



Published in final edited form as:

Ann N Y Acad Sci. 2015 March ; 1337(1): 26–31. doi:10.1111/nyas.12629.

Neuronal oscillations as a mechanistic substrate of auditory temporal prediction

Benjamin Morillon^{1,2} and Charles E. Schroeder^{1,3}

¹Department of Psychiatry, Columbia University Medical Center, New York, New York

²McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

³Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, New York

Abstract

Neuronal oscillations are comprised of rhythmic fluctuations of excitability that are synchronized in ensembles of neurons and thus function as temporal filters that dynamically organize sensory processing. When perception relies on anticipatory mechanisms, ongoing oscillations also provide a neurophysiological substrate for temporal prediction. In this article we review evidence for this account with a focus on auditory perception. We argue that such “oscillatory temporal predictions” can selectively amplify neuronal sensitivity to inputs that occur in a predicted, task-relevant rhythm and optimize temporal selection. We elaborate this argument for a prototypic example, speech processing, where information is present at multiple time scales, with delta, theta, and low-gamma oscillations being specifically and simultaneously engaged, enabling multiplexing. We then consider the origin of temporal predictions, specifically the idea that the motor system is involved in the generation of such prior information. Finally, we place temporal predictions in the general context of internal models, discussing how they interact with feature-based or spatial predictions. We propose that complementary predictions interact synergistically according to a dominance hierarchy, shaping perception in the form of a multidimensional filter mechanism.

Keywords

perception; sensorimotor; expectation; neurophysiology; active sensing

Neuronal oscillations underpin temporal predictions

Seminal theories in neuroscience consider sensory processing as an inference problem: given the noisy sensory and neural signals, the brain estimates which stimuli caused these observations.¹ Accordingly, neuronal representations start to be viewed as continuous top-

Address for correspondence: Benjamin Morillon, McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, QC, Canada H3A 2B4. bnmorillon@gmail.com.

Conflicts of interest

The authors declare no conflicts of interest.

down prediction signals, that is, internal models of causal dynamics in the world shaped by one's perceptual, behavioral, and emotional experiences.² Such theories have two far-reaching corollaries: (1) predictions are reflected in the brain's intrinsic activity;³ and (2) sensory-evoked neuronal activity corresponds to the perturbation of ongoing circuit dynamics by input signals, initiating the generation of an updated representation of the world.⁴⁻⁷

However, it remains unclear how predictions are represented in the brain. One problem resides in the fact that an event can be anticipated with regard to its features (content), location (space), and/or moment (time) of occurrence.⁸⁻¹⁰ Focusing on this latter dimension, dynamic attending theory suggests that during perception attention is modulated dynamically to optimize sensory processing at expected points in time.¹¹⁻¹³ Critically, this framework capitalizes on the fact that many stimuli and actions are rhythmically organized (e.g., speech, music, walking). By extracting such temporal regularities, the brain is able to predict the occurrence of subsequent events of interest and to optimize their processing.

Behavioral studies have shown that temporal expectations indeed optimize perception by dynamically modulating the gain of sensory information.¹⁴⁻¹⁶ The signal-to-noise ratio of the predicted sensory signal is increased, which leads to improved detection performance. This effect is associated with a reduction of the large-scale sensory evoked neuronal response, as observed in both visual^{17,18} and auditory modalities.¹⁹⁻²¹ At the neurophysiological level, low-frequency neuronal oscillations are the hypothesized substrate of temporal predictions.⁵ Neuroelectric oscillations correspond to temporal fluctuations of the excitability state of neuronal ensembles²² and make it possible to selectively amplify neuronal input responses occurring at predicted moments.^{23,24} 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.²⁵

Speech processing at multiple time scales

Speech, together with music, is arguably the most interesting stimulus to study the role of neuronal oscillations in perception. Speech perception arises from the dynamic sampling of acoustic information at multiple time scales simultaneously, referred to as multiplexing. Indeed, although syllabic modulations (1–7 Hz) are critical to speech comprehension,²⁶ acoustic modulations above 12 Hz also contribute to speech identification, as they are associated with some phonetic feature dimensions.^{27,28} In recent years neuronal entrainment of cortical oscillatory dynamics to the stimulus acoustic envelope has been described, suggesting a prominent role of auditory delta (1–4 Hz), theta (4–8 Hz), and low-gamma (25–45 Hz) frequencies in speech parsing.²⁹ In addition to syllabic-based sampling at theta rate, fine-grained sampling (~ 30 Hz) preserves the detailed temporal structure of sounds necessary for phonemic contrasts extraction and articulatory routines execution;³⁰ slower sampling (<4 Hz) is suggested to underline more steady cues such as formants at the basis of vowel recognition but also vocal analysis (e.g., speaker identification, emotional prosody).

Most studies of speech perception emphasize that auditory responses phase-lock to speech syllabic modulations in the theta range.^{31,32} However, in left auditory regions neuronal

Active sensing in the auditory domain

The involvement of sensorimotor loops in perception is often referred to as active sensing.^{47,57} This framework proposes that perception is shaped by the motor system in two distinct ways. First, the motor system directs sensing organs (e.g., finger squeezes, ocular saccades, olfactory sniffs) toward relevant stimuli, thereby structuring the content of bottom-up sensory information inflow. Second, it modulates the processing of sensory information via top-down internal corollary discharge signals, that is, copies of movement commands sent to associated sensory structures.⁵⁸ This latter mechanism corresponds to an active sampling of sensory information by marking the timing (and location) of relevant sensory events according to current motor plans. This predictively modulates sensory processing according to the temporal (and spatial) patterns of motor activity patterns, thus providing “when” (and “where”) predictions at a minimum.⁵⁹

Active sensing was never described in the auditory domain because of its relative bottom-up disconnection from the motor system. Indeed, whereas bottom-up and top-down motor influences are contingent in most sensory modalities, audition is the exception because we cannot selectively move our ears. However, sensitivity to temporal information is at its best in the auditory modality, and music exemplifies the strong relationship existing between rhythm and auditory and motor systems.^{53,60} This suggests that (1) the motor system could exert a specific top-down predictive influence on auditory processing, and (2) top-down motor influences, often conflated with bottom-up ones (the directing of sensing organs), could play a distinct and fundamental role.

Motor contributions to the temporal precision of auditory attention

We recently developed a mechanistic behavioral account of auditory active sensing.⁶¹ This work was based on four theoretically driven parameters. First, we focused on the auditory modality because of its relative bottom-up disconnection from the motor system. Second, we measured the ability of human participants to extract relevant auditory information interleaved with distracting information, in a cocktail party–like fashion, and embedded in rhythmic streams. Thus, performance critically depended on the implementation of temporal expectation. Third, we used a perceptual decision-making task to study the dynamics of evidence accumulation and its modulation by temporal expectation. This led us to the following experiment. Participants were asked to categorize sequences of eight pure tones as higher or lower pitched, on average, than a reference frequency. In order to drive rhythmic fluctuations in attention, the tones (targets) were delivered in phase with a reference beat, presented at the beginning of each trial, and in antiphase with irrelevant yet physically indistinguishable tones (distractors). Fourth, participants were instructed either to use their motor system overtly, by pressing in rhythm with the reference beat, or to keep the rhythm covertly, in both cases as a way to optimize the allocation of temporal attention (predictions) in a purely top-down fashion.

Our findings show that producing a rhythmic movement engages a top-down modulation that sharpens the temporal selection of auditory information. It improves the segregation between relevant and distracting information, facilitating perception of relevant items and

suppressing perception of irrelevant ones. We additionally observed that the impact of overt motor tracking on perception parametrically scales with the temporal predictability of the auditory sequence. Finally, motor contributions depend on the temporal alignment between motor and attention fluctuations. This indicates that this rhythmic top-down process requires an exquisite temporal alignment among motor, attention, and auditory systems during the sampling of sensory information. Altogether, our results suggest that motor-driven temporal predictions optimize perception, and a most likely substrate for this effect is a slow oscillation timed by rhythmic motor activity.

Previous behavioral experiments on the top-down role of the motor system in auditory perception showed that self-initiated head movements can reset the perceptual organization of an auditory scene⁶² or influence the way an auditory rhythmic pattern is perceived.⁶³ Here we went several important steps further by firmly isolating top-down motor influences on auditory stimulus processing and by showing that the engagement of the motor system improves (1) the precision of temporal attention and, thus, (2) the quality of sensory selection.⁶¹ Our work confirms the operation of active sensing in the auditory domain, which emphasizes the prominent role of motor activity in sensory processing through its close interaction with attention.

Hierarchical organization of predictive filters in auditory perception

A first necessary step toward a Cartesian approach⁶⁴ is to study temporal expectations independently of other dimensions (such as location or feature). However, the drawback in most paradigms is that these attributes of noninterest are kept constant across trials and so are highly predictable. These strong (e.g., spatial or feature-based) expectations might in turn impact perception and bias results or limit interpretations.⁶⁵ Indeed, in primary auditory cortex the combination of temporal and spectral expectations results in both the amplification and sharpening of neuronal responses, in essence acting as a spectrotemporal filter mechanism.²³ Such a combination is not restricted to what and when dimensions in audition but can also integrate spatial⁶⁶ and intensity⁶⁷ information at higher levels. These data suggest that predictions constantly fuse into a dynamically evolving model of upcoming events, which enhances the representation of predicted stimuli.

However, predictions might not combine linearly. For example, although temporal expectations alone influence behavior during a visual task,⁶⁸ their impact is increased by concurrent spatial expectations and is erased if spatial expectations are invalid.^{68,69} This indicates that the impact of temporal expectations on perception is optimized by the co-occurrence of spatial expectations. Because the visual cortex is retinotopically (i.e., spatially) organized, temporal predictions might not be able to efficiently modulate the first cortical stages of visual processing unless they are part of a spatiotemporal predictive filter that taps onto retinotopically specific regions. This suggests that the influence expectations exert on perception might rely on a hierarchy of predictive filters that trails the structure of the concerned sensory system. Spatial predictions would be necessary in the retinotopic visual system, whereas spectral predictions would be key in the tonotopic auditory system.

Overall, current research suggests that perception is a rhythmic process⁷⁰ and features neuronal oscillations and the motor system as key operators in sensory processing.⁴⁷ Recent evidence in developmental neurosciences accordingly pinpoints temporal sampling deficits as a potential cause of developmental dyslexia.^{71,72} It moreover suggests that music, which is exquisitely rhythmic and linked to motor activity,⁵³ might be used as a rehabilitation strategy: by boosting temporal prediction mechanisms, musical rhythmic training could ultimately benefit language processing and language acquisition.⁷³

Acknowledgments

Supported by the Fyssen foundation (B.M.) and by NIH (DC011490) (C.E.S.). We thank our reviewers for their helpful comments on the manuscript.

References

1. Helmholtz, H von. Handbook of Physiological Optics. Leipzig: Leopold Voss; 1856.
2. Friston K. A theory of cortical responses. *Phil. Trans. R. Soc. B: Biol. Sci.* 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. Ringach DL. Spontaneous and driven cortical activity: implications for computation. *Curr. Opin. Neurobiol.* 2009; 19:439–444. [PubMed: 19647992]
5. Arnal L, Giraud AL. Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 2012; 16:390–398. [PubMed: 22682813]
6. Fiser J, Berkes P, Orbán G, Lengyel M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* 2010; 14:119–130. [PubMed: 20153683]
7. Winkler I, Denham SL, Nelken I. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends Cogn. Sci.* 2009; 13:532–540. [PubMed: 19828357]
8. Nobre, A.; Rohenkohl, G.; Stokes, M.; Nobre, M. Nervous anticipation: top-down biasing across space and time. In: Posner, MI., editor. *Cognitive Neuroscience of Attention*. 2nd ed.. New York: Guilford; 2012. p. 159-186.
9. Olson IR, Chun MM. Temporal contextual cuing of visual attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 2001; 27:1299–1313. [PubMed: 11550756]
10. Kingstone A. Combining expectancies. *Q. J. Exp. Psychol. A.* 1992; 44:69–104.
11. Large EW, Jones MR. The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 1999; 106:119–159.
12. Jones MR. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychol. Rev.* 1976; 83:323–355. [PubMed: 794904]
13. Jones MR, Boltz M. Dynamic attending and responses to time. *Psychol. Rev.* 1989; 96:459–491. [PubMed: 2756068]
14. Jones MR, Moynihan H, MacKenzie N, Puente J. Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol. Sci.* 2002; 13:313–319. [PubMed: 12137133]
15. Rohenkohl G, Cravo AM, Wyart V, Nobre AC. Temporal expectation improves the quality of sensory information. *J. Neurosci.* 2012; 32:8424–8428. [PubMed: 22699922]
16. Jaramillo S, Zador AM. The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat. Neurosci.* 2011; 14:246–251. [PubMed: 21170056]
17. Praamstra P, Kourtis D, Kwok HF, Oostenveld R. Neurophysiology of implicit timing in serial choice reaction-time performance. *J. Neurosci.* 2006; 26:5448–5455. [PubMed: 16707797]
18. Alink A, Schwiedrzik CM, Kohler A, et al. Stimulus predictability reduces responses in primary visual cortex. *J. Neurosci.* 2010; 30:2960–2966. [PubMed: 20181593]
19. Lange K. Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain Cogn.* 2009; 69:127–137. [PubMed: 18644669]

20. 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]
21. Yaron A, Hershenhoren I, Nelken I. Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron.* 2012; 76:603–615. [PubMed: 23141071]
22. Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 2009; 32:9–18. [PubMed: 19012975]
23. Lakatos P, et al. The spectrotemporal filter mechanism of auditory selective attention. *Neuron.* 2013; 77:750–761. [PubMed: 23439126]
24. 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]
25. Zion Golumbic EM, et al. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron.* 2013; 77:980–991. [PubMed: 23473326]
26. Elliott TM, Theunissen FE. The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.* 2009; 5:e1000302. [PubMed: 19266016]
27. Christiansen TU, Greenberg S. Frequency selective filtering of the modulation spectrum and its impact on consonant identification. *Linguistic Theory Raw Sound.* 2009:119–133.
28. 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]
29. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 2012; 15:511–517. [PubMed: 22426255]
30. Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. *Nature.* 2013
31. Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J. Neurosci.* 2008; 28:3958–3965. [PubMed: 18400895]
32. Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron.* 2007; 54:1001–1010. [PubMed: 17582338]
33. Morillon B, Liégeois-Chauvel C, Arnal L, et al. Asymmetric function of theta and gamma activity in syllable processing: an intra-cortical study. *Front. Psychol.* 2012; 3:248. [PubMed: 22833730]
34. Gross J, et al. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* 2013; 11:e1001752. [PubMed: 24391472]
35. Xiang J, Poeppel D, Simon JZ. Physiological evidence for auditory modulation filterbanks: cortical responses to concurrent modulations. *J. Acoust. Soc. Am.* 2013; 133:EL7–E12. [PubMed: 23298020]
36. Ahissar E, et al. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 2001; 98:13367–13372. [PubMed: 11698688]
37. Peelle JE, Gross J, Davis MH. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex.* 2012
38. Luo H, Tian X, Song K, et al. Neural response phase tracks how listeners learn new acoustic representations. *Curr. Biol.* 2013
39. Giraud AL, et al. Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron.* 2007; 56:1127–1134. [PubMed: 18093532]
40. Morillon B, et al. Neurophysiological origin of human brain asymmetry for speech and language. *Proc. Natl. Acad. Sci. U.S.A.* 2010; 107:18688–18693. [PubMed: 20956297]
41. Arieli A, Sterkin A, Grinvald A, Aertsen A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science.* 1996; 273:1868–1871. [PubMed: 8791593]
42. Cogan GB, Poeppel D. A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *J. Neurophysiol.* 2011; 106:554–563. [PubMed: 21562190]
43. Schroeder CE, Lakatos P, Kajikawa Y, et al. Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 2008; 12:106–113. [PubMed: 18280772]
44. Arnal LH, Doelling KB, Poeppel D. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex.* 2014

45. Fuster JM. Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. *Ann. N.Y. Acad. Sci.* 1990; 608:318–329. discussion 330–336. [PubMed: 2127512]
46. Grahn JA, Rowe JB. Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. *Cereb. Cortex.* 2012
47. Schroeder CE, Wilson DA, Radman T, et al. Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* 2010; 20:172–176. [PubMed: 20307966]
48. Schubotz RI. Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 2007; 11:211–218. [PubMed: 17383218]
49. Bolton TL. *Rhythm.* Am. J. Psychol. 1894
50. Morillon B, Kell CA, Giraud AL. Three stages and four neural systems in time estimation. *J. Neurosci.* 2009; 29:14803–14811. [PubMed: 19940175]
51. Coull JT, Vidal F, Nazarian B, Macar F. Functional anatomy of the attentional modulation of time estimation. *Science.* 2004; 303:1506–1508. [PubMed: 15001776]
52. Grahn JA, Rowe JB. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 2009; 29:7540–7548. [PubMed: 19515922]
53. 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]
54. 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]
55. Coull JT, Nobre AC. 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.* 1998; 18:7426–7435. [PubMed: 9736662]
56. Saleh M, Reimer J, Penn R, et al. Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron.* 2010; 65:461–471. [PubMed: 20188651]
57. Wachowiak M. All in a sniff: olfaction as a model for active sensing. *Neuron.* 2011; 71:962–973. [PubMed: 21943596]
58. Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* 2008; 9:587–600. [PubMed: 18641666]
59. Sommer MA, Wurtz RH. Brain circuits for the internal monitoring of movements. *Annu. Rev. Neurosci.* 2008; 31:317–338. [PubMed: 18558858]
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. Morillon B, Schroeder CE, Wyart V. Motor contributions to the temporal precision of auditory attention. *Nat. Commun.* 2014; 5:5255. [PubMed: 25314898]
62. Kondo HM, Pressnitzer D, Toshima I, Kashino M. Effects of self-motion on auditory scene analysis. *Proc. Natl. Acad. Sci. U.S.A.* 2012; 109:6775–6780. [PubMed: 22493250]
63. Phillips-Silver J, Trainor LJ. Feeling the beat: movement influences infant rhythm perception. *Science.* 2005; 308:1430. [PubMed: 15933193]
64. Descartes R. *Discours de la méthode (Pour bien conduire sa raison et chercher la vérité dans les sciences).* 1637
65. Morillon B, Barbot A. Attention in the temporal domain: a phase-coding mechanism that controls the gain of sensory processing. *Front. Hum. Neurosci.* 2013; 7:e480.
66. 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]
67. 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]
68. 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]
69. Rohenkohl G, Gould IC, Pessoa J, Nobre AC. Combining spatial and temporal expectations to improve visual perception. *J. Vis.* 2014; 14
70. VanRullen R, Zoefel B, Ilhan B. On the cyclic nature of perception in vision versus audition. *Phil. Trans. R. Soc. B: Biol. Sci.* 2014; 369 20130214.

71. Goswami U. A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 2011; 15:3–10. [PubMed: 21093350]
72. Lehongre K, Ramus F, Villiermet N, et al. Altered low- γ sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron.* 2011; 72:1080–1090. [PubMed: 22196341]
73. Schön D, Tillmann B. Short- and long-term rhythmic interventions: perspectives for language rehabilitation. *Ann. N.Y. Acad. Sci.* In press.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript