Published in final edited form as:

Curr Opin Neurobiol. 2015 April ; 31: 230–238. doi:10.1016/j.conb.2014.12.005.

Predictive motor control of sensory dynamics in Auditory Active Sensing

Benjamin Morillon^{1,2,†}, Troy A. Hackett³, Yoshinao Kajikawa⁴, and Charles E. Schroeder^{1,4,†}

¹Department of Psychiatry, Columbia University Medical Center, New York, NY 10032, USA

²McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, QC H3A 2B4

³Department of Speech and Hearing, Vanderbilt University Medical Center, Nashville, TN 37232, USA

⁴Translational Cognitive Neuroscience Program, Nathan Kline Institute, Orangeburg, NY 10962, USA

Abstract

Neuronal oscillations present potential physiological substrates for brain operations that require temporal prediction. We review this idea in the context of auditory perception. Using speech as an exemplar, we illustrate how hierarchically organized oscillations can be used to parse and encode complex input streams. We then consider the motor system as a major source of rhythms (temporal priors) in auditory processing, that act in concert with attention to sharpen sensory representations and link them across areas. We discuss the anatomo-functional pathways that could mediate this audio-motor interaction, and notably the potential role of the somatosensory cortex. Finally, we reposition temporal predictions in the context of internal models, discussing how they interact with feature-based or spatial predictions. We argue that complementary predictions interact synergistically according to the organizational principles of each sensory system, forming multidimensional filters crucial to perception.

Keywords

auditory perception; active sensing; oscillation; prediction; sensorimotor

Neuronal oscillations as substrates of temporal predictions

Sensory processing can be viewed as an inference problem: given the noisy sensory and neural signals, the brain estimates which stimuli caused these observations [¹]. Accordingly, neuronal representations may be viewed as continuous top-down prediction/expectation

[†]**Correspondence** Dr. Benjamin Morillon, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, QC H3A 2B4, (+1) 347-746-5432, ; Email: bnmorillon@gmail.com, Dr. Charles E Schroeder, Cognitive Translational Neuroscience, Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd, Orangeburg, NY 10962, 845-398-6546, ; Email: cs2388@columbia.edu

Conflicts of interest statement

The author declares no conflicts of interest.

signals, that is, internal models of causal dynamics in the world shaped by one's perceptual, behavioral and emotional experience [2]. Such theories have two far-reaching corollaries: 1- predictions are reflected in the brain's intrinsic activity [3,4], and 2- sensory evoked neuronal activity corresponds to the modulation of ongoing circuit dynamics by input signals, initiating the generation of an updated representation of the world [5**,6].

How are predictions encoded in the brain? *Dynamic attending theory* suggests that during perception attention is modulated dynamically to optimize sensory processing at expected points in time [7]. Crucially, this framework capitalizes on the fact that many stimuli and actions are rhythmically organized (*e.g.* speech, music, walking, among others) [8]. By extracting such temporal regularities, the brain is able to predict the occurrence of subsequent events of interest and optimize their processing.

Behaviorally, temporal predictions optimize perception by dynamically modulating the gain of sensory information [9*,10*]. Improved performance is associated with a reduction of the sensory evoked neuronal response [11-13]. Neurophysiologically, low-frequency neuronal oscillations are the hypothesized substrate of temporal predictions [4,5**,14]. Neuroelectric oscillations correspond to temporal fluctuations of the excitability state of neuronal ensembles (Figure 1C) [4]. Using these as 'instruments' [15,16] the brain can selectively amplify neuronal responses to events occurring at predicted moments, while suppressing responses to events that do not adhere to the prediction [17,18,19**]. This mechanism results in a selective neuronal tracking of attended sensory streams, as exemplified by the *cocktail party* effect, where listeners have to 'tune in' to one conversation in a noisy scene [20].

A prototype: the hierarchy of time scales in speech processing

Speech perception arises from the dynamic sampling of acoustic information at multiple time-scales. Specifically, *entrainment* of cortical oscillatory dynamics to the acoustic envelope has been described for delta (1-4 Hz), theta (4-8 Hz), and low-gamma (25-45 Hz) frequencies (Figure 1 A-B) [14,21*,22*]. And top-down predictive influences on speech perception are amply supported by multiple lines of evidence [14,22*,23-25].

The most basic level of the proposed speech processing hierarchy is delta-band mediated parsing of the prosodic phrasal rate of speech [14,21*]. A right-lateralized delta entrainment is observed during speech processing, co-occurring with theta entrainment, but carrying complementary information [22*]. Left-lateralized involvement of delta oscillations is on contrary hypothesized to reflect motor-driven top-down modulation [26,27]. The idea that motor influences contribute to sensory speech processing places it squarely in the domain of '*active sensing*' [28,29], as outlined below.

Although phrasal and syllabic modulations (1-7 Hz) are crucial to speech comprehension [30], acoustic modulations above 12 Hz also contribute to speech identification, as they are associated with some phonetic feature dimensions [31]. The speech processing hierarchy, as it stands now, does not explicitly deal with the question of how the auditory system encodes the higher frequency information, densely packed in consonant sounds. An obvious

possibility is that the neuronal representation of speech utilizes a spike phase code, like that described for naturalistic auditory scene processing [32]. In such a scheme, entrainment of a delta-theta-gamma hierarchy in auditory cortex to the nested low, middle and higher frequency rhythms inherent in speech provides a dynamical structure against which neuronal firing in response to speech sounds can be indexed (i.e. a reference frame). As was shown by Kayser and colleagues (Figure 1E) [32], concurrent implementation of complementary stimulus coding schemes presents several benefits that would be particularly useful in speech comprehension. In particular, such a scheme would increase the amount of information that can be packed into a neuronal firing pattern and make the aggregate representation more robust to noise [32]. In order for higher order neuron assemblies to act as efficient readers they would have to share the reference frame, likely by entraining at the same set of frequencies as the neuron assemblies performing the encoding step in auditory cortex. Interestingly in this regard, widespread delta and theta band entrainment in frontal and prefrontal regions complements that occurring in auditory cortical areas during speech tracking [20].

The proposition of a hierarchical system of brain dynamics that can parse and decode the information that is embedded within at least four nested temporal scales in speech has wide-ranging implications for non-auditory, as well as auditory processing. The fastest scale in this scheme corresponds best to neuronal firing. Firing is essential for the transmission of information through the system, but decoding that information by central readers depends on the fact that sensory evidence is parsed by a nested hierarchy of oscillations and information is systematically encoded according to oscillatory phase within each scale.

Motor origin of temporal predictions in *active sensing*

Capitalizing on predictive coding, and other Bayesian inference frameworks, recent papers suggest that in addition to generating movements, the motor system may be essential in generating the temporal predictions that shape perception (Figure 2A and 4A) [5**, 29,33,34]. The motor system has been shown to be a key structure in the precise estimation of short durations [35,36], and in beat and rhythm processing [33,37,38]. The motor system is likely involved in a wide variety of temporal-predictive functions, and may be a major causal agent in both ambient and task-related rhythms in sensory pathways. As discussed above, these rhythms would imbue the motor system's predictions.

Active sensing [29] mainly refers to acquiring sensory inputs by overt motor sampling behaviors, such as hand movements [39*,40] or eye movements [41,42], and whisking or sniffing [28,43,44]. These sampling movements are generally rhythmic, but can be aperiodic, as in an isolated hand-reach [39*]. In either case, attention is an essential component of the process, and its top-down influence helps to impose the motor sampling pattern on the relevant sensory stream; in addition to driving of activity in sensory areas by volleys of ascending input, there is top-down (corollary discharge and/or attentional) modulation [45], all yoked to the movements [29]. 'Covert' attentional sampling is a part of active sensing that takes over where overt motoric sampling leaves off. When holding still and attending covertly to one of several rhythmic sensory streams (e.g. one conversation at a cocktail party), the brain 'selects' by enforcing oscillatory entrainment of sensory activity to

one task-relevant stream at the expense of all others. This is true whether attention is selecting between modalities [17,46], within a modality [19**,47], or integrating across modalities [20]. Motor systems appear integral to covert-attentional, as well as overt-motoric sampling of sensory input. Motor and premotor cortices consistently emerge as the most active regions in both scalp EEG studies [48] and intracranial studies of low frequency rhythmic attentional selection [20,49]. Low frequency (delta) rhythms in M1 predict the timing of informative cues in a motor planning task [50] and facilitate coherence between M1/M2 and parietal regions [51]. These observations reinforce the more general conclusion that, while the brain's time-keeping networks likely involve cerebellum, basal ganglia, insular cortex and thalamus [48,52], motor and premotor cortices are critical nodes in these networks [53,54], especially when the task is difficult [55], and rhythms are used to anticipate and predict [26,49,50]. According to the *Pre-motor Theory of Attention* [56], covert shifts of spatial attention are governed by the same FEF-centric network that underlies saccades of similar dimensions. Although one can argue specific tenets of the Theory [57], it clearly underscores a fundamental relationship between motor systems and active/attentive sensory processing, one which comes to the fore when event rhythms allow predictions [10*,29].

Active sensing in the auditory domain

The active sensing framework proposes that perception is shaped by the motor system in two distinct ways. First, the motor system directly causes sensory inflow as a consequence of the motor acts it directs (*e.g.* finger squeezes cause somatosensory stimulation, ocular saccades cause specific visual stimulation, sniffs cause stimulation of the olfactory epithelium), thereby structuring both the specific content and the temporal/rhythmic context of bottom-up sensory information inflow. Second, it modulates the processing of sensory information via top-down attentional control and/or corollary discharge signals, that is, copies of movement commands sent to associated sensory structures [58]. Top-down signals predictively modulates sensory processing according to the temporal (and spatial) patterns of attentional and motor sampling plans, thus providing 'when' (and 'where') predictions at a minimum [45].

Although motor influences over auditory cortices have been reported [59*], active sensing has not been described in the auditory domain, because, in contrast to the other modalities, bottom-up auditory processing is generally disconnected from movement. The bulk of auditory inputs are like the subsets of somatosensory inputs that fall on a still hand, or visual events that occur during fixation. Even in species that can move their ears flexibly, such as rabbits, cats and monkeys, these movements function like gross head movements that are used to maximize binaural cues. Crucially, motor acts do not cause inflow of auditory input, excepting when movements make noise. Nonetheless, temporal sensitivity is at its best in audition. The striking relationship between rhythm in the auditory and motor systems [37,60] in music perception points to a more general role that the motor system could play in exerting top-down predictive influences on auditory processing. In the speech example, top-down motor influences would structure both the parsing of complex input streams and the construction of the reference frame for the phase coding of specific acoustic elements in spiking.

Although direct projections from secondary motor cortex to auditory cortex have been observed in mice [^{59*}], no such projections have been as yet reported in primates. There are, however, a number of routes by which activity in the motor system as broadly construed to include primary motor, premotor and motor-related prefrontal regions, basal ganglia, and surprisingly the somatosensory system, may modulate the gain of auditory processing and perception (Fig 2).

Notably absent are projections from motor (Area 4) and premotor (Area 6) regions although there are some projections from Areas 45 and 8 to Temporo-parietal (Tpt), caudal belt and parabelt areas. Prefrontal cortical connections with rostral belt/parabelt areas are reciprocal, but would be an exceedingly indirect relay of motor corollary discharge. There are topographic projections to the striatum from both rostral and caudal belt/parabelt regions of auditory cortex, which could inform motor planning and movement generation, but the return connection by which motor rhythms would be injected into the auditory processing stream is, again, indirect at best.

One interesting possibility outlined here is that motor influences may be conveyed to auditory cortex through adjacent somatosensory areas in the Lateral Sulcus. Movement related signals clearly would access belt regions of auditory cortex through these connections. In this regard, it is intriguing to note that in a task entailing rhythmic presentation of attended audiovisual stimuli, multiunit firing and beta band power in second somatosensory (S2) ensembles (Figure 3) display systematic fluctuations, reflecting modulation of local neuronal excitability in anticipation/prediction of *potential* target stimulus appearance. The predictive component of the activity is like that shown in motor cortex itself [⁵⁰]. The Beta power drop and associated excitability changes are like those that attend movement onset in motor cortex [⁶¹]. These dynamics reflect a complex interdependence of motor and somatosensory physiology surrounding goal directed movement and likely occur in S2 and other somatosensory cortices by virtue of the somatosensory system's extensive connections with the motor system. Auditory belt cortices in turn, would have access to these signals due to their extensive interconnections with the somatosensory system [62]. It is not yet clear whether these S2 effects manifest in auditory cortex.

Motor contributions to the temporal precision of auditory attention

We recently developed a mechanistic behavioral account of auditory active sensing [^{10*}]. We measured participants' ability to extract relevant auditory information interleaved with distracting information and embedded in rhythmic streams; thus, performance depended crucially on temporal predictions. We used a perceptual decision-making task, to study the dynamics of evidence accumulation and its modulation by temporal predictions. During the task, participants were either allowed to use their motor system, by noiselessly pressing in rhythm with the relevant auditory information, or prevented from doing so.

Findings (Figure 4B) show that rhythmic movements engage a top-down modulation that sharpens the temporal selection of auditory information [^{10*}]. It improves the segregation between relevant and distracting information, facilitating perception of relevant items and

suppressing perception of irrelevant items. The impact of overt motor tracking on perception parametrically scales with the temporal predictability of the auditory sequence, and depends on the temporal alignment between motor and attention fluctuations.

Beyond the sensory level: hierarchical organization of predictive filters in auditory perception

Ideally, one would study temporal predictions independently from other dimensions (e.g. location or feature). However, in most paradigms these other attributes are held constant and predictable across trials, and thus may constrain perception. Indeed, in primary auditory cortex the combination of temporal and spectral predictions results in both the amplification and sharpening of expected neuronal responses, acting as a spectrotemporal filter mechanism [19**]: The top-down modulation of neuronal excitability is dynamically and simultaneously adjusted at the temporal and spectral level. Such combination is not restricted to *what* and *when* dimensions in audition, but can also integrate spatial [63] and intensity [64] information at higher levels. Thus predictions seem to constantly fuse into a dynamically evolving model of upcoming events, which enhances the representation of predicted stimuli.

However, predictions do not seem to combine linearly. For example, temporal and spatial predictions interact to improve the quality of visual information, but the benefit of temporal predictions may disappear in lieu of valid spatial predictions [65,66*]. Thus, the impact temporal predictions have on perception may be *conditional* upon the co-occurrence of spatial predictions. The lower visual areas being retinotopically (*i.e.* spatially) organized, temporal predictions might not be able to modulate the first cortical stages of visual processing unless they are part of a spatiotemporal predictive filter that taps onto retinotopic organization. The same logic applied to the lower levels of the auditory system would argue that the most effective predictions would address tonotopic as well as temporal dimensions of stimulation [19**].

It is likely that the hierarchical progression of processing starts with the most fundamental attributes at the lower sensory levels, rhythm and sensory primitives (e.g. frequency in auditory cortex). *Salient* stimuli can drive rhythmic entrainment in A1 in a bottom-up fashion without the subject's attention [67], in which case, the very tendency of local pyramidal/interneuron ensembles to oscillate and nest hierarchically in delta, theta and gamma frequencies at rest, provides a system of predictions. As discussed above, the entrained oscillations provide initial parsing of the input stream and a reference framework for phase coding of information carried in spike trains across the population of A1 neurons. The untested assumption here is that at the level of A1, oscillatory parsing with spike-phase coding creates 'packets' of information that are 'multiplexed', by combining coding across different scales in the hierarchical oscillatory structure. As these packets progress to higher order areas, additional predictions/filters are applied with increasingly longer time constants [3], promoting integration over sensory input stream segments of increasing length. Though speech is an appealing illustration, in principle this would apply to any complex, naturalistic sensory stream.

In active sensing, top-down control becomes determinative, and the system's evolved capabilities combine with the subject's experience to allow stronger predictions. In the 'Motor Theory' of speech perception for example [68], the motor system provide the filter. In the 'Premotor Theory' of attention [56], the premotor regions would do the same for covert visual active sensing. When actively engaged, motor and premotor regions would reinforce the natural tendency for hierarchical entrainment of A1 neuron ensembles to the rhythmic structure of the input stream. These effects are the basis for additional transformation of the input representation in belt areas and beyond. Successively higher-order premotor/prefrontal areas modulate motor cortex to impose hierarchical (predictive) oscillatory structure like that we suppose is operating in language, and this helps to group the processing of primitives like stimulus pitch/frequency and Interaural time difference into higher order structures ('a language-like' stimulus streams) and also combines across primitive code categories to form auditory objects. This thinking is consistent with the idea (e.g. [3]) that the time constant of activity increases at higher order regions. Such progressive abstraction sets the stage for covert (mental) processes that are largely disconnected from the temporal patterns of the ongoing sensorium.

Acknowledgments

Supported by the Bettencourt Schueller Foundation and by the NIH (MH060358, DC011490 and DC012918).

References and recommended reading

1. von Helmholtz H. Handbook of physiological optics. Leipzig: Leopold Voss. 1856 no volume.
2. Friston K. A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2005; 360:815–836.
3. Sadaghiani S, Kleinschmidt A. Functional interactions between intrinsic brain activity and behavior. *NeuroImage*. 2013; 80:379–386. [PubMed: 23643921]
4. Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*. 2009; 32:9–18. [PubMed: 19012975]
- *5. Arnal L, Giraud AL. Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*. 2012; 16:390–398. This review paper reconciles predictive coding and neuronal oscillations frameworks. It specifically describes how what and when predictions could be computationally implemented by distinct neuronal rhythms, and addresses the causal role of the motor system in temporal predictions. [PubMed: 22682813]
6. Fiser J, Berkes P, Orbán G, Lengyel M. Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*. 2010; 14:119–130. [PubMed: 20153683]
7. Large EW, Jones MR. The dynamics of attending: How people track time-varying events. *Psychol Rev*. 1999; 106:119–159.
8. Ding N, Patel AD, Poeppel D. Temporal modulations in speech and music. under review. [no date], [no volume].
- *9. Rohenkohl G, Cravo AM, Wyart V, Nobre AC. Temporal expectation improves the quality of sensory information. *J. Neurosci*. 2012; 32:8424–8428. By combining behavioral measures and modeling, the authors highlight the specific influence of temporal predictions on visual perception. They show that valid temporal predictions improve the signal-to-noise gain of sensory information. [PubMed: 22699922]
- *10. Morillon B, Schroeder CE, Wyart V. Motor contributions to the temporal precision of auditory attention. *Nature Communications*. 2014; 5:5255. This study specifically investigates active sensing in the auditory domain. The authors find that rhythmic movements sharpen the temporal selection of auditory stimuli by facilitating the perception of relevant stimuli, while actively suppressing the interference from irrelevant stimuli.

11. Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. Stimulus predictability reduces responses in primary visual cortex. *J. Neurosci.* 2010; 30:2960–2966. [PubMed: 20181593]
12. Costa-Faidella J, Baldeweg T, Grimm S, Escera C. Interactions between "what" and "when" in the auditory system: temporal predictability enhances repetition suppression. *J. Neurosci.* 2011; 31:18590–18597. [PubMed: 22171057]
13. Yaron A, Hershenhoren I, Nelken I. Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron.* 2012; 76:603–615. [PubMed: 23141071]
14. Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A. Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences.* 2008; 12:106–113. [PubMed: 18280772]
15. Buzsáki G. Hippocampal sharp waves: their origin and significance. *Brain Research.* 1986; 398:242–252. [PubMed: 3026567]
16. Buzsáki G. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron.* 2010; 68:362–385. [PubMed: 21040841]
17. Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science.* 2008; 320:110–113. [PubMed: 18388295]
18. Cravo AM, Rohenkohl G, Wyart V, Nobre AC. Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J. Neurosci.* 2013; 33:4002–4010. [PubMed: 23447609]
- *19. Lakatos P, Musacchia G, O'Connell MN, Falchier AY, Javitt DC, Schroeder CE. The Spectrotemporal Filter Mechanism of Auditory Selective Attention. *Neuron.* 2013; 77:750–761. This outstanding study describes for the first time how selective attention operates in the auditory modality. By recording neuronal activity in the primary auditory cortex of monkeys, the authors show that temporal and spectral predictions are combined into a spectrotemporal filter mechanism that modulates perception. They also beautifully disentangle the contribution of internally driven top-down modulations and stimulus driven responses on neuronal activity. [PubMed: 23439126]
20. Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, et al. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron.* 2013; 77:980–991. [PubMed: 23473326]
- *21. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci.* 2012; 15:511–517. In this seminal review, Giraud and Poeppel propose an oscillatory-based model of speech processing. A key characteristic of this model is that sensory information is simultaneously processed at multiple time-scales. [PubMed: 22426255]
- *22. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLoS Biol.* 2013; 11:e1001752. Using cutting-edge methods of analysis, this paper beautifully confirms that during speech processing gamma and theta neuronal oscillations respectively dominate in left and right hemispheres. [PubMed: 24391472]
23. Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature.* 2012 doi:10.1038/nature11020.
24. Peelle JE, Gross J, Davis MH. Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral cortex (New York, N.Y. : 1991).* 2012 doi:10.1093/cercor/bhs118.
25. Morillon B, Liégeois-Chauvel C, Arnal L, Bénar C-G, Giraud AL. Asymmetric Function of Theta and Gamma Activity in Syllable Processing: An Intra-Cortical Study. *Front. Psychology.* 2012; 3:248.
26. Arnal LH, Doelling KB, Poeppel D. Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral cortex (New York, N.Y. : 1991).* 2014 doi:10.1093/cercor/bhu103.
27. Morillon B, Lehongre K, Frackowiak RSJ, Ducorps A, Kleinschmidt A, Poeppel D, Giraud AL. Neurophysiological origin of human brain asymmetry for speech and language. *Proc. Natl. Acad. Sci. U.S.A.* 2010; 107:18688–18693. [PubMed: 20956297]
28. Kleinfeld D, Ahissar E, Diamond ME. Active sensation: insights from the rodent vibrissa sensorimotor system. *Current Opinion in Neurobiology.* 2006; 16:435–444. [PubMed: 16837190]

29. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P. Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*. 2010; 20:172–176. [PubMed: 20307966]
30. Elliott TM, Theunissen FE. The modulation transfer function for speech intelligibility. *PLoS Comput Biol*. 2009; 5:e1000302. [PubMed: 19266016]
31. Drullman R, Festen JM, Plomp R. Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* 1994; 95:2670–2680. [PubMed: 8207140]
32. Kayser C, Montemurro MA, Logothetis NK, Panzeri S. Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. *Neuron*. 2009; 61:597–608. [PubMed: 19249279]
33. Grahm JA, Rowe JB. Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral cortex (New York, N.Y. : 1991)*. 2012 doi: 10.1093/cercor/bhs083.
34. Schubotz RI. Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*. 2007; 11:211–218. [PubMed: 17383218]
35. Coull JT, Vidal F, Nazarian B, Macar F. Functional anatomy of the attentional modulation of time estimation. *Science*. 2004; 303:1506–1508. [PubMed: 15001776]
36. Morillon B, Kell CA, Giraud AL. Three stages and four neural systems in time estimation. *J. Neurosci*. 2009; 29:14803–14811. [PubMed: 19940175]
37. Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory–motor interactions in music perception and production. *Nat Rev Neurosci*. 2007; 8:547–558. [PubMed: 17585307]
38. Teki S, Grube M, Kumar S, Griffiths TD. Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci*. 2011; 31:3805–3812. [PubMed: 21389235]
- *39. Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV. Neural population dynamics during reaching. *Nature*. 2012; 487:51–56. Churchland et al. show here that motor control depends on population dynamics. Reaching movements are associated with neuronal oscillatory activity in the monkey motor cortex, despite there being no periodic component to this behavior. [PubMed: 22722855]
40. Saleh M, Takahashi K, Hatsopoulos NG. Encoding of coordinated reach and grasp trajectories in primary motor cortex. *J. Neurosci*. 2012; 32:1220–1232. [PubMed: 22279207]
41. MacEvoy SP, Hanks TD, Paradiso MA. Macaque V1 activity during natural vision: effects of natural scenes and saccades. *Journal of Neurophysiology*. 2008; 99:460–472. [PubMed: 18077668]
42. Rajkai C, Lakatos P, Chen C-M, Pincze Z, Karmos G, Schroeder CE. Transient cortical excitation at the onset of visual fixation. *Cerebral cortex (New York, N.Y. : 1991)*. 2008; 18:200–209.
43. Young TA, Wilson DA. Frequency-dependent modulation of inhibition in the rat olfactory bulb. *Neurosci. Lett*. 1999; 276:65–67. [PubMed: 10586976]
44. Crochet S, Petersen CCH. Correlating whisker behavior with membrane potential in barrel cortex of awake mice. *Nat Neurosci*. 2006; 9:608–610. [PubMed: 16617340]
45. Sommer MA, Wurtz RH. Brain circuits for the internal monitoring of movements. *Annu. Rev. Neurosci*. 2008; 31:317–338. [PubMed: 18558858]
46. Lakatos P, Connell MNO, Barczak A, Mills A, Javitt DC, Schroeder CE. The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron*. 2009; 64:419–430. [PubMed: 19914189]
47. Henry MJ, Obleser J. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U.S.A.* 2012; 109:20095–20100. [PubMed: 23151506]
48. Nobre A, Correa A, Coull J. The hazards of time. *Current Opinion in Neurobiology*. 2007; 17:465–470. [PubMed: 17709239]
49. Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE. Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *J. Neurosci*. 2011; 31:3176–3185. [PubMed: 21368029]
50. Saleh M, Reimer J, Penn R, Ojakangas CL, Hatsopoulos NG. Fast and Slow Oscillations in Human Primary Motor Cortex Predict Oncoming Behaviorally Relevant Cues. *Neuron*. 2010; 65:461–471. [PubMed: 20188651]

51. Nácher V, Ledberg A, Deco G, Romo R. Coherent delta-band oscillations between cortical areas correlate with decision making. *Proc. Natl. Acad. Sci. U.S.A.* 2013; 110:15085–15090. [PubMed: 23980180]
52. Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.* 2013; 36:313–336. [PubMed: 23725000]
53. Renoult L, Roux S, Riehle A. Time is a rubberband: neuronal activity in monkey motor cortex in relation to time estimation. *Eur. J. Neurosci.* 2006; 23:3098–3108. [PubMed: 16820000]
54. Kilavik BE, Confais J, Ponce-Alvarez A, Diesmann M, Riehle A. Evoked potentials in motor cortical local field potentials reflect task timing and behavioral performance. *Journal of Neurophysiology.* 2010; 104:2338–2351. [PubMed: 20884766]
55. Erb J, Henry MJ, Eisner F, Obleser J. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J. Neurosci.* 2013; 33:10688–10697. [PubMed: 23804092]
56. Rizzolatti G, Riggio L, Dascola I, Umiltà C. Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia.* 1987; 25:31–40. [PubMed: 3574648]
57. Smith DT, Schenk T. The Premotor theory of attention: time to move on? *Neuropsychologia.* 2012; 50:1104–1114. [PubMed: 22306518]
58. Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. *Nat Rev Neurosci.* 2008; 9:587–600. [PubMed: 18641666]
- *59. Schneider DM, Nelson A, Mooney R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature.* 2014; 513:189–194. By revealing a direct projection from secondary motor cortex to auditory cortex, the authors highlight an anatomical pathway that could optimally transmit corollary discharge signals in mice. [PubMed: 25162524]
60. Repp BH, Penel A. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol Res.* 2004; 68:252–270. [PubMed: 12955504]
61. Donoghue JP, Sanes JN, Hatsopoulos NG, Gaál G. Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements. *Journal of Neurophysiology.* 1998; 79:159–173. [PubMed: 9425187]
62. Smiley JF, Hackett TA, Ulbert I, Karmas G, Lakatos P, Javitt DC, Schroeder CE. Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. *J. Comp. Neurol.* 2007; 502:894–923. [PubMed: 17447261]
63. Rimmele J, Jolsvai H, Sussman E. Auditory target detection is affected by implicit temporal and spatial expectations. *J Cogn Neurosci.* 2011; 23:1136–1147. [PubMed: 20146603]
64. Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* 2006; 26:4970–4982. [PubMed: 16672673]
65. Doherty JR, Rao A, Mesulam MM, Nobre AC. Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 2005; 25:8259–8266. [PubMed: 16148233]
- *66. Rohenkohl G, Gould IC, Pessoa J, Nobre AC. Combining spatial and temporal expectations to improve visual perception. *J Vis.* 2014; 14 This behavioral study orthogonalized the influences of temporal and spatial predictions on visual perception. The authors highlight that temporal predictions are conditional upon the co-occurrence of spatial predictions.
67. Lakatos P, Chen C-M, O'Connell MN, Mills A, Schroeder CE. Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron.* 2007; 53:279–292. [PubMed: 17224408]
68. Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. *Psychol Rev.* 1967; 74:431–461. [PubMed: 4170865]
69. Lakatos P. An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology.* 2005; 94:1904–1911. [PubMed: 15901760]
70. Kajikawa, Y.; Haegens, S.; Schroeder, CE. Entrainment of beta band power in Macaque S2. *Soc Neurosci Annual Meeting*; Wasjington, DC: Society for Neuroscience; 2014. Poster 828no volume
71. Blakemore SJ, Wolpert D, Frith C. Why can't you tickle yourself? *Neuroreport.* 2000; 11:R11–6. [PubMed: 10943682]

Figure 1. Multiple time scales of speech and auditory brain rhythms

(A) Time-frequency decomposition of a sentence envelope (a 1/f detrending on the signal's amplitude was applied for visualization purpose). (Upper inset) Sentence waveform and its corresponding envelope (black thick line). (Right inset) Modulation spectrum. (B) Time-frequency decomposition of the spontaneous activity recorded in the supragranular primary auditory cortex of an awake macaque monkey. (Upper inset) Current source density (CSD) raw data. (Right inset) Modulation spectrum. Adapted from [69]. (C) Relationship between excitability, as indexed by the action potential firing rate (red), and the phase of oscillation, as indexed by a local field potential (blue); neuronal oscillations have optimal (high excitability) and non-optimal (low excitability) phases. (D) Schematic highlighting the typical 'nested' nature of oscillations. The (top) green trace illustrates a typical recording of oscillations. The traces below illustrate the individual oscillatory components in the delta (1.5Hz), theta (7Hz) and low gamma (35Hz) bands that comprise the composite waveform. Hierarchically organized phase-amplitude coupling between frequencies is present. (C-D) Adapted from [14]. (E) From top to bottom: **a.** Low-frequency LFP, with phase angle color-coded. **b.** Binned LFP phase during 30 trials. **c.** Corresponding spike raster plot. **d.** Mean firing rate across trials: The colored bars indicate the typical phase at each peak. Gray lines mark pairs of peaks with similar mean firing rate but different phase angle. The two instances in each pair (e.g. a1, a2) can be distinguished based on LFP phase but not based on the firing rate. Adapted from [32].

Figure 2. Regional connections of auditory cortex in the macaque brain

(A) Topographic connections of belt/parabelt auditory areas with areas of prefrontal, posterior parietal, superior temporal, and visual cortex (red: rostral areas; blue: caudal areas). All connections are reciprocal. Line thickness denotes relative strength. Primary (core) areas (not shown) have extremely limited connections outside of auditory cortex. (B) Schematic diagram summarizing connections of belt/parabelt areas with sensory, multisensory, and motor areas. Note that connections between auditory and premotor/motor areas have not been established in primates, and are either weak or absent. (C) Schematic diagram of principal routes of information flow between major regions of cortex, basal ganglia, and thalamus. Corticostriatal projections from most areas are topographically organized and not reciprocal. (D) Summary of topographic projections from belt/parabelt auditory cortex to caudate and putamen. Topographical relationships indicated by color gradient. Primary (core) auditory cortex does not project to the striatum in primates, or projections are limited. Abbreviations: CB/CPB, caudal belt and parabelt; RB/RPB, rostral belt and parabelt; M1, primary motor cortex; S2, second somatosensory area; PV, parietoventral somatosensory area; Ri, retroinsular area; STGr, rostral superior temporal gyrus; TPO, temporal polysensory area; Tpt, temporal parietotemporal area; VIP, ventral intraparietal area; V1/V2, visual areas 1 and 2; ProS, area prostriata (visual); VA, ventral anterior nucleus; VL, ventral lateral nucleus; MD, medial dorsal nucleus; GP, globus pallidus; NB, nucleus basalis.

Figure 3.

Summary of the average temporal patterns of MUA (blue) and the CSD-derived beta band power (red) in the supragranular layer of the second somatosensory area (S2) recorded from one monkey during performance of an audio-visual (AV) oddball task [70]. Although the monkey held a lever, a series of audiovisual stimuli were presented. Data are aligned to the onset (0 sec) of AV stimuli during the task. Stimuli were repeating 500 ms vocal movie clips (SOA = 1.4 sec; static-to-movie face) with random oddballs (20%) in which face or voice differed from the standard. Upon detection of an oddball, monkey released the lever for a reward. The black dashed line indicates the mean reaction times on the adjoining target trials. For each signal, mean and 95% confidence intervals are shown (n=16, bootstrap). Each signal was normalized by the standard deviation estimated from periods of -2 to 2 sec from the onset of stimuli. MUA was sampled at 2 kHz, CSD power was sampled at 0.5 kHz.

Figure 4. Auditory active sensing

(A) A forward model (corollary discharge/efference copy) predicts the sensory consequences of a movement based on the motor command. When a movement is self-produced, its sensory consequences can be accurately predicted and this prediction can be used to attenuate the sensory effects of the movement. Adapted from [71]. (B) Example of auditory active sensing paradigm. Participants listen to a sequence of pure tones and then estimate the average pitch of targets, while ignoring distractors. First row: rhythmic motor tracking in phase with the reference beat throughout the sequence. Second row: references indicating the beat. Third (fourth) rows: targets (distractors) presented in phase (antiphase) with the reference beat. Dark (light) grey lines indicate the temporal distance between the motor act and the onset of the target (distractor). Fifth row: sensory gains assigned to successive targets and distractors. The shorter the temporal distance between a tone and a motor act, the stronger the temporal prediction, and the higher the sensory gain assigned to the tone. Adapted from [10*].