximal within a cluster of anterior left sensors overlapping with both auditory and motor sensors. Figure 3F represents the time course of this delta–beta coupling effect averaged across the cluster of sensors defined in Figure 3E, separately for the correct and incorrect conditions.

Interestingly, the delta–beta coupling profile across time appeared to be similar (while being shifted in time) between correct and incorrect conditions. We therefore tested the cross-correlation between these 2 temporal profiles and found that the correlation value was maximal when shifting the incorrect time course by 80 ms. Correcting for this time-shift considerably increased the correlation value between these 2 time courses (before time-shifting: $r^2 = 0.13, P < 10^{-4}$; after time-shifting: $r^2 = 0.51, P < 10^{-15}$). This suggests that the magnitude as well as the accurate temporal alignment of the coupling influences the behavioral outcome.

**Poststimulus Correlates of Sensory Evidence and Decision Processes**

We finally hypothesized that while Accuracy effects should occur during the prestimulus period and unveil the neural mechanisms underlying temporal predictions, decisional correlates should mainly affect poststimulus activity.

We therefore sought to determine how the decision about the presence/absence of the delay is reflected in the neural activity (Decision level, see Fig. 1C). In this analysis, we contrasted the time–frequency profiles corresponding to subjective “delayed” versus “normal” responses. This contrast revealed that targets reported as “delayed” induced a larger alpha/low-beta (10–14 Hz) poststimulus suppression than those reported as “normal” (Fig. 4A). Subsequent to target occurrence, subjects update their decisions in accordance with the magnitude of the alpha-band response. This is consistent with other results (Palva et al. 2005; Wyart and Tallon-Baudry 2009) showing that poststimulus alpha-band suppression reflects subjective reports, i.e., the “conscious” detection of delayed targets. Finally, in order to characterize the link between the neural signals observed at the decisional level, we tested how the parametric effect of the Delay interacted with subjects’ Accuracy (Fig. 4B). Interestingly, the Delay-by-Accuracy time–frequency interaction profile was very similar to the Decision effect profile in Figure 4A. This revealed that poststimulus alpha power was negatively correlated (alpha suppression increased) with the Delay for correct responses whereas it was positively correlated (alpha suppression decreased) with the Delay for incorrect responses (Fig. 4B, left and right panels). This suggests that alpha poststimulus suppression could be used as a proxy (i.e., as internal evidence) to make judgments about how much the target is delayed with regard to the beat.

**Discussion**

Our results support the notion that neural oscillations are exploited to determine whether an event occurs at the expected time. We demonstrate that this ubiquitous and important perceptual ability relies on the implementation of coupled delta–beta oscillatory mechanisms. Our experimental design further permits us to distinguish between the neural correlates of predictive and decisional stages. By comparing neural responses as a function of (1) the Accuracy of subjects’ reports, and (2) the actual report (Decision) of the subjects, we found that specific oscillatory mechanisms and functional networks determine how the brain (1) accurately anticipates when something will happen and (2) decides whether it happened at the predicted time.

We assumed that the contrast between correct and incorrect trials should reveal the predictive strategies used to accurately predict “when”. We found that the power of oscillations in the delta (1–3 Hz) frequency band just before the target
occurrence co-varied with the accuracy of temporal judgments, regardless of the Delay factor. Increased prestimulus delta power presumably corresponds to the CNV that reflects temporal expectancy (Praamstra et al. 2006; Stefanics et al. 2010; Cravo et al. 2011). Previous behavioral assessment showed that the phase of delta oscillations can modulate reaction times to stimulus detection (Lakatos et al. 2008; Stefanics et al. 2010) or bias sensory weighting (Wyart et al. 2012; Cravo et al. 2013). Here we demonstrate that the temporal dynamics of delta-oscillations constituted a crucial determinant of response accuracy. If the target occurred in the optimal delta phase, subjects were more accurate at making temporal judgments than when the target occurred in the opposite phase. This supports the conjecture that in order to accurately perceive incoming sensory information, delta oscillations not only need to be entrained to the beat, but also need to be in a specific phase when the stimulus occurs.

Beta-band activity is classically considered to be related to motor functions (Baker 2007) and more recently to top-down control (Engel and Fries 2010; Wang 2010; Arnal and Giraud 2012). During rhythmic perception, beta-band rebound adapts to the beat not only in the auditory, but also the motor system, even if attention is directed away from the auditory stimulus (Fujioka et al. 2009, 2012). Whether this mechanism is involved in sensory sampling remains unclear. We first observed that correct responses were associated with higher prestimulus beta power. This suggests that accurate sensory sampling requires beta power to be maximal at the occurrence of the target. This is consistent with the proposal that the beta power rebound predictively adapts to the beat so that beta activity is maximal at the occurrence of the predicted sound (Fujioka et al. 2012), therefore facilitating its processing. Interestingly, the prestimulus beta-band divergence between correct and incorrect responses observed on auditory sensors was reversed in motor sensors (see Fig. 3C). This beta power reversal—consistent with previous findings showing opposite beta-band dynamics in time between auditory and motor systems (Fujioka et al. 2012)—might suggest that beat perception is processed through sensorimotor loops that entrain beta power at the rate of the stimulation. Furthermore, the ability to accurately judge the target requires the prestimulus auditory/motor beta power “balance” (i.e., high beta power in auditory vs. low beta in motor sensors) to be temporally aligned with the occurrence of the stimulus. This particular role of beta oscillations in mediating top-down connections is also consistent with current formulations of predictive coding; predictions are thought to arise from pyramidal cells in deep cortical layers that accumulate evidence (prediction error) at higher frequencies ( Arnal et al. 2011; Bastos et al. 2012). This is further consistent with the neuroanatomical and physiological evidence that deeper cortical layers oscillate at beta frequencies and are the primary source of descending or top-down projections.

Building on previous observations from intra-cranial recordings in the primary motor cortex ( Saleh et al. 2011), we hypothesized that coupled delta–beta oscillations synchronize to the occurrence of behaviorally relevant events. We found that prestimulus beta power dynamics nest within the phase of low-frequency delta oscillations, suggesting that these signals might actually originate from a cross-frequency mechanism. Such a mechanism could optimize sensory sampling by aligning the excitability of ongoing auditory activity with upcoming targets, while temporally “rationing” the allocation of neural resources to relevant temporal windows ( Schroeder and Lakatos 2009; Saleh et al. 2011; Ng et al. 2012). Note that while it is possible that the increased coupling in correct trials results from larger amplitude delta oscillations, the precise timing of these mechanisms is crucial for delay detection accuracy. This supports the notion that coupled oscillatory mechanisms need to be deployed at the right time to efficiently sample sensory inputs. While previous work suggested that coupled delta–beta dynamics align with the upcoming event in the motor cortex ( Saleh et al. 2011), this is the first demonstration that coupled delta–beta oscillations are instrumental in rhythmic processing, that is, that subjects’ performance depends on whether these mechanisms are accurately implemented. These cross-frequency dynamics involved both auditory and motor sensor groups, suggesting that predictive timing mechanisms are controlled through sensorimotor synchronization with the beat.

Following the concept of active sensing in the auditory domain ( Schroeder et al. 2010), we previously proposed that the efferent (motor) signals that are generated when tapping in synchrony to a beat are also generated during beat perception, which possibly reflects some form of tapping “simulation” by the motor system ( Arnal 2012; Arnal and Giraud 2012). As a consequence, the motor system would periodically propagate efferent signals towards sensory regions to predictively synchronize ongoing oscillations with the occurrence of upcoming sounds. In the absence of source localization of these signals, our data do not allow us to directly test the anatomical origin of such signals. However, together with other recent evidence showing (1) that the motor system’s delta–beta oscillations are recruited during rhythmic perception ( Saleh et al. 2011; Fujioka et al. 2012), and (2) that motor efferent signals modulate activity in the auditory cortex though direct connections ( Nelson et al. 2013), these results support the notion that, sensorimotor interactions temporally arrange processing by controlling ongoing excitability of neuronal populations across time. This is consistent with predictive coding schemes in which time-structured delta waves modulate beta-mediated precision or excitability in attended sensory streams.

While substantially impacting subjects’ performance, this mechanism may be necessary but cannot be sufficient to explain how subjects distinguished whether a target was delayed or not. In order to perform this task, the brain needs to index how much time elapsed since the last event. We propose that the brain solves this problem by accumulating phase evidence about ongoing delta oscillations—which in turn provides a means for subjectively estimating elapsed time. We previously hypothesized that during beat perception, the sensorimotor system predicts the occurrence of future events by synchronizing “simulated” movement with external events ( Arnal 2012). Consistent with other findings showing that the motor system is involved in duration perception at the seconds time-scale ( Ivry and Schlerf 2008; Morillon et al. 2009), we speculate that evaluating short durations could be achieved through motor simulation. We propose that short durations (on the order of seconds) can be accurately evaluated by deriving the length of the trajectory of the movement simulated during this duration, that is by converting time elapsed into a (simulated) distance. In this context, phase-evidence accumulation could be used to keep track of the simulated distance, which in turn would provide an index of duration. Due to the large number of variables and conditions, our dataset does not
allow us to confidently assess this hypothesis. However, while this idea remains highly speculative, we believe that it offers a plausible and simple mechanism to explain how the brain accurately measures duration at the seconds time-scale.

Importantly, our hypothetical model also sits comfortably within current thinking about predictive coding in the brain (Friston 2005; Bastos et al. 2012). In this framework, oscillations play a dual role in optimizing both the precision and accuracy of sensory predictions. On one hand the amplitude of oscillations is associated with a modulatory role in the excitability of neuronal populations involved in temporal prediction. On the other hand, the phase is associated with the accuracy of time-referenced predictions. As a consequence, temporal predictions modulate both the amplitude and phase of ongoing oscillations, which determine the accuracy of sensory processing.

Finally, as anticipated, we found that neural correlates of the decisional stages were mainly observed after the occurrence of the target. We observed that poststimulus alpha-band suppression predicted participants’ subjective reports, that is, “delayed” versus “normal” responses. Interestingly, this effect further revealed that poststimulus alpha suppression parametrically followed the Delay in opposing directions for correct and incorrect responses. While accurate responses coincided with a positive relationship between alpha suppression level and the Delay, incorrect responses showed a reversed effect. This suggests that the brain exploits the level of poststimulus alpha suppression as internal evidence to determine how well the stimulus matched the prediction. Consistent with other results (Palva et al. 2005; Wyart and Tallon-Baudry 2009), poststimulus alpha suppression indexed subjective reports, that is, the “conscious” detection and categorization of targets.

To conclude, we found that the brain uses distinct oscillatory mechanisms in separate frequency bands to (1) accurately anticipate the occurrence of events and (2) make a decision about whether it occurred at the right time. We found that predictive strategies rely on the cross-frequency organization between delta and beta bands, which are possibly controlled through long-range interactions between motor and sensory systems. Importantly, we propose a dual oscillatory mechanism to generate predictions about when an auditory event is expected to occur and to index time online. In this model, delta and beta oscillations are controlled (possibly through synchronization of the motor system with the beat) to predictively allocate neural resources at the right time. Following the occurrence of the target, the brain is argued to use the level of stimulus-induced alpha suppression to update its decision. The presence of these mechanisms determines the participants’ ability to make accurate temporal judgments. While it may be rather hard to specify how much each of these predictive mechanisms depends on or control each other, these mechanisms are not mutually exclusive and can synergistically contribute to optimize the extraction of sensory information across time. Importantly, while delta and beta oscillations were incidentally observed during temporal processing in earlier studies, we demonstrated that their predictive deployment is actually critical for information sampling and determines participants’ behavioral accuracy. Taken together, these results illuminate how the brain predicts and evaluates temporal structure and support a model in which sensorimotor oscillatory dynamics could be exploited to accurately track time.

Notes
This work was supported by National Institutes of Health R01 DC05660 to D.P. L.H.A is supported by a postdoctoral research grant from the Fyssen Foundation. We would like to thank Jeffrey Walker for assistance with MEG acquisition. Conflict of Interest: None declared.

References


Tian X, Poeppel D. 2010. Mental imagery of speech and movement implicates the dynamics of internal forward models. Front Psychol. 1:166.


