Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics
Loic Damm, Déborah Varoqui, Valérie Cochen de Cock, Simone Dalla Bella, Benoît Bardy

To cite this version:
Loic Damm, Déborah Varoqui, Valérie Cochen de Cock, Simone Dalla Bella, Benoît Bardy. Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. Neuroscience and Biobehavioral Reviews, Elsevier, 2019, 10.1016/j.neubiorev.2019.12.024. hal-02434120

HAL Id: hal-02434120
https://hal.umontpellier.fr/hal-02434120
Submitted on 9 Jan 2020

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.
Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics

Loïc Damm\textsuperscript{1,*}, Déborah Varoqui\textsuperscript{1, 15}, Valérie Cochen De Cock\textsuperscript{1, 3, 4}, Simone Dalla Bella\textsuperscript{1, 2, 6, 7} and Benoît Bardy\textsuperscript{1}

\textsuperscript{1}EuroMov, University of Montpellier, Montpellier, France.
\textsuperscript{2}International Laboratory for Brain, Music, and Sound Research (BRAMS), Montreal, Canada
\textsuperscript{3}Department of Neurology, Beau Soleil Clinic, Montpellier, France
\textsuperscript{4}Clinical Investigation Centre (CIC), 1411, University Hospital of Montpellier and Inserm, Montpellier, France
\textsuperscript{5}Current address: Open Mind Innovation, Caen, France
\textsuperscript{6}Dept. of Psychology, University of Montreal, Montreal, Canada
\textsuperscript{7}University of Economics and Human Sciences in Warsaw, Warsaw, Poland

Address for correspondence: EuroMov, University of Montpellier, Montpellier, 34090, France
Email address: loic.damm@umontpellier.fr
Tel: +33 4 34 43 26 42 Fax: +33 4 34 43 26 97

Running title: Moving to the beat

Disclosure of funding: The authors were partly supported by the European Union via the following grants: (i) FP7-ICT-2013-10 “Beat-Health” awarded to Benoît Bardy and Simone Dalla Bella, (ii) H2020-FETPROACT-2018-01 “EnTimeMent” awarded to Benoît Bardy. Simone Dalla Bella was also supported by the Natural Sciences and Engineering Research Council of Canada with a Discovery grant n. ((LOIC >> does “n.” mean number???) RGPIN-2019-05453. The authors declare no conflicts of interest.

* Corresponding author: Loïc Damm
4.2 Auditory Cueing During Walking of PD Patients Positively Influences Cortical Networks ..... 56
4.3 Running Performance in the Presence of External Auditory Pacing .................................. 64
5 OPEN RESEARCH QUESTIONS AND CONCLUSIONS ........................................................................ 68
6 ACKNOWLEDGEMENTS .................................................................................................................. 75
7 APPENDIXES ..................................................................................................................................... 76
7.1 Box 1: Information processing theory provides an alternative theoretical framework to capture auditory-movement synchronization ........................................................................................................ 76
7.1 Box 2. The cerebellar architecture ensures the mapping of sensory information into temporally relevant motor commands. ........................................................................................................... 78
7.2 Box 3: Motivation, musical pleasure and physical activity ......................................................... 82
8 REFERENCES ........................................................................................................................................ 86
9 FIGURE CAPTIONS ................................................................................................................................ 139
9.1 Figure 1. Models of oscillators. ........................................................................................................ 139
9.1 Figure 2. Interplay between neural and mechanical activities underlying Auditory Motor Synchronization (AMS). .................................................................................................................. 140
9.2 Figure 3. Functional connectivity of brain structures during AMS. ............................................. 142
9.3 Figure 4. Sensory and motor performance during AMS ............................................................. 144
9.4 Figure 5. A timekeeper account for the variability of AMS. ....................................................... 147
10 TABLES ............................................................................................................................................... 148
10.1 Table 1. Glossary of basic concepts related to auditory stimulation, gait and dynamical systems. 148
11 FIGURES ............................................................................................................................................... 152
11.1 Figure 1........................................................................................................................................... 152
11.2 Figure 2........................................................................................................................................... 153
11.3 Figure 3........................................................................................................................................... 154
11.4 Figure 4........................................................................................................................................... 155
11.5 Figure 5........................................................................................................................................... 156
Abstract

Humans’ ability to synchronize movement with auditory rhythms relies on motor networks, such as cortical areas, basal ganglia and the cerebellum, which also participate in rhythm perception and movement production. Current research has provided insights into the dependence of this action-perception coupling upon the entrainment of neuronal activity by external rhythms. At a physical level, advances on wearable robotics have enriched our understanding of the dynamical properties of the locomotor system showing evidences of mechanical entrainment. Here we defend the view that modelling brain and locomotor oscillatory activities as dynamical systems, at both neural and physical levels, provides a unified theoretical framework for the understanding of externally driven rhythmic entrainment of biological systems. To better understand the underlying mechanisms of this multi-level entrainment during locomotion, we review in a common framework the core questions related to the dynamic properties of biological oscillators and the neural bases of auditory-motor synchronization. Illustrations of our approach, using personalized auditory stimulation, to gait rehabilitation in Parkinson disease and to manipulation of runners’ kinematics are presented.

Key Words

Rhythm, synchronization, music, auditory cueing, walking, running, cadence, dynamical systems, oscillators, beat, prediction
Highlights

• What does lie behind the ergogenic benefit of music?

• Locomotion and a large number of natural movements consist of neuronal and mechanical oscillations.

• Evidences exist that predictable beat-based auditory stimulations entrain neuronal activity and engage the motor structures of the brain.

• Modelling external and internal rhythms as oscillators provides a unified theoretical framework in which sensory and motor structures are bound to each other during auditory motor synchronization.

• Neural mechanisms associated with beat tracking ability support gait rehabilitation of Parkinson’s disease patients and contribute to stabilize spatio-temporal gait parameters of runners.

• Exogenous rhythms, in order to entrain locomotion, need to comply with specific properties of biological oscillators, at neural as well as physical levels.
1 Introduction

Auditory rhythms are a strong movement incentive. The potential of music as a driving force for movement translates into the ubiquitous human tendency to move to the beat (Repp, 2005), the most salient rhythmical event. In daily life, people spontaneously or deliberately move to the beat of rhythmic sound sequences via finger or foot tapping, body swaying or walking (Leman et al., 2013). This ability, which appears to be natural and universal, develops early in humans (Zentner & Eerola, 2010). What lies behind this ergogenic benefit of music? The coupling between perception and action contributes to the emergence of the compelling link between rhythm and movement. A growing body of evidence shows that merely listening to a musical beat activates motor regions of the brain (Grahn & Rowe, 2009; Zatorre et al., 2007). Due to its peculiar and unique rhythmic features - i.e. a regular and prominent beat structure (Fitch, 2013) - music is thus ideally suited to drive the motor system (Dalla Bella et al., 2013; Madison, 2014).

In this article, we propose to analyse the ability of music to entrain movement, and we do so by reviewing experimental and modelling evidences of its physical and neuronal underpinnings. A particularly targeted context is human locomotion, e.g. walking and running. If virtually any rhythmical movement can be guided by an external rhythm, the intrinsic rhythmicity of locomotion indeed makes it a relevant target. Some structures, such as central pattern generators, have been phylogenetically selected to generate locomotor rhythms (Guertin, 2012). Among the hard-wired factors contributing to gait rhythmicity, some are specific to human lineage and relate to our proficiency in walking and running (Bramble & Lieberman, 2004). The relative extension of leg joints, typical of human gait, makes the maintenance of locomotor cycles possible at a low energetic cost (Sockol et al., 2007) despite the inertia of the lower limb. The widely
accepted mechanical models of walking and running, i.e. the inverted pendulum system (Cavagna et al., 1963; Cavagna & Margaria, 1966) and the mass spring system (Cavagna et al., 1964; Alexander, 1988; Blickhan, 1989) respectively (Cavagna et al., 1977), account for the execution of these periodic motions. Biological oscillators successfully combine energy dissipation and injection, i.e. damping and actuation exerted by the muscles (Gurfinkel et al., 1998; Gerasimenko et al., 2010), to ensure the robustness of their trajectories (Hurmuzlu & Moskowitz, 1986). These complementary mechanisms ensure gait stability. Being stable means that the rejection of perturbation is a core capability of the system. However, real time manipulation of the dynamical environment during walking demonstrates the possibility to mechanically entrain gait patterns (Ahn & Hogan, 2012b). In other words, despite gait steadiness, interactions with external rhythmical events are possible, and contribute to the stabilizing or destabilizing of walking and running.

Locomotion does not rely on simple mechanics. If the possibility to entrain gait with auditory rhythms exists, it requires the engagement of supervising structures. During finger tapping, imagery studies have revealed the contribution of motor structures of the brain during the extraction of an auditory rhythm, e.g. the pulse underlying a temporal pattern (Zatorre et al., 2007; Bengtsson et al., 2009; Grahn, 2012; Chen et al., 2006; Chen et al., 2008a). Contrasting auditory-paced movement vs. self-paced dance steps (Brown et al., 2006) showed activations specifically related to beat information during bipedal movements. Sustained periodic cortical response tuned to the frequency of auditory rhythms, as shown in electroencephalographic (EEG) studies (Nozaradan et al., 2011), can be interpreted as a manifestation of cortical entrainment. It has been suggested that oscillations, which are naturally observed both at cellular and neuronal levels, contribute to the central representation of event timing (Large et al., 2015). In more dynamical terms, an appealing theoretical model consists in viewing
neurodynamics as the support of auditory rhythm perception, and its corollary, motor synchronization.

In this contribution, we propose that sound-induced locomotor entrainment is distributed throughout interacting and reciprocally influencing neural and physical levels of behaviour. Furthermore, we argue that this multi-level entrainment organization promotes auditory-motor synchronization (AMS) as a factor of motor performance. Appropriate spatio-temporal activation of muscles is necessary for movement production, and auditory afferences can positively influence the temporal structure of motor activities when they match the requirements of the task and the intrinsic properties of neurophysiological and biomechanical oscillators. Multiple interconnected biological oscillators, although different in their structure, can be entrained to the same external rhythm. Despite the singularities of brain structures (governed by neurophysiological principles) and locomotor structures (obeying mechanical laws), modelling both structures as (coupled) dynamical systems provides one unifying functioning principle. From perception to action and from action to perception, tracking the influence of external rhythms onto biological rhythms can enable a better understanding of how rhythmical auditory stimulations find their way to motor commands and movement control, and can benefit from our multi-level entrainment framework.

Specifically, the main goal of the present review is to show how the interplay between neuronal, physiological and biomechanical mechanisms can generate the behavioural findings reported in the literature, and affect performance. Because AMS blends physical and neural processes, the boundaries of the present review span across disciplines beyond the strict scope of neuroscience. In the next sections, we introduce the theoretical framework of entrainment and then explore its ramification at biomechanical, physiological and neurophysiological levels. We detail how the chain of oscillators at these different levels are prone to
entrainment during locomotor activities, and how AMS can contribute to the
timing of action and associated performance. The use of AMS for the gait
rehabilitation of PD patients and for entraining runners’ cadence will provide
examples of the potential benefits of AMS delivered during locomotion. These
applications will finally feed the research agenda that we propose for future work
in this fast-changing area of research.

2 Rhythmic Movement and auditory-motor synchronization as Coupled Oscillators: the
Dynamical Systems Approach

The dynamical systems theory provides a framework to model movement
characteristics in the context of AMS. Moving in synchrony with an external
rhythm (a metronome or music) implies that predictable auditory events pace the
movement. The beat corresponds to the regular time interval that we can tap
along to when listening to music (Cooper & Meyer, 1960; Large, 2008). It
corresponds to an internal pulse resulting from the temporal organization of
durations in the musical piece, including different embedded periodicities (metre;
London, 2012). Whereas metronomic ticks provide unambiguous temporal cues,
in a musical stream, the beat is not necessarily a strict property of the stimulus.
Due to the complexity of musical structure, the possibility to interpret multiple
periodicities often exists. Beats being “inferred from but not identical to the
patterns of accentuation at the musical surface” (Lerdahl & Jackendoff, 1983),
variations of acoustic parameters such as pitch (Prince et al., 2009; Ellis & Jones,
2009; Prince, 2014) or intensity (Drake et al., 2000b) can also contribute to beat
perception. The metrical hierarchy can modulate the saliency of some beats with
respect to others, every other beat being accented, e.g. in a march rhythm.
Studying the temporal alignment of actions related to these auditory landmarks
has enabled the investigation of AMS.
Rhythmic movements - such as those occurring during tapping or locomotion - provide the moving limb(s) with the properties of a self-sustained oscillator, characterized by a natural frequency, i.e. the frequency of the system measured when isolated, not determined by the initial conditions at movement onset but by its internal properties, with the ability to be restored after a transient perturbation. Synchronization occurs when two oscillators adjust their rhythms by virtue of their interaction. The nature of the interaction can be of different origin: the synchronization of two clocks hanging from a common support observed by Huygens (Huygens, 1665) is a famous example of mechanical coupling. How weak or strong the interaction indicates that coupling strength is not easy to measure (Pikovsky et al., 2003). In the case of Huygens’ clocks, it depends on the ability of the board supporting the clocks to move, coupling strength being the first factor affecting the synchronization of two oscillators. The second factor is the frequency detuning, which refers to the distance between the preferred frequency of the two oscillators. When the distance between the two natural frequencies is large, their interaction does not result in the establishment of a common oscillation frequency. On the contrary, for a certain range of detuning, which delimits a synchronization region, the frequency of each oscillator is entrained towards an equal frequency, and synchronization occurs. The width of the synchronization region in frequency relates to coupling strength and detuning: the wider the region, the stronger the coupling.

When oscillation frequencies lie within the synchronization boundaries and the coupling between oscillators is strong enough, the oscillators’ frequencies converge to the same value, i.e. are entrained, and synchronization occurs. However, by looking at the phase of each oscillator, defined as their instantaneous position on the waveform cycle, different synchronization regimes can be distinguished. In agreement with the description of frequency entrainment or
frequency locking between oscillators, von Holst (1973), observing the
coordination of fish fins, emphasized the contribution of two properties of
neuronal oscillators. The *beharrungstendenz*, or maintenance tendency, describes
the steadiness of oscillations, whereas the *magneteffekt*, or magnet effect,
accounts for the influence that one oscillator can exert onto another one differing
by its eigenfrequency, resulting in the progressive drawing and coupling to its own
frequency. Both effects are in competition and contribute to relative coordination.
Von Holst interpreted fins moving at the same frequency with a constant phase
relation as the dominance of the second effect. The distinction between
synchronization regimes thus benefits from the measure of the relative phase,
defined as the phase difference between the two oscillators.

Having introduced the properties of oscillatory systems, we can now outline the
reasons why the coupled-oscillators model has been proposed as a model of
auditory-motor synchronization. It involves an adjustment of the motor rhythm to
ensure synchronization with the tempo of external sensory events. This is why
tempo has been considered as one of the most significant determinants of AMS-
related effects on rhythmic movements. Coupling strength can be measured by
the ability of the participant to reach and maintain his/her movements in
synchrony with the beats of the auditory stimulation. When the frequency of
external stimulations is fixed, coupling is unidirectional. Hence, the auditory beats
are used as a forcing oscillator. When movement frequency is entrained to the
frequency of auditory stimulations, the moving limb often oscillates at the
externally driven tempo, as long as the detuning is not too pronounced, giving rise
to frequency locking. In-phase synchronization (equivalent to absolute
coordination reported by von Holst, 1973) occurs when both oscillators reach the
same extreme positions simultaneously (the rightmost or leftmost positions in the
case of two pendulums, tapping on the beat in the case of auditory tapping)
whereas anti-phase synchronization (one specific case of relative coordination) is
characterized by movements in the opposite direction of each oscillator (one pendulum being in the rightmost position when the other one is in the leftmost position or tapping at half the period of the auditory stimulus, which results in syncopation instead of synchronization). Entrainment is the process through which two oscillators are attracted to each other by virtue of their interactions. The term entrainment is often used to refer not only to the case of perfect period and phase synchronization, but also to the tendency towards that state, perfect synchronization being just one specific case of entrainment (Trost et al., 2017). Studying sensorimotor synchronization therefore consists in characterizing frequency and phase anchoring between movement cycles and auditory stimuli.

An appropriate tempo of auditory stimulation is not sufficient to ensure sensorimotor synchronization. Since rhythmical voluntary movements involve bringing the effector into a specific position at a specific time, movement phase coordination with external sensorimotor events is required to ensure their representativeness in time and space. Plotting movement kinematics in a state space such as the space of flexion-extension angles of opposite lower limb joints makes the invariance visible in coordination patterns. A specific form of the state space - the phase portrait - where the angle and angular velocity of an oscillatory system are plotted against each other reveals the presence of attractors (Beek, 1989), which correspond to the regions where the dynamics tend to converge onto a relatively limited number of values. In other words, attractors are characterized by the reduction of the spatio-temporal variability (Byblow et al., 1994; Fink et al., 2000; Maslovat et al., 2009). The influence of auditory-motor synchronization on their emergence (Byblow et al., 1994; Fink et al., 2000) ascertains the contribution of auditory input to motor coordination, and is visible in the phase portrait as the rapid convergence towards the aforementioned values. This influence legitimates the modelling of the motor system and
rhythmical auditory stimulations as coupled oscillators, a theoretical framework in which the coordination of the former can be governed by the latter.

2.1 Attractors Determine Global Movement Dynamics and Affect AMS Stability

In the dynamical approach, the control of motor timing by central structures is considered as an emergent property of dynamical movement principles (Yue et al., 2000; Kelso, 1995; Turvey, 1990) as opposed to the information processing theory which postulates the existence of a central timekeeper (cf. box 1, section 7.1). Support to the dynamical systems approach can be found in the kinematic structure of rhythmic movements.

As presented in the above paragraph, sensorimotor synchronization is characterized by phase and frequency matching between external stimulations and movement. The former ensures timing accuracy when the latter yields appropriate time intervals. By varying the frequency of bilateral index finger oscillations, considered as a representation of coordinated movements, Kelso et al. (1981) revealed two stable modes of bimanual coordination: (i) the in-phase pattern where the relative phase between each finger’s position is zero, and (ii) the anti-phase pattern characterized by a +/- pi relative phase. These patterns can be mapped onto point attractors. Haken et al. (1985) proposed a dynamical model accounting for the observed phase diagram. In the Haken-Kelso-Bunz (HKB) model, the relevant control parameter is the frequency of oscillations (Schöner & Kelso, 1988).

Stable convergence to a dual state is not the only behaviour of a damped mass-spring system. Robustly sustained oscillations having a stable amplitude and frequency with respect to perturbation can also characterize a second-order dynamical system exhibiting limit-cycle behaviours, which appears to be a
landmark of biological oscillators. Kay et al. (Kay, 1988; Kay et al., 1987) interpreted rhythmic finger movements as the combination of noise and limit-cycle dynamics. Fluctuations around a cyclic orbit in the phase plane subsequent to perturbations ascertain the stability of the limit cycle attractor (Kelso et al., 1981). In fact, any path in the state space can be considered as an attractor, contributing to the stability of movement kinematics (Shadmehr & Mussa-Ivaldi, 1994; Won & Hogan, 1995; Burdet et al., 2001). Trajectory robustness and attractors are tightly linked because the latter ensure the re-emergence of the former after perturbation. The correspondence between periodic attractors and rhythmical movement shows that an attractor can act as a deterministic factor of movement dynamics (Rabinovich et al., 2006). From the study of phase transitions, i.e. discontinuities in coordinated movements, Haken et al. (1985) unveiled how the apparent bimodal instability fosters synergies, defined as “functional groupings of structural elements (e.g. neurons, muscles, joints) that are temporarily constrained to act as a single coherent unit” (Kelso, 2008).

Identifying synergies sustained by perceptual events can therefore address the question of movement stabilization by sensorimotor synchronization. The shape of the limit cycle of cyclical movements can be altered by rhythmical stimulations, or more exactly by certain properties of the sensorimotor coupling. In addition, the cognitive intent appears to shape the asymmetry between flexion and extension phases (Carson & Kelso, 2004). For instance, Wachholder and Altenberger (1926) (cited in Sternad, 2001) reported during wrist flexion-extension alternations that the instruction could put the emphasis on one of the phases. However, flexion-extension alternations of the finger entailed a systematic accentuation of the flexion phase. Instructions to accentuate the opposite phase altered the level of coordination. Poor synchronization performances have also been associated with off-the-beat synchronization (syncopation) at high tempo (Kelso, 1995), compared to on-the-beat synchronization. Finally, constraints arising from the neuromuscular properties
can also contribute to shaping the phases of synchronized movements. The frequent transition from an extend-on-the-beat pattern to a flex-on-the-beat one at high tempo also evidences the influence of the neuroanatomical characteristics of recruited muscles (Carson & Riek, 1998). As anti-gravity muscles, flexors are stronger than extensors, require a smaller proportion of motor unit recruitment for the same force production (Vallbo & Wessberg, 1993), receive more facilitatory projections from the cortex (Palmer & Ashby, 1992) and require less cortical activity (Yue et al., 2000). In short, the intention and differential neuroanatomical organization of flexion and extension, among other factors, contribute to AMS stability (Carson & Kelso, 2004). A deep understanding of structural and functional properties of biological effectors is required to catch how their emergent dynamical properties contribute to a successful auditory-motor synchronization.

2.2 Low-Level Movement Parameters Affect Rhythmic Entrainment

Zooming down towards the lower level of movement parameters, such as the resonant frequency of the end effector and its muscular activities, allows us to explore the complementary effects of auditory stimulation on movement production. First, any moving segment is characterized by its eigenfrequency. Auditory-motor frequency locking occurs when the frequency of the stimulus and the eigenfrequency of the system under scrutiny are not too far apart (Leman, 2016), or are at multiples (Roerdink et al., 2009) or fractions (Bouvet et al., 2019) of the natural frequency of the system. When participants perform underarm swings while keeping in cadence with audio-visual stimuli, better synchronization accuracy at a preferred tempo (i.e. the tempo spontaneously adopted by participants, which is at or close to the resonant frequency of the moving limb (Goodman et al., 2000; Abe & Yamada, 2003)), rather than at a faster or slower tempo, is usually reported (Smoll & Schutz, 1982). This finding is in agreement with
the theory of dynamical systems. It confirms the existence of a basin of entrainment, resulting in convergence towards the limit cycle, partly determined by the inertia of the moving limb (Hatsopoulos & Warren Jr, 1996).

Second, entrainment is also visible at the muscular level. Safranek et al. (1982) were among the first to relate electromyographic (EMG) activities with auditory cueing. In their study, participants were instructed to achieve a sequence of reaching movements between three targets in silence, while listening to even or uneven rhythms. EMG variability decreased under the influence of even rhythm while it increased with uneven rhythm. Thaut et al. (1991) noticed less co-contraction of biceps and triceps during a flexion-extension task of the elbow eliciting up- and down-swings of the forearm when the tempo of auditory cues was matching the natural frequency of the movement compared to the same movement performed in silence. An external rhythmical pacemaker can thus elicit changes in onset, duration, and variability of muscular activities during simple flexion-extension movements or during more complex movements such as reaching. In a more recent study, Yoles-Frenkel et al. (2016) clarified the influence of auditory context on the kinematics and corresponding patterns of muscular activity during a finger tapping task. They confirmed the reduction of EMG variability in response to auditory cues and noticed the concomitant reduction of variability in the velocity profiles. More specifically, the variability of both acceleration and muscle activity was reduced when tapping was auditorily cued, during the pre-tap period, but not during the actual tap. Taken together, these findings witness auditory entrainment to the beat at the level of the neuromuscular machinery, from muscular activities to limb movement.
2.3 The Contribution of Nonlinear Oscillators to Locomotor Control

An enormous literature, partly cited in the present review, has analysed AMS-related human behaviour during isolated cyclical movements of the upper limb, such as during tapping to the beat, which is not the most representative movement of our everyday life repertoire. In contrast, locomotion, together with the way the two lower limbs have functionally been coupled throughout evolution, has been less studied in the AMS context. Locomotion is probably the most phylogenetically preserved rhythmical physical activity in humans and other animals. Beyond striking differences between species, e.g. quadrupedalism vs. bipedalism, evolutionary conservation has shown similarities in the development of the neural command of locomotion (Grillner, 2011). All vertebrates possess spinal neuronal networks called Central Pattern Generators (CPG), which can produce meaningful functional patterns of motor activities in the absence of sensory inputs (Brown, 1914). CPG activity has been evidenced by fictive locomotion when the spinal cord is isolated from the periphery in non-human vertebrates (Brown, 1911; Grillner & Wallén, 1985; Grillner et al., 1998). In humans, the evocation of locomotor movements with spinal electromagnetic stimulations provides an indirect indication of the presence of CPG in the spinal cord (Gerasimenko et al., 2010). Nonlinear oscillators used to model CPG (Matsuoka, 1987; Collins & Richmond, 1994; Rybak et al., 2006) possess distinctive properties such as stability, that enable input from feedback pathways with a high entrainment potential. Locomotion would be supported by lower-level command modules, possessing the characteristics of nonlinear oscillators, engaged in specific rhythm generation (a configuration which is more unlikely to underlie tapping for example). However, the gait of patients suffering from spinal cord injury also suggests the contributing role of the supraspinal command to locomotion. Indeed, patients following body-weight support training after recovery exhibit a foot trajectory close to that of healthy participants, with a
different lower-limb coordination when evaluated through the phase relation of their angular motion (Grasso et al., 2004).

High-level and low-level commands are thus not mutually exclusive in producing entrainment, and their complementary nature has been studied, for instance by Ahn and Hogan (2012b). In a striking study, the authors attempted to assess their respective contributions during walking. They used an innovative therapeutic robot able to mechanically perturb the ankle joint. The command of the robot, designed as a torsional spring compensating for its intrinsic inertia, applied periodic mechanical perturbations. Periodic square plantar-flexor torque pulses were applied at frequencies different to the walker’s preferred cadence. This set-up was designed to test the dynamic entrainment hypothesis. The torque perturbation was delivered at constant frequency throughout one trial and consequently the relative phase between its occurrence and gait could vary. The evolution of relative phase across successive strides after the application of low-to high-frequency perturbations was used to assess the basin of entrainment. Walking was found to be entrained by the mechanical perturbation, although only for a limited range of perturbation frequencies. Phase locking occurred during the push-off phase such that the perturbation assisted the movement. These behavioural results, in agreement with the prediction of the nonlinear coupled oscillators theory, show the possible contribution of neuro-mechanical oscillators to walking. However, these effects appear to be modest, as witnessed by the slow convergence of steps and perturbation phases. Ahn and Hogan’s (2012b) conclusion calls for the coexistence of spinal semi-autonomous oscillatory activities and central supervision, a hierarchical organization they refer to as the “episodic supervisory control of a semi-autonomic periphery”. This engineering terminology resonates with the functional anatomy of locomotion. CPGs, which are distributed spinal networks responsible for the sequential activation of antagonist muscle groups (Grillner, 1981), are not the only subsystems involved in
the control of locomotion (Figure 2E). Brainstem command centres activate CPGs, silent at rest, via reticulospinal neurons, which integrate afferent sensory feedback (Zehr & Duysens, 2004; Nielsen, 2003) and vestibular signals (Zelenin et al., 2007). Integration of posture and locomotion is a prerequisite for ensuring propulsion, but adaptation of the building locomotor blocks to a behavioural context is also necessary. Experiments in animals have demonstrated the contribution of the forebrain in movement adaptation (Bjursten et al., 1976; Whelan, 1996). But the fine-tuning of behavioural repertoire benefits from cortical input, the fast-direct corticospinal pathway projecting on most human muscles (Mills, 1999). Interestingly, the direct micro-stimulation of the motor cortex in autonomous walking animals was found to reset step cycle during stance, whereas it increased activity level and duration of flexor muscles during the swing phase (Bretzner & Drew, 2005). The motor cortex thus appears to contribute to both the structure and the timing of locomotor patterns.

The functional organization described above raises the question of the possible influence of periodic auditory stimulations on locomotor activities. A number of studies have shown the influence of rhythmical auditory cues on spatiotemporal parameters of gait. Participants instructed to synchronize their footsteps with a metronome 22.5% faster than their natural cadence increased their walking velocity by increasing their cadence (Leow et al., 2014). Metronome clicks, when 10% faster than a healthy elderly cadence, have recently been found to elicit the increase of both stride length and cadence in the Dotov et al. (2019) experiment. Roerdink et al. (2011) showed that the efficacy of auditory stimulations to elicit gait changes weakened as a function of the gap between pacing tempo and participants’ cadence, an observation in line with the dynamics of coupled oscillators. Beat perception has also been found to accompany the modification noticed in gait performance when synchronizing footsteps with auditory cues. For instance, Leow et al. (2014) reported an association between strong beat
perception and gait performance in the presence of a metronome or music, whereas weak beat-perceivers exhibited slower, more cautious gait, particularly in the presence of musical excerpts with non-salient beats. Ready et al. (2019) confirmed the alteration of gait, as indicated by the increase of double-limb support time, in weak beat-perceivers instructed to synchronize their steps with rhythmical cues, either a metronome or low-beat saliency music. From these converging results, we can conclude that there is an interaction between beat perception ability and auditory-based gait control, illustrating the circular causality between perception and action during AMS. We now focus on the neural processes allowing the brain to make use of the temporal regularities of rhythmical auditory stimulations during motor activities.

3 The Neural Bases of Auditory-Motor Synchronization

In the previous section (2), we clarified the physical principles producing entrainment according to the dynamical systems perspective, and how they could contribute to coupling movement characteristics with auditory beats. Considering the oscillatory nature of locomotor activities, walking and running represent a natural target for periodic auditory-cueing, which can be achieved only if information finds its way into the motor system. We have demonstrated various levels of auditory entrainment, from localized muscular activity to the movement of one joint, to a more global entrainment of the locomotion system when external rhythmic auditory events take place. We now focus on the brain structures involved in that biological entrainment to the beat. Despite our interest in locomotion, this class of movement does not occupy the largest place in AMS studies. In particular, the simple task of finger tapping to an auditory sequence of tones has been a widespread approach in laboratory studies (for a review see Repp, 2005 and Repp & Su, 2013). This task, which requires moving only one finger, facilitates the listening of stimuli and the design of experimental set-ups. The precision of data acquisition from neuroscience methods, particularly brain
imaging techniques, certainly benefits from the participants’ stationarity. We will therefore often refer in the present section to finger tapping studies without which our understanding of neural substrates of AMS would be less advanced. Beyond the identification of brain structures that are involved in the synchronization between movement and audition, the goal of this section is to introduce the most convincing experimental evidence of brain entrainment by auditory cues, which represent a cornerstone of our hypothesis: multi-level entrainment.

3.1 AMS Couples Sensory and Motor Cortical Areas

Listening to periodic auditory streams gives rise to a sense of pulse, more commonly referred to as beat (Nguyen et al., 2018), which determines the spontaneous tapping rate of listeners. Action-perception coupling appears to functionally define the beat. In the present section, we therefore consider beat perception as a temporally relevant detection of the auditory signal likely to be coupled with a goal-directed movement. In order to uncover the specificity of beat perception, it is necessary to complement the widespread notion of psychological event with the clarification of underlying neural mechanisms, particularly the relation between sensory and motor structures. How does the endogenous neural shaping of auditory rhythms drive movement onset? We now review the most compelling evidence in the literature supporting the view that identical neural structures are activated during the encoding of temporal stimuli and during movement control.

3.1.1 Activation of the striato-thalamo-cortical network during beat perception and AMS

Protocols entailing rhythm perception while listening to auditory rhythms have isolated several brain structures which respond to the beat. Functional magnetic resonance imaging techniques (fMRI) clearly reveal the activation of dorsal
premotor cortex, supplementary motor area, pre-supplementary motor area, lateral cerebellum and basal ganglia when participants are listening to beat-based rhythms (Grahn & Brett, 2007; Chen et al., 2008a; Grahn, 2009; Bengtsson et al., 2009; Grahn & Rowe, 2009; Grahn & Rowe, 2013; Kung et al., 2013; Merchant et al., 2015; Chauvigné et al., 2014). The same structures appear to be associated with beat-driven actions (Rao et al., 1997, Schubotz et al., 2000, Rao et al., 1997, Dhamala et al., 2003, Lewis et al., 2004, Kung et al., 2013, for a meta-analysis, see Chauvigné et al., 2014 and Merchant et al., 2015).

### 3.1.2 Movement Sequence Organization in the Motor Cortex

The traditionally accepted functions of premotor (PMC) and supplementary motor (SMA) cortices include the organization of sensory information relevant for the guidance of the movement towards the goal (Weinrich & Wise, 1982), and for the planning of ordered sub-movements, respectively (Mushiake et al., 1991; Tanji & Shima, 1994; Shima & Tanji, 1998). PMC contributes to the production of isochronous rhythms but is not specifically related to sensorimotor synchronization. For instance, during an auditory-paced hand movement, Pollok et al. (2017) reported the absence of modulation of synchronization accuracy while stimulating dorsal PMC with transcranial direct current stimulation (tDCS).

In contrast, the pre-supplementary motor area (pre-SMA) appears to be more directly related to temporal processing. SMA is divided into two different parts rostrocaudally (Picard & Strick, 2001): (i) the pre-SMA, whose phasic activation is related to spatial parameters of the forthcoming movement, and (ii) the SMA-proper, which exhibits phasic movement-related activity time-locked to movement onset (Matsuzaka et al., 1992; Tanji, 2001). During repetitions, the temporal structure of the forthcoming movement is shaped by pre-SMA, whereas its implementation depends on SMA-proper activity (Kotz & Schwartze, 2011). Jantzen et al. (2007) illustrated the functional difference between SMA-proper and pre-SMA by modulating the context of the motor task. When the pacing
stimulus is suspended following a sustained auditory-paced tapping task, contrary to pre-SMA, SMA proper is equally activated irrespective of the pattern of coordination, whether syncopated (on the beat) or synchronized (off the beat). Contrary to cortical areas exhibiting context-dependent activations, the SMA proper would have “a more ubiquitous role in motor timing” (Coull et al., 2011), a hypothesis supported by the activation of that structure during non-motor timing tasks such as duration comparison (Coull et al., 2008). In sum, among those areas related to the control of movement, in which initiation and termination phases structure neuronal activity, and which are also activated by rhythmic movement-based sensory afferences, the SMA appears to be crucially linked to AMS.

3.1.3 Functional Connectivity Between Auditory and Motor Areas

Investigating the structural connectivity of premotor areas provides clues on how their functional role is achieved during AMS: being at the interface between sensory inputs from the thalamus, sensory cortices, and the primary cortex exerting a direct control on simple movements, they are the perfect candidates for regulating the temporal relations between afference-driven actions. Action–auditory perception coupling has long been evidenced, for instance by Chen et al. (2006; 2008b) who investigated the engagement of different parts of the premotor system, by contrasting distinctive action–perception conditions, passive listening, listening with anticipation to tap (listen then tap) and tapping, with subtraction analyses conducted on fMRI data (figure 3 A & B). Functional connectivity between auditory and dorsal premotor cortices was found to be effective during rhythmic tapping, as opposed to passive listening or listening with anticipation to tap, and was a function of metric saliency when subjects tapped in synchrony to isochronous rhythms. The ventral premotor cortex activity did not appear to be sensitive to the temporal complexity of the stimulus. However, it appeared to be directly engaged in the processing of action-related sounds, its activity being significant only during listening with anticipation and during tapping
(Chen et al., 2008a). Chen et al. (2009) concluded that auditory and motor functions, as well as their coupling, can be tracked within the premotor cortex. The auditory cortex also contributes to auditory-motor transformations. Within the auditory regions, the processing of spectro-temporal auditory patterns, initiated in the ascending auditory pathways and the primary auditory cortex (Nelken et al., 1999; deCharms et al., 1998; Fritz et al., 2003), is continued in the planum temporale (Obleser et al., 2007; Nourski et al., 2009; Belin et al., 2000; Griffiths & Warren, 2002), which occupies the posterior part of the superior temporal plane. The connections of the planum temporale with the dorsal premotor cortex (Schmahmann et al., 2007), and their role in the disambiguation of complex sounds, as well as the effects of lesions (Liégeois-Chauvel et al., 1998), and the neuroanatomical characteristics of musicians and their associated rhythmical proficiency in AMS (Meyer et al., 2012), argue for the primary role of the planum temporale in auditory-motor coupling.

The segregation of auditory afferences in two processing streams was hypothesized decades ago (Rauschecker, 1997). The role of the ventral auditory pathway would be to support sound categorization, in opposition to the dorsal pathway, emanating from the planum temporale (temporo-parietal junction area and ventral premotor cortex, Romanski et al., 1999), whose activity appears to be tuned to spatial parameters of the sound (Tian et al., 2001). The faster neural latencies of the dorsal stream (Camalier et al., 2012) argue nonetheless for the hypothesis of its role in processing time (Rauschecker, 2018). The dorsal pathway would transform acoustic information into motor representations (Warren et al., 2005) by providing the motor system with time-tracked relevant information (Karabanov et al., 2009).
3.1.4 Short-Term Memory Engagement During Ensemble Performance

Musical performance in an ensemble is representative of auditory-motor synchronization. Musicians have to deal with the inherent complexity of the musical stream which can make beat identification challenging. Matching the central spectro-temporal representations of the stimuli with the predicted or learned ones ‘stored’ in the working memory has been proposed as a mechanism to decipher musical rhythms. Keller (2008) formulated the hypothesis that, in an orchestra, individual ongoing musical performance could be guided by predictions about the co-performers’ ongoing actions, and supporting behavioural data was provided (Pecenka & Keller, 2009). Participants who perform best when asked to imagine the continuation of a five-beat sequence, and to judge the temporal acuity of a consecutively presented probe beat, were also those who demonstrated good performance in sensorimotor synchronization. The execution performance of musical pieces is logically altered by working memory load (Maes et al., 2015). When tapping in synchrony with progressively more complex and less metrically structured auditory rhythms, greater activity in the PFC of musicians ascertained the contribution of working memory in the deconstruction of complex temporal structure (Chen et al., 2008b). A temporally accurate hierarchical planning of action (Koechlin & Jubault, 2006; Koechlin et al., 2003), and the involvement of working memory (Petrides, 2005), mediated by PFC, have been proposed as explanatory factors for the superior sensorimotor synchronization ability of musicians (Chen et al., 2009). The functional division mentioned in the previous section supports this contribution. The duality of the SMA connectivity - pre-SMA having extensive connections with prefrontal cortex (Wang et al., 2005; Luppino et al., 1993), and SMA-proper projecting directly to motor and premotor cortices (Johansen-Berg et al., 2004; Luppino et al., 1993) as well as spinal cord (Luppino et al., 1994; He et al., 1995) - indeed bears this functional organization. The engagement of short-term memory processes of the PFC is a function of the complexity of the rhythm sequence as revealed by the comparison of cerebral
activities while listening to isochronous sequences vs. metrical or non-metrical sequences (Bengtsson et al., 2009).

In summary, a clear coupling exists between the brain areas that organize movement sequences and those in charge of extracting relevant auditory structure during AMS, beat identification in our specific case. The functional connectivity between auditory and motor areas, specifically reinforced during beat processing, allows the construction of temporally tracked motor representations in pre-motor areas. Behavioural and electrophysiological data, as well as the analysis of connectivity, do indeed support the sequencing of multiple movements in the SMA. On the one hand, specific contributions of pre-SMA receive input from working-memory-related structures and define the temporal relations of events. On the other hand, SMA-proper, close to the descending pathways, is logically assigned with an implementation role (Kotz & Schwartze, 2011). A structure dedicated to the spectral and temporal analysis of complex sounds - the planum temporale - projects onto both dorsal premotor and prefrontal regions. This suggests a tight coupling between auditory and motor areas, with input to and from working memory. The present section has also outlined the whole functional neural loop including basal ganglia and the thalamus. The functional role of premotor areas is achieved through the involvement of the cortical outputs to the basal ganglia and the thalamus, which connects both subareas of SMA (Johansen-Berg et al., 2004), and the output pathway from the basal ganglia back to the cortex. We are now therefore focussing our interest on subcortical structures of the brain. The known contribution of cortico-basal ganglia loops in the contextualization of sensory information for motor learning (Graybiel, 1995) makes their involvement coherent in auditory-motor synchronization.
3.2 Basal Ganglia Fine Tune Movement Timing

3.2.1 AMS and Basal Ganglia Dysfunction

The putative role of basal ganglia in action timing is first supported by pharmacology. The impairment of the execution speed of a tapping task by antagonists of dopamine (Rammsayer, 1997) indeed points towards the association between dysfunctions of the dopamine system and motor synchronization impairment. Our understanding of AMS mechanisms has largely benefitted from the investigation of the abilities of patients suffering from Parkinson’s disease (PD), characterized by the degeneration of dopaminergic neurons in the substantia nigra pars compacta inducing a decrease in striatal dopamine content (Hornykiewicz & Kish, 1987; Kish et al., 1988). The pathology alters the balance between direct and indirect pathways which have opposite net effects (excitatory vs. inhibitory respectively) on thalamic target structures. The loss of the nigrostriatal pathway, which excites the direct pathway and inhibits the indirect pathway, consequently favours the latter at the expense of the former.

PD patients exhibit selective deficit in discriminating beat-based rhythms (Grahn & Brett, 2009). Compared to controls, patients at an early stage of the pathology, i.e. when dopamine depletion is restricted to the putamen and dorsal caudate nucleus, do not benefit from the presence of beats in the sequence to discriminate rhythms. PD patients completing auditory-paced repetitive flexion-extension of the wrist are known to exhibit less accurate inter-response intervals than controls for high frequencies (Pastor et al., 1992). Freeman et al. (1993b) reported larger variability of tapping rhythms among patients in addition to slower and higher rates than the reference below and above 3 Hz respectively. The fMRI data of patients during a finger tapping task (Elsinger et al., 2003) demonstrated a decreased activation within the sensorimotor cortex, cerebellum, and medial premotor system in PD patients compared to controls. Dopamine
supplementation restored brain activation patterns in patients but led to moderate if no performance improvement during task completion, evaluated in terms of accuracy and variability. The lack of performance increase with dopamine, also reported by O’Boyle et al. (1996), in conjunction with brain activation restoration, is intriguing. In an attempt to conciliate these findings, Cameron et al. (2016) asked PD patients and controls to complete a rhythm discrimination test. The performance of the PD patients, which was below that of the controls, was more affected by the complexity of stimulus metricality when they were off medication: in other words, medication increased their sensitivity to simple rhythmic changes and decreased their sensitivity to complex rhythm changes. Even if this appears as a confirmation of basal ganglia function, which is supposed to be partially restored by dopamine, in beat perception, clear explanations of medication-related performance alteration in the presence of complex stimuli are still needed. (LOIC > CHECK THAT THIS CHANGE IS OK))

3.3 Cerebellum Circuitry Predicts Sensory Input in AMS

When reviewing behavioural data, we found a general agreement for the main contribution of the cerebellum to duration-based timing, evidenced by the alteration of the timing abilities of patients affected by cerebellar lesions (Ivry et al., 1988; Moberget et al., 2008; Grube et al., 2010a), or by the disruption of cerebellar function with rTMS (Grube et al., 2010b). However, recent research has extended the implication of the cerebellum to beat-based timing. For instance, Paquette et al. (2017) used voxel-based-morphometry, a neuroimaging technique which reveals the relation between behavioural variables and gray matter signals (Ridgway et al., 2008). They aimed to clarify the relation between inter-individual variations in gray matter volumes across the entire brain and beat interval discrimination skills evaluated with Harvard Beat Assessment Tests. The data supported the involvement of the cerebellum in the beat finding and interval test. Measures of performance in cerebellar patients during temporally non-adaptive
(isochronous pacing) and adaptive (tempo-changing pacing) AMS and a perceptual tempo judgement confirmed a global temporal processing dysfunction (Schwartze et al., 2016). The fact that noticed impairments translated into increased asynchronies and impaired error correction, Penhune et al. (1998) and Schwartze and Kotz (2013) reinforced the conception of the cerebellum as a structure dedicated to the reduction of time relevant sensory input, i.e. the encoding of the sensory flow into temporal discrete events.

The sensory coordination feature of the cerebellum (Gao et al., 1996) also enables the cerebellar networks to act in a feed-forward way, anticipating the sensory consequences of the planned action (Bower, 1997; Mauk et al., 2000; Courchesne & Allen, 1997; Ito, 2005; Ramnani, 2006; Pisotta & Molinari, 2014). The inherent delay of information processing and motor response would spoil AMS with unbearable temporal errors if mere reaction were the brain’s dominant functioning mode. The existence of delays in the sensorimotor system, as well as its limited resources, are compensated by anticipation, a growingly recognized brain function (Bar, 2007; Friston, 2005).

The cerebellar circuitry thus seems to contribute to inserting the timing of sensory inputs into the motor output (box 2, section 7.2). The discrimination of events according to their temporal saliency, associated with temporally specific learning, builds a temporal representation of events in the sub-second range, a key determinant of auditory driven actions. The description of the cerebellum as a brain structure in which online afferences and previous patterns of activations are compared and recognized, i.e. the very sensory coordination feature of the cerebellum (Gao et al., 1996), thus considers the whole cerebellar network as a feed-forward structure, anticipating the sensory consequences of the planned action (Bower, 1997; Mauk et al., 2000; Courchesne & Allen, 1997; Ito, 2005; Ramnani, 2006; Pisotta & Molinari, 2014). Electrophysiological data support this
role. In extra-cerebellar rat structures (superior colliculus and thalamus), for instance, Crispino and Bullock (1984) reported a modulation of the relative weight of auditory afferent inputs when the cerebellum was electrically stimulated. Tesche and Karhu (2000) measured evoked MEG responses to somatosensory stimulations in humans. In their study, the median nerve or the finger was stimulated at a steady interstimulus interval. Random omission of the stimulus did not alter the sustained evoked activity of the cerebellum. Moreover, the existence of a refractory period in the cerebellar response, caused by preceding stimulus, suggests the persistence, within neural networks, of the temporal pattern of somatosensory stimulations during a few seconds. Therefore, sensorimotor coordination appears to benefit from a short-term template of expected afferences, offering the possibility to overcome the delays inherent to feedback control (Wolpert et al., 1998a).

The PPC maintains cerebellar prediction until task completion (MacDonald & Paus, 2003; Wolpert et al., 1998b). In AMS tasks, Pollok et al. (2008) identified the cerebellum and the PPC as a functional unit: measures of tap-related phase synchronization in conjunction with neuromagnetic activity evidenced functional connectivity within a cerebello-diencephalic-parietal network before tap onset when the stimulus was isochronous. When the stimulus onset was randomized, an increase of the parietal–cerebellar functional connectivity was found after tap onset, and was attributed by the authors to the mismatch detection. The analysis of functional connectivity among professional musicians further suggests a positive influence of the functional interaction within a PMC–thalamus–PPC loop on timing precision (Krause et al., 2010). Interestingly, the role of the cerebello-diencephalic-parietal loop in anticipatory motor control was also confirmed by transcranial direct current stimulation (tDCS: direct stimulation with low current via electrodes contacting the head; Anodal tDCS increases neuronal excitability and cathodal tDCS decreases it) applied over the PPC while participants were
synchronizing their finger taps with auditory cues (Krause et al., 2014). In this study, PPC tDCS was found to affect finger synchronization, with performances being degraded by anodal tDCS and improved by cathodal tDCS, a result that was replicated in a population of musicians (Pollok et al., 2017). Repetitive transcranial magnetic stimulation (in rTMS, the application of a magnetic field in a targeted region of the brain elicits repeated electric pulses) over PPC ascertained the same dependence of the auditory and visual synchronization modalities on PPC activity, and confirmed the anticipatory processes taking place in this brain area (Krause et al., 2012). The fact that PPC tDCS affected synchronization, but not continuation, is also in agreement with the view of an afference-efference matching process within the PPC.

3.4 Thalamo-Cortical Network Sustains Rhythmical Movement-Auditory Couplings

We have so far emphasized the role of basal ganglia, thalamus, SMA and cerebellum as key structures of the functional loop implementing rhythmical automatized movements (Kotz & Schwartze, 2010; Kotz & Schwartze, 2011). It appears that sequential movement preparation depends on SMA activity (Mushiake et al., 1990; Shima & Tanji, 1998), particularly when the task requires the update of motor plans (Shima et al., 1996). Once cycles are engaged and competing motor programmes eliminated (Mink, 1996), basal ganglia discharges are related to the encoding of the learned afference-action relationship and sub-movement completion (Mushiake et al., 1990). The whole ganglia-thalamo-cortical network is thus engaged in the self-generation of action and the control of its timing (Jueptner & Weiller, 1998). Since their discovery, direct versus indirect pathway models of basal ganglia have been refined in light of anatomical variations and have revealed intrinsic and extrinsic connections (Jahanshahi et al., 2015). This is important in order to understand the contribution of the ganglio-
thalamo-cortical loop to AMS. Crucially, SMA controls the activation of STN (directly connected to the substantia nigra pars reticulata, the principal output nuclei of the basal ganglia) via the hyperdirect pathway. This ensures the connection between motor cortical areas and the globus pallidus (Monakow et al., 1978; Kitai & Deniau, 1981) with shorter latency than direct and indirect pathways going through the striatum and the internal globus pallidus (Nambu et al., 2002). This cortico–subthalamo–pallidal pathway would allow the cortical activity to favour the emergence of the structure of the upcoming movement (Nambu et al., 2002).

In rhythmic tasks, matching the movement phase to exogenous cues relies on the capacity of the cerebellar-thalamo-cortical loop to encode an event-based temporal structure (Kotz & Schwartze, 2011; Coull et al., 2011). Temporal relations of task-relevant events and the timing of anticipated future events are conveyed by the cerebellar projections on pre-SMA and SMA via the thalamus. The “sensory coordination” hypothesis, which postulates the control of temporal parameters of the cortical activity by the cerebellum, adds a complementary role to the cerebellum, that of implementing “computational subroutines capable of entraining the firing rates of different neural populations” (Molinari et al., 2007). Hence, the cerebellum seems to contribute to the coupling between auditory stimuli and neural activities, in agreement with the two fundamental functions of cerebellar processing identified by D’Angelo and De Zeeuw (2009): timing and sensorimotor adaptation.

The extraction of the temporal features of a predictable auditory sequence, the development of temporal predictions, and the coupling of action to salient auditory events thus require, in a complementary manner, the integrity of the ganglia-thalamo-cortical and cerebellar-thalamo-cortical networks (Dalla Bella et al., 2015). In addition, recent investigations of cerebellum connectivity with other
brain structures (Bostan et al., 2013) have evidenced an anatomo-functional
interrelation between these two networks. Both subareas of the SMA are
connected to the cerebellum: the non-motor domain of the dentate nucleus
projects to the pre-SMA while its motor domain is related to the SMA proper
(Akkal et al., 2007; Dum & Strick, 2003). There is a convergence of subcortical
projections on the SMA from both the basal ganglia and the cerebellum, the
former ones being more represented than the latter ones (Parthasarathy et al.,
1992; Inase et al., 1999; Akkal et al., 2007). Interaction between these subcortical
structures, which still remains to be clarified, would make short latency influence
of the cerebellum on basal ganglia possible. Chen et al. (2014) have indeed shown
in mice the existence of a disynaptic pathway, originating from the cerebellum,
able to alter the activity of striatal neurons. A substantial disynaptic connection
from the subthalamic nucleus to the cerebellar cortex was also found in monkeys
(Bostan et al., 2010). In sum, the rapid, two-way communication between basal
ganglia and cerebellum captures a functional unit. The highly interactive circuitry
between the cerebellum, the basal ganglia, and the cortex (for a review see
Caligiore et al., 2017) supports the transmission of cerebellar timing predictions
to the thalamo-striato-cortical processing streams.

3.5 Coupling sensory and motor structures during AMS by
considering external and internal rhythms as oscillators

In our effort to understand the neural bases of AMS, we have summarized four
decades of neurophysiological data, and we have built a rough scheme of the
underlying operating networks. The global network - reviewed above - of brain
structures and their interconnections contributing to the synchronization of
movement with auditory cues (see figure 2B for a visual summary) constitutes the
neurophysiological foundations on which functional and reciprocal influences can
be understood. The classical view of brain area specialization - if it satisfies a rough
understanding of structure-function relationships - does not account for the
complete dynamics of sensory-motor coupling, in which sensory and motor boundaries tend to vanish. Localizationism can only partially unveil AMS mechanisms, not only because of the interleaving loops, but also because of the exact nature of the coupling between auditory afferences and efferences. The main focus of interest needs to be on the functional connections between the brain structures which are bound together during AMS. For example, the contribution of the sensory cortex to motor control has been evidenced by the evocation of motor responses following stimulations of the primary somatosensory cortex (Matyas et al., 2010). This has shown how prominent the appraisal of functional connections is for the understanding of externally driven movements.

A hint at auditory-motor linkages can also be found in premotor cortex activation based on individual tempo preferences. In participants who passively listened to tone sequences embedding weak or strong beats, TMS of the primary motor cortex revealed larger motor-evoked potentials (MEPs) in ankle-driving muscles, in the presence of metrically strong sequences (Cameron et al., 2012). After having expressed a preference for slow or fast tempo, participants exhibited larger BOLD activity in the premotor cortex when they were presented with the tempo they most highly rated (Kornysheva et al., 2010). Moreover, inhibitory repetitive TMS over premotor cortex altered tempo preference (Kornysheva et al., 2011). Michaelis et al. (2014) hypothesized that preferred tempo could depend on an endogenous oscillator, which would translate into spontaneous motor tempo and preferred perceptual tempo. The authors experimentally identified the latter as the natural tapping tempo of participants, and the former by asking participants to rate their preference for tone sequences in a series differing by their pace. Both were correlated. Moreover, when participants were listening to their preferred tempo, corticospinal excitability was modulated by the TMS of motor cortex during the tempo perception task (motor-evoked potentials being either
increased or decreased across individuals). These results show how rhythmical perceptual and motor tasks translate into action-perception synergies. Interestingly, protocols entailing measurements of behavioural and brain activations point towards the dependence of coupling magnitude on the matching between external and preferred rhythms.

The coupling between dynamical systems through resonance has been proposed as a theoretical framework to account for this relation. Large and Kolen (1994) regarded the perception of rhythmical structures as a dynamical process: their model of entrainment — phase and frequency locking of oscillating networks with periodic events of auditory rhythms — emphasizes the temporal structuring of motor rhythms by incoming ones (Large, 2000). In the case of musical performance, internal self-sustained oscillations covering different periods would entrain to the rhythms of music (Large & Palmer, 2002). Before discussing the experimental evidence for and against this theoretical approach, we first describe the basic mechanisms of endogenous neural oscillatory activities.

3.5.1 Oscillatory Neuronal Activity and Network Architecture Allow

Large Scale Neuronal Interactions

The distinctive properties of the neuronal membrane ensure sharp variations of ionic conductance, which result in a depolarization followed by a hyperpolarization of the membrane potential. The waveform and features of this elementary membrane event, an action potential, are accurately predicted by the Hodgkin and Huxley model, which describes membrane conductance as a function of its potential. An action potential, often referred to as a spike due to its shape in electrophysiological recordings, occurs when a stimulus exceeds the depolarization threshold according to the all-or-none law. Synapses functionally fill the physical gap between neurons. In synapses, the electrical signal is converted into a chemical signal, in the form of a neurotransmitter release. The
subsequent binding of the transmitter to the postsynaptic receptors switches the signal back again into an electrical form. The type of coupling varies however according to the nature of the released transmitter, synapses being either excitatory or inhibitory.

However, it would be misleading to think of a neuron as a passive integrator, only ensuring the transfer of upstream information towards downstream connected cells in the form of action potentials. The neuron doctrine, already predominant before the 1980s, indeed put the emphasis on the connectivity of neurons modelled as bi-stable gates, with some computational ability (Llinás, 1988).

Membrane conductance to ions depends on the regulation of the opening of ion channel proteins. This regulation, referred to as gating, depends on the dynamical cooperation between several processes that we will not describe here in detail. The diversity of channels included in the membrane, the regulation of their expression, and their opening kinetics all give rise to a complex integration ability at the neuron level, beyond its activation threshold and the associated bi-stability.

A clocklike neuronal activity lies in the range of patterns of a single cell, but it also includes transient and sustained firing, as well as subthreshold oscillations, depending on the physiological conditions such as the ionic concentration and the presence of neuromodulators. In other words, the neuronal excitability exhibits a large repertoire of possibilities (Connors & Regehr, 1996), among which oscillations, when the time occurrence of spikes is periodic (König, 1994; Kaneoke & Vitek, 1996). This is the first organizational level of brain oscillatory activity (figure 1B, Stiefel & Ermentrout, 2016).

Experimentally, the dynamical properties of neurons have been investigated by recording the firing of single cells in response to alterations of their potential. Puil et al. (1986) used frequency-domain techniques to model the contribution of passive and active neuronal membrane properties to input–output relationship.
They varied the frequency of the small amplitude perturbations of the membrane voltage, observed the neuronal voltage response and calculated the neuronal impedance, i.e. the frequency dependent resistance. The membrane (modelled as a capacitor), separate charges, and ion channels (modelled as a resistor) allow charged particles to flow (or to leak): the combination of a capacitor and a resistor, characterized by a time constant which represents the time necessary to reach a new steady state when voltage across the membrane is modified, passive properties of the membrane low pass filter voltage variations. Voltage-gated ion channels, modelled as an inductance, yield an active process ensuring the responsiveness for high-frequency variations of membrane potential. The combination of low- and high-pass filtering properties gives distinct frequency preference to neurons (Pike et al., 2000) and qualifies them as a resonant system (Hutcheon & Yarom, 2000).

Given that the hallmark of neurons is their connections, the collective functioning of cells should be considered. The examination of the local synaptic circuitry in the brain and in the spinal cord reveals one widespread pattern, the reciprocal connections between excitatory and inhibitory neural populations: action potentials originating from the former excite the latter, triggering the inhibition of the former (Hoppensteadt & Izhikevich, 1996). By measuring how the discharge of one neuron can be temporally shifted by perturbations such as post excitatory/inhibitory potentials, scientists have been able to draw its phase-response curve (PRC), which consists in plotting the normalized spike time shift against the perturbation time (Achuthan et al., 2011). Knowing the PRC of coupled neurons allows theoretical predictions about their phase difference (Stiefel & Ermentrout, 2016). Those predictions, in addition to the PRC, should consider the intrinsic firing rate of the neurons and the coupling strength. In the case of two identical coupled neurons, if PRC is characterized by a positive slope at zero crossing, the phase difference between neurons is progressively reduced to zero.
This functional unit responsible for periodic activity is called a neural oscillator. Earlier, we mentioned the motor generation systems governing locomotion, which also depend on such functional organization (Marder & Calabrese, 1996). The phase convergence described between two coupled neurons can account for the fixed phase relation between CPG cycles (Marder et al., 2005). In the lamprey, the intersegmental time delay represents 1% of the cycle and is constant, irrespective of the swimming frequency (Cohen et al., 1992).

Oscillation is thus a key characteristic of neuronal activity, at the level of a single cell or for several neurons. The cellular organization of the cortex, which combines six layers of neurons of different types, is, however, much more complex, and loosens causal links between elementary neuronal processes and cortical oscillations. The observation of stochastic discharge in cortical cells (Softky & Koch, 1993; Stevens & Zador, 1998; Harsch & Robinson, 2000) appears to contradict the general shaping of brain activity by synchronous rhythms. Techniques such as EEG or EcoG, which aggregate signals from many cortical cells, nevertheless provide access to the measure of identifiable rhythms despite their limitations. EEG signals are recorded from the scalp and hence suffer from poor localization. Invasive measurements obtained by electrodes placed on the surface of the cortex are not spatially diverted by bone structure, dura and cerebrospinal fluid in EcoG. In animals, the insertion of deep microelectrodes in the brain can provide access to mesoscopic local field potential. Among the above-cited electrophysical techniques, rhythms recorded in LFP, spatially the most precise technique by summing signals with a radius of a few hundred \( \mu \text{m} \) (Xing et al., 2009), cannot be intuitively deduced from the stochastic spike trains of single
cells (Mureșan et al., 2008; Jarvis & Mitra, 2001). This discrepancy represents a challenge for “spike-to-spike synchrony”. The phase response characteristics of neurons modelled as periodic oscillators indeed entail synchronization with or without weak noise. In contrast, Brunel and Hakim (1999) modelled a population of inhibitory interneurons randomly connected by delayed synapses whose activity is dominated by a large background noise. Beyond a specific coupling strength, the amplification of perturbations in a network of strongly coupled neurons yields the emergence of high frequency collective oscillations. The “firing rate synchrony” framework (Brunel & Wang, 2003) proposes to explain network oscillations in terms of network architecture and intrinsic single neuron properties (Geisler et al., 2005; Moca et al., 2014). It can account for fast cortical oscillations which contrast with random and sparse spike trains of single cells (Fries et al., 2001b). Establishing the exact causal links between cell properties and the emergent dynamics of neuronal networks is a broad area of research, beyond the scope of the present review. The accumulation of knowledge on brain electrophysiological properties at multiple scales hints towards the conjunct role of the dynamical properties of elementary elements, and the diversity of neuronal connections (Wang, 2010). The development of techniques giving access to the simultaneous recording of multiple single cells (Buzsáki, 2004; Le Van Quyen & Bragin, 2007; Kobayashi et al., 2019; Jun et al., 2017; Mitz et al., 2017) will probably lead to a breakthrough in our understanding of the emergence of network dynamics (LOIC is this change ok?) from the properties of individual neurons.
Electrophysiological techniques nevertheless reveal oscillations in several bands (covering frequencies from 0.05 Hz to up to 500 Hz in the case of EcoG) during diverse behavioural states (Berger, 1929; Kahana et al., 1999). The frequency range measured in human EEG is generally divided into frequency bands: delta (0 – 4 Hz), theta (4 – 7.5 Hz), alpha (8 – 13 Hz), beta (13 – 30 Hz), gamma (30 – 100 Hz). These rhythms have been linked with behavioural correlates, despite the absence of a precise definition of their functional role. Alpha waves, recorded over sensorimotor cortex and prone to be phase-locked to stimuli (Sabate et al., 2011), are often associated with the facilitation of cortical information processing and working memory (Başar et al., 1997; Palva et al., 2005). They decrease or desynchronize with movement (Niedermeyer, 1997). Beta- and delta-waves, both prevalent in the motor cortex, have been shown to reflect the sensibility to relevant sensory cues in a motor context (Saleh et al., 2010): they are prominent during the preparation of the movement, cease at its onset (Jasper & Penfield, 1949), and increase if the movement is withheld (Pfurtscheller, 1981). The stimulus-induced response in the beta band depending on anticipation (van Ede et al., 2014), beta oscillations also seem to contribute to a top-down modulation of sensory processing (Caras & Sanes, 2017). Gamma band would convey the sensory processing of auditory information (Gurtubay et al., 2004; Steinschneider et al., 2008) and would participate in the selection of salient stimulus. Fast gamma oscillations of the motor cortex also appear to engage in the active part of the voluntary motor action (Cheyne et al., 2008), even if their variability with respect to the ongoing movement challenges their association with specific aspects of the motor command (Donoghue et al., 1998). The gamma-band power is larger during the first repetition of a series of movements (Muthukumaraswamy, 2010). Consequently, gamma activity seems to be uncorrelated to movement sensory feedback. A more global role of gamma rhythms has been proposed, binding the activity of remote populations of neurons (Womelsdorf et al., 2007; Fries, 2009).
which could be involved in the formation and maintenance of motor plans (Donner et al., 2009; Pesaran et al., 2002). This hypothesis would explain the alteration of gamma-band power when participants face competing motor responses (Gaetz et al., 2013). Low frequency theta oscillations would contribute to the behavioural adjustments following the detection of an error. For example, theta oscillations elicited by sound in the fronto-temporo-parietal network are modulated in phase and power when detecting deviations in a sequence of standard tones (Hsiao et al., 2009; Recasens et al., 2018).

Today, experimental data converge towards a crucial role of cortical oscillations in the binding features of sensory modalities. However, their role in performing an audio-triggered movement is debatable (van Wijk et al., 2012). During auditory-cued tapping, a significant coherence was observed in the alpha band within the extensive brain network contributing to AMS, i.e. in the motor and premotor cortices, posterior parietal cortex, auditory cortex, thalamus, and cerebellum (Pollok et al., 2005). Being broadly distributed across cortices, oscillations would recruit functional networks in the service of behavioural control (Narayanan et al., 2013). The long-distance spreading of oscillations between brain structures raises the question of the nature of the relation between network fluctuations and ongoing behavioural dimensions. Looking more closely at how afferences influence brain waves may help to better understand the possible embedding of synchronization-related information within neuronal oscillations. Electrophysiological measures have indeed unravelled the relation between sensory processes and neuronal activity. The excitability of a neuronal population, which synchronizes at a given frequency (Buzsáki & Draguhn, 2004), depends on the phase of the oscillation. The perceptual framing of vision could be partly explained by the phase of the cortical alpha rhythm. The probability of perceiving two visual stimuli as being simultaneous is indeed minimal when the stimulus onset is locked to the positive peak of the alpha cycle. It is maximal when the
stimulus onset is locked to its negative peak (Varela et al., 1981; Gho & Varela, 1988). Using the target auditory oddball paradigm, which consists in identifying task-relevant auditory tones intermixed with irrelevant ones, Haig and Gordon (1998) reported a dependency on the pre-stimulus alpha phase synchronicity of the participants' reaction time (figure 4A). An accumulation of evidence highlights the effect of ongoing central oscillations on sensory processing (Arieli et al., 1996; Fries et al., 2001a; Kisley & Gerstein, 1999; Kruglikov & Schiff, 2003). The importance of this relation is reinforced by the nested modulation of the phase of each frequency band by the lower one in the auditory cortex (Lakatos et al., 2005). This hierarchy opens the possibility for a stimulus to structure the whole temporal activity of the cortex.

In summary, the present description of brain rhythmicity from single cells to interconnected networks legitimates the use of dynamical modelling. However, despite the functional importance of rhythmic activities within the brain, we need to clarify the influence of external rhythms on brain activity in the context of AMS.

### 3.5.2 Evidence of Cortical Entrainment by Auditory Rhythms

The experimental validation of the coupling between auditory and endogenous rhythms proposed by Large and Kolen (1994) currently remains challenging. If there is no doubt that neural networks can spontaneously oscillate (Hoppensteadt & Izhikevich, 1996; Llinás, 2014; Yuste et al., 2005), establishing their entrainability in the presence of auditory stimulations raises methodological difficulties. Neuronal entrainment, defined as the “alignment of ongoing neuronal activity to the temporal structure of external rhythmic input stream” (Lakatos et al., 2019), supposes (i) the restriction of frequency relationships among endogenous and external oscillators to harmonics (e.g. 2:1, 3:1), subharmonics (e.g. 1:2, 1:3),
or integer ratios (e.g. 3:2, 4:3; Large, 2008), and (ii) phase synchronization in the presence of stimulations matching the frequency range of neural oscillations. Auditory cortical activity satisfying (i) and (ii) in the presence of rhythmic patterns has proven to be common to different types of stimulation such as music, amplitude modulated noise or structured tones (Doelling & Poeppel, 2015; Henry & Obleser, 2012; Arnal et al., 2015). An alternative explanation to the action of a neural oscillator is the so-called repetitive evoked response, according to which a rhythmic stimulus can give rise to a rhythmic input in the signal of the brain region processing it (Shah et al., 2004). We now briefly review the pros and cons, which have recently fed a lively debate on the possibility of directly assessing auditory entrainment in the brain.

By recording the EEG activity of participants listening to musical samples including beats, Nozaradan et al. (2011) found a sustained periodic EEG response tuned to the frequency of the beat. The authors also evidenced the dependency of EEG response on metre interpretation. Specifically, when participants were asked to judge whether samples corresponded to a binary or a ternary metre, additional periodic signals at the corresponding subharmonics of the original tempo emerged in the frequency spectrum. The adjustment of neuronal activities within a limited frequency range, evidenced by steady state-evoked potentials observed in the EEG spectrum at frequencies matching the rhythmic pattern envelope (Nozaradan et al., 2012), opens the possibility to track the external rhythms back in the cortex. Similarly, in a neuromagnetic study (Fujioka et al., 2015), beta-band oscillations were found to depend on whether beats were perceived as accented or not, regardless of the origin of the accentuation, physically embedded in the stimulus or imagined. Tal et al. (2017) directly assessed the correspondence between the spectrum of neural activities while
listening and the modulation spectrum of the stimuli. Neural responses in the auditory cortex were enhanced at beat frequency, and, when the beat was omitted, remained phase-locked to the times of the missing beat. In a magnetoencephalography experiment using a pitch distortion detection task, Doelling and Poeppel (2015) showed that (i) the cortical oscillations were phase locked over a range of musical tempo, at least for tempo above one note per second in non-musicians, and (ii) entrainment was correlated with performance. In contrast, musicians’ recordings showed entrainment across the whole range of the tested tempo. Neural plasticity, at the origin of the better performance of musicians compared to non-musicians, evidences the bidirectional nature of the relationship between behavioural and cortical entrainment. In short, the combination of high sensorimotor synchronization performance and high cortical entrainment among musicians supports the view that neural entrainment contributes to sensorimotor synchronization. A longitudinal study targeting the concomitant brain development and behavioural performances of musicians could clarify how neural entrainment and perception shape each other (Nozaradan et al., 2016a).

The results of these studies, in which the neural response mirrored the spectrum of the stimulus, were interpreted as evidences of cortical entrainment. This interpretation is in line with the neural resonance theory which assumes that beat perception involves the emergence of the internal periodic representation of a pulse through nonlinear coupling between neural oscillators (Large & Snyder, 2009; Large et al., 2015). One specific resonance frequency range would translate into the observed selective neuronal activity adjustment. However, as per the definition of physical entrainment, we carefully stated in the first section of the present review that the occurrence of peaks of activity in the neural response at frequencies matching those embodied in the stimulus, while being consistent with neural entrainment, does not represent one self-sufficient evidence. Novembre
and Iannetti (2018) raised the possibility that steady state responses (SSRs), noticed in EEG responses when listening to auditory sequences (Nozaradan et al., 2011; Nozaradan et al., 2012; Doelling & Poeppel, 2015), might be driven by auditory event-related potentials (ERPs). Mancini et al. (2018) characterized the effect of prolonged and repeated stimulation on the main components of somatosensory ERPs and reported small amplitude ERPs despite habituation over repetition. One can argue that the amplitude of the ERPs measured by Nozaradan et al. (2011) may not have appeared to be measurable beyond the first stimulus of the sequence. But the contribution of ERPs, particularly the biphasic negative–positive (N–P) vertex wave and the contingent negative variation (CNV), cannot be fully discarded (Novembre & Iannetti, 2018). These ERPs are indeed both modulated by non-periodic attention and expectation (Nobre & van Ede, 2018). CNV in particular has been associated with auditory motor synchronization timing error correction (Jang et al., 2016).

The identification of prediction mechanisms is central for the understanding of neural processing which prevails during AMS. Breska and Deouell (2017) proposed to disentangle the predictions based on rhythm-dependent or rhythm-independent mechanisms by comparing temporal predictions based on rhythms to those based on memory. The ability of exogenous rhythms to entrain endogenous oscillations should depend on the periodicity of the stimulation: a periodic stream of beats is expected to be the most efficient input. Conversely, an alteration of the periodicity of the stimulation, obtained by jittering the inter-onset intervals, would be detrimental to the necessary period and phase adjustments of neural oscillations. Memory tracking of the time intervals, which can include knowledge of temporal contingency, would be less prone to be affected by the non-periodicity of the stimulation. In order to decipher the contribution of each type of temporal prediction, whether rhythm-based or memory-based, Breska and Deouell (2017) presented visual stimuli with either
periodic or jittered intervals to participants instructed to press a button on target appearance. The phase modulation of delta-band activity did not show dissimilarities between periodic and non-periodic conditions. This is the first outcome of this study: the limitation of EEG activity alignment as an indicator of entrainment despite the correlations with perceptual and behavioural performances (Snyder & Large, 2005; Stefanics et al., 2010; Besle et al., 2011; Henry & Obleser, 2012). A preparatory ERP analysis revealed similar waveforms across conditions, but differences were noticed when one expected stimulus was omitted. The long-negative component of the response, the CNV, known to be altered by the unexpected suppression of the imperative stimulus (Walter et al., 1964), was immediately resolved after the omission of one anticipated event for rhythm-based predictions, as if the stimulus was delivered. The vanishing of the negativity, despite the omission of the target, can be seen as the electrophysiological equivalent of the maintenance of cortical oscillatory activities after stimulus termination (Lakatos et al., 2013; Bartolo et al., 2014). When one self-sustained oscillator is coupled to a phase-resetting signal, upon the termination of this influential input, several oscillatory cycles are indeed expected. Breska and Deouell (2017) were, to our knowledge, among the first to report specific electrophysiological indicators of entrainment in the context of periodic rhythms. They were followed by Doelling et al. (2019) who attempted to specifically disentangle the two distinct mechanisms - neuronal entrainment and ERPs - in the presence of a structured auditory stream. They hypothesized that a neural oscillator would be a better prediction and would have more computation ability. They analysed the phase lag between neural response and acoustic input as a function of the stimulus rate. They predicted a larger phase lag as the stimulus rate increased for the evoked model, in contrast to a stable phase lag for the oscillator model. In agreement with the oscillator model, the MEG activity of the participants’ auditory cortex revealed constrained phase regime while listening to music of varying note rates (1 to 8 Hz).
3.5.3 Auditory Entrainment Spreads Through Various Brain Structures

The recent experimental evidence of auditory cortical entrainment in both sensory and motor areas and at a specific tempo, as well as the hypothesis of an encoding of the time of synchronization in neuronal oscillations, elegantly legitimizes the use of dynamical models for the understanding of the interaction of brain-auditory rhythms. However, neuronal oscillations do not only encode information. They can also spread throughout connected populations (Buzsáki & Draguhn, 2004), in such a way that the influence of behaviourally irrelevant stimuli is lowered and the impact of relevant ones is reinforced. The technique of the event-related potential (ERP), which consists in averaging the post-stimulus EEG over a large number of trials, has been extensively used to disentangle the relation between the underlying electrical activity of the brain and various cognitive functions. Using this method in the context of AMS, significant differences between accented and unaccented events related to metronome tones in perceived and imagined metre have been reported (Schaefer et al., 2011) and associated with top-down cerebral mechanisms of rhythm processing. This type of control has also been evidenced with magnetoencephalography. Whilst listening to identical click stimuli, Fujioka et al. (2010) contrasted the auditory evoked responses from hippocampus, basal ganglia, and auditory and association cortices related with "March" and "Waltz" metric conditions. In the former condition, participants tapped every second click, while they tapped every third click in the latter one. The neuromagnetic activity of several brain structures was tracked during the listening period following the tapping task. In basal ganglia, the larger response was found for March. Metric conditions elicited different hippocampus activation latency: 80 ms for the march and 250 ms for the waltz. The right temporal lobe activation also differed according to the metric contrast. In the auditory cortex, the modulation of evoked response occurred at latencies between 80 and 200 ms. Together, these results indicate that the hippocampal memory system and the temporal processing system in the basal ganglia modulate
auditory cortical activity in this latency range, even in the absence of movement, but according to the specificities of the prescribed task. They reveal that neural activities induced by the metre context of musical samples facilitate the auditory perception of temporally relevant stimulus for the completion of a related motor task (Iversen et al., 2009; Grube & Griffiths, 2009; Pablos Martin et al., 2007; Snyder & Large, 2005; Brochard et al., 2003).

Gain modulation of event-related potentials and top-down control of rhythm perception thus represent convincing clues for the influence of external rhythm on the binding of sensory and motor areas of the brain. However, the shaping process of afferences by sensory pathways, e.g. brainstem response, during auditory synchronized movements, and the concomitant influence of corticofugal projections on this response have recently become a research focus. In order to clarify the encoding of sensory inputs, Nozaradan et al. (2016b) conducted an interesting study using the frequency tagging approach to simultaneously record brainstem and cortical EEG activity (figure 3 C, the frequency tagging method consists in using a stimulus with specific frequencies that can be easily isolated in the neural response). Participants, while listening to periodic amplitude-modulated tones, were instructed not to move or to tap in synchrony on every second beat. (LOIC do you mean here that the participant must neither move nor tap???) The frequency spectrum of the stimulus was set to elicit frequency-locked responses in the brainstem and in the cortex. The harmonic chord with partials at 200, 400, and 600 Hz, and amplitude modulated at the frequency of 2.4 Hz, was likely to evoke a steady-state brainstem nuclei response at the frequency of the partial, and a steady-state cortical response modulated at 2.4 Hz. The authors compared two conditions: tapping at half the frequency of the amplitude modulation (1.2Hz) vs. passive listening. The increased cortical response at beat frequency during sensorimotor synchronization, compared to passive listening, was expected and measured, in accordance with the previous results of the same
The maximal activity was measured at 2.4 Hz despite the movement frequency set at 1.2 Hz, which can be interpreted as a specificity of the AMS task. However, differentiating motor from auditory response remains subject to caution in this set of data. First, the existence of a peak of EMG activity at 2.4 Hz, in addition to the expected peak at 1.2 Hz, could explain EEG activity at 1.2 Hz and harmonics. The second potential source of motor overlap confound resides in the difficulty to separate sources in scalp-recorded EEG. The overlap of cortical motor and auditory activities in EEG can be addressed by alternative recording techniques. Ten Oever et al. (2017) used electrocorticography (EcoG) from one grid of electrodes implanted on the lateral frontal, temporal, and parietal lobes of one participant, ensuring a more robust mapping of the signal. Their protocol entailed detecting beeps embedded in white noise, with their intensity increasing monotonically over the trial presented at 1.5 Hz. Participants were instructed to react the moment they perceived the sounds by pressing a button. The authors compared inter-trial coherence, which was expected to be high at the frequency of the stimulus presentation compared to a random sequence. Phase alignment indeed occurred before participants reported the stimulus as audible. On the contrary, auditory evoked responses were detectable after the beeps became audible. The same results were noticed in an experiment based on the same protocol with MEG recordings in 16 participants.

As a solution for the limitation of low spatial resolution of EEG, the EEG-fMRI fusion method has been used by Li et al. (2019) to distinguish the neuronal populations that entrain to beat or metre. The spectra of auditory steady-state evoked potentials provided an indication of the entrainment ability of brain structures to beat (2.4 Hz) or metre (1.2 Hz). The thalamus, basal ganglia and SMA all exhibit peaks of activity both at beat and metre frequencies. The SMA was more prone to frequency locking at beat frequency than at metre frequency. Finally, with different measuring methods, Ten Oever et al. (2017) and Li et al. (2019) legitimate the conclusions of Nozaradan et al. (2016b).
The novelty of the Nozaradan et al. (2016b) study lies in the brainstem response. Auditory brainstem response is usually characterized by a transient response to a click, and a frequency-following response (FFR) to sustained periodic stimulations (Skoe & Kraus, 2010). The latter would reflect the basic processing of auditory information (Skoe & Kraus, 2010). Nozaradan et al. (2016b) reported a concomitant boosting of the brainstem responses, and the specific amplification of lower-pitched tone of the chord with no correlation with EMG tapping data. The observed gain modulation of the auditory response of the brainstem supports the view of an online modification of sound encoding during movement synchronization. The enhancement of the brainstem encoding during AMS could contribute to amplifying behaviourally relevant sound features. The efferent corticofugal projections (Musacchia et al., 2007; Chandrasekaran et al., 2009; Lehmann & Schönwiesner, 2014) could mediate the shaping of brainstem response by the cortex. Descending projections from the sensory cortex to the inferior colliculus appear to be crucial for the calibration of the auditory system as a function of experience (Bajo et al., 2010). Response properties of colliculus neurons are modified by the alteration of the cortico-collicular projection (Bajo & King, 2012). In bats, Ma and Suga (2001) set off shifts in the frequency-response curves of collicular and cortical neurons with electrical stimulation of the auditory cortex. The electrical stimulation of the somatosensory cortex extended the recovery period of auditory cortical neurons. Corticofugal projections can reorganize the auditory response of collicular and cortical neurons to specific features (frequency in the study cited above) of sounds (see also Yan & Zhang, 2005). The contribution of non-auditory cortex shows that the reorganization could put the emphasis on behaviourally meaningful stimulus. Polley et al. (2006) demonstrated the possibility of eliciting a parameter specific plasticity in the auditory cortex by training rats to selectively attend to either the frequency or the intensity of auditory stimuli. The role of the corticofugal descending pathway in
driving experience-dependent plasticity, and more specifically musical processing ability, has also been evaluated in humans. In a MEG study, Coffey et al. (2016) revealed the cortical origin of FFR. They also reported a positive correlation between the strength of FFR and cumulative hours of musical training as well as between the strength of FFR and pitch discrimination thresholds. Over the past decade, researchers have gathered evidence of the joint contribution of cortical and subcortical components to FFR. The upcoming challenge lies in the disentanglement of their respective contribution to sound processing.

3.5.4 Oscillatory Activities of Interconnected Brain Structures Support Interdependence of Beat Tracking

The perception of the auditory events to be temporally matched with movements appears as an intuitive prerequisite of AMS. We first opened this section with the neural substrates of beat tracking ability. We then introduced the experimental evidence of neural entrainment spreading. The detailed top-down supervision of the saliency of motor-relevant sensory information reveals the tight coupling between action and perception during AMS. Particularly, the fact that the voluntary organization of musical beats - in the form of metre imagery - can be tracked at the cortical level suggests its influence on perceptual ability. Selective enhancement of motor-salient frequencies by the cortex was supported by EEG recordings of cortical and subcortical activities (Nozaranad et al., 2018b): the former exhibited greater dissociation from the frequencies of the input than the latter by putting an emphasis on metre-related frequencies. The selective enhancement of metre-related frequencies does not fully rely on subcortical auditory properties, but is critically shaped at the cortical level, possibly through functional connections between the auditory cortex and other movement-related brain structures.
Action could therefore contribute to sharpening perception. Active sensing has been proposed in cognitive sciences and robotics, among other disciplines, to describe sensors which have not for sole input the energy from the sensed source: active sensing would also adequately describe most sensory processing (Schroeder et al., 2010). Being active involves perceptual exploration made possible, for example, by motor routines. Active sensing is a mechanism of the selective amplification of task-relevant auditory input. An example is the work by Morillon et al. (2014) who hypothesized that rhythmical movements would contribute to temporally sharpen the selection of auditory information. Participants tracking reference tones interleaved with distractors were found to perform better in target extraction when they were required to finger tap rather than during listening only (figure 4B). In the same vein, participants, when asked to listen to rhythms perceivable either as a binary metre or a ternary metre, exhibited neural activities larger at the frequency corresponding to the ternary metre if trained before listening (Chemin et al., 2014). These examples legitimate the vision of active sensing as a generalization of sensory-motor binding elicited by AMS. Sensory-motor experience thus appears to tune the temporal precision of predictions, and consequently shape perception with top-down processes. Descending processes have been evidenced in non-human primates with electrophysiological recordings. In non-human primates exposed to streams of pure tones and noise bursts interleaved with random acoustic patterns, Barczak et al. (2018) hypothesized that the detection of repeating patterns would require the entrainment of internal rhythms to demarcate the repeated rhythmical structures. They reported delta-phase locking, and the modulation of neuronal firings occurred earlier in the thalamus than in the primary auditory cortex. The recording of the sequence of endogenous rhythm alteration across central structures reveals the top-down processes during a task which necessitates the grouping of complex auditory patterns. This result also contradicts the definition of entrainment as an acoustically driven reflexive brain response (Novembre &
Iannetti, 2018). It indicates that the ability to identify rhythmic stimulus sequences
without defined boundaries, solely characterized by their statistical regularities,
*preludes* entrainment.

### 4 Using Auditory Entrainment and AMS to Enhance Locomotor Performance

The previous section (3) presented evidence for the functional role of movement-based neural loops articulating ganglia-thalamo-cortical and cerebellar-thalamo-cortical networks, together with the dynamical properties of the locomotor system such as the vectors of synchronization of endogenous (i.e. biological) and exogenous (e.g. auditory) rhythms. Brain activations during AMS reveal that “motor-relevant” auditory cues functionally connect neural regions dealing with movement execution. This particular sensorimotor coupling legitimizes the use of auditory cues in order to influence the timing of motor actions and the success of its achievement. Appropriate spatio-temporal coordination of muscular activations is indeed necessary for movement production, and because temporal consistency often decides the success or failure of motor actions, it is a primary factor of performance. If the relation between locomotor performance and auditory stimulation is behaviourally substantiated, for example in the context of gait rehabilitation in PD patients, investigating the specific underpinning mechanism still only represents a partially explored research field. In this last section (4.1), we propose to investigate how the two types of entrainment, physical and neural, can be coupled to increase locomotor performance in walking patients and healthy runners.

#### 4.1 Biomechanical Entrainment as a Factor of Performance

From the beginning of the nineties, bipedal robots - initiated by McGeer (1990), designer of passive dynamic walkers - have been built to walk downhill with
humanlike gaits, without actuation or perceptual control of any kind. Despite the
absence of control of every joint angle, which makes complex computations
unnecessary, passive walkers show a robust periodic limb trajectory, which
restores itself if perturbed slightly. They are also categorized as limit-cycle
walkers. The interaction between the inertial and gravitational mechanics of the
two inverted pendulums which constitute the walker ensures a rather stable
coordination (Collins et al., 2001). To enable ground-level walking, Collins et al.
(2005) added small active power sources, in the form of ankle or hip actuation, as
a substitute for gravity. In humans, the spring-like behaviour of the leg is not fully
energetically conservative. While energy is stored in muscles and tendons during
the impact absorption of the ground contact phase, and is reutilized during the
supporting phase, some energy is also dissipated by muscle fibres when they
lengthen. The maintenance of a limit-cycle behaviour is achieved at the energetic
cost of muscular activation. The episodic supervisory control we introduced earlier
thus appears as an efficient component when modelling locomotion control. The
relative hierarchical organization of the CNS is combined with (i) the autonomy of
peripheral networks in pattern generation, and (ii) the use of peripheral feedback
to adapt movement kinematics. This fundamental feature of biological systems
can find some resonance in this engineering model.

The observation of a limit cycle behaviour in the locomotor system, without
supraspinal control (Gurfinkel et al., 1998; Gerasimenko et al., 2010), makes the
passive dynamic model with reduced control a plausible model of CPG-driven
walking (Miyakoshi et al., 1998; Aoi & Tsuchiya, 2007). Bipedal stepping stability
of robots also benefits from peripheral sensors which can be used as state
detectors, and provides input for the actuation (Miyakoshi et al., 1998). Similarly,
in humans, beyond the evocation of fictive locomotion with peripheral
stimulations, load afferences are required to sustain the emergence of a
functional locomotor pattern (Dietz & Harkema, 2004; Van Wezel et al., 1997).
Finally, with behavioural measures showing gait entrainment towards a narrow range of periodic perturbations, the relative contribution of limit cycle oscillators in walking has been evidenced (Ahn & Hogan, 2012b). However, in a subsequent article, the same group proposed an alternative simple state-determined model, based on afferent feedback processes, independent of supraspinal control or oscillatory networks (Ahn & Hogan, 2012a). The associated phase locking of the perturbation within the locomotor cycle ensured the assistance of the propulsion phase. The primary benefit of limit cycle walking is indeed energy efficiency. Both artificial and biological systems are energetically optimal at their mechanical resonant frequency. Since oscillatory systems respond at greater amplitude when the frequency of their oscillations matches their resonant frequency, the cost of transport can be optimized by walking at the cadence at which periodic movement can be sustained with minimal contribution of actuators (Ahlborn & Blake, 2002).

For a given speed, the walking cadence lies in an interval ensuring less energy spending (Holt et al., 1991; Minetti et al., 1995; Umberger & Martin, 2007; Zarrugh et al., 1974; Bertram & Ruina, 2001). However, it is only recently that the ability of human adults to continuously optimize the energetic cost of walking has been demonstrated. Selinger et al. (2015) used robotic exoskeletons during normal walking to manipulate the metabolic cost landscape across the potential combination of stride length, stride frequency, speed, and associated energetic cost. The exoskeleton shifted the participants’ optimal step frequency to higher or lower values than the one initially preferred by applying resistive torque. Participants were found to adapt their cadence to match the new optimal cadence despite the small amount of energetic gain associated with the updated cadence (< 5%). Energy efficiency thus does not appear only as an emergent property of the movement but can also be considered as an influencing input variable.
These findings provide the essential clues into understanding the benefits of auditory cueing during physical activities, for instance, when running in synchronization with music. If recommendations have to be formulated, the tempo should be chosen according to the properties of the limit cycle oscillator to maximize entrainment. This choice has at least two consequences. First, in the case of healthy walking or running, the normal and therefore most efficient cadence naturally represents the tempo value to match. We have indicated above that people are able to select their natural cadence, but obviously not in situations where other factors such as fatigue or pathology alter this ability. In the last section of the present review, we show how pathological gait (4.2) and running performance (4.3) can benefit from appropriate stimulations to counteract these side factors. The second consequence, somewhat linked to the first, is an “out of the lab” consideration. The practice of locomotor activities supposes different environments, durations, speeds and equipment according to the individuals or the training sessions: all these factors being prone to influence the natural cadence, the adaptability of the stimulation represents both a necessity and a technological challenge. The adaptability of the stimulation also leads us to consider its personalization as a promising way to address individual (non-)responsiveness.

4.2 Auditory Cueing During Walking of PD Patients Positively Influences Cortical Networks

The PD patients’ ability to initiate movement is impaired, due to reduced functionality of the ganglia-thalamo-cortical network accompanying the progressive loss of dopaminergic nigro-striatal neurons (Galvan & Wichmann, 2008). This often gives rise to a freezing of gait, which is the transient difficulty in gait initiation or progression when approaching an obstacle or a turn (Giladi et al., 1992). The automatization of learned sequences is also altered (Wu & Hallett, 2005). It was early hypothesized that the explanation for such deficits lie in the
impairment of internal cues generation to trigger the subsequent sub-movement in a sequence. During the execution of a visually cued sequence, preparation and execution movement times are known to co-vary with the reduction of the external cues in PD patients (Georgiou et al., 1994). Their difficulties in movement sequence completion, which can be reversed by external information, provide evidence of the role of basal ganglia in the release of relevant internal cues for the transition of movement stages.

Abnormal sustained activity in the beta band across the cortex and basal ganglia has been associated with the loss of dopamine neurons (Cassidy et al., 2002; Levy et al., 2002; Sharott et al., 2018; Kühn et al., 2005). Despite the mechanistic link between impairment in movement initiation and the alteration of the communication between brain structures (Engel & Fries, 2010), few studies have correlated objective measures of motor impairment and the amplitude of local field potential in the beta band (Brown, 2007). PD patients, who were implanted with neurostimulators in the STN, and who performed repetitive index finger to thumb taps, exhibited modulation of local field potential in the beta band (Androulidakis et al., 2008). As the movement became more bradykinetic over trial time, the amplitude modulation of the STN activity decreased. Steiner et al. (2017) also recorded local field potentials simultaneously with motor performance several months after the implantation of deep brain stimulation electrodes. They assessed bradykinesia during a self-paced continuous alternating pronation and supination task which was completed as quickly as possible and with the largest amplitude possible. Oscillatory activity was suppressed in the beta band during repeated movements. The diminution of the beta band suppression was accompanied by the decline of the frequency and amplitude of movements. Fischer et al. (2018) recently assessed the role of beta activity in stepping by measuring the neural dynamics in the STN of PD patients implanted with DBS electrodes. They reported the suppression of oscillations in the beta band, time
locked to the gait cycle, following ipsilateral heel strike and contralateral foot raising. This modulation was increased when auditory cues assisted patients in timing their steps. A simultaneous reduction of step timing variability was noticed. Interestingly, in the presence of auditory cues, beta decreased before the ipsilateral heel strike became faster. The possible relation between the degree of step-related modulation of oscillations, influenced by auditory cueing, and movement performance in PD patients highlights the role of beta oscillations in the healthy brain. The beta reactivity of the STN, which is related to the saliency of cues with respect to future motor actions (Williams et al., 2003; Oswal et al., 2012), supports the role of basal ganglia in weighing the behavioural relevance of environmental information.

Providing an external clock would be prone to compensating for self-initiated and self-paced movement timing difficulties due to basal ganglia dysfunction. This would explain the missing internal cue at the origin of PD patients’ difficulty to initiate and maintain cyclic movements such as walking (Rochester et al., 2010). If self-initiated and predictably cued movements indeed elicit the same levels of SMA activation in healthy participants (Jenkins et al., 2000), multiple arguments call for the dysfunction of the supplementary motor cortex in PD (Nachev et al., 2008) such as the decreased activity of pre-SMA and SMA observed in PD patients (Playford et al., 1992; Grafton, 2004). Caligiore et al. (2017) recently hypothesized a causal relation between alterations of the cerebellum to SMA circuit, the direct link between the SMA and the subthalamic nucleus (Nambu et al., 1996), and the impaired action sequencing of PD patients. The altered activity of SMA reported in PD patients would be associated with the decrease of the anticipatory activation of the STN by the hyper-direct pathway, a mechanism possibly prone to alter subsequent movement preparation in a sequence. Considering the overlapping of brain structures activated during movements relying either on internal or external cues, this model can be of interest in the context of AMS.
The deficiency of the basal ganglia–cortical circuitry necessitates some compensatory mechanisms that tend to maintain the temporal structure of actions to achieve coordinated movements such as those responsible for forward progression. Two explanatory hypotheses logically apply to patients' improved gait performance under the influence of auditory stimulation: (i) the residual activation of the basal ganglia by auditory cueing, and (ii) the compensatory mechanisms which originate from SMA (Eckert et al., 2006; Lewis et al., 2007) and/or from the cerebellum (Lewis et al., 2007). In the presence of an altered striato-thalamo-cortical loop which characterizes PD, compensatory timing mechanisms based on cerebello-thalamo-cortical circuits have been claimed (Dalla Bella et al., 2015; Lewis et al., 2007; Galvan & Wichmann, 2008; Kotz & Schwartz, 2011; Nombela et al., 2013). The overactivation of the cerebellum (Rascol et al., 1997) and motor cortex (Playford et al., 1992) measured with neuroimaging in PD patients during sensory-motor coordination is accompanied by functional changes, notably enhanced connectivity within the cerebello-thalamo-cortical loop (Palmer et al., 2010). Moreover, the dynamics of this connectivity appears to depend on the progress of the disease (Sen et al., 2010; Wu et al., 2011).

However, the balance between striato-thalamo-cortical and cerebello-thalamo-cortical circuits in PD remains an open research question. Miller et al. (2013) proposed to directly assess whether the striatal dopaminergic denervation, measured by reduced dopamine binding potential, would be associated with sensorimotor synchronization impairment. Patients finger tapped in synchrony with an isochronous tone sequence (500, 1000, 15000 ms inter-onset interval). Regression analyses conducted on the relation between synchronization accuracy, or variability, and denervation, did not appear very strong. Their conclusions, less straightforward than expected, finally call for a subgroup definition of patients: a
cluster analysis, yielding 3 subgroups of patients, substantiated the parallel degradation of synchronization accuracy, for the 500 ms target time interval, and the pattern of dopaminergic denervation. Patients with an asymmetric pattern of denervation between left and right hemispheres showed degraded timing ability for the corresponding affected hand, and patients with symmetric denervation showed similar degradation for both hands. This finding supports the contribution of striatal denervation to sensorimotor synchronization impairment. However, the subgroup of patients who had the largest proportion of striatal denervation exhibited the best synchronization performance compared to the other subgroups. The differences between subgroups suggest the existence of potential compensatory mechanisms: patients with advanced striatal denervation could have progressively switched to an alternative timing strategy over the course of the disease. Miller et al. (2013) related these results with the greater cerebellum activation. During a similar task, Jahanshahi et al. (2010) noticed finger tapping in synchrony with an isochronous tone at 1 Hz. Jahanshahi et al. (2010) found that cerebellar hyperactivation in patients was negatively correlated with the activation of the contralateral putamen during auditory-paced movements (Yu et al., 2007) and progresses with the disease (Sen et al., 2010).

The existence of compensatory mechanisms raises the question of their evocation by AMS. In addition to the activation cerebellar networks, bypassing or facilitating the striato-thalamo-cortical loop has been proposed as explanatory mechanisms of gait improvement triggered by AMS. However, very few studies have attempted to identify the most prominent compensatory mechanism. Majsak et al. (1998) noticed bradykinesia when PD patients were performing maximal speed arm reaches toward a stationary ball, but not when they reached for a moving ball. This difference has been interpreted as a demonstration of the impairment of PD patients to move their limbs on a self-initiation basis compared to external stimulus-driven movements (see also Freeman et al., 1993a and Azulay et al.,
1999), in accordance with the hypothesis formulated by Goldberg (1985): the
cerebellum, the parietal lobe, and the lateral premotor cortex represent the
primary control loop for externally guided movements, whereas the basal ganglia
and the supplementary motor area are more predominant in self-generated
actions. Schenk et al. (2003) used a similar experimental paradigm encompassing
reach-to-grasp movement in patients exposed to controlled deep brain
stimulation (DBS). Patients implanted with a neurostimulator in the internal
globus pallidum and the subthalamic nucleus show good improvement in tremor
and bradykinesia (defined by Queen Square Brain Bank criteria as the “slowness
of initiation of voluntary movement with progressive reduction in speed and
amplitude of repetitive action” (Gibb & Lees, 1988)), and l-Dopa induced
dyskinesias (Krack et al., 1998). In the same vein, Schenk et al. (2003) compared
an externally timed condition (target moves away) and an internally timed
condition (target is stationary) and found that auditory cueing used to initiate
movement contrasted with the pursuit of a mobile target. This provided a cue for
both movement initiation and execution. The second type of cue was more prone
than the first type to influence movement execution parameters, such as the
velocity profile. This conclusion calls for the use of more complex auditory
stimulations than just a simple metronome. As expected, the effects of DBS on
movement parameters were more pronounced in the internally timed condition
than in the externally timed condition. The structures activated during visually
guided movements include the superior parietal cortex, the premotor cortex, the
thalamus, and cerebellar lobule VI. Conversely, activation of the basal ganglia, the
supplementary motor area, the cingulate motor cortex, the inferior parietal, the
frontal operculum, and the cerebellar lobule IV-V/dentate nucleus dominates
when movements are internally generated (Debaere et al., 2003). Elaborating on
these results, one can predict a more restricted influence of basal ganglia on
auditory-cued movements than on self-initiated movements. This prediction was
tested (Nowak et al., 2006) in PD patients with chronic stimulation of the
subthalamic nucleus, during a grip–lift task performed either at self-selected speed or in response to an auditory cueing signal. When DBS was off, auditory cues improved akinesia, restored grasp formation and compared to healthy controls. DBS was associated with excessive grip forces, irrespective of the type of movement initiation. In other words, auditory cueing contributed to movement initiation in patients, and DBS facilitated movement execution when lacking fine control. The premotor cortex, which is tightly connected with the cerebellum (Hoover & Strick, 1999), is responsible for the integration of environmental-related afferences, i.e. sensory mapping (Wise, 1985), and indirectly receives projections from the auditory cortices via the parietal areas (Pandya et al., 1969).

Auditory motor coupling can consequently emerge from this alternative network (Weeks et al., 2001) as a substitute pathway to the damaged basal ganglia – supplementary motor area network. The restoration of fine motor control when AMS is applied to PD patients is also ascertained by muscular activation in the arms and legs. Their analysis revealed an increase in their symmetry and a decrease in their timing variability (Miller et al., 1996; Fernandez del Olmo & Cudeiro, 2003).

There is now convincing experimental evidence supporting the activation of compensatory pathways by AMS. However, the persistency of the contribution of impaired pathways cannot be wiped out, particularly under medication. The engagement of reward mechanisms in healthy participants completing physical exercise (see box 3, section 7.3) could also apply to PD patients, and help to maintain dopamine release. If phasic activity of SN DA neurons is considered as a reward prediction error signal (Schultz, 1998; Schultz, 2007; Schultz, 2016), Jin and Costa (2010) also showed its role in the signal initiation and termination of action sequences. The dual role of DA neuron supports the hypothesis of an effect of pleasurable auditory stimuli on PD motor deficits, mediated by the reward system.

One consequence of dopamine depletion therapy, “compulsive singing” (Bonvin
et al., 2007), suggests the reciprocal influence of musical reward and dopamine release, which is the consequence of the involvement of basal ganglia in both rhythm and emotion processing (Trost et al., 2014). This behaviour is congruent with the beat-based timing impairment of patients, and its improvement under dopaminergic medication (Cameron et al., 2016).

The supervision of peripheral activities by higher levels of the CNS enables the contribution of alternative sensory inputs to the control of gait. This assumption opens the possibility of the manipulation of gait parameters by auditory cueing. In a seminal work, Thaut et al. (1996) demonstrated the benefits of auditory cueing on the improvement of parkinsonian patients’ gait. Patients used auditory stimulations as a pacemaker during a 3-week training programme. Temporally predictable auditory cues had a post-training effect on gait by increasing gait velocity by 25%, stride length by 12%, and step cadence by 10% in comparison to self-paced patients who improved their velocity by only 7%. Benefits can generalize to non-cued gait after an extensive period of training with auditory cues (de Bruin et al., 2010; Frazzitta et al., 2009; Ford et al., 2010). However, Arias and Cudeiro (2008) argued that the grade of the patients’ disability must be considered in order to observe gait facilitation. Other factors should also be considered because the success of this type of training varies significantly between individuals.

Beyond the patients’ grade, Leow et al. (2014) showed that beat perception ability influences footstep synchronization with music. In recent studies, we proposed to identify the specific rhythmical abilities of patients, which may help to shed light on the functional mechanisms underlying the effects of AMS (Dalla Bella et al., 2017). The effect on gait parameters observed during auditory cueing in patients with PD was linked to their rhythmical abilities (figure 4C), especially their history of musical training and associated rhythmic skills (Cochen De Cock et al., 2018). In
another group of patients submitted to AMS for four weeks, where synchronization abilities were assessed before and after the training with auditory paced hand tapping and walking to auditory cues, individual responses to the training were either positive, neutral or negative, as measured by the distance covered during a six-minute walk test. A positive response to AMS was predicted by the synchronization performance in hand tapping and gait tasks. More severe gait impairment, low synchronization variability, and a prompt response to a stimulation change foster a positive response to AMS training. These findings experimentally demonstrate the role of general-purpose perceptual timing and sensorimotor abilities for gait rehabilitation in PD patients with auditory cueing. They can potentially be generalized as evidence of the functional link between such mechanisms and training strategies relying on neuronal plasticity. This link is supported by the recently unveiled link between cortical entrainment and synchronization accuracy. In the third section of the present review (3.5.1), we have indeed presented data showing the entrainment of EEG cortical responses. In the Nozaran et al. (2016a) study, movement synchronization accuracy with a rhythmic beat could be explained by the amplitude of neural activity selectively locked with the beat period. Moreover, the strength of the endogenous neural entrainment was correlated with better temporal prediction abilities. These findings plead for the facilitation of AMS by neural entrainment.

4.3 Running Performance in the Presence of External Auditory Pacing

The use of synchronized music has also been proposed to runners, so that they hit the ground in time with the beats of a metronome or music. The example of the Ethiopian athlete, Haile Gebreselassie, who famously broke the indoor 2000 metre world record in February 1998 while the Scatman song was being played in the arena, is often cited to support the positive influence of auditory-motor synchronization on running performance (Simpson & Karageorghis, 2006; Bood et
al., 2013). As appealing as this interpretation may be, it should be considered with caution for at least two reasons. First, beat-step synchronization cannot be claimed in the absence of appropriate measures at high sampling rate. Second, the gap between Haile Gebreselassie’s cadence (about 180 - 200 steps per minute) and the Scatman tempo (136 beats per minute) would only have made synchronization possible if he had adopted a syncopated rhythm (off the beat). Considering the ratio of the musical tempo to the runner’s cadence - ideally 0.66 in the case of syncopated rhythm - this type of synchronization would have remained stable if Haile Gebreselassie’s cadence was close to 204 steps per minute. Higher cadence being associated with better performance, the music could have contributed to maintain the runner within the basin of highly efficient cadence. However, distinguishing the effects of auditory-motor-synchronization from those related to motivation is necessary because both coexist when motivational music is used. Haile Gebreselassie, who judged the “rhythm [of Scatman] perfect for running”, could have benefited from motivational and/or attention diversion effects (see box 3, section 7.3).

The specific use of musical rhythm to maintain an optimal cadence nearing exhaustion has recently been clarified. Runners who synchronize their steps with the beat of a metronome see their time of exhaustion postponed during a high-intensity exercise compared to the same session conducted without stimulations (Bood et al., 2013). The associated decrease of the stride variability encouraged the authors to consider the consistency of the cadence as the explanatory factor of performance. There is indeed other evidence of the association between performance and stride rate variability, such as the decrease of the latter elicited by endurance training (Slawinski et al., 2001), or its increase with fatigue (Williams et al., 1991; Dutto & Smith, 2002). Similarly to Bood et al., Simpson and Karageorghis (2006) showed that synchronous oudeterous music (i.e. considered as neutral in terms of motivational qualities) can improve the performance of non-
elite runners in a 400 metre sprint. The musical beats could have acted as an ideal

cadence maker, the stride frequency of recreational runners dropping
dramatically well before the end of a sprint (Girard et al., 2016). In the experiment
of Terry et al. (2012), the time to exhaustion of elite triathletes was also longer
when they were listening to music considered as neutral in terms of motivation,
compared to silence. Neutral music also decreased the perceived effort. The
authors reported an improvement of running economy, which should be
considered with care, the decrease of oxygen consumption [1.3-2.7%] remaining
mostly within the range of measurement errors of the metabolic cart (2.5%,
Medbø et al., 2012).

Morin et al. (2007) demonstrated the dependency of contact time and associated
cadence on leg stiffness, as predicted by modelling the leg as a spring-mass
system. The stiffness of the leg-surface combination is supposed to remain
constant (Ferris et al., 1999; Ferris et al., 1998) in order to preserve the dynamics
of the locomotor cycle (Farley & Gonzalez, 1996). However, Dutto and Smith
(2002) observed changes in the stiffness characteristics of the leg during a run
leading to fatigue. We can reasonably conceive that auditory beats delivered at
the appropriate tempo could counteract the deleterious effects of fatigue on
kinematics. This assumption could motivate further investigations. This approach
would be an innovative strategy for the prevention of overuse injuries (Chumanov
et al., 2012; Heiderscheit et al., 2011; Souza & Powers, 2009; Brindle et al., 2003;
Fredericson et al., 2000; Geraci & Brown, 2005; Edwards et al., 2009; Wellenkotter
et al., 2014; Williams et al., 2004; Williams et al., 2001; Granata et al., 2002). Stride
variability is rarely characterized in the above-mentioned study despite its
predictive value in the context of injuries. As revealed by detrended fluctuation
analyses, step-to-step timing exhibits long-range correlations (Jordan et al., 2006),
which, however, decrease over time during a prolonged run (Meardon et al.,
2011). Interestingly, participants who suffered from previous injuries exhibited
lower long-range correlation than non-injured participants. The predictability of
the stride interval decreasing with fatigue or injuries, together with the real-time
measure of runners’ stride time, could open interventional possibility through
rhythmical auditory stimulation in an attempt to restore the kinematic complexity.

The contribution of auditory stimulation could be extended beyond the locomotor
system, with the synchronization of two underlying biological rhythms, for
instance locomotion and respiration. Locomotor-respiratory coupling (LRC) has
been initially reported in quadruped animals as a mechanism transferring the
mechanical deformations associated with locomotion onto the rib cage via the up
and down movements of the visceral mass (Bramble & Carrier, 1983): the visceral
piston mechanism results on locomotor dependent loading and unloading of the
diaphragm, and alters the pulmonary volume. Animal studies revealed that the
coupling is also neurological. Connections between locomotor and respiratory
centres were evidenced by the increase of respiratory rhythm evoked by fictive
locomotion (Waldrop et al., 1986) and the stimulation of peripheral muscular
mechanoreceptors (Iscoe & Polosa, 1976). In quadruped mammals, the ratio of
locomotor and respiratory frequencies is 1:1 (number of strides:number of
respiratory cycles). However, bipedalism has significantly reduced the mechanical
linking between both systems, the movement of the spine being nearly non-
existent. The direct consequences are the multiplicity of possible ratios of
frequencies (usually 3:2, 2:1, 5:2, 3:1) and potential independence of the phases
of both systems (e.g. Bardy et al., 2015; Hoffmann & Bardy, 2015). The inspiratory
and expiratory volumes ascribed to the visceral piston in quadrupeds reaches 20% of
the tidal volume (Lee & Banzett, 1997), whereas in humans, estimates vary
between 2% (Banzett et al., 1992) and 10-13% (Daley et al., 2013). How critical
LRC is for energy efficiency is a matter of debate for additional reasons. First,
evaluating the coupling of two systems, whose eigenfrequencies are different,
remains challenging. Methods are debatable (Zelic et al., 2017) and the variables
which are used to appraise coupling are not obviously the ones most representative of the coupling. Most authors used the impact of the foot as the locomotor event of reference (McDermott et al., 2003), but the loading of the visceral mass, affecting lung volume, is initiated with a large delay with respect to foot landing. Second, the real contribution of LRC to pulmonary air exchange is debatable. In a recent study about LRC, Daley et al. (2013) postulated that the synchronization of the peak acceleration of the centre of mass, with the inspiratory and expiratory transitions, would be the preferred pattern used by runners to contribute significantly to the breathing mechanics. Further studies should challenge this specific hypothesis in order to provide a better understanding of the nature of the LRC in bipedal locomotion. More generally, how music as an exogenous AMS flows into multiple and mechanically coupled biological systems exhibiting various natural frequency ratios, such as the locomotor system and the respiratory systems, remains wide open for further investigation.

5 Open Research Questions and Conclusions

Explaining how physics and neuroscience contribute to the emergence of stable yet adaptable movements is one aim of behavioural science. The present review is an attempt to exploit this complementarity between the two approaches towards our understanding of AMS. Our first step consisted in identifying a relevant theoretical framework. If empirical studies are valuable to identify structures and networks in the brain involved in perception and action, modelling neural activity with dynamical systems provides a universal mechanism for the transformation of auditory information into efficient motor commands through oscillator-based coupling and entrainment principles. We demonstrated the appropriation of time-related information in the brain, and the relative segregation in different structures of the afferences according to their rhythmical characteristics and their salience in an action perspective. The notion of motor
salience of a given rhythm is empirically illustrated by the dependence of rhythm perception on anthropometric parameters of the listener (Todd et al., 2007). The dynamical properties of the locomotor system, which partly relies on oscillatory spinal networks, putting into motion a chain of articulated segments characterized by their eigenfrequencies, have only recently been clarified. Dynamical entrainment appears to be possible at both central and biomechanical levels, validating the limit cycle oscillator model as a relevant model of cyclical locomotor activities. In the presence of rhythmical auditory stimulations, the reciprocal influence of neuronal and behavioural entrainment substantiates the relevance of the dynamical system approach as an explanatory framework for SMS.

In this review, we have presented the multi-level distribution of auditory entrainment, from neural to biomechanical structures, or in other words, how auditory stimulations find their way towards the rhythmic locomotor system. The cited studies show the complexity of the functional attributes of the auditory stimulations. There is no doubt that we have learned an immense amount about neural responses over the past two decades. However, it is evident that much remains to be established about the precise mechanisms sustaining locomotor performances. Viewing sensory and motor functions in isolation probably does not help to discover the whole potential of auditory stimulations on motor performance, considering their tight coupling. Uncovering the whole chain of coupling factors between perception and action presents a surmountable challenge, and, inspired by the state of the art synthesized above, some recommendations for future research can now be made.

As legitimate as the analysis of living systems is, through the prism of dynamical systems, this theoretical framework does not explain all emergent properties of biological oscillators. The characteristics of movement kinematics during a synchronization task in particular are not fully explained. Some kinematic data
challenge the theoretical frameworks used to model synchronization. For instance, during a finger-tapping task paced by auditory beats, Balasubramaniam et al. (2004) noticed that the trajectory was roughly sinusoidal, but movements to and from the beat could be distinguished, breaking some predictions of the dynamical system theory. Specifically, the symmetry in position and velocity were not perfect, and movements towards the temporal target were found to be more rapid than movements ensuring the return phase. To demonstrate the link between timing and asymmetry, Balasubramaniam et al. correlated timing accuracy with movement kinematics, and indeed found a positive relation: the relative asynchrony, i.e. the early or late tap with respect to the auditory cue was negatively correlated with the subsequent return phase. The duration of the movement towards the cue was, on the contrary, relatively constant. The strong negative correlation between each phase bolsters the existence of a closed-loop error-correction principle, visible in the trajectory asymmetry. This set of data is not in complete agreement with open-loop or pure limit-cycle oscillator models accounting for repetitive movements, since these models do not account for movement asymmetry. A unifying model, reconciling consecutive event correlation and movement asymmetry, is still to be proposed.

If evidence of cortical entrainment by musical tempo has been now established, only a systematic investigation of the relation between sound inputs and neural outputs will reveal the neural mechanisms of sound perception. The shaping of cortical activity by sensory inputs indeed raises the question of the preceding processing stages which ensure the selective entrainment to metre-related frequencies (Nozaradan et al., 2018b). The embodiment mechanisms of auditory rhythms will be better understood after how rhythm complexity affects temporal selection has been clarified. In other words, the investigation of input–output nonlinear transformation of auditory rhythms and their contribution to the emergence of perceived beats should be continued. In this regard, if the mapping
of stimulus inputs and EEG outputs provides a macro description of the relation between stimulus properties and brain activity, it fails to reveal specific neural processes (Nozaradan et al., 2018a). ECoG recordings, which offer better signal-to-noise ratio than scalp-recorded EEG, could contribute to clarify the link between single stimulus presentations and cortical activity (Sturm et al., 2014).

Auditory rhythm perception can shape movement, but the opposite is also true. If moving one finger, or oscillating our body in time to music helps to feel the beat, can we consider this common practice as an evidence of dynamic attending? Chemin et al. (2014) showed that in participants listening to a rhythmetrical sequence before and after moving their body to this rhythm, the electroencephalographic responses to the rhythm were superior at the specific rhythm frequencies after body movement. Phillips-Silver and Trainor (2008), by demonstrating how influential passive motion of the head can be on rhythmic metric encoding, posited the vestibular input as a determining factor of rhythm perception, and outlined the cross-modal nature of the coupling between rhythm perception and movement. Vestibular and proprioceptive systems, as complementary sources of movement-related inputs, connect body movement and auditory rhythm processing (Phillips-Silver & Trainor, 2007; Phillips-Silver & Trainor, 2005). If disentangling the specific mechanisms of a reciprocal influence remains intrinsically challenging, the influence of movement on the perception of rhythmical cues has not only motivated the search for experimental evidence, but has also nourished a renewed theoretical approach of sensory-motor coupling. In a protocol entailing the disentanglement of two streams of sound on the basis of endogenous temporal cues, Morillon and Baillet (2017) evidenced the dependence of predictive timing on the activity of the motor system. They noticed beta neural oscillations directed towards auditory regions and the sharpening of the quality of temporal predictions, as well as better temporal selection of relevant auditory information with overt rhythmic movements. These results
reinforce the active sensing hypothesis which represents a relevant framework to
investigate the interdependence between action and perception.

Top-down influences on rhythm perception should not be neglected either. In
musicians presented with either a sequence of tones or blinking visual cues at the
same frequency on which they were instructed to project a ternary metre, the
electrophysiological responses exhibited an enhancement of the amplitude at the
ternary subharmonic (one third of the stimulus frequency) across sensory
modalities (Celma-Miralles et al., 2016). Recent studies precisely attempted to fill
the gap between central and behavioural evidence of entrainment, by establishing
a link between rhythmical perception and synchronization abilities with
rhythmical auditory stimulations (Drake et al., 2000a), or their influence on gait
rehabilitation (Nombela et al., 2013; Dalla Bella et al., 2015; Dalla Bella et al., 2017;
Sihvonen et al., 2017). Cochen De Cock et al. (2018) reported a positive response
to cueing, measured as an increase of walking speed, in 17 out of 39 non-
demented patients with PD. These patients aligned their steps more often to the
rhythmic cues, and showed better music perception than the 22 patients with
non-positive response. In six of the patients, gait performance was worsened with
rhythmic cues. Rhythmic and musical skills, which can be modulated by musical
training, may increase beneficial effects of rhythmical auditory cueing in Parkinson’s
disease. These results point out how the personalization of auditory samples, as a
function of rhythmical perception ability, will ensure the best therapeutic value
for PD patients. The definition of the characteristics of rhythmical auditory
stimulations, such as frequency matching and variability, is another venue of
research to maximize the therapeutic effects of cueing. The temporal properties
of variability are particularly important in gait which involves movement
repetition. The temporal structure of cueing appears to influence the temporal
dynamics of gait. Dotov et al. (2017) demonstrated that cueing with isochronous
or randomly varying inter-stimulus/beat intervals removed the LRC in the stride
cycle. In contrast, persistent correlation in gait was unaffected with stimuli which mimicked the properties of variability found in healthy gait. The long-term consequences of biological variability embedded in stimuli should be investigated. Addressing this research question would also tackle the associated dilemma concerning the most beneficial type of cueing, synchronizing while scarifying natural variability, or sparing natural variability but diminishing the positive effects of synchronizing with the cue. Interestingly, an extensive analysis of the spectrum of western music revealed that rhythm fluctuations follow the fractal relation (Levitin et al., 2012) and appear to contribute to the pleasure experienced by listeners (Hennig et al., 2011).

Despite the common belief associating rhythmical auditory stimulations, and particularly music which boosts sports performance, further research is required to clarify the physiological mechanisms likely to support this relation. The main challenge lies in the interconnection between physiological and neurophysiological factors, since music reinforces the connection between perception and movement, and probably alters the activity of motor networks and their associated physiological processes. In contrast with the developing literature, which aims to clarify the link between the rhythmic skills and therapeutic potential of auditory gait training, similar systematic investigations of the runners’ gait, under the influence of music as a function of their perceptual ability, are still missing. A prerequisite could be the search for direct evidence of central entrainment during sports activities. Similarly assessing the entrainment potential of the locomotor system during running would complement the knowledge acquired on walking (Ahn & Hogan, 2012b). Beyond the use of periodic mechanical perturbations which entails the use of apparatus prone to alter the execution of fast movements, we believe that an innovative set up allowing the manipulation of the dynamical properties of the environment could bring some answers. In the presence of a time-dependent variation of ground stiffness (which
is a key determinant of the cost of transport and consequently governs running kinematics), how would runners explore this new dynamical environment? Would they adapt more quickly to the dynamics of the new environment in the presence of auditory cues? Such investigation of the balance between central and biomechanical factors would challenge our initial proposition of the existence of multi-level entrainment.

In conclusion, refining our understanding of how body movements are shaped by rhythm perception amounts to a better understanding of how action and perception are coupled. This dialogue is beautifully illustrated by the dependence of rhythm perception on body movement. Brain responses to rhythm are enhanced by previous movement performance (Chemin et al., 2014). The influence of anthropometric factors on preferred tempo (Todd et al., 2007) can also be seen as an embodiment of our movement repertoire within perception. The mechanism by which the brain favours task-relevant inputs (Lakatos et al., 2008) logically opens the possibility to manipulate movement performance with relevant sensory inputs. Action-based relevance of the stimulation appears as the primary factor to consider when the goal is to foster auditory-locomotor coupling.
This research was supported by the European Union with the following grants: (i) FP7-ICT-2013-10 “Beat-Health”, and (ii) H2020-FETPROACT-2018-01 “EnTimeMent”. We thank Stefan Janaqi for his comments on the manuscript, and Julien Lagarde and Charlotte Roy for providing the data plotted in figure 5A.
7 Appendixes

7.1 Box 1: Information processing theory provides an alternative theoretical framework to capture auditory-movement synchronization

The present review mainly focuses on biological and physical phenomena for which dynamical models are suitable. Information processing theory, which provides an alternative theoretical framework, is also relevant to analyse variability and error correction related to AMS tasks. Timekeepers represent the core of that theory, which supposes an explicit generation of discrete time intervals. Through this approach, the emphasis is consequently put on cycle-to-cycle measure and correction of motor timing. Error correction is crucial to sustain SMS because of the inherent variability of movements (Woodworth, 1899; Fitts, 1954). When plotted in the phase plane, the kinematics of biological movements, for instance finger flexion/extension during tapping, or bimanual coordination, appears as the superposition of partially recovering closed curves which describe a band (figure 5A).

Linear timekeeper models have been particularly relevant to explain data obtained in the continuation tapping paradigm, in which participants are instructed to tap in synchrony with a metronome and continue to tap at the same tempo after the auditory pulses have been turned off. This paradigm reveals the ability of human participants to maintain a mean tapping rate close to the reference value while the variability increases with the duration of the interval (Stevens, 1886). Moreover, inter-tap interval duration during the continuation phase is negatively correlated with the interval duration at lag one, i.e. there is an alternation of longer and shorter intervals compared to the original prescribed period. Wing and Kristofferson (1973b) proposed a model (W-K model, figure 5 B)
to account for the mean and the variance of the inter-response intervals. In the W-K model, a central timer generates intervals subject to temporally variable motor implementation due to motor delays. Both the clock and the motor delays possess their own variance, which can be estimated (Wing & Kristofferson, 1973a). Vorberg and Wing (1996) established the increase of clock variability with decreasing movement frequency, whereas the variability related to motor delay was found to be constant (Wing & Kristofferson, 1973b; Vorberg & Wing, 1996).

Hence this combination associates an increase of the total variance, i.e. the variance of the intervals’ durations, with lower movement frequency. The increase of perceptual variability with interval duration, predicted by the model, has been measured experimentally (Friberg & Sundberg, 1995).

In addition to drift, the existence of long-period fluctuations has also been reported, as early as 1886 by Stevens (1886), who noticed “larger and more primary waves” as opposed to “constant zig-zag”. Many biological rhythms follow a mathematical law (1/f) (Gilden, 2001; Van Orden et al., 2003). In the case of rhythmical motor activities, it is expressed as long-range correlations or long memory of the duration of intervals, which tend to be correlated with the duration of intervals that occur immediately after, as well as with the duration of intervals that follow much later in the series (Torre et al., 2011; Yamada, 1995; Roberts et al., 2000). Fluctuations in tempo, even if they are more marked during tapping synchronization continuation, are still present during sensorimotor synchronization (Chen et al., 1997). Power spectra of interval fluctuations of participants who are instructed to reproduce temporal intervals, exhibit linear negative slopes at low frequencies and linear positive slopes at high frequencies, representative of 1/f and white noises respectively (Gilden et al., 1995). The mixture of noises embedded in the data, as well as numerical simulations, encouraged the authors to consider the central time keeper of the W-K model as a source of 1/f noise and the motor command as a source of white noise. With
some adaptation of its parameters, the W-K model is still considered as a solid
paradigm by the advocates of the existence of a central time keeper.

7.1 Box 2. The cerebellar architecture ensures the mapping of
sensory information into temporally relevant motor
commands.

Imagery studies associate the cerebellum with AMS (Pollok et al., 2005).
Moreover, the transient alteration of cerebellar function with TMS entails an
impairment of synchronization performance during paced finger tapping,
evidenced by an increase of the variability of the intertap interval (Del Olmo et al.,
2007). These experimental clues should not hide the debate about the cerebellar
contribution to beat based timing.

Since cerebellum has been under the scrutiny of researchers, its neuronal
architecture has been associated with timing. The anatomy and functional
properties of cerebellar networks have indeed justified neurophysiological models
of cerebellar timing. Cerebellar circuitry is characterized by one sole output, the
Purkinje cells, located in the cerebellar cortex (Ito, 2000), which receive multiple
inputs from granule cells relaying mossy fibres via parallel fibres, and numerous
synaptic junctions from one climbing fibre (Eccles et al., 1967). Early on,
Braitenberg theorized the role of parallel fibres as "delay lines", activating
sequentially different Purkinje cells (Braitenberg, 1961; Braitenberg et al., 1997).
Another model, based on the oscillatory properties of the olivo-cerebellar
network has been proposed (Lampl & Yarom, 1993; Yarom & Cohen, 2002). Olivary
cells, whose climbing fibres project on the Purkinje cells, are characterized by
potential rhythmic sub-threshold membrane oscillations whose apex potentiates
spike occurrence by putting the neuron potentially closer to the threshold (Llinás
Oscillations enable the synchronization of clusters of neurons, as revealed by multiple-microelectrode recordings in rodents (Welsh et al., 1995), or more recently by dual whole-cell recordings in primates (Turecek et al., 2016). Subthreshold oscillations have given rise to the "timing" hypothesis (Ivry & Keele, 1989), which associates cerebellar activity with the temporal relationship between relevant events. This would entail the existence of a specific cerebellar timing unit (Ivry, 1997). In other words, the cerebellum would invoke an explicit representation of time (Ivry et al., 2002), such as equal intervals during an auditory-paced isochronous finger-tapping task.

The integration of sensory information with timing function would be the mechanism by which the cerebellum contributes to AMS. Jacobson et al. (2008) proposed that, within the cerebellar cortex, there is an association of the contextual inputs carried by the mossy fibres with a temporal pattern. In other words, temporal patterns supported by the cerebellar networks would not necessarily be the byproducts of an inflexible clock, but could be generated “upon request” (Jacobson et al., 2008). In vivo recordings revealed low threshold and short latency response of olivary cells to sensory stimulations (Gellman et al., 1983; Gellman et al., 1985) and also weak periodicity (Lang et al., 1999; Keating & Thach, 1995; Chorev et al., 2007). Whole cell-recording demonstrates the correlation between subthreshold oscillation frequencies and the frequencies of preferred spiking. Khosrovani et al. (2007) revealed the existence of a repertoire of different neurons characterized by stable frequency settings. These findings associate different phases of the oscillations with olivary cell discharge, and consequently legitimize the concept of an internal clock tuned by subthreshold oscillations. Mathy et al. (2009) proposed to characterize the subthreshold oscillation phase - axonal discharge relationship with a combination of direct patch-clamp recordings. A modulation of the relative phase between the synaptic
input and the subthreshold oscillations evidenced the dependence of the number of olivary cells spikes in the phase of the oscillation. Recordings of climbing fibres from Purkinje cells confirmed their activation by olivary bursts and their role in triggering plasticity mechanisms. Olivary cell bursts would convey more information than an on-off state (Najafi et al., 2014; Najafi & Medina, 2013). Mathy et al. (2009) proposed that the phase dependence of the number of axonal bursts maximizes the transmission of in-phase information from the olive without abolishing the out of phase input. Subthreshold oscillations, in addition to their role as a timekeeping device, would assign a level of saliency to stimulus as a function of their phase, in-phase stimulus being the most prone to enhance plasticity mechanisms in Purkinje cells. This mechanism could meet the premises of both timing, and classical error prediction theories (Albus, 1971; Marr, 1969). If research now emerges to challenge the cerebellar mapping of proprioceptive information into a temporal framework during motor learning processes (Kimpo et al., 2014; Yang & Lisberger, 2014), similar experimental evidence of its contribution to SMS is needed. The synchronous discharge of inferior olive neurons at a frequency matching those of tremors, following the injection of harmaline (Llinás & Volkind, 1973), or time locking of Purkinje cells activity to skilled movements (Welsh et al., 1995), supports the online influence of the olivo-cerebellar complex on motor control (Lang & Blenkinsop, 2011). The alteration of motor responses, spinocerebellar reflexes following sensory perturbations, and Purkinje cells activity in mice lacking electrical coupling in inferior olive cells, recently highlighted the direct relation between timing of movement and synchrony of olivary discharges (De Gruijl et al., 2014): the role of the coupling between olivary cells, which is a condition for their contribution to motor control (Lang & Blenkinsop, 2011), is not limited to motor learning (as postulated by Van Der Giessen et al., 2008 e.g.), but seems to have a significant effect on online temporal processing of sensory information related to motor control.
The time-pacing ability of cerebellar networks, associated with sensory integration, still represents a challenge for neuroscientists looking for a functional definition of cerebellar networks. When reviewing behavioural data, an agreement on the main contribution of the cerebellum to duration-based timing, evidenced by the alteration of the timing abilities of the patients affected by cerebellar lesions (Ivry et al., 1988; Moberget et al., 2008; Grube et al., 2010a), or disruption of cerebellar function with rTMS (Grube et al., 2010b), was dominant until recently. But the last beat-based timing protocols extended the implication of the cerebellum to this type of rhythm. Paquette et al. (2017) used voxel-based-morphometry, a neuroimaging technique which reveals the relation between a behavioural variable and gray matter signal (Ridgway et al., 2008), to clarify the relation between inter-individual variations in gray matter volumes across the entire brain volume and beat interval discrimination skills evaluated with Harvard Beat Assessment Tests. The data supported the involvement of the cerebellum in both beat interval test (BIT) and beat finding and interval test (BFIT), although the scores of the latter test correlated strongly with gray volume data, whereas the correlation for the scores of the former test was only marginally significant. If both tasks require perceiving and producing a gradual tempo change and are consequently related to absolute timing, the widely accepted function of the cerebellum, the BFIT, compared to the BIT, adds a beat finding component: this appears in contradiction with the vision of non-involvement of the cerebellum in beat perception. The experimental data, in agreement with the compensation of basal ganglia timing deficits by cerebellar loops in PD patients (see main text), also legitimizes their contribution to beat based timing. Recent timing models no longer restrict beat-based timing ability to basal ganglia. Despite the distinction between neural substrates supporting duration-based and beat-based auditory timing that Teki et al. (2011b) helped to establish, they subsequently suggested that neither network may be independent (Teki et al., 2011a). The deficits in both duration and beat-based timing evaluated in patients suffering from Huntington
disease (Cope et al., 2014), which is a pure striatal disorder, support the validity of this unified timing model (Teki et al., 2011a).

In sum, there is a general agreement on the existence of some kind of temporal encoding in the cerebellar granular layer (Manto et al., 2012; D'Angelo & De Zeeuw, 2009), but a unified model capturing the whole sum of experimental data is still missing. Measures of performances in cerebellar patients during temporally non-adaptive (isochronous pacing) and adaptive (tempo-changing pacing) AMS confirm a global temporal processing dysfunction (Schwartze et al., 2016). The fact that noticed impairments, which translated into increased asynchronies and impaired error correction, appeared to be directly linked to the tens-of-millisecond range preceding the auditory events, reinforced the conception of Penhune et al. (1998) and Schwartze and Kotz (2013) of the cerebellum as a structure dedicated to the temporal reduction of time relevant sensory input, i.e. the encoding of the sensory flow into temporal discrete events.

7.2 Box 3: Motivation, musical pleasure and physical activity

What lies behind the ergogenic benefit of rhythmical auditory stimulations? Other factors than the underlying dynamical processes reviewed in this article must be considered. For instance, rhythmical stimulations act as a pain distractor. Directing attention to music or to a metronome during physical activity may alleviate unpleasant sensations related to fatigue and pain (Nethery, 2002; Nethery et al., 1991; Edworthy & Waring, 2006; Karageorghis & Priest, 2012). By focusing on auditory inputs, people may thus be more willing to push themselves beyond their mental and physical limits, thereby increasing their running distance as ascertained by the extra time taken to reach exhaustion. Fritz et al. (2013) clarified this attentional diversion effect of music by modulating musical agency during a
strenuous task which relied on the experience of body proprioception. Because musical agency significantly decreased perceived exertion, the observed benefits cannot solely be attributed to an effect of diversion from proprioceptive feedback. Bigliassi et al. (2017) argued that moderate exercise could promote a combination of specific mechanisms down-modulating the exercise-related consciousness, i.e. a reallocation of attentional focus toward auditory pathways, with a concomitant reduction of communication across somatosensory regions. As a support of this hypothesis, they reported more positive perceptual and affective responses from participants when they were listening to music. The down-regulation of the EEG amplitude at ~700 ms after the onset of muscle bursts corresponded to the resynchronization time of neurons. According to the authors’ interpretation, by reallocating attention to external influences and dampening fatigue-related sensations, music would appear to favour more autonomous control of movements (Bigliassi et al., 2016). These results call for the standardization of motivational aspects of music when used experimentally. Karageorghis et al. (Karageorghis et al., 1999; Karageorghis et al., 2006; Karageorghis, 2008) developed a questionnaire to assess the motivational aspects of musical samples in the context of physical exercise.

A glimpse into contemporary hypotheses uncovering the origin of musical pleasure appears as a prerequisite to deconstruct the satisfaction elicited by AMS. Musical pleasure would be given through the interplay between prior learning and changes in the structure of the stimulus (Huron, 2008). The repeated exposure to a specific style of music, such as Western tonal music, develops our implicit knowledge of musical rules and regularities, e.g. mental representation of tonal relationships (Tillmann et al., 2000). When listening to music, the progressive unfolding of the rhythmical structure, and beats in particular, meet timing predictions (Bharucha & Stoeckig, 1987), which can be confirmed or violated if the beat is perceived earlier or later than expected (Rohrmeier & Koelsch, 2012; Bailes
et al., 2013; Pearce & Wiggins, 2012). Early on, Meyer (1956) indeed hypothesized that there was a relation between the degree of expectation fulfilment and emotion. Acting as a pleasant stimulus, music evokes complex emotions, and the associated central activations encompass networks involved in memory, sensorimotor processes, and reward (Blood & Zatorre, 2001; Peretz, 2010; Pannese et al., 2016).

Salimpoor et al. (2013) identified the activity in the mesolimbic striatal regions, especially the nucleus accumbens, as the best predictor of the amount of time listeners desire to spend on previously unheard music. The activities of auditory cortex, amygdala, and prefrontal regions were also increased when valuation was required, but did not appear as a good predictor of reward value contrary to the functional connectivity of these regions with the nucleus accumbens. As a whole, ventral striatum (nucleus accumbens) and dorsal striatum (caudate, putamen, pallidum) appear to contribute respectively to emotion processing derived from music (Berridge & Robinson, 1998) and rhythm perception and production (Grahn & Brett, 2007). The interconnection between both parts of the striatum and their influence by the dopaminergic system represent neuroanatomical clues of the role of the reward system in associating entrainment with pleasure. If there is no doubt about the coding of expected reward value by dopamine neurons (Montague et al., 1995; Schultz, 1998), the activity of single dopamine neurons in awakened monkeys indicates that they are also involved in coding reward probability and perceptual salience (Zink et al., 2004; Berridge, 2007; Schultz, 2010): dopamine neuron activity would code for the precision of prediction error (Friston, 2009), and would be representative of predictive coding (Schultz, 2016). High-precision predictions would possess large motivational salience. How this “revisited” dopamine function fits with the classically admitted role of basal ganglia in movement selection (Redgrave et al., 1999) is a key question. The control of dopamine release would optimize proprioceptive predictions, a key determinant
of good SMS performance. Based on the interaction between musical appreciation and prediction error, this still debated theoretical framework could provide an account for the interplay between SMS and musical enjoyment (see Haile Gebreselassie’s testimony in the ‘main text). More generally, this interaction provides a rationale for the power of music emotional salience, or “groove” (Michaelis et al., 2014) to elicit movement, as evidenced by its influence on motor excitability (Kornysheva et al., 2010; Giovannelli et al., 2013; Stupacher et al., 2013).
8 References


Hoover, J. E., & Strick, P. L. (1999). The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal


Parthasarathy, H. B., Schall, J. D., & Graybiel, A. M. (1992). Distributed but convergent ordering of corticostriatal projections: analysis of the frontal eye field
and the supplementary eye field in the macaque monkey. *Journal of Neuroscience, 12* (11), 4468-4488.


lactate threshold training on the relationship between mechanical stride
descriptors and aerobic energy cost in trained runners. *Arch Physiol Biochem, 109*
(2), 110-116.
response to preferred and nonpreferred tempos. *Journal of Human Movement
Studies, 8* (3), 123-138.
energetics and the origin of human bipedalism. *Proceedings of the National
Academy of Sciences of the United States of America, 104* (30), 12265-12269.
5 Softky, W. R., & Koch, C. (1993). The highly irregular firing of cortical cells is
inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience,
13* (1), 334-350.
strength, and muscle activation between subjects with and without
Phase entrainment of human delta oscillations can mediate the effects of
expectation on reaction speed. *Journal of Neuroscience, 30* (41), 13578-13585.
bradykinesia months after neurostimulator implantation. *Mov Disord, 32* (8),
1183-1190.
of evoked and induced electroencephalographic responses in primary auditory


van Wijk, B. C., Beek, P. J., & Daffertshofer, A. (2012). Neural synchrony within the motor system: what have we learned so far. *Front Hum Neurosci, 6*, 252.


9 Figure Captions

9.1 Figure 1. Models of oscillators.

A. Mechanical oscillators.

a. A simple gravity pendulum experiences sinusoidal oscillations around its equilibrium point at a constant frequency, with a constant amplitude. Periodic oscillation is represented by a closed curve in the phase plane.

b. The damping of the mass-spring system described here results in the progressive decay of the amplitude of its oscillations. Biological rhythms are governed by self-sustained oscillators, which possess a source of energy able to compensate for the dissipation in the system.

B. Neural Oscillatory processes.

a. The measure of the voltage of a spiking neuron describes periodic variations. These variations are characterized by an attractive limit cycle in their phase portrait.

b. Phase-response curve (PRC), adapted from Stiefel and Ermentrout (2016), is when spike time shift is induced by a perturbation of the neuron’s voltage ((LOIC check that this addition of “is” is ok)). The perturbation results in a delayed spike. The curve describing the PRC is negative before becoming positive: spikes can be advanced (first half of the cycle) or delayed (second half of the cycle). In the case of two coupled neurons, the discharge of one neuron can phase shift the spikes of the other neuron, and vice versa, and deviations from synchrony are corrected.

c. Schematic representation of neural oscillators. Neural oscillations can arise from the interaction between excitatory and inhibitory neural populations (above adapted from Hoppensteadt & Izhikevich, 1996). Periodic auditory stimuli would entrain auditory cortical oscillations (below, adapted from Large et al., 2015) which couple a distributed network of motor and sensory structures.
9.1 Figure 2. Interplay between neural and mechanical activities underlyng Auditory Motor Synchronization (AMS).

A. Evidence of modulation of cortical activity with auditory stimulations. **Sound envelope of an excerpt of music and associated beats (a).** The auditory system extracts beats as salient perceptual events of a musical piece. (b) **Beat-and metre-related steady-state EPs recorded in a single representative subject** (adapted from Nozaradan et al., 2011). Nozaradan et al. recorded the electroencephalogram while participants listened to a musical beat and imagined a binary or a ternary meter on this beat. Top: the topographical maps of EEG signal amplitude at 0.8, 1.2, 1.6, and 2.4 Hz, obtained in beat/beat+binary imagery/beat+ternary imagery conditions. Bottom: the EEG amplitude spectrum (in microvolts) within a frequency range comprising the frequency of the beat (2.4 Hz) and the frequency of the imagined binary and ternary metres (1.2 and 0.8 Hz, respectively). In all three conditions, the auditory stimulus elicited, at 2.4 Hz, a clear beat-related steady-state EP. In the binary metre imagery condition a metre-related steady-state EP emerged at 1.2 Hz, and at 0.8 and 1.6 Hz in the ternary metre imagery condition.

B. **Dynamical properties of the locomotor chain.** Harmonic mechanical oscillators are commonly used to model walking (c: inverted pendulum) and running (f: mass spring system). The amplitude of the oscillations depends on the initial perturbation. Biological oscillators, in contrast, tend to have not only a characteristic period, but also a characteristic amplitude. In the phase space their trajectories correspond to a limit cycle. If a perturbation is exerted, they will automatically come back to their normal behaviour, i.e. to their limit cycle. They
indeed incorporate a dissipative mechanism to damp large oscillations and a source of energy to prevent oscillations becoming too small. In passive walkers (c) motion is maintained by gravity. Left: after footstrike the swing leg swings forward past the stance leg until the swing leg hits the ground and a new step begins. $\theta$ is the angle between the two legs (adapted from Garcia et al., 1998). Right: nominal cyclic motion trajectory of the simplest walking model in phase space (adapted from Hobbelen & Wisse, 2007).

C. Dynamic entrainment of locomotion to periodic mechanical perturbations (adapted from Ahn & Hogan, 2012b). Ahn and Hogan applied periodic torque pulses to the ankle at periods that were different from preferred cadence. Entrainment occurred when the perturbation period was close to the participants’ cadence, demonstrating the existence of a narrow basin of entrainment, in agreement with the predictions based on limit cycle oscillators. Entrainment at both cortical and biomechanical levels exemplifies the possibility to entrain multiple structures associated with movements. This is so, as long as the external driving force lies within the basin of entrainment of the biological structure.

D. Neural Bases of auditory motor synchronization. Motor areas recruited during AMS tasks mark out the time boundaries of sequentially executed motor actions. As opposed to the execution of a cueing-free movement sequence, AMS requires transforming auditory information to extract relevant temporal events regulating the action. The planum temporale, involved in spectral and temporal analyses of sounds, projects both to dorsal premotor and prefrontal regions, suggesting a tight coupling between auditory and motor systems, with input to and from working memory. Multiple movements are sequenced in SMA and pre-SMA areas, receiving input from working-memory-related structures, and implementing the motor command. The functional role of premotor areas is also achieved through the involvement of cortical outputs to the basal ganglia and the thalamus,
connecting both subareas of SMA and the output pathway from the basal ganglia back to the cortex. The cerebellum contributes to encode the flow of auditory afferences into motor relevant sensory information. See main text for details.

E. Schematic diagram of the control of locomotion. Adapted from Grillner et al. (2008).

9.2 Figure 3. Functional connectivity of brain structures during AMS.

A. Functional link between auditory and motor areas during tapping (Chen et al., 2006). (a) Manipulating metrical structure during tapping in synchrony with an external beat, with concomitant scanning of the participants’ brain (fMRI). Louder tones were associated with longer tap durations (b) and concomitant increases in the BOLD response (c) at auditory (part of the superior temporal gyrus, STG) and dorsal premotor (dPMC) cortices. This functional connection is supported by direct connection between posterior STG and PMC, PMC appearing as the relay for the transmission of auditory information to the primary motor cortex (Chavis & Pandya, 1976; Luppino et al., 2001; Petrides & Pandya, 2006).

B. Neural correlates of perception vs. perception action during tapping (Chen et al., 2008a). a. Depiction of stimulus. b. Participants listened in anticipation to the upcoming tapping task, and then tapped in synchrony with the rhythm. Activations of the supplementary motor area (SMA), mid-premotor cortex (midPMC), and cerebellum were found during listening with anticipation. When participants naively listened to the beat (e.g. without foreknowledge of the later (LOIC is it later or latter??) tapping task), the same motor regions were activated. However, in contrast, dPMC - which is sensitive to higher-order information such
as the metrical structure - was only engaged during tapping. Ventral premotor
cortex (vPMC) was engaged whenever the sounds were functionally linked to a
motor response, even though delayed. This provides evidence on the
complementary roles of PMC pathways during perception and perception-action
situations (Chen et al., 2009).

C. Simultaneous recording of brainstem and cortical electroencephalographic
activities during tapping (Nozaradan et al., 2016b). a. Stimulus: participants,
while listening to a chord - three tones periodically amplitude-modulated inducing
a steady beat - were instructed not to move or to tap on every second beat. The
frequencies spectrum of the stimulus was set to elicit frequency-locked responses
in brainstem and cortical areas. The harmonic chord with partials at 200, 400, and
600 Hz, likely to evoke brainstem nuclei response, was amplitude modulated at
2.4 Hz to induce a beat, prone to elicit a cortical response. b. As expected, cortical
response at beat frequency during sensorimotor synchronization was increased
compared to passive listening. c. brainstem partial steady-state evoked potentials
(SS-Eps) showing enhanced subcortical response in tapping condition compared
to auditory one. d. Spectra of each of the brainstem partial SS-Eps which were also
increased during tapping. For all three types of responses (cortical, brainstems
partials, brainstems sidebands) the increased amplitude in the tapping condition
compared to the auditory condition suggests the contribution of cortical and
subcortical activities modulation in the encoding of sounds during AMS. Brainstem
response enhancement could involve top-down modulation from cortical to
subcortical structures (Bajo & King, 2012).
9.3 Figure 4. Sensory and motor performance during AMS.

A. Prestimulus EEG alpha phase synchronicity during a standard auditory "oddball" paradigm (Haig & Gordon, 1998).

a. The synchrony of alpha phase with N100 was measured (ERP, event-related potential arising between 80 and 120 milliseconds after the onset of a stimulus and generated by the primary and association auditory cortices) while listening to predictable auditory sounds. Stimulus consisted in regular tones of 1000 Hz at an interval of 1.3 s. Participants were instructed to ignore these tones and to respond to a second target tone of 1500 Hz which was intermixed with the lower tone by pressing two buttons.

b. In order to clarify the relationship between prestimulus alpha activity and N100 components, the authors proposed the alpha phase synchronicity as a new measure. It corresponds to the circular variance of the alpha phase at stimulus onset across the parieto-occipital sites. Simulated alpha-phase synchronicity (five parieto-occipital sites) at stimulus onset is represented, one with high-phase synchronicity (top) and one with low-phase synchronicity (bottom).

c. Topographic maps showing the group averages of the conventional average N100 amplitude (ERP epochs were conventionally averaged together) and the high- and low-phase synchronicity N100 amplitude (alpha activity phase across the recording sites allowed the identification of high and low synchronicity groups). N100 amplitude was significantly greater in the high than the low phase synchronicity subaverages.

d. Significance probability map (SPM) for the comparison between high- and low-phase synchronicity N100 amplitudes. N100 significantly differed in all 15 sites. The reduction of reaction time in the high prestimulus alpha phase synchronicity cases shows the interplay between external rhythms and brain state related to subsequent stimulus processing.
B. Motor activity contributes to sensory selection (Morillon et al., 2014).

a. Experimental protocol and motor-tracking-locked rhythmic gain model.
Rhythmic sequences of 20 pure tones were used. Four reference tones indicated
the beat, followed by an alternation of eight targets and eight distractor tones of
variable frequencies presented in a quasi-rhythmic manner. Targets and
distractors occurred in phase and antiphase with the preceding references.
Participants initially tracked reference beat. They used the reference beat to
maximize the integration of relevant sensory cues while minimizing the relative
weight of irrelevant ones in the sequence of interleaved targets and distractors.
First row: rhythmic motor tracking of musical beats. Second row: references. Third
row: targets in phase with the reference beat. Arrows indicate the temporal
distance between the movement and the target. Fourth row: distractors
presented in antiphase with the reference beat. Arrows indicate the temporal
distance between movements and distractor onsets. Fifth row: gains assigned to
targets and distractors in the model.

b. Experimental validation of the model. Target and distractor gains as a function
of their temporal distance to taps (dashed lines stand for the “listen” condition).
The mean frequency of distractors was always equal to the reference frequency
whereas the mean frequency of targets was varying. Participants indicated
whether the mean frequency of targets was higher or lower than the reference
frequency. Sensory gains were estimated for each target and distractor tone using
a multivariate logistic regression of choice against a weighted sum of the
information provided by each tone, expressed in relative distance from the
reference frequency. In the “listen” condition, participants stayed still. In the
“motor-tracking” condition, they performed the task while pressing a noiseless
pad with their index finger in phase with the reference beat.

c. Left: categorization performance in the “motor-tracking” and “listen”
conditions. Right: contributions of targets and distractors to the decision in the
“motor-tracking” and “listen” conditions. Overt rhythmic motor activity enhanced
specifically the sensitivity to target tones. This effect relies on cyclic fluctuations in sensory gain time-locked to individual movements. Top-down influences would sharpen sensory processes, giving sense to the concept of “active sensing”.

C. Rhythmic abilities and musical training in Parkinson's disease (Cochen De Cock et al., 2018).

a. Patients’ individual gait response to various rhythmic stimuli was tested, as well as their motor and non-motor rhythmic performance. The differences between patients with positive response and non-positive response to cueing were examined. The rate of auditory cues was set to be 10% faster than each participant’s preferred cadence, measured at pretest.

b. Individual responses to rhythmic cueing, expressed by the difference in gait speed between cueing and the baseline, in patients with Parkinson's disease and in controls. Patients who aligned their steps to the beat also increased their speed; this was not the case in the controls.

c. Beat perception
9.4 Figure 5. A timekeeper account for the variability of AMS.

A. Phase plane trajectories of the right index during synchronization with a metronome. One trial consisting in a sequence of 80 stimuli presented with an inter-onset interval of 800 ms (1.25 Hz). The partial recovering of closed curves describes a band. Data from Roy et al. (2017).

B. Synchronization–continuation paradigm and Wing–Kristofferson two-level timing model (adapted from Wing, 2002). Participants synchronized their tapping with the pacing stimuli. During the following unpaced phase (continuation), they were instructed to continue tapping at the same tempo. Inter-response intervals (I) are defined by timekeeper intervals (C) and motor implementation delays (M). Average (I) is matching (C), with variations in (I) reflecting both (C) and (M). The dashed lines show how negatively correlated (I) (short and long intervals alternation) can be explained by variation in (M).
## Table 1. Glossary of basic concepts related to auditory stimulation, gait and dynamical systems.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rhythm</strong></td>
<td>Regular temporal patterns embedded in music in the form of the alternation of strong and weak sounds, in movements when they are cyclical, or in any temporal series, which exhibit periodicity, such as electrical signals.</td>
</tr>
<tr>
<td><strong>Beat</strong></td>
<td>When listening to musical pieces, the perceived beat is the most regularly recurring psychological event in response to a rhythm (Cooper &amp; Meyer, 1960; Large, 2008). More practically, the beat is the regular time interval that we can tap along to when listening to music.</td>
</tr>
<tr>
<td><strong>Dynamical System</strong></td>
<td>An ideal model governed by a function which describes its time-dependent evolution, with future states being unambiguously determined (Pikovsky et al., 2003). For example: a simple harmonic oscillator which undergoes sinusoidal oscillations around its equilibrium point (figure 1A).</td>
</tr>
<tr>
<td><strong>Oscillator</strong></td>
<td>A system which exhibits rhythm(s). The time-varying amplitude of the variable which characterizes the system (angle of a pendulum, tension in an electronic circuit, membrane potential of a neuron, figure 1A and B), describes variations about a central value or between different states. Oscillators can be periodic (the time between consecutive corresponding points of the same amplitude is constant: e.g. harmonic oscillators whose output is governed by the sine function) or chaotic. Models referred to in the present review are built upon harmonic oscillators.</td>
</tr>
<tr>
<td><strong>Natural frequency</strong></td>
<td>The frequency (number of beats in a unit of time) at which a system oscillates when it is stable and isolated.</td>
</tr>
<tr>
<td><strong>Linear vs. nonlinear systems</strong></td>
<td>A linear system of differential equations, defined by affine functions of the variable $x$, $f(x) = ax + b$, describes the state of linear physical systems. Nonlinear systems of equations, in which the unknowns are variables of a polynomial of degree higher than</td>
</tr>
</tbody>
</table>
one, or arguments of a function which is not affine, apply to nonlinear systems behaviour. In a nonlinear system, the solution to the equations is not a linear combination of the unknown variables or functions. The coexistence of many solutions for the same initial parameters is possible.

Within a narrow range of initial conditions and for a short interval of time, a linear equation might describe, with acceptable accuracy approximation, the output of a nonlinear system. The simple swinging pendulum governed by the equation

\[
\frac{d^2x}{dt^2} + \frac{g}{L} \sin x = 0
\]

(in which \(x\) is the angle of the pendulum from vertical, \(g\) is the acceleration due to gravity, and \(L\) is the length of the pendulum) is a model for the linear oscillator if the term \(\sin x\) is approximated by \(x\). This is valid for small angles and the numerical solution of the equation is straightforward, the equation now being in the form of a second-order linear differential equation. However, when \(x\) is large, the equation remains in the nonlinear form as depicted above.

<table>
<thead>
<tr>
<th><strong>Driven Oscillator</strong></th>
<th>An oscillator whose position is influenced by an external time dependent force. If the motion of the system is entirely sustained by the time varying force, oscillations are forced.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-sustained Oscillator</strong></td>
<td>In the simplest models of harmonic oscillators, such as a mass-spring system or a pendulum, energy dissipation is neglected. The associated equation of motion predicts the kinetic energy transformed into a potential energy and vice versa, the total energy of the system remaining constant. However dampening mechanisms act against the mass or pendulum motions. If a source of energy compensates for the gradual decay of the movement amplitude, we are in the presence of a self-sustained oscillator. In the pendulum clock the potential energy of the lifted weight is converted into oscillatory motion to prevent the inherent decay of the angle of the pendulum with respect to the vertical. The locomotor system possesses the properties of a self-sustained oscillator as is the case for numerous biological systems such as cardiovascular and respiratory systems.</td>
</tr>
</tbody>
</table>
Limit-Cycle Oscillator

To describe the state of an oscillator, the value $x(t)$ of its output process is not sufficient since, for a given value of $x$, the associated function can increase or decrease. A second variable is often called to provide an unambiguous description of the state of the system (Pikovsky et al., 2003). In the case of a pendulum clock, the time evolution of the pendulum angle ($x$) with respect to the vertical and its angular velocity $x'(t)$ can be used as pairs of coordinates in the phase space, the plot of $x$ vs. $x'$ being the phase portrait. The pendulum clock, which is a periodic self-sustained oscillator, exhibits a phase portrait which consists in superposed closed curves: it satisfies the definition of limit cycle as a closed curve in phase space towards which the trajectories converge.

Resonance

Occurs when an oscillating system is forced to vibrate at its natural frequency. At its resonant frequency, a minimal energy is required to get large amplitude oscillations.

Oscillator Coupling

Coupling between oscillators associates some transfer of energy between them. Two similar clocks on the same support get synchronized due to vibration transmission. Coupling strength describes how weak or strong this interaction is. In the example above, this is determined by the rigidity of the supporting structure and its ability to vibrate (LOIC is this what you mean?).

Synchronization

“Adjustment of rhythms of oscillating objects due to their weak interaction” (Pikovsky et al., 2003).

Entrainment

When coupled, oscillators with differing and somewhat close natural frequencies exert a mutual influence on each other. This drives the emergence of a common frequency of oscillation: they are frequency entrained, which can conduct to phase synchronization. However, perfect period and phase synchronization is only one specific case of entrainment. This term also describes partial period and phase adjustment without perfect matching.

For a given oscillator, its natural frequency, as well as its frequency associated with the presence of another oscillator, must be measured in appropriate experimental conditions, i.e. with the system to characterize in isolation vs. in specific coupling conditions.
Neural or brain oscillation

Rhythmic electrical activity observed at different levels of organization, within individual neurons, or within populations of neurons (recorded with commensurate techniques, e.g. with single-unit recording and EEG/MEG respectively). The combination of excitatory and inhibitory activities of different neural populations can give rise to nonlinear neural oscillations (Wilson & Cowan, 1973; Hoppensteadt & Izhikevich, 1996). Neuronal projections can spread oscillations across multiple brain regions. As for any other oscillation, they are characterized by their frequency. Beta waves (13-32 Hz) are modulated by movements in healthy subjects and their suppression is correlated to the predictive value of the cue prior to movement (Gilbertson et al., 2005; Engel & Fries, 2010). It has been proposed that beta activity reinforces postural state at the expense of new movement onset (van Wijk et al., 2012; Doyle et al., 2005; Hammond et al., 2007; Moran et al., 2008). Gamma waves (25-140 Hz) are observed in the sensorimotor areas that are involved during movement (Crone et al., 1998). Motor cortex, as well as subcortical structures (basal ganglia), exhibit transient bursts of high frequency gamma oscillations (60-90 Hz) during movement (Muthukumaraswamy, 2010; Gaetz et al., 2013; Jenkinson et al., 2013). However, on the basis of the irregularity of the gamma phase over multiple oscillation cycles (Xing et al., 2012), Nikolić et al. (2013) proposed that the phase of the oscillation cycle could be locked in a stimulus-dependent manner. The dynamic change of the phase of the oscillation cycle as a function of the spiking activity of neurons preferentially firing for a specific stimulus (Havenith et al., 2011), supports this hypothesis. Globally, action entails event-related beta desynchronization and movement-related gamma synchronization.

<table>
<thead>
<tr>
<th>Cadence</th>
<th>Number of steps per time unit (usually minutes) in walking or running.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running economy</td>
<td>Quantity of energy spent per distance unit during running.</td>
</tr>
<tr>
<td>tDCS</td>
<td>entails direct stimulation through low current via electrodes connected to the head. Anodal tDCS, which decreases the membrane potential, increases neuronal excitability. Cathodal tDCS, which hyperpolarizes the membrane potential, decreases the excitability of neurons.</td>
</tr>
</tbody>
</table>
11 Figures

11.1 Figure 1

A Mechanical oscillators
   a Pendulum

   b Damped mass-spring system

B Neural oscillatory processes
   a Spiking neuron

   b Phase-response curve

   c Neural oscillators
11.2 Figure 2
11.3 Figure 3
11.4 Figure 4
11.5 Figure 5

![Diagram showing synchronization and continuation processes with labels for pacing stimuli, motor responses, timekeeper interval, motor delay, and interresponse interval.]