



HAL
open science

Proactive sensing of periodic and aperiodic auditory patterns

Johanna M Benjamin Rimmele, Benjamin Morillon, David Poeppel, Luc Arnal

► **To cite this version:**

Johanna M Benjamin Rimmele, Benjamin Morillon, David Poeppel, Luc Arnal. Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences*, 2018, 22 (10), pp.870-882. 10.1016/j.tics.2018.08.003 . hal-03583582

HAL Id: hal-03583582

<https://hal.science/hal-03583582>

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Trends in Cognitive Sciences

Proactive sensing of periodic and aperiodic auditory patterns

--Manuscript Draft--

Manuscript Number:	TICS-D-18-00056R2
Article Type:	Opinion
Corresponding Author:	Johanna M. Rimmele, Ph.D. Frankfurt am Main, GERMANY
First Author:	Johanna M. Rimmele, Ph.D.
Order of Authors:	Johanna M. Rimmele, Ph.D. Benjamin Morillon, PhD David Poeppel, PhD Luc H. Arnal, PhD
Abstract:	<p>The ability to predict when something will happen facilitates sensory processing and the ensuing computations. Building on the observation that neural activity entrains to periodic stimulation, leading neurophysiological models imply that temporal predictions rely on oscillatory entrainment. While they provide a sufficient solution to predict periodic regularities, these models are challenged by a series of findings that question their suitability to account for temporal predictions based on aperiodic regularities. Aiming for a more comprehensive model of how the brain anticipates “when” in auditory contexts, we emphasize the capacity of motor and higher order top-down systems to prepare sensory processing in a proactive and temporally flexible manner. Focusing on speech processing, we illustrate how this framework leads to new hypotheses.</p>

1 **Highlights**

- 2 • We review research on temporal predictions (particularly in auditory contexts) and
3 consider how recent empirical evidence challenges standard oscillatory entrainment
4 models.
- 5 • Top-down phase resetting mediated by the motor system or higher order language or
6 attention related systems facilitate bottom-up perceptual processing.
- 7 • A model that considers neuronal oscillations as intrinsic temporal constraints (rather than
8 subserving a specific function) and incorporates top-down phase resetting is suggested as
9 a parsimonious solution for both periodic and aperiodic temporal predictions.
- 10 • Assuming that temporal predictions prospectively control oscillatory constraints, the
11 proposed perspective sets a new frame for the interpretation of neurophysiological
12 responses to auditory and speech streams.

13

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Proactive sensing of periodic and aperiodic auditory patterns**

2 Rimmele, J.M.^{1*}, Morillon, B.², Poeppel, D.^{1,3}, Arnal, L.H.⁴

3

4 1 Department of Neuroscience

5 Max-Planck-Institute for Empirical Aesthetics

6 Grüneburgweg 14

7 D - 60322 Frankfurt am Main, Germany

8

9 2 Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst

10 27, Boulevard Jean Moulin

11 13005 Marseille, France

12

13 3 Department of Psychology and Center for Neural Science

14 New York University

15 6 Washington Place

16 New York , NY 10003

17

18 4 Department of Fundamental Neuroscience

19 University of Geneva - Biotech Campus;

20 Geneva 7 1202; Switzerland.

21

22 Correspondence: j.rimmele@ae.mpg.de (J.M. Rimmele), luc.arnal@unige.ch (L.H. Arnal)

23

24 **Keywords:** oscillation; prediction; auditory perception; speech; motor; entrainment

25

26

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Abstract**

2 The ability to predict when something will happen facilitates sensory processing and the ensuing
3 computations. Building on the observation that neural activity entrains to periodic [stimulation](#),
4 leading neurophysiological models imply that temporal predictions rely on oscillatory entrainment.
5 While they provide a sufficient solution to predict periodic regularities, these models are challenged
6 by a series of findings that question their suitability to account for temporal predictions based on
7 aperiodic regularities. Aiming for a more comprehensive model of how the brain anticipates
8 “when” [in auditory contexts](#), we emphasize the capacity of motor and higher order top-down
9 systems to prepare sensory processing in a proactive and temporally flexible manner. [Focusing on](#)
10 [speech processing](#), we illustrate how this framework leads to new hypotheses.

11

1 **Temporal predictions in auditory contexts**

2 We extract **temporal regularities** from the sensory environment to anticipate upcoming events.
3 Although there exists an infinite number of possible temporal patterns or cues that one can exploit
4 for prediction, the **periodic** (isochronous, often informally referred to as **rhythmic**; see glossary)
5 occurrence of a sound is perhaps the first regularity that comes to mind. The human brain is
6 captivated by periodic streams, and human listeners –as well as **some** other species– seem
7 compelled to synchronize gestures to predictable rhythmic beats [1]. One possible explanation of
8 this drive is that proactively ‘**tracking** the future’ optimizes processing of incoming information
9 and the adaption of behavior [2]. The majority of existing data **is** compatible with the notion that
10 the **entrainment** of neural activity to periodic stimulations is instrumental to sensory selection,
11 facilitating the processing of events that occur around the entrained, and therefore expected beat
12 [3]. Although periodicity is the most obvious form of temporal regularity and is used in various
13 sonic contexts (**speech**, music), limiting the scope of temporal processing to this type of pattern (or
14 studying it in isolation from other types of patterns) biases and restricts the generalizability of
15 current models. **Speech**, for instance, contains quasi-periodic energy fluctuations at the syllabic
16 rate [4]. **However**, aperiodic cues, such as provided by syntactic or semantic information, can also
17 be used to predict upcoming speech events, e.g. the pauses that typically follow prepositional
18 phrases [5]. **Although** generally periodic, music often contains aperiodic but trackable, regular
19 meter at the beat or subdivision level, as featured in music from the Balkan or Africa, amongst
20 others [6]. **Whether** there exists a principled reason to assume different neuronal mechanisms for
21 periodic compared to other types of regularities is unclear. At any rate, invoking multiple models
22 to account for the same process comes at the expense of **parsimony**.

23 We briefly review the main results that have contributed to neuronal (oscillatory) entrainment (Box
24 1) as the mainstream mechanistic explanation for periodic temporal predictions. We then argue that
25 recent evidence of perceptual facilitation by aperiodic predictions [7,8] is not compatible with
26 classical models of oscillatory entrainment. **Crucially**, we suggest that this tension calls for a new
27 model of how temporal predictions **modulate** intrinsic oscillations to enhance sensory processing.
28 First, we **advance the view that** neuronal oscillations constitute a ‘processing **constraint**’ (Box 2)
29 rather than a specific temporal prediction mechanism **operating through entrainment**. **We argue that**
30 **this view requires introducing the notion of a flexible (e.g. aperiodic) top-down phase-reset** of
31 neuronal oscillations into the stimulus-**driven** model. **Concerning** the putative sources of top-down

1 modulation, we suggest that in the context of temporally structured auditory sequences, the motor
2 system constitutes a plausible candidate to generate flexible (periodic and aperiodic) temporal
3 predictive signals and to influence perception in a top-down manner [9–13] (Box 3). We also
4 consider other sources of temporal predictions related to distinct cognitive processes [10,14–16].
5 Finally, focusing on speech as an informationally complex but predictable signal (containing quasi-
6 periodic and aperiodic regularities), we exemplify how the framework we advance improves the
7 explanatory power of classical entrainment models and generates novel testable hypotheses.

8 The classical entrainment model and temporal predictions

9 *Entrainment aligns behavior with periodic events*

10 Temporal predictions have been studied in paradigms where predictions are set-up in a bottom-up,
11 stimulus-driven manner by periodic stimulation. Typically, auditory targets, are detected more
12 easily when presented on-time in an isochronous sequences than when presented early or late, or
13 preceded by a temporally random tone sequence [17]. A by-now classical view, the ‘**Dynamic**
14 **Attending Theory**’ (DAT; [2,18]), proposed that attention is directed in time through the
15 entrainment of neuronal oscillations to such periodic stimulation (Box 1). Oscillatory entrainment
16 (as measured using electrophysiological recording) refers to the **phase alignment** and frequency
17 adaptation of endogenous neuronal oscillators, to a periodic stimulation (Fig. 1 A2, Key Figure).
18 In the DAT view, entrainment narrows the time-based attentional focus and optimizes perception
19 [18]. This theory has strongly influenced visual, auditory, attention, music, and speech processing
20 research [19]. However, while the DAT –and more generally oscillatory entrainment theories–
21 provide an efficient and plausible neural mechanism to serve periodic temporal predictions, it
22 cannot account for the brain’s capacity to generate –and benefit from– temporal predictions
23 inferred from aperiodic patterns.

24 *Aperiodic temporal predictions: aperiodic but predictable streams*

25 The notion of temporal predictions is often confounded with the idea of entrainment to periodic
26 sensory inputs. The latter mechanism, however, cannot account for temporal predictions inferred
27 from heterochronous streams of events [8,20], **symbolic cues** (‘memory-based’ predictions, e.g. a
28 yellow traffic light that indicates a switch to red in a few seconds), or **hazard functions**
29 (‘probability-based’ predictions, e.g. the increasing conditional probability over time that an event

1
2
3
4 1 will occur given that it has not already occurred) [21,22]. For example, aperiodic temporal
5
6 2 predictions occur in language (cf. last section), or in music, where many musical traditions employ
7
8 3 non-isochronous (aperiodic) meter, which builds up temporal expectations and can be embodied in
9
10 4 dance [6]. Comparing the ability of listeners to extract temporal regularities from periodic as well
11
12 5 as aperiodic but –implicitly– predictable sound streams, a recent study [8] showed that performance
13
14 6 is equally enhanced for predictable conditions (whether periodic or not) as compared to
15
16 7 unpredictable ones. Another study [7] compared the neural responses to periodic and aperiodic
17
18 8 predictions and demonstrated that periodic and aperiodic temporal predictions induce similar phase
19
20 9 alignment of low-frequency activity prior to an expected target (see also: [23]). These observations
21
22 10 contrast with entrainment models, which hypothesize that reducing stimulus periodicity should
23
24 11 result in reduced phase-alignment. **Importantly**, these findings suggest that a single neural substrate
25
26 12 could account for both **periodic and aperiodic** predictions at the sensory level (for criticism see:
27
28 13 [24]).

29 14 **The ‘rhythmic attending mode’** proposed by the DAT (where attention is directed in time through
30
31 15 the entrainment of neuronal oscillations; Fig. 1 A2) can be extended by multiple oscillators being
32
33 16 hierarchically organized (nested), and thus indirectly entrained to complex, multi-scale, rhythms
34
35 17 [18]. **Furthermore, under sustained vigilance task demands, the DAT** proposes an ‘**analytic**
36
37 18 **attending mode’**. In this mode, in the absence of a periodic stimulation, low frequencies are
38
39 19 suppressed and high frequency activity subserves continuous processing by keeping the system in
40
41 20 a high-excitability state [17]. Such models, however, cannot provide a mechanistic solution with
42
43 21 respect to aperiodic temporal predictions. It has been proposed that aperiodic predictions rely on
44
45 22 memory based ‘interval-timing traces’, to which new intervals are compared [7], an idea
46
47 23 compatible with intrinsic models of timing [25]. Such a mechanism does not require periodicity
48
49 24 but instead relies on anticipatory processes based on the learning of temporal contingencies. At the
50
51 25 neurophysiological level, this has been associated with ‘**ramping-based models’** (for review [25])
52
53 26 where the build up of expectations is reflected in linearly increasing neuronal firing rates (**ramping**
54
55 27 **or ‘Climbing Neuronal Activity’,** CNA). In ramping-based models, recurrent excitatory
56
57 28 connections within neuronal networks integrate the spiking information to estimate interval
58
59 29 duration. Although involved in interval timing, **ramping activity** seems to encode stimulus
60
61 30 probability rather than absolute time [26]. **However, because these paradigms typically require a**
62
63 31 **behavioral response, it is difficult to exclude the possibility that ramping activity reflects** response
64
65

1 preparation, or time-dependent decision-making processes.

2 To summarize, temporal predictions are classically explained by oscillatory entrainment to periodic
3 sensory inputs. The facilitating effects of aperiodic predictions, however, cannot be accounted for
4 by entrainment, raising questions about the neural substrate of periodic and aperiodic predictions.

5 In the following, we outline an alternative approach to reconcile the findings and clarify the
6 connection between neuronal oscillations and predictive processing.

7 *Criticisms and refinements: the entrainment hypothesis revisited*

8 Whether distinct neuronal mechanisms exist for mediating periodic and aperiodic temporal
9 predictions, and whether the entrainment of low-frequency (<10 Hz) oscillations plays a **functional**
10 **role** for temporal predictions remains a matter of active debate (Box 2). One criticism about
11 attributing a function to the entrainment of neuronal oscillations is that correlation does not entail
12 causation (or specificity). That periodic auditory stimulation (as featured in speech or music, see
13 [4]) invariably entrains low-frequency neural oscillations does not imply that neural oscillations
14 play a functional role in temporal predictions.

15 We propose that neural oscillations primarily constitute an intrinsic processing constraint rather
16 than a dedicated temporal prediction mechanism operating through entrainment (Box 2). In this
17 perspective, the entrainment model (Fig. 1 A) can be refined to account for empirical observations
18 involving periodic and aperiodic predictions. The ‘constraint view’ extends the stimulus-driven
19 model by integrating the temporally flexible, internally-driven top-down alignment (**top-down**
20 **phase reset**) of neuronal oscillations to anticipated events. It assumes that neuronal oscillations do
21 not constitute a temporal prediction mechanism by itself, but are part of the cortical architecture
22 and support a more generic, canonical function, such as the message-passing of information
23 throughout the cortical hierarchy [27,28]. As a consequence, neuronal oscillations reflect
24 processing time constants, i.e. physiological (**intrinsic**) constraints due to the oscillatory
25 excitability cycles of neuronal ensembles. The preferred oscillatory frequency of these ensembles
26 determines optimal temporal frames for processing. Importantly, depending on both the networks
27 underpinning a cognitive process and their inter-connectivity and dynamics, specific temporal
28 constraints emerge [29–31]. The processing constraint view, contrasting with non-oscillatory
29 models [25,26], also accounts for empirical observations suggesting that perception is a discrete –

1
2
3
4 1 rather than continuous– process [32]. For example, auditory temporal processing is optimal within
5
6 2 certain time windows in the delta- and theta-ranges [33–35], which reflects endogenous brain
7
8 3 rhythms of auditory cortex [36–38]. Temporal constraints are also reflected in our ability to
9
10 4 understand time-compressed speech, which breaks down at compression rates exceeding the upper
11
12 5 limit of the theta range [39]. Further research will elucidate whether and under what conditions
13
14 6 plasticity can overcome these constraints.

15
16 7 Temporal predictions (periodic, aperiodic) operate upon endogenous constraints by predictively
17
18 8 aligning neuronal excitability in time to facilitate the processing of anticipated events (Fig. 1 B).
19
20 9 Top-down influences correspond to an anticipatory phase-reset (originating from higher-level
21
22 10 processes; not directly driven by low-level stimulus features) of ongoing oscillations in those
23
24 11 neuronal subpopulations involved in the processing of the expected event (e.g. beat or syllable)
25
26 12 [40]. While distinct neuronal mechanisms have been assumed for periodic (i.e., oscillatory
27
28 13 entrainment) and aperiodic (e.g., memory-based mechanisms) temporal predictions (for review:
29
30 14 [21]), we propose that entrainment of neuronal oscillations is one possible manifestation of a
31
32 15 common temporal prediction mechanism, in which stimulus-driven and top-down phase-resets co-
33
34 16 occur synchronously (Fig. 1 B). While it may not be possible to dissociate between stimulus-driven
35
36 17 and top-down influences during periodic stimulation, the evidence of phase-alignment in aperiodic
37
38 18 or contextual situations [7,40] argues in favor of the existence of internally guided, top-down
39
40 19 predictive phase-resets.

41 20 In the following, we discuss the evidence for top-down effects originating in different brain areas
42
43 21 and their role in (a)periodic temporal predictions (in the form of top-down phase resets).

44 45 22 **Top-down sources of temporal predictions**

46
47
48 23 By enabling the efficient allocation of processing resources, predicting upcoming inputs reduces
49
50 24 sensory uncertainty, thereby enhancing the processing of noisy (weak or ambiguous) inputs. To
51
52 25 reduce both external and internal noise, the brain can arguably exploit any available source of top-
53
54 26 down priors. Consequently, all neural systems that contribute to precise, time-related computations
55
56 27 might inform sensory processing in a top-down manner, without being selectively dedicated to this
57
58 28 function [25]. The next section describes the neural sources and mechanisms involved in reducing
59
60 29 sensory uncertainty in the time-domain.

1
2
3
4 1 *The motor system's expertise in temporal predictions*

5
6
7 2 Auditory perception is carried out within functionally dissociable pathways, the ventral and dorsal
8
9 3 **streams** [41,42], comparable to visual perception [43]. Following predictive processing theories
10
11 4 [44], the dorsal pathway is related to the processing of spatial and temporal features and establishes
12
13 5 predictions that pertain to 'how' and 'when' something will happen.

14
15 6 The capacity of the dorsal stream for sequence processing has been associated with the timing
16
17 7 capabilities of the motor system [45]. Precise temporal anticipation at the time-scale of tens to
18
19 8 hundreds of milliseconds is critical, as a large number of phenomena pertaining to perception and
20
21 9 action (i.e., speech **and music** perception and production, movements, etc.) occur at this scale.
22
23 10 Several studies have shown that the motor system is involved, as part of a cortico-striatal network,
24
25 11 in the processing of events and sequences at these time-scales (~above 300 ms) [46,47]. This is not
26
27 12 surprising, given the expertise of this system in monitoring the sensory feedback resulting from
28
29 13 motor outputs. The motor system is not only involved in motor-timing, however, but also recruited
30
31 14 during rhythm perception in passive listening tasks, even when attention is directed away from the
32
33 15 auditory stream [48–50]. These effects directly speak to the universal inclination to automatically
34
35 16 synchronize our movements to external rhythms [47]. Additionally, evidence for the role of the
36
37 17 motor system in timing is that the predictive alignment of neuronal excitability depends on sensory-
38
39 18 motor expertise [51] and improves with periodic-motor priming (vocally) or audio-motor training
40
41 19 [52].

42 20 It is therefore plausible that predictive timing in perception and action rely on similar systems,
43
44 21 computations, and neurophysiological mechanisms [45,53]. On this view, temporal predictions in
45
46 22 the auditory domain correspond to a covert form of **active sensing** [12,54]. In other words, the
47
48 23 **efferent motor signals** that are generated when producing audible actions are also generated
49
50 24 during the passive perception of temporally structured auditory streams [47,48,50,55,56]. These
51
52 25 signals propagate downward toward sensory systems to predictively phase-reset and align ongoing
53
54 26 oscillations in auditory cortex with upcoming events [11,15]. By comparing simulated and actual
55
56 27 sensory outcomes, prediction error responses can be computed and used to learn, correct, and
57
58 28 improve temporal predictions. At the neurophysiological level, temporal predictions might be
59
60 29 carried by beta-band (~20 Hz) motor related signals, that are also related to the representation of
61
62 30 temporal information [11,49,57–59] (Box 3).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 Importantly, the motor system relies on a flexible internal repertoire of actions (i.e. we can perform actions with diverse temporal dynamics) to emulate the temporal trajectory of upcoming inputs. Because of this flexibility, this model does not require distinct mechanisms for periodic versus aperiodic temporal predictions. While we assume that the specific role of motor-driven efferent signals lies in aligning the phase of ongoing oscillations in sensory regions, the operational mechanism arguably involves many other structures. For instance, it has been proposed that the specific computations underlying time estimation are carried out by subcortical structures, probably the basal ganglia or the cerebellum, depending on the specifics of the duration to be estimated, with the putamen involved in relative or beat-based duration estimation and the cerebellum in absolute or discrete duration estimation [60–62]. The inferior parietal cortex [63,64] is also possibly involved at the interface between motor-driven internal predictions and sensory processing, i.e. redirecting the modulatory efferent signals according to stimulus- and task-relevance.

Altogether, the experimental data suggest that the motor system is automatically recruited to track temporal regularities (whether periodic or aperiodic) and proactively enhance the processing of upcoming events. Whether the motor system’s recruitment is actually necessary for perception or just useful in noisy or adverse listening condition remains debated. Moreover, whether temporal predictions exploit efferent motor signals [65], and the underlying pathways remains to be determined [12].

Alternative sources of temporal predictions

The motor system, even in a broad sense, is unlikely to be the sole system that contributes to the predictive facilitation of perception in the time domain. For instance, recent findings in non-human primates demonstrate the involvement of the pulvinar in the top-down phase reset of neuronal oscillations in the primary auditory cortex [40]. Importantly, these findings provide evidence for phase-alignment to statistical regularities (pattern repetition) in the absence of relevant periodicity in the stimulation.

Humans exploit other forms of aperiodic temporal predictions e.g. symbolic cues that activate associated memories and infer target timing [21]. Such types of predictions –based on working or long-term memory– arguably involve the hippocampus [66]. Whether these systems predictively shape perception through top-down phase reset of neuronal oscillations has, to our knowledge, not been systematically investigated. Support for the notion of top-down phase reset comes from

1
2
3
4 1 research on symbolic cues used in higher-level language processing (cf. following section) and
5
6 2 selective attention to certain points in time [21]. Top-down expectations, manipulated using
7
8 3 predictive cues or stimulus probability of occurrence, enhance performance and predictively reset
9
10 4 intrinsic sensory rhythms - such as alpha and theta oscillations in vision [67–69] and delta
11
12 5 oscillations in audition [23]. In crowded auditory settings (e.g. cocktail party experiments),
13
14 6 selective attention facilitates temporal tracking [70,71] by aligning excitability with stimuli in the
15
16 7 attended stream. Tracking the speech of one particular talker enhances the alignment between the
17
18 8 stimulus and low-frequency oscillations in auditory cortex and higher processing areas [71].
19
20 9 Similarly, prior knowledge of sound patterns that are inserted within a sound sequence, can be used
21
22 10 to attend selectively and align oscillatory activity to those patterns [40,72].

23
24 11 Additional brain regions not belonging to the core motor system are involved in cueing paradigms.
25
26 12 In these settings, temporal predictions are typically reflected in anticipatory electrophysiological
27
28 13 components [25,73]. Such signals have been observed in various regions, including the anterior
29
30 14 cingulate cortex, premotor cortex, posterior parietal cortex, supplementary motor area (SMA), and
31
32 15 pre-SMA [26,74], all of which arguably contribute to proactive simulation processes. Whether
33
34 16 these electrophysiological signals reflect anticipatory computations locally, or whether they
35
36 17 actually propagate top-down to subserve the predictive modulation of sensory excitability remains
37
38 18 to be determined (Fig. 1 B).

39 19 *A/periodic predictions - Speech as relevant case*

40
41
42 20 Temporal processing is typically studied using highly simplified stimuli (e.g. tone sequences).
43
44 21 Although useful, this does not allow us to fully capture the essence of predictive processes, namely
45
46 22 to enhance the analysis of complex, ecological signals. Speech constitutes an informationally rich
47
48 23 signal from which temporal regularities at many time-scales can be used to facilitate processing. It
49
50 24 provides a theoretical challenge and a valuable experimental assay to investigate the substrates and
51
52 25 mechanisms underlying predictive processing.

53 54 26 *The entrainment view on speech processing*

55
56
57 27 Despite controversy about the extent of rhythmic structure in the speech signal [75], most
58
59 28 languages exhibit remarkably similar temporal modulations at around 4-8 Hz reflecting quasi-
60
61 29 periodic fluctuations of the amplitude envelope and implicating the syllabic scale [4,76]. Current

1
2
3
4 1 data suggest that entrainment of intrinsic neuronal oscillations (theta-band, 4-8 Hz) in auditory
5
6 2 cortex [36,38,77–79] aligns high neural excitability phases in the output layers of auditory cortex
7
8 3 [77]. In this framework, entrainment to acoustic landmarks (sharp amplitude modulations; [80])
9
10 4 sets up temporal frames that allow the brain to segment the signal at the syllabic scale [77].
11
12 5 Coupling of theta and (low, 25-35 Hz) gamma rhythms [81,82] reflects the hierarchical
13
14 6 organization between syllabic and phonemic cues and allows for simultaneously segmenting and
15
16 7 extracting the syllabic and phonemic information [77]. Conclusive evidence that the observed
17
18 8 activity reflects oscillatory entrainment rather than iterated evoked responses or anticipatory
19
20 9 components remains sparse (Box 1).

10 *Temporal constraints in speech processing*

11 Recent experimental observations set oscillatory entrainment in a different light, intimating that a
12
13 12 new interpretative framework is needed to account for these phenomena. For example, that theta-
14
15 13 band entrainment occurs with natural as well as with time-reversed, unintelligible speech [83]
16
17 14 questions the initially assumed functional role of theta oscillations in comprehension, and suggests
18
19 15 an involvement at lower processing levels. On the other hand, although not sufficient, theta
20
21 16 entrainment is likely a necessary precondition to ensure speech comprehension [80]. Using
22
23 17 (unintelligible) compressed speech, studies have shown that introducing periods of silence (80 ms)
24
25 18 between periods (40 ms) of compressed speech restores comprehension [35]. This suggests that
26
27 19 speech comprehension is limited by the amount of time needed to decode the information into
28
29 20 exploitable internal representations [84].

30
31
32 21 Another criticism of a stimulus-driven entrainment account for speech is that phase alignment to
33
34 22 acoustic landmarks [80] needs to be continuously adjusted, because of the quasi- (but not strictly)
35
36 23 periodic nature of speech. Thus, the observed speech tracking requires a more flexible mechanism
37
38 24 [33,85].

39
40
41 25 These observations are compatible with the view that intrinsic rhythms impose temporal constraints
42
43 26 on sensory processing [84]. This idea predicts that to be optimally processed, connected sound
44
45 27 sequences (e.g. speech, music) should be provided in temporal chunks of adequate length to match
46
47 28 hierarchical oscillatory temporal scales [77,78]. Importantly, although oscillatory constraints may
48
49 29 have a limited impact when stimuli are easy to process, they are relevant in noisy or adverse
50
51 30 listening conditions, where top-down processes are recruited to operate upon these constraints to

1
2
3
4 1 optimize sensory processing.
5
6

7 2 *Top-down phase reset in speech processing*
8
9

10 3 In light of the syntactic and semantic variability of language, the timing of high-level linguistic
11 4 information is typically more variable than that of lower-level input features [86]. Although
12 5 syntactic/semantic information is not necessarily provided in a periodic manner, it allows
13 6 generating probabilistic predictions of upcoming speech [87,88] and thus constitutes an aperiodic,
14 7 symbolic source of temporal predictions. The capacity of syntactic-semantic contexts to permit
15 8 temporal predictions and facilitate processing has been shown in early reaction time [89] or eye-
16 9 tracking studies [90]. Linguistic corpora-based computational models showed that syntactic
17 10 information, e.g. whether the last constituent was a noun or verb phrase, predicts pause timing in
18 11 continuous speech [5]. That syntactic predictability crucially relies on timing abilities, is supported
19 12 by basal ganglia patients deficits in using syntactic predictions in a temporally irregular context
20 13 [91].
21
22
23
24
25
26
27
28
29
30

31 14 Accordingly, several studies suggest that during speech perception, motor and higher-linguistic
32 15 processing areas predictively reset and align low-frequency activity in the auditory cortex
33 16 [10,15,79,92]. Optimal audio-motor coupling is observed for stimuli presented within the range of
34 17 the natural syllabic rate (i.e., ~ 4 Hz [92]) and during speech perception [10,15,79]. Comparing
35 18 intelligible versus backward speech, these studies further show that speech tracking and
36 19 comprehension are optimized by feedback from higher-level linguistic processes [15,16,70]. These
37 20 top-down predictions arguably enhance the precision of speech processing through anticipatory
38 21 top-down phase-reset of neuronal oscillations, as reflected in phase-phase and phase-amplitude
39 22 coupling measures [10,93]. Similarly, cross-modal enhancement of speech perception by lip-
40 23 reading is possibly mediated by higher-order frontal/motor regions through low-frequency phase
41 24 alignment [13,93,94]. These findings suggest that predictions from multiple levels and networks
42 25 can optimize neural entrainment to speech and intelligibility [10].
43
44
45
46
47
48
49
50
51
52
53

54 26 Despite these advances, whether linguistic predictions feedback to sensory processing levels
55 27 remains unclear [95]. Furthermore, the origins of top-down linguistic predictions –whether at the
56 28 phonological [14,96,97], lexical [96] and/or sentence level [98]– are complicated and poorly
57 29 understood. What roles the motor cortex plays in speech perception is also actively debated [99].
58 30 In addition to generating temporal predictions, speech motor areas presumably represent
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 phonological information [100]. To capture the essence of the motor contributions in speech
2 comprehension, and in line with the constraint view, we argue that it is necessary to consider audio-
3 motor coupling has a key mechanism [101].

4 In summary, predictions based on the motor system or higher-level language processing seem to be
5 effectively used to facilitate perception and enhance speech tracking. Further research is required
6 to elucidate the mechanisms and nature of the top-down signals provided by frontal, motor and/or
7 other areas and how these align temporal constraints at the sensory level.

8 **Concluding Remarks and Future Directions**

9 We propose a parsimonious framework for auditory temporal predictions, including for predictive
10 processing of speech. We suggest a neuronal implementation that simultaneously makes possible
11 bottom-up (stimulus-driven) and top-down (motor-driven, language-driven) phase-reset of low-
12 frequency oscillations in auditory cortex. On the proposed view, neuronal oscillations do not a
13 priori subserve *specific* functions but rather constitute an intrinsic temporal constraint. Top-down
14 phase-reset signals are generated to predictively align these internal temporal frames (excitability
15 cycles) and improve stimulus processing. This hypothesis, meant to synthesize current debates and
16 data on temporal predictions from research in neuroscience and psycholinguistics, is of course far
17 from comprehensive and requires development to evaluate the details of the operations
18 (Outstanding Questions).

1
2
3
4 1 *Box 1 Entrainment of neuronal oscillations*

5
6
7 2 Neuronal oscillations are ubiquitous in the brain of vertebrates and are seen even in the absence of
8
9 3 external stimulation. Brain rhythms most likely emerge from neurons acting as self-sustained
10
11 4 oscillators with a natural frequency [102,103] and thus can act as pacemakers. Spontaneous
12
13 5 oscillations have been described in several brain areas, including delta (1-4 Hz), theta (4-10 Hz)
14
15 6 and gamma (30-50 Hz) oscillations in primary auditory cortex [78]. Natural frequencies arguably
16
17 7 arise from subthreshold oscillations in the membrane potential of neurons, or short-term depression
18
19 8 and facilitation at the synaptic level (which can be different for synapses of the same presynaptic
20
21 9 cell, allowing selective network communication) [103,104]. Importantly, band-limited oscillatory
22
23 10 dynamics pre-exist in auditory cortical regions [36–38] and **resonate** after the end of an auditory
24
25 11 rhythmic stimulation [52,105–107] supporting the notion of intrinsic oscillators.

26
27 12 Building upon the successful application of oscillatory entrainment models to the study of circadian
28
29 13 rhythms [108,109], subsequent research extended such models to investigate a putative role of
30
31 14 neuronal oscillations in temporal attention [2]. Entrainment typically denotes that a quasi-periodic
32
33 15 external stimulation aligns the phase of slow (~1-10 Hz) endogenous neuronal oscillations in the
34
35 16 same frequency range (as measured with magneto- or electro-physiological recordings) and often
36
37 17 modulates behavior accordingly [17,110]. There exists a wealth of evidence that brain rhythms
38
39 18 coincide with the rate (~1-10 Hz) of certain cognitive domains (e.g. speech, music; [4]), and
40
41 19 preferred temporal rates of sensory processing [111]. Thus, it is often assumed that oscillatory
42
43 20 entrainment plays an instrumental role in temporal processing [17,77], and sensory selection
44
45 21 [3,78,112].

46
47 22 These assumptions need careful methodological consideration. First, the use of sinusoidal
48
49 23 transformations (e.g. Fourier and wavelet transforms) to analyse neural time-series results in
50
51 24 displaying apparently –but possibly spurious– oscillatory features, thereby raising potential
52
53 25 interpretative issues [113]. In addition, in the context of periodic stimulation, it is challenging to
54
55 26 distinguish oscillatory entrainment from repetitive, event-related potentials [114,115], or
56
57 27 endogenous anticipatory activity (e.g. Contingent Negative Variation, CNV; [73,116]). To
58
59 28 disentangle these processes, it is important to develop novel experimental paradigms isolating
60
61 29 exogenously driven from endogenously driven neural effects (see [7,11,19]).

1
2
3
4 1
5
6
7 2 *Box 2 Neuronal functions and processing constraints*

8
9
10 3 One striking feature of current neuroscience research is that the operationalization of cognitive
11 4 functions (e.g. consciousness, memory) coincides with the identification of dedicated oscillatory
12 5 processes. Thus it has been implicitly assumed these processes are causally linked – and therefore
13 6 a functional relevance has been attributed to neural oscillations. The term ‘function’, however, is
14 7 controversial and usually undefined in cognitive neuroscience. Here we favor a ‘causal role theory’
15 8 of physiological function, wherein a process is called a function when it causally contributes to the
16 9 behavior of a system [117]. We distinguish between neurophysiological phenomena that underlie
17 10 generic neuronal processes versus phenomena that causally subserve specific cognitive processes.
18 11 The first are necessary but not involved in one specific function. They ‘constrain’ cognitive
19 12 processes by being part of the neurophysiological architecture. In contrast, the latter are necessary
20 13 for a specific cognitive function.

21
22
23 14 Even without compelling evidence, claims have been advanced that oscillations are functionally
24 15 specific to various cognitive processes, e.g. [118]. Other accounts suggest, however, that oscillatory
25 16 processes support a more generic, canonical function: the message-passing of information
26 17 throughout the cortical hierarchy [27,28]. The canonical architecture imposes ‘cyclic temporal
27 18 constraints’ on cognitive processes, as the frequency of oscillations of a neuronal population
28 19 determines the excitability fluctuations of these neurons, thereby periodically imposing alternating
29 20 phases of high and low excitability [78]. Because cognitive processes imply the transmission of
30 21 information between neuronal populations, they entail specific spectral signatures possibly
31 22 reflecting processing constraints.

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49 23 The constraint interpretation accounts for many observations, including empirical evidence for
50 24 discrete perception [32,33,35], and should perhaps be preferred to functional interpretations, as
51 25 long as it is not possible to provide evidence otherwise. Crucially, [as we propose here](#), it accounts
52 26 for the functional effectiveness of predictive top-down resets on sensory processing for both
53 27 periodic and aperiodic temporal predictions.
54
55
56
57
58

1
2
3
4 1 *Box 3 Beta-band oscillations involved in top-down predictions*

5
6
7 2 Beta-band activity is typically evidenced in sensory-motor areas during temporal tracking as well
8
9 3 as during continuous speech processing [11,49,57,59,119,120]. Recent evidence suggested that
10
11 4 beta-band activity in motor regions proactively tracks the timing of upcoming events and modulates
12
13 5 low-level auditory processes to facilitate behavioral responses [11,49,57] (Fig. I, during selective
14
15 6 attention to target tones, modulated at 1.5 Hz, temporal predictions are encoded in delta-beta
16
17 7 coupled oscillations functionally directed toward auditory regions to modulate the 3 Hz auditory
18
19 8 input, adapted from [11]).

20
21 9 Beta-band signals are also involved in top-down predictions of higher level linguistic features, such
22
23 10 as semantic and syntactic predictions [121,122]. Therefore, it has been suggested that top-down
24
25 11 signals carried by beta-band oscillations indicate more than ‘mere’ temporal predictions but
26
27 12 communicate accurate information about the content of the expected events, as suggested by their
28
29 13 role in lexical/semantic predictions of upcoming words but not of the syntactic category (for
30
31 14 review, see [88]).

32
33 15 However, the functional relevance and specificity of beta-band oscillations for predictive processes
34
35 16 remains to be clarified. One plausible account for current empirical findings is that beta-band
36
37 17 activity is more generically involved in feedback communication between brain areas [28]. In the
38
39 18 ‘predictive routing’ account, descending beta-band signals predictively pre-activate neuronal
40
41 19 population that are relevant for the processing of expected inputs [123]. Consistent with this view,
42
43 20 recordings in the macaque and human early sensory cortices support the involvement of beta-band
44
45 21 signals originating in deep cortical layers in the top-down propagation of information [81,124–
46
47 22 126].

48
49 23 In sum, although the detailed neural mechanisms remain to be elucidated, the data supports the
50
51 24 notion that beta-band signals predictively modulate sensory processing. Moreover, while the nature
52
53 25 of the predictive information seems to depend on the source-region generating beta oscillations,
54
55 26 temporal predictions are specifically associated with motor-related beta activity.

1
2
3
4 1 *Outstanding Questions*
5
6

- 7 2 • If intrinsic neuronal oscillations constitute generic constraints that pre-exist
8 (ontogenetically or phylogenetically), can these oscillations be experimentally evidenced
9 3 prior to the emergence of **specific functions (e.g. language or sequence processing)**, for
10 4 **example** in non-verbal infants or primates?
11 5
12
13
14
15 6 • Is **the efficiency of predictive neural alignment and ensuing** behavioral improvement solely
16 7 conditioned by sensorimotor training and expertise [51]? Alternatively, could other types
17 8 of plasticity, such as repeated exposure to these sequences or sensory deprivation [127]
18 9 result in adapting oscillatory constraints at the sensory level?
19
20
21
22
23
24 10 • Are **ramping** components (e.g. CNA) a mere signature of anticipatory activity or do they
25 11 play a causal role (akin to phase resets) in predictively shaping sensory processes in a top-
26 12 down manner? In the latter case, what would be the nature (modulatory, driving) of the
27 13 underlying mechanism?
28
29
30
31
32 14 • **In line with the notion of intrinsic timing models**, do all **neural systems that perform precise,**
33 15 **time-related computations** (e.g. motor, attentional or memory systems), use similar (or
34 16 shared) mechanisms to prospectively influence perceptual processes?
35
36
37
38
39 17 • **Given that hierarchically structured temporal signals (e.g. speech and music) embed**
40 18 **multiple, nested levels of temporal regularities**, is the motor system involved in proactively
41 19 tracking temporal regularities at all (e.g. syllabic, prosodic or phrasal) or only at specific
42 20 levels?
43
44
45
46

47 21 *Glossary*
48
49

50 22 **Active sensing:** denotes the proactive contribution of sensorimotor processing to perception. One
51 23 implementation relies on top-down predictions based on internal copies of movement commands
52 24 (see also corollary discharge);
53
54
55
56

57 25 **Analytic attending mode:** in sustained vigilance settings, high frequency activity maintains a
58 26 continuous high-excitability state of the system, in order to support attention;
59
60

61 27 **Climbing neuronal activity:** **ramping anticipatory** neuronal activity that increases until –and is
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 resolved after– an expected time point;

2 **Constraint:** in reference to neural processing, denotes aspects of the neuronal architecture that
3 affect neuronal functions but are not specific or instrumental to a certain function;

4 **Dynamic Attending Theory:** describes how entrainment of neuronal oscillations to a periodic
5 stimulation directs attention in time;

6 **Efferent motor signals:** signals generated by the motor system to propagate down to peripheral
7 systems to generate overt –but also possibly covert– movements.

8 **Entrainment:** the process through which two or more self-sustained oscillators become coupled
9 (Box 2);

10 **Functional role:** central but controversial and not well-defined concept in neuroscience; here, a
11 ‘causal role theory’ of function is favored, where a process is called a function if it causally and
12 specifically contributes to the functional capacity of a system. In contrast, ‘selectionist theories’
13 define the function of a process with regard to its contribution to natural selection;

14 **Hazard functions:** describe the probability of occurrence of forthcoming events;

15 **Modulate:** processes that change the responsiveness and effectiveness of driven neuronal
16 responses [128];

17 **Phase alignment:** neuronal oscillatory excitability phases co-occur with energy phases of a
18 periodic stimulation;

19 **Parsimony** (see also Occam’s razor and *lex parsimoniae*): the principle that among competing
20 hypotheses, the one underlying the fewest assumptions should be preferred.

21 **Periodic:** successive events are separated by a constant duration;

22 **Ramping-based models:** monotonic changes in firing rate encode time until a threshold is met;

23 **Resonate:** resonance is a passive phenomenon where neuronal responses reflect the stimulation
24 frequency. This is generally distinguished from oscillatory phenomena, where neurons oscillate at
25 a preferred frequency in the absence of a stimulation and can entrain to an external stimulation;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Rhythmic:** in music (and language) terminology acoustics that contain a regular pattern with
2 respect to timing. Although periodic patterns are always rhythmic (and rhythmicity and periodicity
3 are often used synonymously) the reversal is not true; [129]; Oscillatory brain activity is called
4 rhythmic (Box 1), whereas most measures analyze periodicity.

5 **Rhythmic attending mode:** entrainment of neuronal oscillations to a periodic stimulation directs
6 attention in time;

7 **Symbolic cues:** cues that activate associated memories about the target timing;

8 **Temporal regularities:** a statistical rule that can be extracted to predict a temporal pattern;

9 **Tracking:** the neural sensory response profile reflects the energy fluctuations in the acoustic (e.g.
10 the envelope of speech or the stimulation rate of sounds);

11 **Top-down phase-reset:** endogenous, non stimulus-driven process that aligns the phase of neuronal
12 oscillations with an expected event to facilitate its processing;

13

14 *Acknowledgements*

15 This work was funded by the Max-Planck-Institute for Empirical Aesthetics (JR, DP) and the
16 BrainCom Project, Horizon 2020 Framework Programme (732032) (LHA). We thank Felix
17 Bernoulli and Sarah Brendecke for graphics support.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **References**

2 1 Merchant, H. *et al.* (2015) Finding the beat: a neural perspective across humans and non-
3 human primates. *Philos. Trans. R. Soc. B Biol. Sci.* 370, 20140093
4 2 Jones, M.R. (1976) Time, our lost dimension: toward a new theory of perception,
5 attention, and memory. *Psychol. Rev.* 83, 323–355
6 3 Schroeder, C.E. and Lakatos, P. (2009) Low-frequency neuronal oscillations as
7 instruments of sensory selection. *Trends Neurosci.* 32, 9–18
8 4 Ding, N. *et al.* (2017) Temporal modulations in speech and music. *Neurosci. Biobehav.*
9 *Rev.* DOI: 10.1016/j.neubiorev.2017.02.011
10 5 Nguyen, T.T. *et al.* (2015) A latent variable model for joint pause prediction and
11 dependency parsing, in *INTERSPEECH 2015, 16th Annual Conference of the*
12 *International Speech Communication Association, Dresden, Germany, September 6-10,*
13 *2015*, pp. 2719–2723
14 6 Polak, R. (2017) Non-isochronous meter is not irregular. A review of theory and
15 evidence. In *Proceedings of the Annual Meetings of the Gesellschaft für Musiktheorie*
16 *2015* (Aydintan, M. *et al.*, eds), Olms
17 7 Breska, A. and Deouell, L.Y. (2017) Neural mechanisms of rhythm-based temporal
18 prediction: Delta phase-locking reflects temporal predictability but not rhythmic
19 entrainment. *PLOS Biol.* 15, e2001665
20 8 Morillon, B. *et al.* (2016) Temporal prediction in lieu of periodic stimulation. *J. Neurosci.*
21 36, 2342–2347
22 9 Arnal, L.H. (2012) Predicting “when” using the motor system’s beta-band oscillations.
23 *Front. Hum. Neurosci.* 6, 225
24 10 Keitel, A. *et al.* (2017) Auditory cortical delta-entrainment interacts with oscillatory
25 power in multiple fronto-parietal networks. *NeuroImage* 147, 32–42
26 11 Morillon, B. and Baillet, S. (2017) Motor origin of temporal predictions in auditory
27 attention. *Proc. Natl. Acad. Sci.* at
28 <<http://www.pnas.org/content/early/2017/09/26/1705373114.abstract>>
29 12 Morillon, B. *et al.* (2015) Predictive motor control of sensory dynamics in auditory active
30 sensing. *Curr. Opin. Neurobiol.* 31, 230–238
31 13 Park, H. *et al.* (2016) Lip movements entrain the observers’ low-frequency brain
32 oscillations to facilitate speech intelligibility. *eLife* 5, e14521

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 14 Mai, G. *et al.* (2016) Delta, theta, beta, and gamma brain oscillations index levels of
2 auditory sentence processing. *NeuroImage* 133, 516–528
3 15 Park, H. *et al.* (2015) Frontal top-down signals increase coupling of auditory low-
4 frequency oscillations to continuous speech in human listeners. *Curr. Biol.* 25, 1649–1653
5 16 Peelle, J.E. *et al.* (2013) Phase-locked responses to speech in human auditory cortex are
6 enhanced during comprehension. *Cereb. Cortex* 23, 1378–1387
7 17 Henry, M.J. and Herrmann, B. (2014) Low-frequency neural oscillations support dynamic
8 attending in temporal context. *Timing Time Percept.* 2, 62–86
9 18 Large, E.W. and Jones, M.R. (1999) The dynamics of attending: How people track time-
10 varying events. *Psychol. Rev.* 106, 119–159
11 19 Haegens, S. and Zion Golumbic, E. (2017) Rhythmic facilitation of sensory processing: a
12 critical review. *Neurosci. Biobehav. Rev.* DOI: 10.1016/j.neubiorev.2017.12.002
13 20 Cope, T.E. *et al.* (2012) Temporal predictions based on a gradual change in tempo. *J.*
14 *Acoust. Soc. Am.* 131, 4013–4022
15 21 Nobre, A.C. and van Ede, F. (2018) Anticipated moments: temporal structure in attention.
16 *Nat. Rev. Neurosci.* 19, 34–48
17 22 Coull, J.T. (2009) Neural substrates of mounting temporal expectation. *PLoS Biol.* 7,
18 e1000166
19 23 Wilsch, A. *et al.* (2015) Slow-delta phase concentration marks improved temporal
20 expectations based on the passage of time. *Psychophysiology* 52, 910–918
21 24 Obleser, J. *et al.* (2017) What do we talk about when we talk about rhythm? *PLOS Biol.*
22 15, e2002794
23 25 Paton, J.J. and Buonomano, D.V. (2018) The neural basis of timing: Distributed
24 mechanisms for diverse functions. *Neuron* 98, 687–705
25 26 Janssen, P. and Shadlen, M.N. (2005) A representation of the hazard rate of elapsed time
26 in macaque area LIP. *Nat. Neurosci.* 8, 234–241
27 27 Bosman, C.A. and Aboitiz, F. (2015) Functional constraints in the evolution of brain
28 circuits. *Front. Neurosci.* 9, 303
29 28 Fries, P. (2015) Rhythms for cognition: Communication through coherence. *Neuron* 88,
30 220–235
31 29 Palmigiano, A. *et al.* (2017) Flexible information routing by transient synchrony. *Nat.*
32 *Neurosci.* 20, 1014–1022

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 30 Pillai, A.S. and Jirsa, V.K. (2017) Symmetry breaking in space-time hierarchies shapes
2 brain dynamics and behavior. *Neuron* 94, 1010–1026
3 31 Womelsdorf, T. *et al.* (2014) Dynamic circuit motifs underlying rhythmic gain control,
4 gating and integration. *Nat. Neurosci.* 17, 1031–1039
5 32 VanRullen, R. (2016) Perceptual cycles. *Trends Cogn. Sci.* 20, 723–735
6 33 Ghitza, O. (2016) Acoustic-driven delta rhythms as prosodic markers. *Lang. Cogn.*
7 *Neurosci.* DOI: 10.1080/23273798.2016.1232419
8 34 Teng, X. *et al.* (2017) Concurrent temporal channels for auditory processing: Oscillatory
9 neural entrainment reveals segregation of function at different scales. *PLOS Biol.* 15,
10 e2000812
11 35 Ghitza, O. and Greenberg, S. (2009) On the possible role of brain rhythms in speech
12 perception: Intelligibility of time-compressed speech with periodic and aperiodic
13 insertions of silence. *Phonetica* 66, 113–126
14 36 Giraud, A.-L. *et al.* (2007) Endogenous cortical rhythms determine cerebral specialization
15 for speech perception and production. *Neuron* 56, 1127–1134
16 37 Keitel, A. and Gross, J. (2016) Individual human brain areas can be identified from their
17 characteristic spectral activation fingerprints. *PLoS Biol* 14, e1002498
18 38 Morillon, B. *et al.* (2012) Asymmetric function of theta and gamma activity in syllable
19 processing: an intra-cortical study. *Front. Psychol.* 3, 248
20 39 Nourski, K.V. *et al.* (2009) Temporal envelope of time-compressed speech represented in
21 the human auditory cortex. *J. Neurosci.* 29, 15564–15574
22 40 Barczak, A. *et al.* (2018) Top-down, contextual entrainment of neuronal oscillations in the
23 auditory thalamocortical circuit. *Proc. Natl. Acad. Sci.* DOI: 10.1073/pnas.1714684115
24 41 Kaas, J.H. and Hackett, T.A. (2000) Subdivisions of auditory cortex and processing
25 streams in primates. *Proc. Natl. Acad. Sci.* 97, 11793
26 42 Rauschecker, J.P. and Tian, B. (2000) Mechanisms and streams for processing of “what”
27 and “where” in auditory cortex. *Proc. Natl. Acad. Sci.* 97, 11800–11806
28 43 Goodale, M.A. and Milner, A.D. (1992) Separate visual pathways for perception and
29 action. *Trends Neurosci.* 15, 20–25
30 44 Friston, K. (2005) A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*
31 360, 815–836
32 45 Schubotz, R.I. (2007) Prediction of external events with our motor system: towards a new

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 framework. *Trends Cogn. Sci.* 11, 211–218

2 46 Coull, J.T. (2015) A frontostriatal circuit for timing the duration of events. In *Brain*

3 *Mapping* (Toga, A. W., ed), pp. 565–570, Academic Press

4 47 Merchant, H. and Yarrow, K. (2016) How the motor system both encodes and influences

5 our sense of time. *Time Percept. Action* 8, 22–27

6 48 Chen, J.L. *et al.* (2008) Listening to musical rhythms recruits motor regions of the brain.

7 *Cereb. Cortex* 18, 2844–2854

8 49 Fujioka, T. *et al.* (2012) Internalized timing of isochronous sounds is represented in

9 neuromagnetic beta oscillations. *J. Neurosci.* 32, 1791–1802

10 50 Grahn, J.A. and Rowe, J.B. (2013) Finding and feeling the musical beat: striatal

11 dissociations between detection and prediction of regularity. *Cereb. Cortex* 23, 913–921

12 51 Doelling, K.B. and Poeppel, D. (2015) Cortical entrainment to music and its modulation

13 by expertise. *Proc. Natl. Acad. Sci.* 112, E6233–E6242

14 52 Cason, N. *et al.* (2015) Bridging music and speech rhythm: rhythmic priming and audio-

15 motor training affect speech perception. *Acta Psychol. (Amst.)* 155, 43–50

16 53 Coull, J.T. *et al.* (2011) Neuroanatomical and neurochemical substrates of timing.

17 *Neuropsychopharmacol.* 36, 3–25

18 54 Schroeder, C.E. *et al.* (2010) Dynamics of active sensing and perceptual selection. *Cogn.*

19 *Neurosci.* 20, 172–176

20 55 Tian, X. and Poeppel, D. (2013) The effect of imagination on stimulation: the functional

21 specificity of efference copies in speech processing. *J. Cogn. Neurosci.* 25, 1020–1036

22 56 Tian, X. and Poeppel, D. (2010) Mental imagery of speech and movement implicates the

23 dynamics of internal forward models. *Front. Psychol.* 1, 166

24 57 Arnal, L.H. *et al.* (2015) Delta-beta coupled oscillations underlie temporal prediction

25 accuracy. *Cereb. Cortex* 25, 3077–3085

26 58 Iversen, J.R. *et al.* (2009) Top-down control of rhythm perception modulates early

27 auditory responses. *Ann. N. Y. Acad. Sci.* 1169, 58–73

28 59 Saleh, M. *et al.* (2010) Fast and slow oscillations in human primary motor cortex predict

29 oncoming behaviorally relevant cues. *Neuron* 65, 461–471

30 60 Breska, A. and Ivry, R.B. (2016) Taxonomies of timing: Where does the cerebellum fit

31 in? *Curr. Opin. Behav. Sci.* 8, 282–288

32 61 Nozaradan, S. *et al.* (2017) Specific contributions of basal ganglia and cerebellum to the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 neural tracking of rhythm. *Cortex* 95, 156–168

2 62 Teki, S. *et al.* (2011) Distinct neural substrates of duration-based and beat-based auditory

3 timing. *J. Neurosci.* 31, 3805–3812

4 63 Bolger, D. *et al.* (2014) Metrical rhythm implicitly orients attention in time as indexed by

5 improved target detection and left inferior parietal activation. *J. Cogn. Neurosci.* 26, 593–

6 605

7 64 Patel, A.D. and Iversen, J.R. (2014) The evolutionary neuroscience of musical beat

8 perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front.*

9 *Syst. Neurosci.* 8, 57

10 65 Crapse, T.B. and Sommer, M.A. (2008) Corollary discharge across the animal kingdom.

11 *Nat. Rev. Neurosci.* 9, 587–600

12 66 Aly, M. and Turk-Browne, N.B. (2017) Flexible weighting of diverse inputs makes

13 hippocampal function malleable. *Neurosci. Lett.* DOI: 10.1016/j.neulet.2017.05.063

14 67 Bonnefond, M. and Jensen, O. (2012) Alpha oscillations serve to protect working memory

15 maintenance against anticipated distracters. *Curr. Biol.* 22, 1969–1974

16 68 Cravo, A.M. *et al.* (2011) Endogenous modulation of low frequency oscillations by

17 temporal expectations. *J. Neurophysiol.* 106, 2964–2972

18 69 Samaha, J. *et al.* (2015) Top-down control of the phase of alpha-band oscillations as a

19 mechanism for temporal prediction. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8439–8444

20 70 Rimmele, J. *et al.* (2015) The effects of selective attention and speech acoustics on neural

21 speech-tracking in a multi-talker scene. *Cortex* 68, 144–154

22 71 Zion Golumbic, E.M. *et al.* (2013) Mechanisms underlying selective neuronal tracking of

23 attended speech at a “cocktail party.” *Neuron* 77, 980–991

24 72 Costa-Faidella, J. *et al.* (2017) Selective entrainment of brain oscillations drives auditory

25 perceptual organization. *NeuroImage* 159, 195–206

26 73 Walter, W.G. *et al.* (1964) Contingent negative variation: An electric sign of sensori-

27 motor association and expectancy in the human Brain. *Nature* 203, 380

28 74 Wittmann, M. (2013) The inner sense of time: how the brain creates a representation of

29 duration. *Nat. Rev. Neurosci.* 14, 217–223

30 75 Nolan, F. and Jeon, H.-S. (2014) Speech rhythm: a metaphor? *Philos. Trans. R. Soc.*

31 *Lond. B Biol. Sci.* 369,

32 76 Pellegrino, F. *et al.* (2011) A cross-language perspective on speech information rate.

1
2
3
4 1 *Language* 87, 539–558
5
6 2 77 Giraud, A.-L. and Poeppel, D. (2012) Cortical oscillations and speech processing:
7
8 3 emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517
9
10 4 78 Lakatos, P. *et al.* (2005) An oscillatory hierarchy controlling neuronal excitability and
11
12 5 stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911
13
14 6 79 Morillon, B. *et al.* (2010) Neurophysiological origin of human brain asymmetry for
15
16 7 speech and language. *Proc. Natl. Acad. Sci. U. S. A.* 107, 18688–18693
17
18 8 80 Doelling, K.B. *et al.* (2014) Acoustic landmarks drive delta–theta oscillations to enable
19
20 9 speech comprehension by facilitating perceptual parsing. *NeuroImage* 85, Part 2, 761–768
21
22 10 81 Fontolan, L. *et al.* (2014) The contribution of frequency-specific activity to hierarchical
23
24 11 information processing in the human auditory cortex. *Nat Commun* 5, 4694
25
26 12 82 Gross, J. *et al.* (2013) Speech rhythms and multiplexed oscillatory sensory coding in the
27
28 13 human brain. *PLoS Biol* 11, e1001752
29
30 14 83 Howard, M.F. and Poeppel, D. (2010) Discrimination of speech stimuli based on neuronal
31
32 15 response phase patterns depends on acoustics but not comprehension. *J. Neurophysiol.*
33
34 16 104, 2500–2511
35
36 17 84 Pefkou, M. *et al.* (2017) Theta- and beta-band neural activity reflect independent syllable
37
38 18 tracking and comprehension of time-compressed speech. *J. Neurosci.* DOI:
39
40 19 10.1523/JNEUROSCI.2882-16.2017
41
42 20 85 Canavier, C.C. (2015) Phase-resetting as a tool of information transmission. *SI Brain*
43
44 21 *Rhythms Dyn. Coord.* 31, 206–213
45
46 22 86 Scontras, G. *et al.* (2015) Syntactic complexity effects in sentence production. *Cogn. Sci.*
47
48 23 39, 559–583
49
50 24 87 Jadoul, Y. *et al.* (2016) Seeking temporal predictability in speech: Comparing statistical
51
52 25 approaches on 18 world languages. *Front. Hum. Neurosci.* 10, 586
53
54 26 88 Meyer, L. (2017) The neural oscillations of speech processing and language
55
56 27 Comprehension: State of the art and emerging mechanisms. *Eur. J. Neurosci.* DOI:
57
58 28 10.1111/ejn.13748
59
60 29 89 Marslen-Wilson, W. and Tyler, L.K. (1980) The temporal structure of spoken language
61
62 30 understanding. *Cognition* 8, 1–71
63
64 31 90 Altmann, G.T.. and Kamide, Y. (1999) Incremental interpretation at verbs: restricting the
65
66 32 domain of subsequent reference. *Cognition* 73, 247–264

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 91 Kotz, S.A. and Schmidt-Kassow, M. (2015) Basal ganglia contribution to rule expectancy
2 and temporal predictability in speech. *Cortex*. 68, 48–60
3 92 Assaneo, M.F. and Poeppel, D. (2018) The coupling between auditory and motor cortices
4 is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Sci. Adv.* 4, eaao3842
5 93 ten Oever, S. and Sack, A.T. (2015) Oscillatory phase shapes syllable perception. *Proc.*
6 *Natl. Acad. Sci.* 112, 15833–15837
7 94 Giordano, B.L. *et al.* (2017) Contributions of local speech encoding and functional
8 connectivity to audio-visual speech perception. *eLife* 6, e24763
9 95 McQueen, J.M. *et al.* (2016) When brain regions talk to each other during speech
10 processing, what are they talking about? Commentary on Gow and Olson (2015). *Lang.*
11 *Cogn. Neurosci.* 31, 860–863
12 96 Keitel, A. *et al.* (2018) Perceptually relevant speech tracking in auditory and motor cortex
13 reflects distinct linguistic features. *PLOS Biol.* 16, e2004473
14 97 Sohoglu, E. and Davis, M.H. (2016) Perceptual learning of degraded speech by
15 minimizing prediction error. *Proc. Natl. Acad. Sci.* 113, E1747–E1756
16 98 Ding, N. *et al.* (2016) Cortical tracking of hierarchical linguistic structures in connected
17 speech. *Nat. Neurosci.* 19, 158–164
18 99 Skipper, J.I. *et al.* (2017) The hearing ear is always found close to the speaking tongue:
19 Review of the role of the motor system in speech perception. *Brain Lang.* 164, 77–105
20 100 Du, Y. and Zatorre, R.J. (2017) Musical training sharpens and bonds ears and tongue to
21 hear speech better. *Proc. Natl. Acad. Sci.* DOI: 10.1073/pnas.1712223114
22 101 Strauß, A. and Schwartz, J.-L. (2017) The syllable in the light of motor skills and neural
23 oscillations. *Lang. Cogn. Neurosci.* 32, 562–569
24 102 Kurths, J. *et al.* (2001) *Synchronization: a universal concept in nonlinear sciences*,
25 Cambridge University Press New York.
26 103 Llinas, R.R. (1988) The intrinsic electrophysiological properties of mammalian neurons:
27 insights into central nervous system function. *Science* 242, 1654–1664
28 104 Izhikevich, E.M. *et al.* (2003) Bursts as a unit of neural information: selective
29 communication via resonance. *Trends Neurosci.* 26, 161–167
30 105 Falk, S. *et al.* (2017) Tuning neural phase entrainment to speech. *J. Cogn. Neurosci.* 29,
31 1378–1389
32 106 Hickok, G. *et al.* (2015) The rhythm of perception: Entrainment to acoustic rhythms

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 induces subsequent perceptual oscillation. *Psychol. Sci.* 26, 1006–1013

2 107 Sanabria, D. and Correa, Á. (2013) Electrophysiological evidence of temporal preparation
3 driven by rhythms in audition. *Biol. Psychol.* 92, 98–105

4 108 Aschoff, J. (1954) Zeitgeber der tierischen Tagesperiodik. *Naturwissenschaften* 41, 49–56

5 109 Pittendrigh, C.S. (1960) Circadian rhythms and the circadian organization of living
6 systems. *Cold Spring Harb. Symp. Quant. Biol.* 25, 159–184

7 110 Ding, N. and Simon, J.Z. (2014) Cortical entrainment to continuous speech: functional
8 roles and interpretations. *Front. Hum. Neurosci.* 8, 311

9 111 Holcombe, A.O. (2009) Seeing slow and seeing fast: two limits on perception. *Trends*
10 *Cogn. Sci.* 13, 216–221

11 112 Lakatos, P. *et al.* (2008) Entrainment of neuronal oscillations as a mechanism of
12 attentional selection. *Science* 320, 110–113

13 113 Cole, S.R. and Voytek, B. (2017) Brain oscillations and the importance of waveform
14 shape. *Trends Cogn. Sci.* 21, 137–149

15 114 Kösem, A. *et al.* (2016) High-frequency neural activity predicts word parsing in
16 ambiguous speech streams. *J. Neurophysiol.* 116, 2497

17 115 Zoefel, B. and VanRullen, R. (2016) EEG oscillations entrain their phase to high-level
18 features of speech sound. *NeuroImage* 124, 16–23

19 116 Herrmann, B. and Johnsrude, I.S. (2018) Neural signatures of the processing of temporal
20 patterns in sound. *J. Neurosci.* DOI: 10.1523/JNEUROSCI.0346-18.2018

21 117 Roux, E. (2014) The concept of function in modern physiology. *J. Physiol.* 592, 2245–
22 2249

23 118 Weiss, S. and Mueller, H.M. (2012) “Too many betas do not spoil the broth”: The role of
24 beta brain oscillations in language processing. *Front. Psychol.* 3,

25 119 Kulashekhar, S. *et al.* (2016) The role of cortical beta oscillations in time estimation.
26 *Hum. Brain Mapp.* 37, 3262–3281

27 120 Wiener, M. and Kanai, R. (2016) Frequency tuning for temporal perception and
28 prediction. *Time Percept. Action* 8, 1–6

29 121 Cope, T.E. *et al.* (2017) Evidence for causal top-down frontal contributions to predictive
30 processes in speech perception. *Nat. Commun.* 8, 2154

31 122 Lewis, A.G. *et al.* (2016) A predictive coding perspective on beta oscillations during
32 sentence-level language comprehension. *Front. Hum. Neurosci.* 10, 85

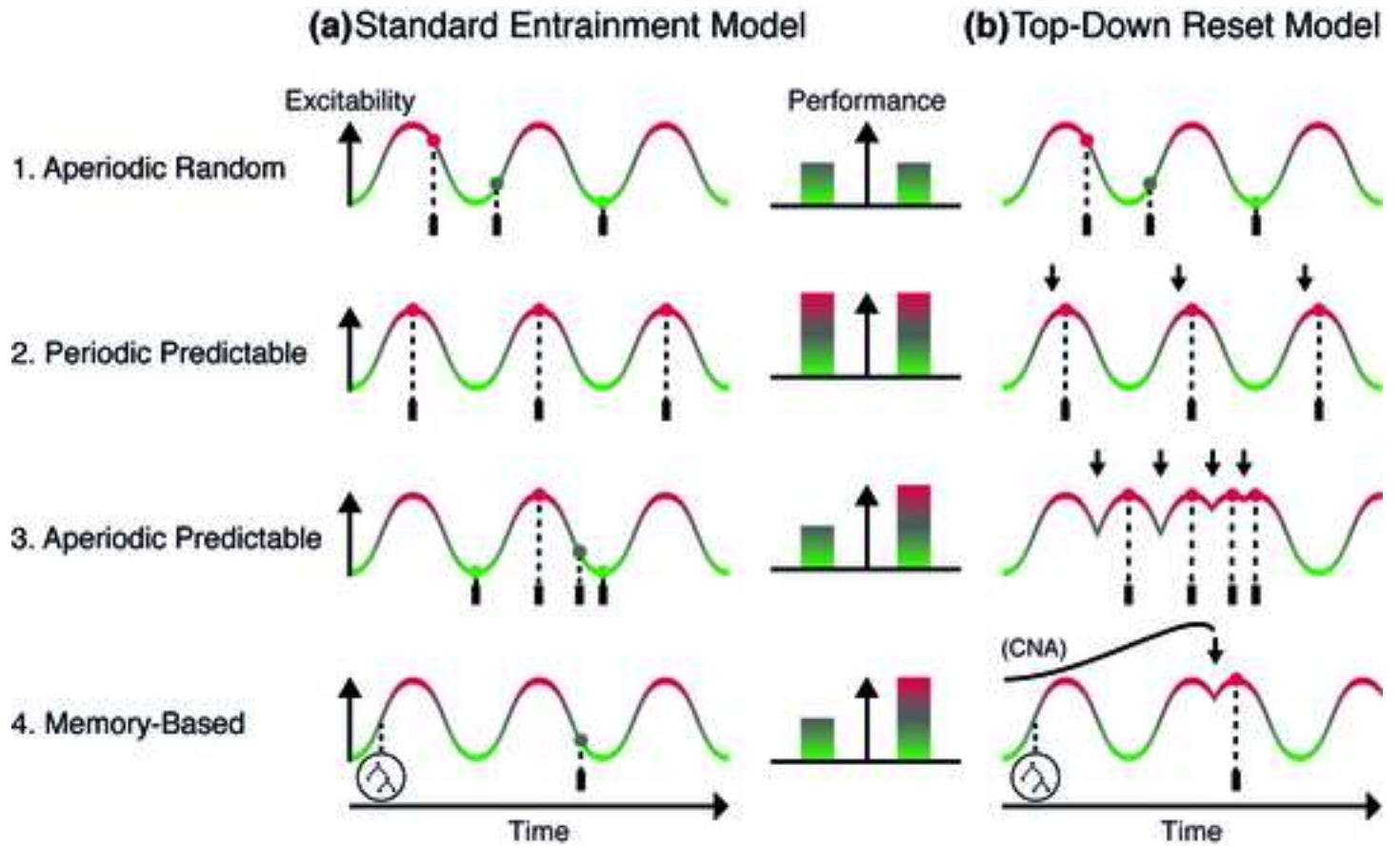
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 123 Spitzer, B. and Haegens, S. (2017) Beyond the status quo: A role for beta oscillations in
2 endogenous content (Re)Activation. *eNeuro* 4, ENEURO.0170-17.2017
3 124 Arnal, L.H. and Giraud, A.-L. (2012) Cortical oscillations and sensory predictions. *Trends*
4 *Cogn. Sci.* 16, 390–398
5 125 Bastos, A.M. *et al.* (2015) Visual areas exert feedforward and feedback influences
6 through distinct frequency channels. *Neuron* 85, 390–401
7 126 van Kerkoerle, T. *et al.* (2014) Alpha and gamma oscillations characterize feedback and
8 feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci.* 111, 14332–
9 14341
10 127 Hertrich, I. *et al.* (2013) Tracking the speech signal – Time-locked MEG signals during
11 perception of ultra-fast and moderately fast speech in blind and in sighted listeners. *Brain*
12 *Lang.* 124, 9–21
13 128 David, O. *et al.* (2006) Mechanisms of evoked and induced responses in MEG/EEG.
14 *NeuroImage* 31, 1580–1591
15 129 Patel, A.D. (2008) *Music, language, and the brain*, Oxford University Press.
16 130 Wyart, V. *et al.* (2012) Rhythmic fluctuations in evidence accumulation during decision
17 making in the human brain. *Neuron* 76, 847–858
18

1 *Outstanding Questions*

- 2 • If intrinsic neuronal oscillations constitute generic constraints that pre-exist
3 (ontogenetically or phylogenetically), can these oscillations be experimentally evidenced
4 prior to the emergence of **specific functions (e.g. language or sequence processing)**, for
5 **example** in non-verbal infants or primates?
- 6 • Is **the efficiency of predictive neural alignment and ensuing** behavioral improvement solely
7 conditioned by sensorimotor training and expertise [51]? Alternatively, could other types
8 of plasticity, such as repeated exposure to these sequences or sensory deprivation [127]
9 result in adapting oscillatory constraints at the sensory level?
- 10 • Are **ramping** components (e.g. CNA) a mere signature of anticipatory activity or do they
11 play a causal role (akin to phase resets) in predictively shaping sensory processes in a top-
12 down manner? In the latter case, what would be the nature (modulatory, driving) of the
13 underlying mechanism?
- 14 • **In line with the notion of intrinsic timing models**, do all **neural systems that perform precise,**
15 **time-related computations** (e.g. motor, attentional or memory systems), use similar (or
16 shared) mechanisms to prospectively influence perceptual processes?
- 17 • **Given that hierarchically structured temporal signals (e.g. speech and music) embed**
18 **multiple, nested levels of temporal regularities**, is the motor system involved in proactively
19 tracking temporal regularities at all (e.g. syllabic, prosodic or phrasal) or only at specific
20 levels?

21



(c) Sources of Temporal Predictions

