



HAL
open science

Motor origin of temporal predictions s in auditory attention

Benjamin Morillon, Sylvain Baillet

► **To cite this version:**

Benjamin Morillon, Sylvain Baillet. Motor origin of temporal predictions s in auditory attention. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114 (42), pp.E8913-E8921. 10.1073/pnas.1705373114 . hal-03583584

HAL Id: hal-03583584

<https://hal.science/hal-03583584>

Submitted on 21 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Motor origin of temporal predictions in auditory attention

Benjamin Morillon^{1, 2} and Sylvain Baillet²

¹Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France

²McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada

Keywords: auditory perception; sensorimotor; magnetoencephalography; psychophysics; rhythm; oscillation; beta frequency; and active sensing

Author contributions: B.M. and S.B. designed the experiment. B.M. collected and analysed the data. B.M. and S.B. wrote the paper.

Corresponding author: Dr. Benjamin Morillon (bnmorillon@gmail.com)

Abstract

In behavior, action and perception are inherently interdependent. Yet, the actual mechanistic contributions of the motor system to sensory processing are unknown. We present neurophysiological evidence that the motor system is involved in predictive timing – a brain function that aligns temporal fluctuations of attention with the timing of events in a task-relevant stream, thus facilitating sensory selection and optimizing behavior. In a magnetoencephalography experiment involving auditory temporal attention, participants had to disentangle two streams of sound on the unique basis of endogenous temporal cues. We show that temporal predictions are encoded by interdependent delta and beta neural oscillations originating from the left sensorimotor cortex, and directed toward auditory regions. We also found that overt rhythmic movements improved the quality of temporal predictions and sharpened the temporal selection of relevant auditory information. This latter behavioral and functional benefit was associated with increased signaling of temporal predictions in right-lateralized fronto-parietal associative regions. In sum, this study points at a covert form of auditory active sensing. Our results emphasize the key role of motor brain areas in providing contextual temporal information to sensory regions, driving perceptual and behavioral selection.

Significance Statement

How the motor system participates to auditory perception is unknown. In a magnetoencephalography experiment involving auditory temporal attention, we show that the left sensorimotor cortex encodes temporal predictions, which drive the precise temporal anticipation of forthcoming sensory inputs. This encoding is associated with bursts of beta (18-24 Hz) neural oscillations that are directed toward auditory regions. Our data also show that the production of overt movements improves the quality of temporal predictions and augments auditory task performance. These behavioral changes are associated with increased signaling of temporal predictions in right-lateralized fronto-parietal associative regions. This study points at a covert form of auditory active sensing, and emphasizes the fundamental role of motor brain areas and actual motor behavior in sensory processing.

Introduction

Active sensing, the act of gathering perceptual information using motor sampling routines, is well described in all but the auditory sense (1-4). It refers to a dual mechanism of the motor system, which both controls the orienting of sensing organs (ocular saccades, whisking, sniffing, etc.) and provides contextual information to optimize the parsing of sensory signals, via top-down corollary discharges (5). In audition, bottom-up sensory processing is generally disconnected from movements. Moreover, motor acts do not cause inflow of auditory input, except when movements produce sounds (e.g., hand claps, musical instruments). However, top-down motor influences over auditory cortices have been well characterized during self-generated sounds (6, 7). Strikingly, auditory and motor systems strongly interact during speech or music perception (8-11). We recently proposed a theoretical framework, whereby a *covert* form of active sensing exists in the auditory domain, through which oscillatory influences from motor cortex modulate activity in auditory regions during passive perception (1, 12). This idea implies that the chief information provided by the motor system to sensory processing is contextual temporal information.

Relevantly, the motor system is the core amodal network supporting timing and time perception (13, 14). It is also involved in beat and rhythm processing during passive listening (9, 15-18). Further, beta (~20 Hz) activity, the default oscillatory mode of the motor system (19, 20), is also related to the representation of temporal information (21-26). Time estimation could thus rely on the neural recycling of action circuits (27) and be implemented by internal, non-conscious simulation of movements in most ecological situations (28, 29). Following these conceptual lines, we may predict that efferent motor signaling would convey temporal prediction information, with behavioral benefits in anticipating future sensory events and the facilitation of their processing.

We recently developed a mechanistic behavioral account of auditory active sensing and showed that during temporal attending, overt rhythmic movements engage a top-down modulation that sharpens the temporal selection of auditory information (30). Crucially, we found that the participants were able to disentangle two streams of sound on the unique basis of temporal cues. In the present study, we recorded Magnetoencephalography (MEG) while participants performed such paradigm to dissociate stimulus- from internally-driven neural dynamics, and quantified the precision of temporal attention during auditory perception. We addressed three complementary questions: How are temporal predictions encoded? How do temporal predictions modulate auditory processing? Why do overt rhythmic movements improve the temporal precision of auditory attention?

Results

Behavioral effects of temporal attention during passive listening

We asked participants to categorize sequences of pure tones as higher- or lower-pitched, on average, than a reference frequency (see Methods and (30)). In order to drive rhythmic fluctuations in attention, we presented the tones (targets) in phase with a reference beat, and in anti-phase with irrelevant, yet physically indistinguishable tones (distractors). Each trial started with the rhythmic presentation of 4 *reference* tones indicating both the reference frequency and beat (1.5 Hz), followed by a melody unfolding at 3 Hz (lasting ~5.5 sec), corresponding to the alternation of 8 targets and 8 distractors of variable frequencies, respectively presented in quasi phase and anti-phase with the reference beat (Fig. 1A). This interleaved delivery of sensory events forced participants to use the reference beat to discriminate between targets and distractors – *i.e.* to maximize the integration of relevant sensory cues (targets) while minimizing the interference from irrelevant ones (distractors). This protocol (30) ensured that their attentional focus was temporally modulated over an extended time period. Thus, during the *3-Hz melody* (our period-of-interest), while sensory events were delivered at a 3-Hz rate, temporal attention fluctuated at 1.5 Hz, effectively dissociating stimulus- and internally-driven temporal dynamics (Fig. 1B). The task consisted of two conditions: in the *listen* condition, participants performed the task while staying completely still during the duration of the trial; in the *tracking* condition, they performed the task while touching rhythmically a pad with their left index finger, in phase with the reference beat.

The task in the *listen* condition was relatively difficult, with an averaged categorization performance of 73 %. We quantified the relative contribution of each tone to the decision (see Methods). We observed that the first target had a considerable influence in the decision, due to its leading position at the beginning of each melody sequence (Fig. 1C). To properly investigate the effects of temporal attention, we thus excluded this particular event – and for the same reason, the last distractor – from all subsequent analyses. Participants were able to assign greater weight to targets than distractors in their decision (repeated-measures ANOVA of 2 stimulus-categories x 7 elements; main effect of stimulus-category: $F_{1,18} = 7.1$, $p = .016$; see Fig. 1C-D). Moreover, as predicted by the logistic regression model, the influence a given target wielded on the subsequent decision scaled parametrically with its absolute distance from the reference frequency. This effect (*i.e.*, the slope of the function) was significantly reduced for distractors (paired t-test: $t_{18} = 5.2$, $p < .001$; Fig. 1E). In other words,

participants were able to sustain and orient their temporal attention at a 1.5-Hz rate to segregate targets from distractors and maximize the integration of targets.

Neural correlates of the 3-Hz melody

The 3-Hz melody corresponded to a stream of tones presented in quasi-rhythm fashion (Fig. 1A). We estimated the neural response to this auditory stream by investigating those regions whose oscillatory activity was synchronous to the melody. We relied on two complementary measures: the phase-phase (PLV) and phase-amplitude (PAC) coupling between MEG source signals from correct trials, and a signal template of the 3-Hz melody (see Methods). The dynamics of the 3-Hz melody principally correlated with neural oscillatory activity in bilateral auditory regions, both at 3 Hz (phase coupling) and in the 5-18 Hz frequency-range (amplitude coupling) (one-tail t-tests against zero, $p < .05$, FDR corrected; Fig. 2A-C top).

Core substrate of covert temporal predictions

While listening to the 3-Hz melody, participants modulated their attention temporally by relying on covert temporal predictions produced by an internal model approximating the 1.5-Hz beat cued from the four reference tones (Fig. 1A-B). We estimated the neural correlates of temporal predictions by examining where neural oscillatory activity fluctuated at the beat frequency (*i.e.*, 1.5 Hz) during presentation of the 3-Hz melody. Thus, we computed PLV and PAC between MEG source signals from correct trials, and a signal template of the 1.5-Hz reference beat of interest (see Methods).

Whole-brain statistical maps of phase-phase coupling revealed that a network of sensorimotor regions – bilateral, over the pre- and post-central areas – tracked internal temporal predictions (one-tail t-tests against zero, $p < .05$, FDR corrected; Fig. 2A bottom). Moreover, phase-amplitude coupling measures revealed the presence of beta bursting (18-24 Hz) at the delta rate in the left sensorimotor cortex (one-tail t-tests against zero, $p < .05$, FDR corrected; Fig. 2B bottom).

Response profile of the right associative auditory (rAA) and left sensorimotor (ISM) cortex during passive listening

To further investigate these results, we analyzed the response profiles of two key regions: the right associative auditory (rAA) and the left sensorimotor (ISM) cortex, where the respective phase-amplitude coupling with the 3-Hz melody and the 1.5-Hz reference beat was found to be maximal. We first confirmed that phase-amplitude coupling with the 3-Hz melody was significant in rAA (5-30 Hz; t-tests against zero, $p < .05$, FDR corrected), but not

ISM (all $p > .2$, FDR corrected), and peaked in the alpha range (7-14 Hz; Fig. 2C-D top). Second, the 1.5-Hz beat was indeed phase-amplitude coupled to beta oscillations in ISM (17-27 Hz; $p < .05$, FDR corrected), but not rAA (all $p > .2$, FDR corrected; Fig. 2C-D bottom). Additional coupling with the 1.5-Hz beat was also observed in the alpha range in both rAA and ISM (at 7-10 Hz and 8-13 Hz, respectively; $p < .05$, FDR corrected; Fig. 2C-D bottom), but these results did not reach significance at the whole-brain level of analysis.

The evoked response to tones was drastically larger in rAA than ISM and, importantly, significantly different between targets and distractors in rAA only (at 30-53 ms, 106-141 ms, 176-195 ms, and 382-396 ms; paired t-tests, $p < .05$, FDR corrected; in ISM all $p > .5$, FDR corrected; Fig. S1A), highlighting the impact of temporal attention on auditory processing. Of note, while the M50, M200 and M300 peaks were larger for targets than distractors, the opposite was true for the M100 peak. In contrast, the time-frequency response to tones was significantly different between targets and distractors in ISM only ($p < .05$, FDR corrected; Fig. S1B). This analysis revealed where 1.5-Hz oscillatory activity originated from, and confirmed the phase-amplitude coupling results (Fig. 2D bottom). The 1.5-Hz beat was not associated with evoked activity in ISM (all $p > .5$, FDR corrected; Fig. S1C bottom), confirming that the amplitude modulations at 1.5 Hz observed in the (alpha and) beta range did not correlate with unexpected motor acts during the *listen* condition.

Altogether these results indicate that covert temporal predictions at 1.5 Hz are represented in the phase of delta oscillations in a distributed sensorimotor network, with phase-coupled amplitude modulations of beta oscillations from the left lateralized sensorimotor cortex (Fig. 2A-B bottom). We emphasize that this neural substrate is distinct from melody-driven neural dynamics, both in terms of oscillatory frequencies and regions involved (Fig. 2), and could not be explained by the presence of evoked motor components (Fig. S1).

Frequency-specific directional interactions between rAA and ISM during passive listening

In our paradigm, two complementary dynamics were occurring simultaneously: the melody-driven bottom-up response and the top-down internal temporal predictions, anticipated to operate at 3 Hz and 1.5 Hz, respectively. We investigated the directionality of functional connectivity between two regions-of-interest (rAA and ISM; see Methods), across the frequency spectrum (1-50 Hz). We found bottom-up directed connectivity at 3 Hz (*i.e.* from rAA to ISM), and top-down directed connectivity in the beta range (18-24 Hz; *i.e.* from ISM to rAA; $p < .05$, FDR corrected; Fig. 3). The results pointed at the role of beta band oscillations in mediating temporal predictions. Of note, using either the left or right AA cortex as the sensory region-of-interest led to similar results (Fig. S2).

Behavioral advantage of overt motor tracking

In sum, we described the role of the left sensorimotor (ISM) cortex in representing and propagating temporal predictions (via beta-band oscillations), which were mandatory to modulate auditory processing in the present study. We further investigated whether overt rhythmic motor activity improved the segmentation of auditory information. At the behavioral level, the comparison between *listen* and *tracking* confirmed our previous finding (29) of a significant increase in categorization performance during overt motor tracking (one-tail paired t-test: $t_{18} = 2.0$, $p = .033$; Fig. 1C). We confirmed that during *tracking*, participants were able to assign greater weight to targets than distractors in their decision (repeated-measures ANOVA of 2 stimulus-categories x 7 elements; main effect of stimulus-category: $F_{1,18} = 21.0$, $p < .001$; paired t-test between targets and distractors slopes of element-weight profiles as a function of tone frequency: $t_{18} = 3.3$, $p = .004$; Fig. S3). Finally, we observed a significant difference across element weights between *listen* and *tracking* (repeated-measures ANOVA of 2 stimulus-category x 2 conditions; interaction: $F_{1,18} = 4.6$, $p = .047$; Fig. 4B): the contribution of targets was larger during *tracking* (post-hoc paired t-test: $t_{18} = 3.5$, $p = .003$), whereas the contribution of distractors was not significantly different across conditions ($t_{18} = 0.2$, $p = .88$). These results show that overt motor tracking further enhances the perceptual sensitivity to target auditory events.

Neural correlates of overt temporal predictions

During *tracking*, the dynamics of the 3-Hz melody principally correlated with neural 3-Hz oscillatory activity in bilateral auditory regions (phase coupling), and in a broad low-frequency range (5-20 Hz; amplitude coupling; Fig. 5A-B top). In addition, and in contrast to the *listen* condition, we found PAC between the 3-Hz melody and the right sensorimotor cortex (rSM) in the beta (14-24 Hz) range (one-tail t-tests against zero, $p < .05$, FDR corrected; Fig. 5C top). Temporal predictions, which were overtly conveyed during *tracking* (through left index movements), were unsurprisingly maximally represented over the right sensorimotor cortex (i.e. contralateral to finger movements), both in 1.5 Hz phase (Fig. 5A bottom) and broadband (4-50 Hz) amplitude modulations, with a main and secondary spectral peak in the beta (~18-24 Hz) and theta (~4-6 Hz) ranges, respectively (one-tail t-tests against zero, $p < .05$, FDR corrected; Fig. 5B-C bottom). Importantly, coupling was also observed in ISM at 1.5 Hz (phase coupling) and within the beta band (amplitude coupling), but not in the theta band (amplitude coupling; Fig. 5 bottom). We then detailed the profile of phase-amplitude coupling with the 3-Hz melody or the 1.5-Hz beat in rAA, ISM and rSM, and qualitatively confirmed the above-mentioned results (Fig. S4A-B). In rAA and ISM, the

evoked and time-frequency response profiles were qualitatively similar than during the *listen* condition (Fig. S4C-E). In rSM (see Fig. S4 bottom), we observed a strong evoked response to the 1.5-Hz beat that slightly anticipated beat onset (-87-14 ms; also reoccurring after 1.5 Hz (*i.e.* 667 ms); $p < .05$, FDR corrected). This pattern was also visible in the evoked response to tones (with a peak slightly prior to target onsets and following the distractors onset by ~ 333 ms). Finally, the time-frequency response profile in rSM confirmed the presence of broadband amplitude modulation at 1.5 Hz (with peaks in the theta and beta ranges).

Frequency-specific directional interactions during overt motor tracking

We then investigated the directionality of functional connectivity between the three regions-of-interest (rAA, ISM, and rSM). We first observed that directed connectivity from rAA to ISM was found only at 3 Hz ($p < .05$, FDR corrected; Fig. 6A). Second, directed connectivity from rSM was directed toward both rAA (1-8 Hz) and ISM (1-4 Hz; $p < .05$, FDR corrected) in the low-frequency range. Finally, we also observed directed connectivity from ISM toward both rAA and rSM in the alpha (8-12 Hz) and beta (18-24 Hz) ranges (and additionally at 1.5 Hz toward rAA; $p < .05$, FDR corrected). Importantly, ISM was neither acoustically stimulated (contrary to rAA; Fig. S4C), nor associated with overt movements (contrary to rSM; Fig. S4E). Nevertheless this region was a source of directed functional connectivity during *tracking* (Fig. 6B).

Taken together, our results from both the *listen* and *tracking* conditions revealed a key role for the left sensorimotor cortex (ISM) in representing and directing toward sensory regions temporal predictions through beta bursting (18-24 Hz).

Extended influence of temporal predictions during overt motor tracking

To further comprehend how temporal predictions modulate sensory processing and drive performance accuracy, we contrasted the neural correlates of temporal predictions, estimated at the 1.5-Hz phase, in *tracking* vs. *listen* correct trials. The data confirmed the qualitative difference in neural encoding of temporal predictions between *listen* and *tracking*. Specifically, we showed that overt motor tracking enhanced phase coupling with the 1.5-Hz beat in a broad set of regions, encompassing the bilateral sensorimotor cortex and right-lateralized temporal, parietal and frontal regions (paired t-tests, $p < .05$, FDR corrected; Fig. 7A). No significant enhancement was observed in *listen* relative to *tracking* ($p > .5$, FDR corrected).

Motor origin of temporal predictions

We further investigated possible accuracy effects by contrasting the neural correlates of temporal predictions in correct vs. incorrect trials, for both conditions. During *tracking*, the phase of 1.5-Hz oscillations in right sensorimotor, inferior parietal and lateral frontal regions, was significantly related to the behavioral outcome, with stronger phase coupling in correct than incorrect trials (paired t-tests, $p < .05$, FDR corrected; Fig. 7B). Importantly, this result could not be explained by a difference in the precision of motor tracking events (*i.e.* the timing of individual motor acts) between correct and incorrect trials (paired t-test, $t_{18} = 0.3$, $p = .29$; Fig. S5). In the *listen* condition, no significant difference between correct and incorrect trials was observed at the whole-brain level of analysis (paired t-tests, $p > .5$, FDR corrected). We note that without correcting for multiple comparisons, the same analysis revealed a single cluster located around ISM (paired t-tests, $p < .05$ uncorrected). Finally, no significant difference between correct and incorrect trials was observed at the whole-brain level of analysis, in the *tracking* or *listen* conditions, for phase-amplitude coupling estimates of temporal predictions (paired t-tests, $p > .05$, FDR corrected).

Together these results therefore reveal that in right sensorimotor, inferior parietal and lateral frontal regions, the phase of neural oscillations at the beat frequency (1.5 Hz) is significant to behavioral outcome: phase coupling was stronger in the condition with the highest categorization score (*i.e.* *tracking*; Fig. 4), and in correct trials of this latter condition (Fig. 7B).

Discussion

The present study aimed at identifying the neural substrates of temporal predictions and the mechanisms by which they impact auditory perception. We hypothesized that temporal prediction information is encoded in sensorimotor cortex oscillatory activity, and that overt movements enhance their representation and further their influence. We designed a paradigm involving endogenous temporal attention, *i.e.*, attention guided by internal predictions (31), and importantly, in which temporal attention and stimulus dynamics were operating at distinct rates (1.5 vs. 3 Hz, respectively, Fig. 1). Schematically, our paradigm involved the maintenance of temporal predictions and their deployment through temporal attention to modulate sensory processing.

Behaviorally, our results show that participants were able to sustain and orient their temporal attention to segregate targets from distractors in both conditions (Fig. 1C-E & S3). At the cortical level, we observed that temporal attention modulated the processing of auditory information down to the auditory cortex (Fig. S1A). This indicates that endogenous temporal attention impacts on stimulus-driven event patterns, tuning sensory processing to its own internal temporal dynamics to drive perceptual and behavioral selection (32).

To perform this paradigm, participants had to capitalize on temporal predictions (*i.e.* an internal model of the 1.5-Hz beat) during presentation of the 3-Hz melody (Fig. 1A-B). While we observed differences between the *listen* and *tracking* conditions, especially in terms of oscillatory frequencies and regions involved (Fig. 2 and 5), a condition-invariant representation emerged as a candidate to encapsulate the core neural substrate of temporal predictions. Indeed, we first observed in both conditions a low-frequency phase reorganization combined with the presence of beta bursting at the delta rate -- a result compatible with previous findings (21, 28, 33, 34). Second, these delta-beta coupled oscillations were present in the left sensorimotor cortex (ISM). Park and colleagues recently identified that this region was playing a pivotal role in the top-down modulation of auditory processing (35, 36). Finally, the functional connectivity conveyed by beta oscillations directed toward the right auditory cortex (rAA), probably to modulate perceptual and behavioral selection (Fig. 3 and 6). It is striking that even during *tracking*, where beta bursting was dominant in the right sensorimotor cortex (rSM; due to left index movements; Fig. 5C), the connectivity of beta oscillations was still directed from ISM, toward both rAA and rSM. Altogether and to the best of our knowledge, these results are first to bridge between recent electrophysiological studies highlighting the role of beta oscillations in representing temporal information (21-26), with studies that emphasized the role of the motor system in timing (9, 13-18, 37). They also fuel the idea that left-lateralized motor structures are more preferentially engaged than their

right counterparts in predicting future events, possibly to optimize information processing (37). While our results underline the left sensorimotor cortex, it is likely that subcortical structures and in particular the striatum may drive the effects we observed at the cortical level with MEG (16-18, 38-42).

Multiple functional MRI studies have reported the involvement of the left inferior parietal cortex in temporal attention (43-46). While our analyses highlight regions implicated in the encoding of temporal predictions and sensory dynamics, we have not investigated the neural bases of temporal attention per se. Rather, we suggest that the left inferior parietal cortex is at the interface between internal predictions and sensory processing, *i.e.* may be driven by the motor system to modulate auditory processing.

The role of beta oscillations in encoding temporal predictions can be interpreted in two complementary manners. First, temporal predictions might be encoded in beta bursting because the motor system is (among other things) dedicated to the encoding of temporal predictions and beta (~20 Hz) activity is the default oscillatory mode of the motor system (19, 20). This would represent an *incidental* hypothesis. Second, we posit that top-down predictions are ubiquitously encoded through beta oscillations, while bottom-up information tends to be reflected in higher-frequency (gamma) oscillations: this perspective has gained substantial attention in the recent years (47-50). This would represent an *essential* hypothesis. In this framework, beta oscillations could mediate the endogenous reactivation of cortical content-specific representations in the service of current task demands (51). Our results also pinpoint at the role of alpha oscillations in both bottom-up and top-down processing (Fig. 2, 5 and 6). However, these alpha-related findings were not consistent across conditions, making it difficult to interpret.

Our behavioral findings also confirm that overt rhythmic movements improve performance accuracy. The additional act of tapping, instead of interfering with the auditory task, favors it. This indicates that motor tracking and temporal attention synergistically interact (29). Further, this beneficial strategy does not capitalize on any additional external information (e.g. light flashes presented at the reference beat), but is purely internally driven. This suggests that embodied representations of rhythms are more stable than internal, cognitive alternatives. This behavioral benefit was driven by an increased modulation of sensory evidence encoding by temporal attention, *i.e.* an increased differential weighting between targets and distractors (Fig. 1) (30). However in our study, motor tracking only enhanced sensitivity to target tones, while it was found it also suppressed sensitivity to distractors in a previous study (30). Such discrepancy is probably due to this study's setting and instructions, which resulted in enhanced attention to targets (see Methods). Overall, our behavioral results confirm that overt tracking sharpens the temporal selection of auditory information.

Motor origin of temporal predictions

Such behavioral advantage has broad consequences. For example, it could directly be exploited in language rehabilitation to help patients suffering from auditory or language deficits understand speech better. Using an overt motor strategy (in the form of rhythmic movements) may indeed help focusing on the temporal modulations of a speech signal of interest, and therefore facilitate speech perception and segmentation, especially in adverse listening conditions (52-54).

This behavioral advantage is mirrored at the cortical level by a task-specific amplified phase reorganization of low-frequency delta oscillations during motor tracking, in an extended network comprising bilateral sensorimotor and right lateralized auditory and fronto-parietal regions (Fig. 5A). This set of right-lateralized regions that spreads well beyond the sensorimotor cortex corresponds to the typical network involved in auditory memory for pitch (55, 56). It plays a major role in the current task, which involves pitch estimation and retention. Importantly, our results thus show that temporal predictions spread selectively toward task-relevant regions and propagate at the network level. Further, in a subset of right-lateralized regions, delta phase reorganization was significant to behavioral outcome (Fig. 7). Our results thus show that during overt tracking, additional portions of cortex were implicated, through delta oscillations, probably to modulate sensory processing, emphasize target selection, and augment behavioral performance. However, our results do not address whether such motor-driven benefit requires a sense of agency (i.e. the intention to produce the movement) (57), and/or is mediated by somatosensory (tactile) feedback (58).

Altogether these results show that the motor system is ubiquitously involved in sensory processing, providing (at least) temporal contextual information. It encodes temporal predictions through delta-beta coupled oscillations, and has directed functional connectivity toward sensory regions in the beta range. We also show that performing overt rhythmic movements is an efficient strategy to augment the quality of temporal predictions, and amplify their influence on sensory processing, enhancing perception.

Methods

Participants

21 healthy adult participants were recruited (age range: 20-48 years; 10 females). The experiment followed ethics guidelines from the McGill University Health Centre (protocol NEU-13-023). Informed consent was obtained from all participants before the experiment. All had normal audition and reported no history of neurological or psychiatric disorder. We did not select participants based on musical training. Two participants were excluded from further analyses, as they did not perform the task properly.

Experimental design

Sequences of pure tones (Fig. 1A) were presented binaurally at a comfortable hearing level, using the Psychophysics-3 toolbox (59) and additional custom scripts written with MATLAB (The Mathworks). Each pure tone was sampled at 44100 Hz and lasted 100 ms with a dampening length and attenuation of 10 ms and 40 dB, respectively.

Each trial (~15 sec) was composed of 20 pure tones, which were qualified as references, targets, and distractors: 4 reference tones initiated the sequence, indicating both the reference tone frequency ($f_0 = 440$ Hz) and beat (1.5 Hz; *i.e.* ISI = 667 ms). They were followed by an alternation of 8 target and 8 distractor tones of variable frequencies (with a standard deviation of 0.2 in base-2 logarithmic units) presented in a quasi-rhythmic fashion (at 3 Hz; *i.e.* ISI = 333 ± 67 ms). The distribution of the 67 ms jitter across the tones of each sequence was approximately Gaussian and shorter than 141 ms. This ensured that no overlap between targets and distractors could occur. Importantly, targets and distractors occurred in phase and anti-phase with the preceding references, respectively, so that participants could use the beat provided by the references to distinguish targets from distractors, which were otherwise perceptually indistinguishable.

Participants performed a two-alternative pitch categorization task at the end of each trial, by deciding whether the mean frequency of targets f^{tar} was lower or higher than f_0 . Participants reported their choice with either their right index (lower) or right middle (higher) finger. The mean frequency of distractors f^{dis} was always equal to f_0 , hence non-informative. The absolute difference between f^{tar} and f_0 was titrated for each participant to reach threshold performance (see below). The task consisted of two conditions, *listen* and *tracking*, the only difference being that participants tapped in rhythm with the relevant sensory cues in the *tracking* condition. To do so, participants were required to follow the beat by moving their left index finger from the beginning of the sequence (the 2nd reference tone), so that their rhythm was stabilized when the 3-Hz melody started. In essence, this condition is a variation of the synchronization-continuation paradigm (10). We instructed participants to use their left hand, in order to spatially dissociate the neural correlates of motor acts (in the right motor cortex) from those of temporal predictions, hypothesized to be lateralized to the left hemisphere (42-45). Their movements were recorded with an infrared light pad. To ensure that participants' finger was properly positioned on the pad and occluded the infrared light when tracking, we added a small fabric patch below the infrared light, that participants had to make sure to touch. Thus, in addition to the task, participants had also to focus on the somatosensory

ry feedback from their finger in the *tracking* condition, which probably resulted in enhanced attention to targets as compared to a previous experiment (30). Participants were asked to stay still in the *listen* condition, not moving any part of their body, so as to minimize the overt involvement of the motor system. Absence of movements during the *listen* condition was monitored visually, using online video capture.

For each participant, the experiment was divided into 3 sessions occurring on different days, each lasting approximately 1 hour. The first session, purely behavioral, started with a short training, where the task was made easier by increasing the relative distance between the mean frequency of targets f^{tar} and f_0 , and by decreasing the volume of distractors. Following this, participants performed a psychophysical staircase with f^{tar} as the varying parameter. The staircase was set to obtain 75% of categorization performance in the *listen* condition. Then, the participants performed 2 separate sessions inside the MEG scanner. 200 trials per condition (*listen* or *tracking*) were acquired using a block-design paradigm, with conditions alternating every 20 trials (~5 min). Feedback was provided after each trial to indicate correct/incorrect responses, and overall feedback on performance indicating the total number of correct responses was given after every 20-trial block, for motivational purposes.

Data acquisition

MEG data were acquired at the McConnell Brain Imaging Centre, Montreal Neurological Institute McGill University (Canada) using a 275-channel MEG system (axial gradiometers, Omega 275, CTF Systems Inc.) with a sampling rate of 2400 Hz. Six physiological (EOG and ECG), one audio, two response buttons, and one infrared light pad channels were recorded simultaneously and synchronized with the MEG. The head position of participants was controlled using 3 coils, which locations were digitized (Polhemus) together with ~100 additional scalp points, to facilitate anatomical registration with MRI. Instructions were visually displayed on a grey background (spatial resolution of 1024 by 768 pixels and vertical refresh rate of 60 Hz) onto a screen at a comfortable distance from the subject's face. Auditory stimuli were presented through insert earphones (E-A-RTONE 3A, Aero Company). On each trial participants had to fixate a cross, located at the centre of the screen, to minimize eye movements. Prior to the first recording session, participants performed a 5-minute eyes-open resting-state session to be used as baseline activity in ensuing analyses. For cortically constrained MEG source analysis, *i.e.* the projection of the MEG sensor data onto the cortical surface, a T1-weighted MRI acquisition of the brain was obtained from each participant (1-mm isotropic voxel resolution).

Behavioral data processing

We estimated the 'element weight', *i.e.* the additive contribution assigned to each target (g_k^{tar}) and distractor (g_k^{dis}), to the subsequent decision (higher- or lower-pitched than f_0). We calculated these parameters across trials via a multivariate logistic regression of choice on the basis of a linear combination of the frequencies of the sixteen tones (8 targets and 8 distractors):

$$P(\text{high}) = \Phi \left[\sum_{k=1}^8 g_k^{\text{tar}} \cdot f_k^{\text{tar}} + \sum_{k=1}^8 g_k^{\text{dis}} \cdot f_k^{\text{dis}} + b \right] \quad (1)$$

where $P(\text{high})$ is the probability of judging the target sequence as higher pitched, $\Phi[\cdot]$ the cumulative normal density function, f_k^{tar} (f_k^{dis}) the frequency of the target (distractor) tone at position k in the sequence (expressed in logarithmic distance to f_0), and b an additive response bias toward one of the two choices. This analysis was based on best-fitting parameters estimates from the logistic regression model, which are not bounded and meet the a priori assumptions of standard parametric tests (30, 60). To estimate the element weight as a function of tone frequency, we binned the data into 100 overlapping quartiles before computing the multivariate logistic regression. We then estimated per participant, and separately for targets and distractors, the slope (linear coefficient) of the resulting element-weight profiles as a function of the tone frequency.

Timing of motor acts in the *tracking* condition

We extracted the timing of individual motor acts to investigate the precision of motor tracking. This analysis was performed separately for correct and incorrect trials of the *tracking* condition, after having matched the number of trials in these two groups for each participant. We then computed the preferred phase orientation (in radian) of motor acts relative to the 1.5-Hz beat (with zero corresponding to perfect simultaneity between a motor act and the 1.5-Hz beat). The mean resultant vector length was finally extracted to statistically compare the distribution of tapping events in correct and incorrect trials.

MEG data preprocessing

Preprocessing was performed with Brainstorm (61), following good-practice guidelines (62). Briefly, we removed electrical artifacts using notch filters (at 60 Hz and its first three harmonics), slow drifts using high-pass filtering (at 0.3 Hz), and eye blink and heartbeat artifacts using source signal projections (SSP). We detected sound onsets with a custom threshold-based algorithm on the simultaneously recorded audio traces. Data were both split into 10-sec trials and 600-ms epochs [-100:500] centered at the onset of each tone. MRI volume data were segmented with Freesurfer and down-sampled into Brainstorm to 15,002 triangle vertices, to serve as image supports for MEG source imaging. Finally, we computed individual MEG forward models using the overlapping-sphere method, and source imaging using weighted minimum norm estimates (wMNE) onto preprocessed data – all with using default Brainstorm parameters – and obtained 15002 source-reconstructed MEG signals (i.e. vertices). wMNE also included an empirical estimate of the variance of the noise at each MEG sensor, obtained from a 2-min empty-room recording done at the beginning of each scanning session.

Neural correlates of the 3-Hz melody

The 3-Hz melody corresponded to a stream of tones presented in quasi-rhythm fashion (Fig. 1A). We estimated the neural response to this auditory stream by means of two complementary measures inspired by (63): phase-phase coupling (indexed by the phase-locking value, PLV) and phase-amplitude coupling (PAC) between MEG signals and the 3-Hz melody. For each trial, we first extracted the MEG signals recorded during presentation of the 3-Hz melody (Fig. 1A; ~2.5-8 sec). We approximated the stimulus dynamics with a sinusoid oscillating around 3 Hz, in which each cycle corresponded to the duration between two successive tone onsets (with tone onset corresponding

to $\phi=0$). For phase-phase coupling, we band-pass filtered the MEG source signals at 2-4 Hz (using an even-order linear phase FIR filter), extracted their Hilbert phase, and computed PLV scores (64) between each resulting MEG signal and the 3 Hz *acoustic* sinusoid over time and trials. For phase-amplitude coupling, we first estimated the amplitude of MEG source signals in 26 frequencies, between 4-50 Hz, by applying a time-frequency wavelet transform, using a family of complex Morlet wavelets (central frequency 1 Hz; time resolution (FWMH) 3 s, hence 3 cycles). Then, we computed the phase-amplitude cross-frequency coupling (65) between the 3-Hz *acoustic* sinusoid (phase) and each resulting MEG signal (amplitude), over time and trials. Finally, parameter estimates (PLV or PAC) were contrasted with similarly processed resting-state data (assigned randomly), projected on a standard default brain, spatially smoothed (30a), and z-scored across vertices (to normalize across participants), prior to group-level statistical tests. Note that z-scoring was applied for main effect (e.g., *listen* condition) but not contrast analyses (e.g., *tracking vs. listen*), as this latter is a normalization procedure in itself.

Neural correlates of temporal predictions

During the 3-Hz melody, participants modulated their attention in time by relying on an internal model, an approximation of the 1.5-Hz beat initiated by the four reference tones (Fig. 1A). We estimated the neural dynamics of these temporal predictions by means of phase-phase (PLV) and phase-amplitude (PAC) coupling between MEG signals and the internal 1.5 Hz reference beat. We approximated the internal reference beat with a pure sinusoid oscillating at 1.5 Hz (with beat occurrence corresponding to $\phi=0$). We then applied the procedure described above, using the 1.5-Hz pure sinusoid as signal template, and band-passed filtered the MEG source signals within 1-2 Hz, for phase-phase coupling.

Regions-of-interest

We defined four regions-of-interest, of 100 vertices each, using functional localizers in each individual: We defined the left and right associative auditory (AA) cortices as respectively the left and right hemispheric vertices having the largest M100 (~110 ms) auditory evoked responses to both target and distractor tones, in both *listen* and *tracking*. The left and right sensorimotor (SM) cortices were defined as the vertices having the largest M50 (~40ms) evoked response, corresponding respectively to the right-handed motor responses produced at the end of each trial (to report participants' choices), and to the left-handed finger movements produced in the *tracking* condition (to track the reference beat). Region-of-interest analyses were carried out by performing principal component analysis across the 100 vertices composing each region-of-interest, to reduce their spatial dimensionality.

Evoked responses

Different categories of evoked responses were computed by averaging MEG signals across trials, for each time sample around targets, distractors or the 1.5-Hz beat onsets. The resulting signal averages were z-scored, using similarly processed resting-state data as baseline prior to group-level statistical tests.

Time-frequency responses

Time-frequency responses were computed by first estimating the amplitude of MEG signals in 30 frequency bins, between 1-50 Hz, by applying a time-frequency wavelet transform (see above) around targets and distractors onsets (-1:1 s). The resulting signals were averaged across trials. We finally contrasted in a normalized manner (across frequencies) targets and distractors by subtracting their respective time-frequency response and dividing the result by their sum.

Directed phase transfer entropy

Directed functional connectivity between regions-of-interest was estimated using directed phase transfer entropy (66). Briefly, it is based on the same principle as Wiener–Granger causality, with time series described in terms of their instantaneous phase (see also (67)). Signals were band-pass filtered in specific frequency bands (covering the 1-50 Hz range) prior to computation of directed phase transfer entropy. We then normalized the results between -0.5 and 0.5, with the sign indicative of the dominant direction of functional connectivity.

Statistical procedures

All analyses were performed at the single-subject level and followed by standard parametric tests at the group level (e.g., paired t-tests, t-tests against zero, repeated-measures ANOVAs). To test for the presence of significant 1.5-Hz beat evoked responses, data were z-scored over time (i.e. normalized around zero) and t-tests against zero were then performed. One-tail t-tests were applied when analyses were constrained with an *a priori* hypothesis. Two-tail t-tests were applied otherwise. In particular, main effects were subjected to one-tail t-tests against zero, to investigate the specific hypothesis of a stronger impact of acoustic stimuli on task-related rather than resting-state recordings. The alternate hypothesis was not investigated, as it was considered unrealistic. FDR corrections for multiple comparisons were applied over the dimensions of interest i.e., time, vertices and frequencies, using the Benjamini–Hochberg step-up procedure. For whole-brain analyses, only clusters >60 vertices in size were reported. For time-frequency analyses, only clusters >400 data points are reported.

Data and analysis pipeline availability

The data reported in this paper together with analysis pipelines can be accessed upon request via: <http://dx.doi.org/10.23686/0088165>.

Acknowledgments

B.M. was supported by a fellowship from the Montreal Neurological Institute. S.B. was supported by a Senior-Researcher grant from the Fonds de Recherche du Quebec-Santé (27605), a Discovery Grant from the National Science and Engineering Research Council of Canada (436355-13), the NIH (2R01EB009048-05) and a Platform Support Grant from the Brain Canada Foundation (PSG15-3755).

We thank Indiana Wollman, Philippe Albouy, Beth Bock, François Tadel and Peter Donhauser for providing comments and suggestions at multiple stages of the study.

References

1. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology* 20(2):172–176.
2. Kleinfeld D, Ahissar E, Diamond ME (2006) Active sensation: insights from the rodent vibrissa sensorimotor system. *Current Opinion in Neurobiology* 16(4):435–444.
3. Wachowiak M (2011) All in a sniff: olfaction as a model for active sensing. *Neuron* 71(6):962–973.
4. Hatsopoulos NG, Suminski AJ (2011) Sensing with the motor cortex. *Neuron* 72(3):477–487.
5. Crapse TB, Sommer MA (2008) Corollary discharge across the animal kingdom. *Nat Rev Neurosci* 9(8):587–600.
6. Schneider DM, Nelson A, Mooney R (2014) A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513(7517):189–194.
7. Flinker A, et al. (2010) Single-Trial Speech Suppression of Auditory Cortex Activity in Humans. *J Neurosci* 30(49):16643–16650.
8. Scott SK, McGettigan C, Eisner F (2009) A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. *Nat Rev Neurosci* 10(4):295–302.
9. Zatorre RJ, Chen JL, Penhune VB (2007) When the brain plays music: auditory–motor interactions in music perception and production. *Nat Rev Neurosci* 8(7):547–558.
10. Repp BH, Su Y-H (2013) Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon Bull Rev* 20(3):403–452.
11. Patel AD, Iversen JR (2014) The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front Sys Neurosci* 8:57.
12. Morillon B, Hackett TA, Kajikawa Y, Schroeder CE (2015) Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology* 31:230–238.
13. Merchant H, Yarrow K (2016) How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*. doi:10.1016/j.cobeha.2016.01.006.
14. Coull JT (2015) A Frontostriatal Circuit for Timing the Duration of Events. *Brain Mapping (Elsevier)*, pp 565–570.
15. Chen JL, Penhune VB, Zatorre RJ (2008) Listening to musical rhythms recruits motor regions of the brain. *Cerebral cortex (New York, NY : 1991)* 18(12):2844–2854.
16. Teki S, Grube M, Kumar S, Griffiths TD (2011) Distinct neural substrates of duration-based and beat-based auditory timing. *J Neurosci* 31(10):3805–3812.
17. Merchant H, Grahn J, Trainor L, Rohrmeier M, Fitch WT (2015) Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1664). doi:10.1098/rstb.2014.0093.
18. Grahn JA, Rowe JB (2012) Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral cortex (New York, NY : 1991)* 23(4):913–921.
19. Murthy VN, Fetz EE (1992) Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proc Natl Acad Sci USA* 89(12):5670–5674.

Motor origin of temporal predictions

20. Niso G, et al. (2016) OMEGA: The Open MEG Archive. *NeuroImage* 124(Pt B):1182–1187.
21. Arnal LH, Doelling KB, Poeppel D (2014) Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral cortex (New York, NY : 1991)*. doi:10.1093/cercor/bhu103.
22. Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. *J Neurosci* 32(5):1791–1802.
23. Fujioka T, Ross B, Trainor LJ (2015) Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *J Neurosci* 35(45):15187–15198.
24. Kulashekhar S, Pekkola J, Palva JM, Palva S (2016) The role of cortical beta oscillations in time estimation. *Hum Brain Mapp*. doi:10.1002/hbm.23239.
25. Iversen JR, Repp BH, Patel AD (2009) Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences* 1169:58–73.
26. Saleh M, Reimer J, Penn R, Ojakangas CL, Hatsopoulos NG (2010) Fast and Slow Oscillations in Human Primary Motor Cortex Predict Oncoming Behaviorally Relevant Cues. *Neuron* 65(4):461–471.
27. Coull JT (2011) Discrete Neuroanatomical Substrates for Generating and Updating Temporal Expectations. *Space, Time and Number in the Brain* (Elsevier), pp 87–101.
28. Arnal L (2012) Predicting “When” Using the Motor System's Beta-Band Oscillations. *Front Hum Neurosci* 6:225.
29. Schubotz RI (2007) Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences* 11(5):211–218.
30. Morillon B, Schroeder CE, Wyart V (2014) Motor contributions to the temporal precision of auditory attention. *Nature Communications* 5:5255.
31. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3(3):201–215.
32. Besle J, et al. (2011) Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *J Neurosci* 31(9):3176–3185.
33. Arnal L, Giraud AL (2012) Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences* 16(7):390–398.
34. Breska A, Deouell LY (2017) Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biol* 15(2):e2001665.
35. Park H, Kayser C, Thut G, Gross J (2016) Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *Elife* 5:e14521.
36. Park H, Ince RAA, Schyns PG, Thut G, Gross J (2015) Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Curr Biol* 25(12):1649–1653.
37. Coull JT, Davranche K, Nazarian B, Vidal F (2013) Functional anatomy of timing differs for production versus prediction of time intervals. *Neuropsychologia* 51(2):309–319.
38. Grahn JA, Rowe JB (2009) Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J Neurosci* 29(23):7540–7548.

Motor origin of temporal predictions

39. Gouvêa TS, et al. (2015) Striatal dynamics explain duration judgments. *Elife* 4:e11386.
40. Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6(10):755–765.
41. Nozaradan S, Schwartz M, Obermeier C, Kotz SA (2017) Specific Contributions of Basal Ganglia and Cerebellum To The Neural Tracking of Rhythm. *Cortex*. doi:10.1016/j.cortex.2017.08.015.
42. Schwartz M, Kotz SA (2013) A dual-pathway neural architecture for specific temporal prediction. *Neurosci Biobehav Rev* 37(10 Pt 2):2587–2596.
43. Coull JT, Nobre AC (1998) Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J Neurosci* 18(18):7426–7435.
44. Bolger D, Coull JT, Schön D (2014) Metrical Rhythm Implicitly Orients Attention in Time as Indexed by Improved Target Detection and Left Inferior Parietal Activation. *J Cogn Neurosci* 26(3):593–605.
45. Wiener M, Turkeltaub PE, Coslett HB (2010) Implicit timing activates the left inferior parietal cortex. *Neuropsychologia* 48(13):3967–3971.
46. Coull JT (2015) Directing Attention in Time as a Function of Temporal Expectation. *Brain Mapping* (Elsevier), pp 687–693.
47. Fontolan L, Morillon B, Liegeois-Chauvel C, Giraud A-L (2014) The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications* 5:4694.
48. Bastos AM, et al. (2014) Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*. doi:10.1016/j.neuron.2014.12.018.
49. van Kerkoerle T, et al. (2014) Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci USA*. doi:10.1073/pnas.1402773111.
50. Buschman TJ, Miller EK (2007) Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science* 315(5820):1860–1862.
51. Spitzer B, Haegens S (2017) Beyond the Status Quo: A Role for Beta Oscillations in Endogenous Content (Re)Activation. *eNeuro* 4(4). doi:10.1523/ENEURO.0170-17.2017.
52. Schön D, Tillmann B (2015) Short- and long-term rhythmic interventions: perspectives for language rehabilitation. *Annals of the New York Academy of Sciences* 1337:32–39.
53. Kotz SA, Schwartz M, Schmidt-Kassow M (2009) Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex* 45(8):982–990.
54. Goswami U (2011) A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences* 15(1):3–10.
55. Zatorre RJ, Evans AC, Meyer E (1994) Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience* 14(4):1908–1919.
56. Peretz I, Zatorre RJ (2005) Brain organization for music processing. *Annu Rev Psychol* 56(1):89–114.
57. Timm J, Schönwiesner M, Schröger E, Sanmiguel I (2016) Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex* 80:5–20.

Motor origin of temporal predictions

58. Reznik D, Ossmy O, Mukamel R (2015) Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *J Neurosci* 35(5):2173–2180.
59. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R (2007) What's new in Psychtoolbox-3. *Perception*.
60. Wyart V, de Gardelle V, Scholl J, Summerfield C (2012) Rhythmic Fluctuations in Evidence Accumulation during Decision Making in the Human Brain. *Neuron* 76(4):847–858.
61. Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience* 2011:879716.
62. Gross J, et al. (2012) Good practice for conducting and reporting MEG research. *NeuroImage* 65C:349–363.
63. Peelle JE, Gross J, Davis MH (2012) Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral cortex (New York, NY : 1991)* 23(6):1378–1387.
64. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208.
65. Özkurt TE, Schnitzler A (2011) A critical note on the definition of phase-amplitude cross-frequency coupling. *Journal of Neuroscience Methods* 201(2):438–443.
66. Lobier M, Siebenhühner F, Palva S, Palva JM (2014) Phase transfer entropy: A novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *NeuroImage* 85:853–872.
67. Hillebrand A, et al. (2016) Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proc Natl Acad Sci USA* 113(14):3867–3872.

Figures

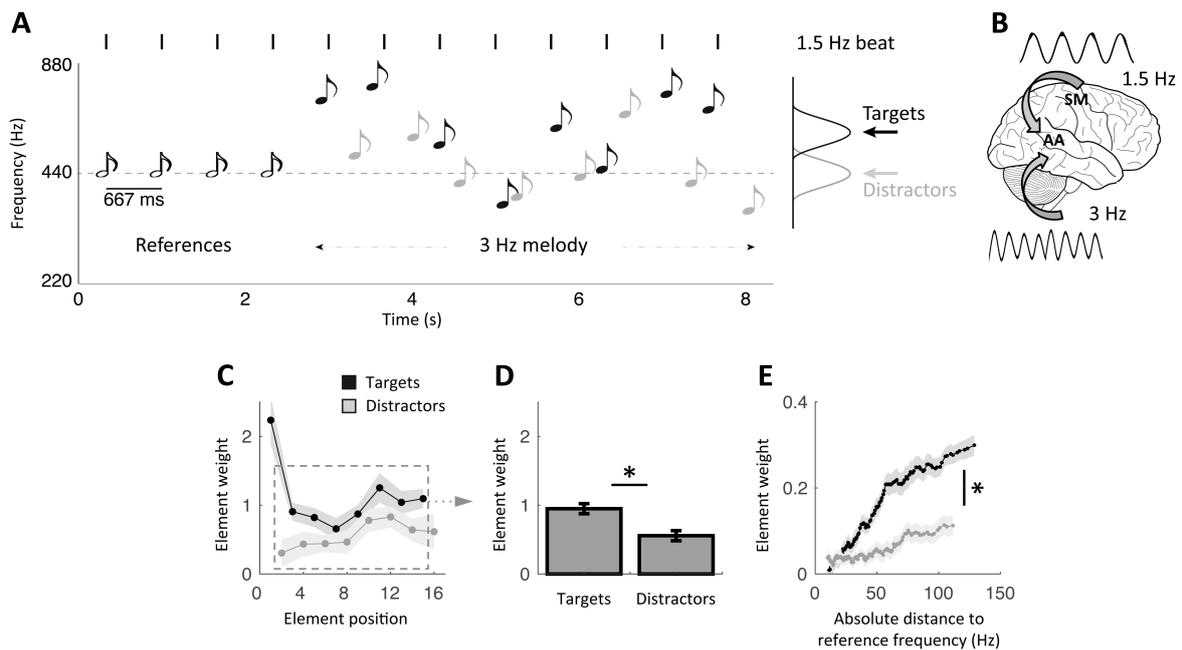


Figure 1. Experimental design and behavioral results in the *listen* condition. (A) Quasi-rhythmic sequences of 20 pure tones were presented binaurally on each trial. 4 reference tones preceded an alternation of 8 target and 8 distractor tones of variable frequencies. Targets occurred in phase with the preceding references, whereas distractors occurred in anti-phase. Participants had to decide whether the mean frequency of targets was higher or lower than the reference frequency. Participants performed the task without moving before the end of the sequence in the *listen* condition, and while expressing the reference-beat by moving their left index finger in the (motor) tracking condition. The 3-Hz melody is the period-of-interest for all analyses. **(B)** In this design, neural activity dedicated to stimulus or temporal predictions processing are dissociable based on their temporal dynamics (3 Hz vs. 1.5 Hz, respectively). **(C-E)** Behavioral results in the *listen* condition: **(C)** Element-weight profiles for targets (black) and distractors (grey) as a function of element position in the sequence. **(D)**. Averaged contribution of targets and distractors to the decision. Element weights were pooled separately across targets and distractors, excluding the first target and last distractor (dashed square in panel C). **(E)** Element-weight profiles for targets (black) and distractors (grey) as a function of tone frequency. (Shaded) error bars indicate s.e.m. Star indicates significant difference ($n = 19$; paired t-tests; $p < .05$).

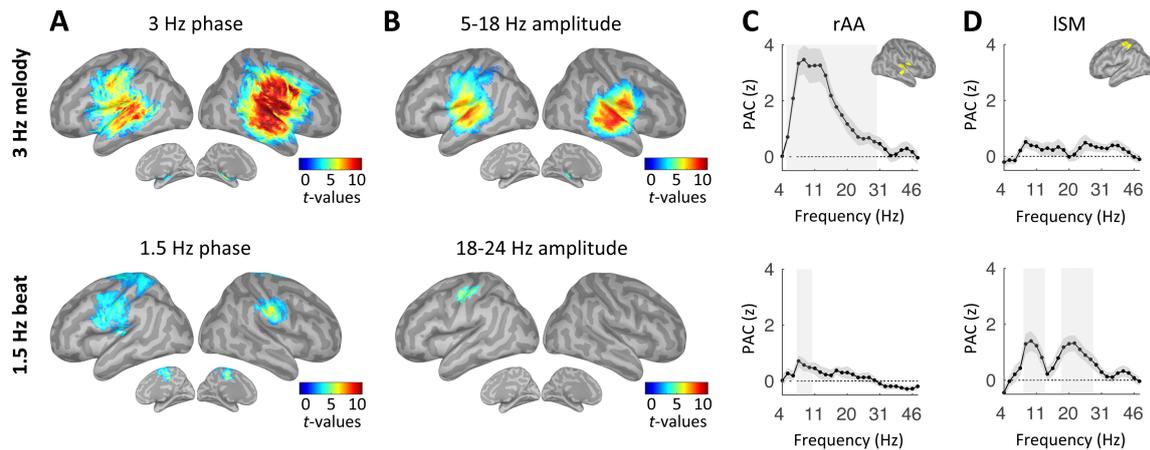


Figure 2. Listen: Neural correlates of (top) the 3-Hz melody and (bottom) covert temporal predictions. (A-B) Whole-brain statistical maps of **(A)** phase-phase (PLV) and **(B)** phase-amplitude (PAC) coupling estimates for *listen* correct trials. Coupling estimates were respectively computed between the phase of the (top) 3-Hz melody or (bottom) 1.5-Hz beat dynamics, and specific brain rhythms (see titles). **(C-D)** Detail of the phase-amplitude coupling estimates in (C) right associative auditory (rAA) and (D) left sensorimotor (ISM) cortex, in the 4-50 Hz range. Shaded error bars indicate s.e.m. Shaded vertical bars highlight significant effects ($n = 19$; one-tail t-tests against zero (A-B) or paired t-tests (C-D); $p < .05$, FDR corrected).

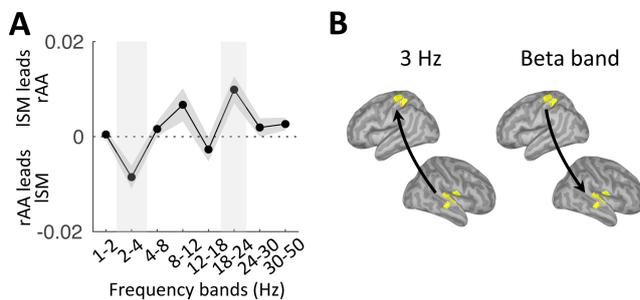


Figure 3. Listen: Frequency-specific directional interactions between right associative auditory (rAA) and left sensorimotor (ISM) cortex. (A) Directed functional connectivity between rAA and ISM in the 1-50 Hz range, for *listen* correct trials. Shaded vertical bars highlight significant effects. **(B)** Schematic of the significant results depicted in (A) ($n = 19$; paired t-tests; $p < .05$, FDR corrected).

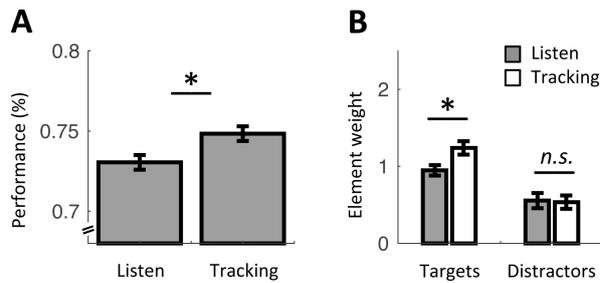


Figure 4. Behavioral differences between *listen* and *tracking*. (A) Average categorization performance in *listen* and *tracking*. (B). Averaged contribution of targets and distractors to the decision in the *listen* (grey) and *tracking* (white) conditions. Element weights were pooled separately across targets and distractors, excluding the first target and last distractor. Error bars indicate s.e.m. Stars/n.s. indicate significant/non-significant differences ($n = 19$; paired t-tests; $p < .05$).

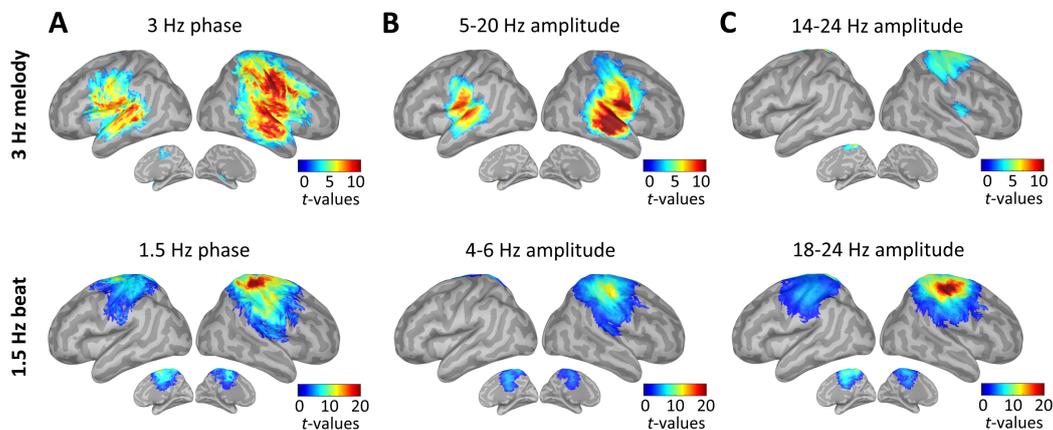


Figure 5. *Tracking*: Neural correlates of (top) the 3-Hz melody and (bottom) overt temporal predictions. Whole-brain statistical maps of (A) phase-phase (PLV) and (B-C) phase-amplitude (PAC) coupling estimates for *tracking* correct trials. Coupling estimates were respectively computed between the phase of the (top) 3-Hz melody or (bottom) 1.5-Hz beat dynamics, and specific brain rhythms (see titles; $n = 19$; one-tail t-tests against zero; $p < .05$, FDR corrected).

Motor origin of temporal predictions

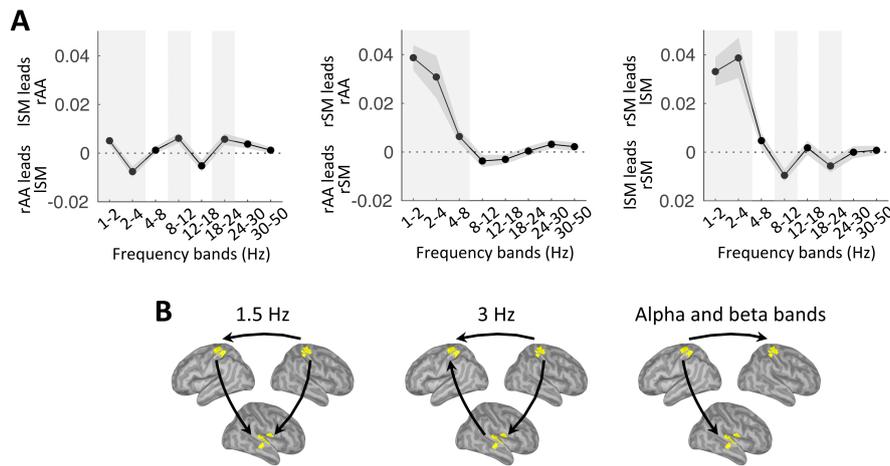


Figure 6. Tracking: Frequency-specific directional interactions between right associative auditory (rAA), left sensorimotor (ISM) and right sensorimotor (rSM) cortex. **(A)** Directed functional connectivity between (left) rAA and ISM, (middle) rAA and rSM, and (right) ISM and rSM, in the 1-50 Hz range for *tracking* correct trials. Shaded vertical bars highlight significant effects. **(B)** Schematic of the main significant results depicted in (A) ($n = 19$; paired t-tests; $p < .05$, FDR corrected).

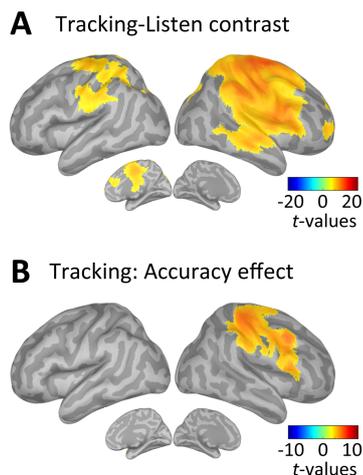


Figure 7. Modulation of the neural correlates of temporal predictions with condition and accuracy. **(A-B)** Whole-brain statistical maps of the contrast between 1.5-Hz beat phase-phase coupling estimates **(A)** in *tracking* versus *listen* correct trials, and **(B)** in *tracking* correct versus incorrect trials ($n = 19$; paired t-tests, $p < .05$, FDR corrected).