Intrinsic and Task-Evoked Oscillatory Dynamics underlying Auditory-Motor Coupling

Oscar Bedford

Integrated Program in Neuroscience

McGill University, Montreal

December 2024

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Doctor of Philosophy

Contents

List of Figures

Study 1

Study 2

List of Tables

Study 1

Study 2

Appendices

Abstract

Understanding the temporal dynamics of oscillatory interactions between auditory and motor cortices is crucial for unravelling the neural basis of coordinated actions necessary for speech and music processing. To this end, this investigation employed electrophysiology to explore intrinsic and task-evoked oscillatory dynamics underpinning auditory-motor coupling. The first study examined intrinsic phase-based functional connectivity between auditory and motor cortices using resting-state magnetoencephalography across a sample of healthy young adults (n=90). As predicted, we observed greater phase-locking values in auditory-motor compared to visuomotor pairings across all frequency bands. Consistent with prior literature, the strongest synchronization was observed between right primary auditory regions and the right ventral premotor cortex, most prominently in the theta, alpha, and beta bands. Directed connectivity estimates confirmed the expected motor-to-auditory preference in the beta band and an auditory-to-motor preference in the alpha band. The second study used electroencephalography to investigate the dynamics of mu suppression, an index of anticipatory motor activity in a melody learning task. Using a novel data-driven approach, we first localized mu suppression preceding movements during motor training and then applied these coordinates to reveal mu suppression during passive listening to the previously learned melody. Crucially, suppression was observed at the single-note level. Together, these findings reveal distinct but complementary time-based oscillatory mechanisms for auditory-motor integration. Overall, this thesis sheds light on the human brain's capacity for managing remarkably low latencies between sounds and movements.

Résumé

La compréhension des dynamiques temporelles des interactions oscillatoires entre les cortex auditif et moteur est cruciale pour élucider la base neuronale des actions coordonnées nécessaires à la parole et à la musique. À cette fin, cette recherche a utilisé l'électrophysiologie pour explorer les dynamiques oscillatoires intrinsèques et induites par la tâche qui sous-tendent le couplage auditif-moteur. La première étude a examiné la connectivité fonctionnelle intrinsèque basée sur la phase entre les cortex auditif et moteur à l'aide de la MEG en état de repos (rs-MEG) sur un large échantillon (n=90). Comme attendu, nous avons observé des valeurs de verrouillage de phase plus élevées dans les paires auditive-motrices que dans les paires visuomotrices, et ce, dans toutes les bandes de fréquence. Conformément à la littérature antérieure, la synchronisation la plus forte a été observée entre le cortex auditif primaire droit (A1) et le cortex prémoteur ventral droit (vPMC), principalement dans les bandes thêta, alpha et bêta. Les estimations de connectivité dirigée ont confirmé la préférence attendue du moteur vers l'auditif dans la bande bêta et de l'auditif vers le moteur dans la bande alpha. La deuxième étude a utilisé l'EEG pour examiner la dynamique de la suppression de la bande mu, un indice d'activité motrice anticipatoire, que nous avons réussi à déclencher chez des non-musiciens pour la première fois. Grâce à une approche novatrice axée sur les données, nous avons d'abord localisé la suppression mu précédant les mouvements pendant l'entraînement moteur, puis appliqué ces coordonnées pour révéler une suppression mu pendant l'écoute passive de la mélodie précédemment apprise. Fait important, la suppression a été observée au niveau de la note individuelle, constituant une première dans la littérature. Ensemble, ces résultats révèlent des mécanismes oscillatoires distincts mais complémentaires pour l'intégration auditive-motrice dans les bandes de fréquence alpha et bêta. Globalement, cette thèse met en lumière la capacité exceptionnelle du cerveau humain à gérer des latences remarquablement faibles entre les sons et les mouvements.

Acknowledgements

I would like to start by extending my warmest gratitude to Profs. Robert Zatorre and Virginia Penhune, my supervisor and co-supervisor, for… well, *everything*. For placing your trust in me, for believing in me, and for supporting me through thick and thin. Also, for teaching me how a pro deals with all aspects particular to science: thinking, reading, theorizing, hypothesizing, experimenting, discussing, listening, presenting, writing, citing, editing, publishing, divulging, mingling, celebrating, and helping. I have taken one page out of your book when it comes to each of these aspects, even if I sometimes fail to do it with half as much style as you guys. Know that I truly aspire to become somebody who combines your larger-than-life strengths, regardless of the professional path I end up choosing when we part ways.

Thank you as well to the two members of my advisory committee, Profs. Julien Doyon and Alexandre Lehmann, for your wealth of conceptual and methodological contributions to my studies throughout the past seven years. To the people from my *alma mater*, most especially to Profs. Marco-Pallarés, Escera, Costa-Faidella, and Mas-Herrero: know that without your teachings, guidance, and reassurance I would never have dared to cross the Atlantic and kickstart this adventure. Thanks as well to everybody I met at Goldsmiths, most especially to Prof. Müllensiefen, Prof. Stewart, Dr. Harrison, Pedro, Gabriela, and Lily: whether you provided me with mentorship, friendship, or both, I treasure it to this day. Thank you also to Prof. Sylvain Baillet, Prof. Philippe Albouy, Dr. Alberto Ara, Dr. Alex I. Wiesman, Marc Lalancette, Marc Bouffard, and Ian Marci for your precious technical and methodological support. Special appreciation goes to Joe Thibodeau for building the piano keyboard used in Study 2, writing the associated code, and helping me set up; as well as to Jennifer Cohen, Sebastian Kolde, and Yingrui He, my research assistants, for helping me collect data.

Acknowledgements

A message to my fellow scientific peers and friends who are further ahead than me, most especially to Alberto, Connor, Melanie, Jérémie, and Étienne: I am perpetually inspired by your many talents and the stellar approaches you each take to science. To the labmates and friends who are on the same boat as me, Isabelle, Alix, Marcel, and Arielle: I cherish all of our discussions and the many ways we have given each other support over these last few years. And to those who will be completing their degrees a bit after me, most especially Miyoung, Alex A, Isaac, Giorgio, and María: you can do this; just keep believing in yourselves the way we all believe in you.

I am eternally grateful to my parents and my brother, most especially for teaching me, re-teaching me, and reminding me of the importance of creativity and following my bliss. Of all the ways in which I am privileged for having been born into this family, I feel especially privileged for this one. Thank you also to everybody on my mother's side of the family, most especially "*a la laia*", Marta, Lourdes, Ester, Marc, Arnau, and Natalia for cheering me on so affectionately. Also, thank you to everyone on my father's side of the family, most especially to Jo, Cally, and Jake for being so loving and kind to me. I look back fondly to the time when you guys helped me get started in London, and I miss laughing together about anything and everything.

To my lifelong or very old friends who are eagerly waiting for me to return home; Laura, Eloi, Gonzalo, Ruben, Núria, André, Alvaro, Sansa, and Alex: your friendship is the kind that keeps on giving, and I am honoured every time I am fortunate to be around you. Thank you as well to the dear friends I have met in Montréal: Noelia, Aarón, Pratap, Irem, Neelima, Shashank, Andrea, and Maeva. You are all amazing and I am so happy to have met you. Finally, a message to all the people who have visited the lab over the years, most especially Lei, Kazuma, Ana, Mario, Sarah, and Suzie: I am still repairing my heart after your departure, but I know it was a farewell, not goodbye.

And, of course, thank you to my partner Lisa, for being my rock and my absolute inspiration in so many more ways than one. Sometimes I wish you could see yourself through my eyes, because frankly I am nothing short of blessed to be so often in the

presence of your otherworldly warmth and enthusiasm, the depth of your curiosity, humanity, kindness, and respect for all lifeforms, as well your disarmingly passionate, earnest, loyal, transparent way of loving me. Know that the trust you have placed in my hands is something I have never taken or will take lightly, and that I still feel the same way as the day I became aware that I had fallen hopelessly in love with you.

Je t'aime, mon amour

Finally, I would like to dedicate this thesis to the following:

To Barcelona, for being my cradle and someday my grave. To London, for proving I was always free. And to Montréal, for the cracks that let the light in. To the loved ones who are no longer with us; Lit, avi, grandpa, nanna Sylvia, nanna Val; I miss you all dearly. If you're looking down, I hope you're proud. To my brother, for choosing to be himself. To my mom, for choosing herself. And to my dad, for choosing to be. I look up to you three, and I'm proud.

To music, for electrifying my synapses and reward areas, my life vision and purpose, and the timeless aspect of my soul. To life, for obliterating me at 27 and granting me the rare privilege of reshaping the puzzle from the ground up, piece by piece, better than it was before. And lastly to Lisa, for being the lighthouse that illuminated the throughlines during those darkest of hours, and for becoming the reason to see it all through to the end and beyond.

Thank you!

Contribution to Original Knowledge

This thesis makes several significant contributions to contemporary neuroscientific understanding, particularly in the fields of music cognition, auditory neuroscience, and functional connectivity. Study 1 utilized magnetoencephalography (MEG) to quantify differences in phase-based resting-state functional connectivity between cortical auditory and motor regions versus cortical visual and motor regions—a novel exploration in the literature. Furthermore, this study employed a cutting-edge metric, Phase Transfer Entropy (PTE; Lobier et al., 2014), to evaluate the preferred intrinsic direction of information flow between auditory and motor regions. This represents the first application of PTE in this specific context, offering valuable insights into oscillatory auditory-motor coupling paradigms.

Study 2 leveraged electroencephalography (EEG) to demonstrate that anticipatory mu suppression, a phenomenon tied to auditory-motor associations, can be detected at the single-note level in non-musicians. This finding is unprecedented in the literature on both fronts. Additionally, this study provided translational value by adapting an advanced methodology for functional localization in time-frequency-channel space (Maris & Oostenveld, 2007) to the field of music cognition. Together, these two studies address critical gaps in the understanding of the temporal dynamics of auditory-motor coupling—a field with direct clinical implications. Notably, this work supports the growing scientific foundation of Neurologic Music Therapy (NMT), a recognized and evidence-based intervention for neurological rehabilitation (Thaut & Hoemberg, 2016).

Contribution of Authors

Chapter 1 – General Introduction

Contributions of the student

Conducted literature review and wrote the chapter.

Contributions of the supervisor and co-supervisor

Prof. Robert Zatorre: Advised student (50%) and edited chapter (50%).

Prof. Virginia Penhune: Advised student (50%) and edited chapter (50%).

Chapter 2 – Study 1

Contributions of the first author

Pre-processed and analysed the data (50%). Performed all statistical analyses. Conducted the literature review and wrote the article. Designed all figures and tables. Researched and wrote all responses to reviewers.

Contributions of middle authors

Alix Noly-Gandon: Pre-processed and analysed the data (50%).

Dr. Alberto Ara: Offered statistical advice (100%).

Dr. Alex I. Wiesman: Offered scientific advice (75%).

Dr. Philippe Albouy: Offered methodological advice (50%).

Prof. Sylvain Baillet: Offered scientific (25%) and methodological advice (50%).

Contributions of the supervisor and co-supervisor

Prof. Robert Zatorre:	Provided idea for study (100%) and all funds. Advised student (50%). Edited manuscript and responses to reviewers (50%).
Prof. Virginia Penhune:	Advised student (50%). Edited manuscript and responses to reviewers (50%).

Chapter 3 – Study 2

Contributions of the first author

Provided idea for study (50%). Recruited all participants, ran all experimental sessions, and trained all research assistants. Pre-processed and analysed all the data. Designed novel methodology (functional localizer). Performed all statistical analyses. Conducted literature review and wrote the chapter. Designed all figures and tables.

Contributions of middle authors

Dr. Alberto Ara:	Offered statistical (100%) and scientific advice (50%).
Dr. Philippe Albouy:	Offered scientific advice (50%).

Contributions of supervisor and co-supervisor

Prof. Robert Zatorre:	Provided all funds. Advised student (50%). Edited chapter (25%).

Prof. Virginia Penhune: Provided idea for study (50%).
Advised student (50%).
Edited chapter (75%).

Chapter 4 – General Discussion

Contributions of the student

Conducted literature review and wrote the chapter.

Contributions of supervisor and co-supervisor

Prof. Robert Zatorre: Advised student (50%) and edited chapter (60%).

Prof. Virginia Penhune: Advised student (50%) and edited chapter (40%).

Chapter 1: General Introduction

1.1 Overview

Auditory-motor coupling is a crucial system underlying speech and music, two vital functions that have been observed across all cultures (Patel, 2007). Beyond elucidating how the brain orchestrates these essential functions, the study of auditory-motor coupling holds significant relevance for neuroscience, offering insights into complex structural and functional connectivity (Rauschecker & Scott, 2009), multimodal integration and higher-order representations (Zatorre et al., 2007), embodied cognition (Gallese & Lakoff, 2005), plasticity and learning (Lahav et al., 2007), theory of mind (Kohler et al., 2002), non-verbal social communication and coordination (Sebanz et al., 2006), real-time error-correction processes (Ito, 2008), predictive processing and anticipation (Friston, 2010), as well as time-keeping and fine-tuned temporal processing in the brain (Merchant & Honing, 2014). While each of these aspects is fascinating, the present thesis focuses specifically on elucidating the last point: auditory-motor coupling's capacity for fine-tuned temporal processing. This is because it is not possible to further interpret the extensive wealth of behavioural, anatomical, and localization-based neuroimaging studies on auditory-motor coupling without developing a deeper understanding of the neural mechanisms that govern this system's sophisticated spectrotemporal dynamics.

The average human brain is capable of successfully linking sounds to arbitrarily complex movements in a remarkably brief period of time. This idea is supported by the fact that nonmusicians display enhanced motor excitability during passive exposure to learned musical sequences almost immediately after motor training (D'Ausilio et al., 2006; Stephan et al., 2018). Moreover, we know that nonmusicians can transcend intuitive sound-to-movement mappings (left-low vs right-high) and successfully acquire non-conventional ones in the lab—both musical (Lega et

al., 2016) as well as non-musical (Van Vugt & Ostry, 2018). Even prior to training, the connectivity between auditory and motor cortices during passive listening is reliably predictive of training success (Wollman et al., 2018), which indicates that the auditory-motor system is primed to learn new associations. Furthermore, well-learned melody sequences can be dissociated from auditory feedback (Highben & Palmer, 2004), which suggests that the auditory-motor system has an additional drive for optimization. Yet, despite being quick to acquire and refine, auditory-motor associations are also surprisingly robust. Evidence can be found in behavioural studies showing that even small perturbations to acquired melody sequences can cause big interferences in playback (Keller & Koch, 2008).

These findings are rendered more astounding when one considers the wide range of motor effectors and movement patterns that auditory-motor coupling can serve. Note that there is little overlap in muscle group recruitment between speaking and playing a string instrument, for example. Nevertheless, while the effectors may vary, there is evidence that the demands on the system and the neural correlates themselves overlap significantly. For instance, there are many accounts of the convergence between speaking and singing coming from both neuroimaging (Hickok et al., 2003; Zarate & Zatorre, 2005) and electrophysiology (Liu & Larson, 2007). Moreover, there is a substantial commonality between the neural correlates for speaking or singing and those of musical instrument playing (Parsons et al., 2005), including the neural correlates supporting the playing of fretless instruments that allow for pitch to be modified in a manner comparable to vocalizing (Segado et al., 2018). Taken together, these studies lend support to the "neural recycling" theory (Dehaene & Cohen, 2007), under which modern auditory-motor tasks, such as playing an instrument, borrow much from pre-existing, phylogenetically ancient auditory-motor mechanisms that developed for vocalization.

Evidence of neural recycling can be gleaned from the fact that auditory-motor error correction has an instinctual basis. Indeed, there are at least two well-known

compensatory reflexes which carry over from vocalization to music-making. One is called the pitch-shift reflex (Houde & Jordan, 1998) and it is an automatic, involuntary adjustment in pitch that occurs when an individual hears an errant shift in the pitch of their vocalization or instrument. The other is called the Lombard effect (Luo et al., 2018) and it consists of an involuntary increase in vocal loudness in proportion to increasing ambient noise. That these two reflexes exist in the first place, paired with the fact that they can cross into music-making, denotes that the system has an impetus to mediate feedforward and feedback loops beyond any specific sound-effector pairing. It follows that a cognitive model describing auditory-motor coupling must encompass the findings from speaking, singing, and music-making studies. Furthermore, it must account for the interaction between diverse processes such as motor planning and execution, auditory and sensory feedback, and auditory-motor transformations (Zatorre et al., 2007).

But an effective model must also address the privileged and unique temporal relationship between auditory and motor areas, the central focus of this thesis. For instance, it should include the auditory-motor system's remarkable capacity for managing very low latencies between sounds and movements (Comstock et al., 2018). Additionally, it should account for the fact that humans of all ages, including babies, instinctively synchronize their body movements to rhythmic sound patterns in the environment (Zentner & Eerola, 2010). Indeed, this innate synchrony has been harnessed in Parkinson's rehabilitation, where simultaneous exposure to periodic sounds naturally improves movement parameters (McIntosh et al., 1997). Conversely, none of these findings has been replicated in other types of sensory-motor coupling, such as visuomotor coupling. The fact that humans do not naturally tap to visual stimuli, for example, suggests that the auditory-motor system has evolved specialized mechanisms for temporal processing that are not shared by other sensory-motor pathways. Such mechanisms would explain why passive exposure to familiar melodies consistently elicits complementary motor activity

(Lahav et al., 2007) and why this activity often anticipates upcoming sounds (Stephan et al., 2018).

While a range of studies has demonstrated the capacity for precise temporal alignment between auditory and motor areas, studies focusing specifically on oscillatory dynamics are poised to reveal critical mechanistic insights. Neuronal oscillations function as intrinsic "clocks" that organize neural activity across various timescales, providing a powerful framework for understanding temporal processes in the brain (Buzsáki & Draguhn, 2004). For example, slow oscillations are known to encode the rhythmic structure found in music (Nozaradan et al., 2012), and beta rhythms naturally couple with auditory beats to facilitate predictions in time (Arnal, 2012; Fujioka et al., 2015). Importantly, oscillations also support functional connectivity between brain regions, ensuring that this communication occurs promptly (Fries, 2005). For instance, compelling evidence of reciprocal functional connections between auditory and motor cortices reveals that information flows from auditory to motor regions via slow oscillations, while top-down influences from motor areas are conveyed through beta-band coupling (Morillon & Baillet, 2017). Yet, despite significant advancements since its inception at the turn of the century, the oscillatory framework remains relatively recent, with much still to be uncovered about the role of oscillations in auditory-motor coupling.

Within this context, the current investigation aims to advance our understanding of auditory-motor coupling by addressing four key questions and gaps in the literature that can only be explored by targeting the spectrotemporal dimensional within an oscillatory framework. First, it seeks to clarify the degree to which common patterns of auditory-motor activity reported in prior fMRI, EEG, and TMS studies are task-dependent or instead reflect intrinsic phase-based connectivity between auditory and motor regions, particularly as compared to visual and motor regions. Building on this, this thesis aims to pinpoint the similarities and differences between the largely underexplored auditory-motor resting state and the phase-based connectivity

patterns typically associated with active auditory-motor engagement. Another crucial question addressed in this thesis is whether auditory-motor co-activity observed in fMRI activation-based and stimulation studies anticipates or follows auditory events across frequency bands, which involves determining whether the nature of this motor activity is predictive or simply reactive. Finally, this thesis examines how the presence or absence of concurrent movements induces time-based fluctuations in the oscillatory processing of sounds associated with a learned motor program. By addressing these questions, this investigation aims to bring novel insights to a system that is uniquely adapted for processing temporal information.

1.2 Background

1.2.1 Auditory-Motor Coupling Models

1.2.1.1 Conceptual Models

There are several theoretical models aiming to explain the properties of auditory-motor coupling outlined above. At a higher level, several theories stem from the grounded cognition framework (Barsalou, 2008), which encompasses and serves as a foundation for the embodied cognition framework (Wilson, 2002). Grounded cognition proposes that abstract and semantic representations are inseparably linked to perceptual and motor representations. Central to this framework is simulation—the process by which the brain partially recreates sensory and motor states associated with past experience to operate with related constructs. Mental imagery is perhaps the most well-understood type of simulation (Kosslyn, 1994) because it involves the deliberate creation of conscious representations in working memory. However, grounded cognition also accounts for simulations that occur unconsciously, as sensory states are thought to be automatically stored in memory during perception and later triggered by similar stimuli. Crucially, this reactivation would allow for perceptual predictions that extend beyond the immediate sensory input along various dimensions, including time. Perception would also be capable of

activating simulations of actions, an idea supported by findings that hearing a word activates the articulatory actions associated with producing it (Pulvermüller et al., 2006), or that pianists' ability to identify auditory recordings of their playing depends on simulating the motor commands underlying it (Repp & Knoblich, 2004).

One important offshoot of the grounded cognition framework is the common coding theory (Hommel, 2015; Prinz, 1990), which is based on the tenet of shared neural codes for perception and action. Specifically, the planning or execution of an action and the perception of its sensory consequences are said to be represented similarly in the brain (Maes et al., 2014). One key corollary is the existence of an internal model linking the body with the external environment, consisting of interacting inverse and forward components (Wolpert et al., 1995). In the context of auditory-motor coupling, inverse models would account for information flowing from audition to action, allowing the system to infer the motor commands that produced a perceived sound. Conversely, forward models would predict the sounds likely to result from planned or executed actions, enabling the system to anticipate auditory feedback based on motor output. While the embodied cognition framework traditionally only used inverse models to explain motor activity on perception—thus attributing a passive role to the motor system—it is presently accepted that available sensory predictions of planned actions are projected onto the internal representation of the auditory stimulus (Halász & Cunnington, 2012). This suggests that one of the roles of the motor system must be to generate forward models capable of streamlining the temporal alignment of actions and sounds, an idea that is central to this thesis.

Another descendant of grounded cognition is the dynamic attending theory (DAT) by Large & Jones (1999), an oscillatory model of how attention operates by aligning itself with the temporal structure of external stimuli. This model incorporates important elements from neural resonance theory (Large, 2000) to explain how oscillatory expectancies occur and develop in real time (Large, 2008). Namely, DAT posits an "attending rhythm", a self-sustaining periodic oscillator in the brain that dynamically

locks onto "external rhythms" of the stimulus at multiple timescales. The external rhythm need not be strictly periodic and is instead seen as a sequence of temporally localized onsets. While speech and music often do contain regular rhythms, DAT allows and expects variations, treating deviations as informative. On the other hand, the self-sustaining nature of the attending rhythm implies that it will naturally settle into a period, with its phase relative to the external rhythm encoding temporal expectations and their violations. Specifically, a phase difference of zero is expected, and any deviations will indicate that the internal oscillator must adjust. Thus, the theory rests on the notion of entrainment, where the external rhythm acts as an attractor for the internal rhythm that drives it to re-synchronize, and the internal oscillator must, in turn, be able to modulate its phase and period to correct for phase and rate drifts, respectively. Crucially, findings that passive listening to musical beats strongly engages the motor system (Kung et al., 2013) and that temporal predictions depend on basal ganglia and cerebellum activity (Grahn, 2012; Merchant & Honing, 2014) suggest that some internal oscillators must inhabit the motor system.

Active Sensing is a more modern theory that builds on key premises from DAT (Schroeder & Lakatos, 2009) and explicitly incorporates motor involvement in sensory processing, proposing that perception depends on motor-driven sampling routines that guide and constrain sensory input (Schroeder et al., 2010). Classic examples include bat echolocation and human visual scene processing, which involves systematic eye movements and fixations (MacEvoy et al., 2008). Like in DAT, active sensing involves two interacting rhythms, though unlike in DAT, both are thought to be internally generated: a bottom-up driving rhythm located in primary sensory areas that encodes stimulus properties, and a top-down modulatory rhythm originating in motor areas that resets the phase of the driving rhythm to amplify or suppress sensory input. In other words, motor cortical inputs are thought to shift neural excitability cycles in A1 to modulate sensory gain and enhance specific features. Importantly, this theory also applies to contexts where motor activity is covert, thereby influencing perception without the need for sampling movements to

be effectuated. Thus, covert oscillatory influences from motor cortices could still be entraining rhythmic activity in A1 in passive listening contexts. While findings like minimal comprehension impairment after M1 disruption challenge this notion (Galantucci et al., 2006), others propose that top-down motor influences could still be reaching A1 via a relay in S2's lateral sulcus (Morillon et al., 2015). Thus, the picture remains complex, with many aspects still unclear due to the recency of these findings. Nevertheless, the present thesis will shed light on the role of covert motor modulations on passive listening.

1.2.1.2 Implementation-Level Models

In addition to the more theoretical models reviewed thus far, several models address how auditory-motor coupling might be implemented in the brain. One example is the dorsal stream framework, which, as in the Hickok & Poeppel (2003) and Rauschecker & Scott (2009) models for speech, adapts the ventral-dorsal dichotomy from visuomotor studies to auditory processing. As such, this framework hinges on the notion that cortical systems for different sensory modalities share functional organizational principles (Rauschecker, 1998). The original proposal by Rauschecker & Tian (2000) mirrored visuomotor accounts by positing parallel anatomical pathways dedicated to processing two distinct aspects of sounds: a ventral stream responsible for encoding "what" information (pitch identification), and a dorsal stream in charge of integrating "where" information (sound localization). However, insufficient evidence of dedicated areas in A1 for spatial information processing (Belin & Zatorre, 2000), along with evidence of dorsal stream engagement during sound identification (Belin et al., 2000), resulted in revisions to the theory. One revision claimed that the dorsal pathway is involved in perceiving the time course of the signals emitted by auditory objects, effectively substituting the "where" for "when", or "where in frequency", as the authors put it (Belin & Zatorre, 2000). This idea was further advanced by the hypothesis that the true function of the dorsal auditory stream is the preparation of timely motor responses based on incoming auditory

information (Warren et al., 2005). That is, dorsal auditory stream areas would be responsible for computing auditory-motor transformations—an idea supported by greater dorsal activity compared to passive listening during the execution of actions that reproduce auditory spatial (Zatorre et al., 2002) and vocal input (Hickok et al., 2003).

Another important model is the Action Simulation for Auditory Prediction (ASAP) hypothesis by Patel & Iversen (2007), which specifically aims to explain musical beat perception in humans. On the one hand, ASAP posits that the motor planning system uses a simulation of periodic movement patterns to entrain its neural activity patterns to the beat period (Patel & Iversen, 2014), in line with grounded cognition. This idea rests on the fact that the motor system naturally generates neural periodicities in the time range that musical beats occupy, such as the intervals between footfalls in normal gait (Styns et al., 2007). Similar to Active sensing, the internal motor simulation is allowed to either consist of an efference copy or instead exist at an abstract level, untied to any specific effector or motor imagery (Schubotz, 2007). On the other hand, ASAP proposes that simulated motor activity travels from motor planning regions to auditory regions, where they serve as a predictive signal for the timing of upcoming beats (Iversen et al., 2009). Thus, the role of the motor system would be to modify audition via temporal predictions (Patel & Iversen, 2014), unlike in Active sensing, where its goal is to modify audition via sensory gating. Importantly, Patel & Iversen (2014) champion the dorsal auditory stream as the likely candidate for supporting ASAP processes due to its capacity for temporally precise, bidirectional signalling between auditory and motor regions (Hickok & Poeppel, 2007). Importantly, given that this thesis examines the bidirectional oscillatory information flow in the dorsal auditory stream, it is well-placed to inform the dorsal stream and ASAP frameworks.

Many of the theories reviewed thus far propose that motor areas play a causal role in the temporal processing of auditory information. This shared idea draws from the

notion of predictive timing, which itself is an extension of the predictive coding framework (Friston, 2010). The core formulation of predictive coding depicts perception as a process of probabilistic, knowledge-driven inference (Clark, 2013). In this view, neural resources are optimally allocated when the brain continuously generates predictions about upcoming sensory input and updates these expectations by comparing them with actual sensory feedback (Friston, 2005). Thus, sensory processing would involve a dynamic interaction between top-down predictions and bottom-up sensory signals (Rao & Ballard, 1999). Predictive timing builds on this idea, and is defined by Arnal & Giraud (2012) as "the process by which uncertainty about when events are likely to occur is minimized to facilitate their processing and detection". In their paper, these authors outline the final model reviewed in this section—that of oscillation-based predictive coding and predictive timing processes in the brain. Because this model encompasses and interconnects key ideas from all the other models introduced thus far, it is vertebral to the present investigation. Therefore, the following paragraphs will provide an examination of this model's mechanisms and its significance to the overarching themes of this thesis.

Oscillations were originally viewed as tools enabling flexible communication between distant neuronal populations (Fries, 2005), but nowadays they are also considered instrumental in predictive processing (Engel et al., 2001). Moreover, while the brain likely generates predictions about "what" and "when" simultaneously, emerging studies suggest that the underlying oscillatory mechanisms for these predictions are different (Schroeder & Lakatos, 2009). In this context, auditory-motor predictive timing is defined as a two-step process: first, A1 adapts by matching its spike rates and oscillatory activity to the rhythmic or quasi-rhythmic aspects of the stimulus, thereby exploiting its probabilities. Then, entrained oscillations become predictive, generating periodic temporal windows that facilitate higher-order processing. This process has support in speech decoding, where the timing of regular delays for higher-order readout is more critical than the quality of the encoded sensory signal itself (Ghitza & Greenberg, 2009). Building on the DAT model, Arnal &

Giraud (2012) propose that the mechanism for implementing predictive timing at the sensory processing level involves the synchronization of cortical delta-theta oscillations to stimulus rates. Evidence includes the fact that the phase of delta-theta activity is reset before stimulus onset when the brain is capable of anticipating sensory events (Lakatos et al., 2008). This in turn accelerates stimulus detection (Stefanics et al., 2010) and either amplifies or reduces sensory responses, depending on whether attention is concurrently present (Lakatos et al., 2008) or absent (Costa-Faidella et al., 2011).

More importantly, delta-theta phase resetting is also thought to be modulated by auditory-motor interactions. Indeed, motor areas are said to use temporal priors to reset the phase of oscillations in A1, thereby aligning neuronal excitability in A1 in a manner similar to attention. This idea finds support in the Contingent Negative Variation, a slow EEG component produced in motor regions during predictive timing (Praamstra et al., 2006). When motor action is additionally paired with perception, response reduction is thought to specifically rely on efference copies that propagate from motor to auditory cortices (Houde et al., 2002), much like in Active sensing. For instance, during speech production efference copies are believed to suppress auditory responses in the high gamma band (Schubotz, 2007). However, in line with the ASAP hypothesis, efference copies may also support anticipation of externally generated auditory inputs, as evidenced by the fact that motor beta-band activity can track the expected timing of beats (Iversen et al., 2009), and is recruited during speech perception (Morillon et al., 2010) even when attention is directed away from the stimulus (Fujioka et al., 2012). Thus, the beta band is thought to play a key top-down modulatory role in ongoing auditory activity during predictive timing, an idea supported by the fact that the phase of delta-theta oscillations when anticipating a stimulus is coupled with beta power modulations in motor areas (Saleh et al., 2010).

That said, the precise role of the beta band in establishing an oscillatory auditory-motor link to support predictive timing remains to be fully established (Arnal &

Giraud, 2012). One related question which will be directly addressed in Study 1 of this thesis is to which degree the oscillations involved in auditory-motor connectivity are altered when the brain is at rest and not engaged in synchronizing with stimulus periodicities. Addressing this question will clarify whether top-down beta modulations from motor areas are task-dependent or whether they belong to an intrinsic mechanism of human auditory-motor coupling. Finally, Arnal & Giraud (2012) propose that predictive timing also operates by dissolving activity in the alpha band, as temporal predictions of event occurrences have been tied to alpha desynchronization at the expected onset of the stimulus (Rohenkohl & Nobre, 2011). However, most studies assessing alpha desynchronization have been conducted in the context of visuomotor tasks (Thut et al., 2006), so it is therefore less clear whether this oscillatory phenomenon also pervades auditory-motor contexts and to which degree it depends on musical training. Accordingly, these two gaps in the literature specifically relating to the alpha band will be assessed in Study 2 of this thesis.

1.2.2 Auditory-Motor Coupling Anatomy

1.2.2.1 Cortical areas and structural connectivity

The following is a cursory overview of the primary cortical areas comprising the dorsal auditory stream and their associated fibres and tracts, collectively referred to as the cortical auditory-motor loop. While hemispheric differences are plentiful within this circuit, cortical and subcortical auditory-motor networks on the whole are generally regarded as bilateral systems (Pranjić et al., 2024). Furthermore, although the somatosensory cortices and the vestibular system play a critical role in auditory-motor coupling, they unfortunately fall outside the scope of this thesis. Consequently, this section will not explore lateralization, nor the somatosensory and vestibular contributions to frameworks like the *sensorimotor theory of rhythm and beat induction* (Todd & Lee, 2015), instead focusing on primary auditory cortex (A1) and its interactions with the premotor and primary motor cortices described below.

A1 belongs to the broader auditory cortex and it is located on the superior plane of the temporal lobe, within the Sylvian fissure, near the medial and central portions of Heschl's gyrus (Da Costa et al., 2011). The larger auditory cortex is subdivided into the core, belt, and parabelt areas, which are organized hierarchically based on their distinct anatomical characteristics as well as their connectivity patterns with the thalamus and with each other (Moerel et al., 2014). As part of the core area, A1 represents the first level of cortical processing and exhibits the clearest tonotopic organization, likely due to its direct and predominant input from the medial geniculate nucleus (MGN) of the thalamus (Winer, 1992). In contrast, the belt and parabelt display weaker or absent tonotopy and are thought to integrate information across frequencies and broad temporal windows (Leaver & Rauschecker, 2016). Evidence of the central role of A1 for auditory-motor processes can be found in the simple fact that transcranial magnetic stimulation (TMS) to right A1 results in decreased resting state functional connectivity in the somatomotor network, including much of the dorsal auditory stream itself (Andoh et al., 2015). As the first node of the auditory-motor circuit, A1 performs spectrotemporal analysis of auditory information relayed from subcortical areas and transmits it to other dorsal auditory stream regions, where it is further integrated into auditory-motor programs (Zatorre, 2024).

Short-range U fibres connect A1 with a triangular surface on the supratemporal plane known as the planum temporale (PT), within posterior Superior Temporal Gyrus (STG; Eggermont, 2010). PT also receives direct input from the MGB of the thalamus, and it is demarcated by the Sylvian fissure, the insula, and the STG's lateral lip (Hickok & Saberi, 2012). The global view of PT put forth by Griffiths & Warren (2002) is that it functions as a 'computational hub' for routing sensory signals into different cortical streams. However, when it comes to PT's specific role in auditory-motor coupling, there are two perspectives worth mentioning. According to Rauschecker & Scott (2009), pSTG as a whole supports the implementation of internal models, which are computational mechanisms that simulate the input-output dynamics of the motor system for the purpose of motor error correction and predicting the

auditory consequences of motor actions. Thus, according to this perspective, the PT would be in charge of orchestrating the inverse and forward models proposed by the common coding theory reviewed above (Maes et al., 2014). The second view, put forth by Hickok & Saberi (2012), is that PT is not homogeneous and is instead composed of subfields that perform different dorsal stream operations, such as spatial or auditory-motor processes. Specifically, the authors highlight subregion Spt (Sylvian–parietal–temporal) as the auditory-motor integration nucleus within PT. Area Spt could be responsible for inverse modelling by performing coordinate transforms between auditory-based and motor-based representations (Hickok et al., 2009). However, a more interesting possibility is that Spt comprises an auditory target map that compares forward model predictions with true auditory feedback and generates an error signal in cases of mismatch (Golfinopoulos et al., 2010), in line with the predictive coding framework.

The Superior Longitudinal Fasciculus (SLF), specifically its temporoparietal tract (SLF-tp), connects pSTG to the Posterior Parietal Cortex or PPC (Gierhan, 2013; Patel & Iversen, 2014), where the intraparietal sulcus (IPS) resides. IPS is situated in the middle of a functional gradient spanning spatial functions in the Superior Parietal Lobule (SPL) to nonspatial functions in the Inferior Parietal Lobule (IPL; Husain & Nachev, 2007). Moreover, IPS is a multisensory region that receives converging anatomical inputs from visual, auditory, and tactile sensory cortices (Lewis & Van Essen, 2000). That said, IPS was originally tied to higher-order visual processing, particularly for spatial operations such as visual mental rotation (Zacks, 2008). Recently, however, it is increasingly viewed as a region that is critically involved in transforming sensory representations across intramodal frames of reference (Zatorre, 2024). Notably, in the context of auditory-motor coupling, activity in this area has been causally linked to the ability to manipulate musical imagery sequences in auditory working memory (Albouy et al., 2017). Moreover, fMRI activity related to melody transposition and melody reversal within the IPS overlaps at the subject level, which indicates a general involvement of IPS in mental transformations across pitch

and time, respectively (Foster et al., 2013). Furthermore, proficiency in melody transposition is associated with greater cortical thickness and grey matter concentration in IPS (Foster & Zatorre, 2010), and this region's activity covaries with transposition difficulty (key distance) but not absolute pitch distance (Foster et al., 2013), which suggests that IPS is capable of applying auditory imagery transformations across higher-order dimensions.

PPC is further connected to premotor cortex (PMC), the first cortical motor region in the dorsal auditory stream and one of the key regions in this thesis. PMC is located in the precentral gyrus of the frontal lobe and can be subdivided into dorsal (dPMC) and ventral (vPMC) aspects, which are respectively located above and below z = 51 of Talairach space (Rizzolatti et al., 2002). Specifically, SLF II links SPL to dPMC, whereas SLF III connects IPL to vPMC and potentially caudal dPMC. In addition, the arcuate fasciculus (AF) directly connects vPMC and caudal dPMC to pSTG, including regions near A1 and PT (Petrides & Pandya, 1988). Notably, these two subdivisions of PMC are the only cortical motor areas that are directly connected with pSTG via the AF, in addition to being indirectly connected to it via the SLF, which is indicative of their central role in auditory-motor coupling (Chen et al., 2009). As such, this thesis will focus extensively on the ventrodorsal aspects of PMC, but there are at least two other important subdivisions of PMC worth noting in passing.

The first involves the latero-medial axis and its potential role in demarcating the 'frontomedial wall' of action initiation and motivation (Seitz et al., 2000). Namely, research appears to converge on the idea that the medial aspect of PMC is more heavily involved in internally guided movement, while the lateral aspect predominates in externally cued movement (Schubotz & Von Cramon, 2003). The second subdivision worth discussing involves the rostrocaudal plane, which is well established for dPMC but less so for vPMC (Picard & Strick, 2001). Specifically, the rostrocaudal gradient in dPMC has been implicated in transitions from complex to simple execution (Schubotz & Von Cramon, 2003), from action intention to execution

(Simon et al., 2002), and from early to late auditory-motor learning stages (Iacoboni et al., 1998).

Both premotor cortices are known to exhibit somatotopy in monkeys, with vPMC containing forelimb and orofacial representations and caudal dPMC representing hindlimbs superior and medial to the forelimbs (Godschalk et al., 1995). In addition, monkey PMC contains a representation of postures, exhibiting a dorsoventral gradient from legs to feet, arms to hand, and face to mouth (Graziano et al., 2002a). As for human PMC, evidence of body maps is substantial and growing (Schubotz & Von Cramon, 2003). For instance, a recent study by Schellekens and colleagues (2022) combined high field fMRI, graph theory, and a novel non-rigid population response field model to identify neuronal populations across motor cortices that selectively respond to 18 different body parts, uncovering separate body maps in vPMC and dPMC, among others. That said, there are also known overlapping hand and digit representations across the ventrodorsal continuum in humans (Dum & Strick, 2005), which has led some to question the usefulness of a somatotopic model altogether (Graziano et al., 2002b).

At any rate, the distinctions between the two subregions are more revealing in the context of representational maps, where some have proposed that vPMC uses an egocentric coordinate system based on one's body, while dPMC would use a more allocentric or visuospatial reference frame (Rijntjes et al., 1999). Indeed, dPMC has been suggested as the most likely area where movement parameters from the non-egocentric reference frame are transformed into limb-centred movement descriptions (Wise et al., 1997), thereby mirroring PT subarea Spt's aforementioned putative role in performing coordinate transforms of sounds to movements (Hickok et al., 2009). In terms of functionality, PMC as a whole is famously involved in planning and organizing movements and actions (Daroff & Aminoff, 2014). However, dPMC is considered to specifically code 'supramodal' action plans or sequences,

while vPMC is thought to code 'surface properties' of those behaviours (Schubotz & Von Cramon, 2003). This distinction will be crucial to the contents of this thesis.

Both vPMC and dPMC have reciprocal connections with the Supplementary Motor Area (SMA; Dum & Strick, 2005), located in the medial aspect of the brain—specifically in the dorsomedial frontal cortex, which sits above the cingulate sulcus and is itself part of superior frontal gyrus (Picard & Strick, 2001). However, anatomical and functional studies suggest that the SMA is comprised of two distinct areas: a more rostral portion, known as pre-SMA, and a more caudal portion known as SMA proper (Nachev et al., 2008). The border between pre-SMA and SMA is considered to be the vertical plane through the anterior commissure (Johansen-Berg et al., 2004), and there are some noteworthy anatomical differences between these two areas. First, the SMA contains a somatotopic body map along an anterior-to-posterior (face-to-legs) gradient (Chainay et al., 2004), whereas motor responses in pre-SMA can only be evoked using high current intensities, and the relationship between stimulation site and movement is more variable (Picard & Strick, 1996). Second, while SMA exhibits connections to the spinal cord and other motor cortices, the pre-SMA mainly connects to prefrontal, anterior premotor, and cingulate areas (Johansen-Berg et al., 2004).

Both pre-SMA and SMA are involved in auditory perceptual and imagery tasks across a wide range of sounds, including speech, nonverbal vocalizations, and music (Lima et al., 2016). Importantly, there is reported evidence of overlapping activity in pre-SMA during perception and action that is detectable even at the subject level (Gazzola & Keysers, 2009). Moreover, pre-SMA and SMA responses occur in tasks without overt motor components, as well as tasks involving the mental generation of speech and music sounds (Lima et al., 2016). Thus, these areas appear crucial for motor imagery and likely rely on mechanisms involved in overt motor execution (Sharma & Baron, 2013). In fact, pre-SMA and SMA are both recruited more strongly for imagery than for passive listening (Halpern et al., 2004), suggesting that motor

programs in these regions support broader mechanisms than other motor cortices, notably those tied to the ASAP hypothesis (Cannon & Patel, 2021). As mentioned, motor information can modulate sensory processing via efference copies, and the SMA has been pinpointed as the likely source of this modulation (Reznik et al., 2015). In terms of the differences between the two, pre-SMA is more associated with planning voluntary movements and processing their sensory consequences (Jo et al., 2014), while SMA proper is more tied to movement generation and control (Nachev et al., 2008). That said, responses in SMC proper are quite complex and not directly related to movement execution, which has led to the view that SMA could be in charge of planning movement sequences (Tanji, 2001), as has also been proposed for dPMC (Janata & Grafton, 2003).

SMA proper, vPMC, and caudal dPMC are all connected via short-range U fibres to primary motor cortex (M1; Zatorre et al., 2007), located in the anterior bank of the central sulcus and on the adjacent caudal portion of the precentral gyrus (Dum & Strick, 2004). Historically, M1 was delineated based on the fact that movement thresholds under electrical stimulation are lower in this region than in the rest of the neocortex (Penfield & Boldrey, 1937). Cytoarchitectonically, M1 corresponds to area 4, which is identified by the presence of giant pyramidal cells famously ascribed to cortical layer V (Brodmann, 1909). As part of broader motor cortex, M1, premotor cortices and SMA proper all contribute fibres to the corticospinal tract (CST), which transmits motor commands to the spinal cord for voluntary movement (Seo & Jang, 2013). However, M1 exhibits a much more detailed somatotopic map than the other motor cortical areas (Dum & Strick, 2004), and activity in M1 is considered to be more closely related to concrete aspects of movement than activity in premotor and SMA cortices (Dum & Strick, 2005). For example, dPMC retains separate representations for pitch and timing from upstream dorsal auditory inputs, but in M1 these representations are combined into a single code amenable to action execution (Kornysheva & Diedrichsen, 2014).

That said, M1 neurons fire 5-100 milliseconds before movement onset and do not tend to control individual muscles directly, instead coordinating movements or sequences of movements that implicate multiple muscle groups (Knierim, 2016). Specifically, M1 is responsible for signalling parameters of movement such as force (Evarts, 1968), direction (Georgopoulos et al., 1982) and speed (Moran & Schwartz, 1999), among others. Thus, it is alpha motor neurons in the spinal cord, which are themselves one level of abstraction lower from M1, that finally close out the cortical auditory-motor loop by encoding the force of contraction needed to engage specific groups of muscle fibres (Farina & Negro, 2015).

1.2.2.2 Subcortical areas, medial areas, and structural connectivity

The following are the main subcortical and medial areas that compose the dorsal auditory stream and several of their anatomic connectivity fibres and tracts. Because this thesis focuses exclusively on areas and interactions at the cortical level, special attention will be given to the tracts and bundles interconnecting subcortical and cortical structures.

The first subcortical area of import is the cerebellum, which is located in the hindbrain, posterior to the occipital lobes, and dorsally with respect to the brainstem (Voogd et al., 2013). This structure consists of two large hemispheres, and each hemisphere is divisible into three lobes along the rostrocaudal plane: the anterior, posterior, and much smaller flocculonodular lobe (Rahimi-Balaei et al., 2023). The cerebellum also consists of at least four nuclei, organized into two distinct aspects: the caudomedial group consisting of the fastigial and globose nuclei, and the rostrolateral group consisting of the emboliform and dentate nucleus (Voogd et al., 2013). The fastigial nucleus projects to the vestibular and reticular nuclei in the brainstem and is accordingly involved in balance and eye movements. The globose and emboliform on the one hand, as well as the dentate nucleus of the posterior lobe on the other, project to the red nucleus of the midbrain and send excitatory input upwards along the cerebellothalamic tract to the ventrolateral nucleus (VL) of the

thalamus, where information is relayed to motor cortices at large (Rahimi-Balaei et al., 2023). Among these, the dorsal aspect of the dentate nucleus is the most relevant to the dorsal auditory stream, as it is involved in fine motor coordination and voluntary control of movement (Matano, 2001) via the modulation of motor neuron activity (de Leon & Das, 2024).

The *internal cerebellar forward model on motor planning and motor control* (Wolpert et al., 1998) proposes that M1 sends efference copies of actions to the cerebellum for the generation of motor-to-somatosensory predictions. Via a corollary discharge mechanism (Sperry, 1950), predictions are then compared with incoming sensations for the main cerebellar purpose of motor error correction (Knolle et al., 2012), in line with the common-coding idea of forward modelling (Maes et al., 2014) and the more recent *forward model theory* linking cerebellum to perception of agency (Welniarz et al., 2021). Notably, a similar mechanism has been found for motor-to-auditory integration, as self-initiated sounds will suppress auditory activity in the form of the N100 event-related potential (ERP) in controls but not patients with cerebellar lesions (Knolle et al., 2012).

Similarly, perturbations in auditory feedback consistently induce compensatory reflexes that correlate with increases in cerebellar activity in speech production (Tourville et al., 2008), and perturbations themselves correlate with increases in beta power during music production (Herrojo Ruiz et al., 2017). Moreover, listening to trained melodies—but not untrained melodies—recruits cerebellar centres (Herholz et al., 2016), and cerebellar activity has been positively and negatively related to metrical complexity and performance during beat-tapping tasks, respectively (Kung et al., 2013). Interestingly, participants with early musical training show reduced white matter density in both cerebellar hemispheres and in specific cerebellar regions, as compared to late starters. Specifically, enhanced timing abilities have been tied to reduced sizes in right lobule VI of the posterior lobe, suggesting that smaller volumes in musicians reflect streamlined auditory-motor integration (Baer et

al., 2015). Thus, there is an intimate relationship between auditory processing and cerebellar structure and function, leading some to claim that the cerebellum receives information about incoming stimuli directly from auditory centres, encodes the temporal relationships between them, and sends this information upward to motor cortices (Kotz & Schwartze, 2010), thereby providing the neural substrate for the predictive timing processes discussed above.

Recent findings suggest that the dentate nucleus of the cerebellum has direct reciprocal connections with the basal ganglia (BG), specifically with its globus pallidus internus (GPi) and substantia nigra pars reticulate (SNr) subregions (Milardi et al., 2019). The BG group is deep-seated in the forebrain and surrounds the ventricles, being situated below the cortex, above the brainstem, and lateral to the thalamus (Singh, 2006). Its input nuclei or striatum consist of caudate nucleus (CN), putamen, and the nucleus accumbens (NAc), and mostly receive cortical, thalamic, and nigral inputs (Lanciego et al., 2012). The output nuclei, composed of the GPi and SNr, mostly transmit inhibitory signals to the ventral anterior (VA) and ventrolateral (VL) aspects of the thalamus (Lanciego et al., 2012). Thus, the dentate nucleus of the cerebellum and the output nuclei of the BG are not only connected but also innervate VL synergistically, respectively providing excitatory and inhibitory inputs to fine-tune movement (Hintzen et al., 2018). Finally, the globus pallidus externus (GPe), the subthalamic nucleus (STN) and the substantia nigra pars compacta (SNc) collectively make up the intrinsic nuclei, which mediate the input and output nuclei (Lanciego et al., 2012). In addition, there are two non-thalamic white matter structures worth mentioning: the striatopallidal pathway, which connects key region putamen to other BG structures and is essential for regulating motor control and action selection (He et al., 2020), and the corticostriatal tract, which connects striatum to A1, PMC, and PPC and is thought to encode/decode temporal structure (Kotz et al., 2016).

Notably, the SMA cortices appear to have a unique relationship with BG, as all of their subregions connect to at least one BG component, with pre-SMA connecting to more

anterior sites of the striatum than SMA (Nachev et al., 2008). Indeed, the number of cells that project from GPi to both the SMA and the pre-SMA via the thalamus is between 3 to 4 times the number that project from the cerebellum, unlike any of the other motor cortical areas (Akkal et al., 2007). In addition, both the SMA and the pre-SMA have a 'hyperdirect' connection to the STN of the intrinsic nuclei (Nambu et al., 2002), which in turn projects into cerebellum via the pontine nucleus (Milardi et al., 2019). Overall, these connections between BG and the SMA cortices are said to belong to the 'direct pathway' of motor control, which facilitates movement by removing inhibition from motor-related thalamic nuclei such as VL (Rocha et al., 2023).

Their co-involvement in this direct pathway may explain why both SMA cortices and putamen have been specifically implicated in beat perception, which can be defined as the perception of time intervals relative to a beat (Merchant et al., 2015). Conversely, the perception of absolute durations of time intervals has traditionally been the province of the cerebellum (Teki et al., 2011), although recent findings indicate this region may also be implicated in relative time interval perception (Andersen & Dalal, 2021). At any rate, beat perception need not be accompanied by beat production for premotor and striatal activations to occur (Chen et al., 2008), and BG affectations such as those of Parkinson's disease are known to impair the extraction of implicit beats from the temporal structure of non-isochronous stimuli (Grahn, 2009). Findings such as these have recently culminated in a notable update to the ASAP hypothesis introduced above, according to which motor simulations are specifically implemented via the interplay between precisely patterned neural time-keeping activity in SMA and sequencing processes in dorsal striatum (Cannon & Patel, 2021).

Both cerebellum and sensorimotor striatum are connected to the cingulate cortex (CC), the former via a pontine relay (Purves et al., 2001) and the latter via the cingulostriatal projection of the aforementioned corticostriatal pathway

(Haber, 2016). The CC is a region that stretches from the paraolfactory sulcus to the marginal ramus and is located in the medial aspect of each cerebral hemisphere, inferiorly bordered by the corpus callosum and superiorly bordered by prefrontal cortex and SMA (Oane et al., 2023). This structure can be functionally divided into four main sub-cortices: anterior cingulate (aCC), midcingulate (mCC), posterior cingulate (pCC), and retrosplenial cortex. Each of these subareas has been implicated in auditory-motor coupling to varying degrees, often in overlapping ways. For instance, primate studies by the Tanji (Ninokura et al., 2004) and Petrides (Amiez et al., 2012) teams have converged on the involvement of aCC in the coding of serial order. Moreover, activity in aCC has been observed before the execution of piano errors (Maidhof et al., 2010), regardless of whether auditory feedback was presented or not (Ruiz et al., 2009), and has been specifically tied to beta enhancement around 100ms before error onset (Ruiz et al., 2011). Similarly, alteration of pitch feedback during piano performance has been shown to modulate BOLD signalling in a network including aCC, SMA, and cerebellum (Pfordresher et al., 2014). As for the mCC subregion, it is characterized by having a particularly well-defined body map that integrates motor signals (Procyk et al., 2014), for processing feedback signals during the early stages of learning (Amiez et al., 2013), and for its connections to all of parietal cortex for multisensory action monitoring (Vogt, 2016). Finally, pCC has also been associated with beta bursting before error onsets in musical performance (Herrojo Ruiz et al., 2017), like aCC, and both pCC and retrosplenial cortex are highly connected with medial and lateral parietal cortices to support actions in space (Rolls, 2019), like mCC. Thus, while the involvement of cingulate cortex in auditory-motor coupling is undeniable, more research is needed to disentangle the exact contributions of each subregion in humans.

1.2.3 The Oscillatory Dynamics of Auditory-Motor Coupling

1.2.3.1 Justifying the Study of Auditory-Motor Coupling with Music

The neuroanatomy of the dorsal auditory stream reviewed above provides a crucial neurophysiological foundation for understanding the functional influences underlying auditory-motor integration. However, such a localizationist view cannot itself capture the dynamic nature of functional auditory-motor coupling, given the sheer complexity of interwoven neural loops and the nuanced interplay between auditory inputs and outputs. To investigate these dynamics, electrophysiological measures such as event-related potentials (ERPs) have been extensively studied, yielding valuable insights over the years. However, a critical limitation of ERPs is their focus on time-locked, phase-averaged responses across frequency bands, which can obscure the rich spectral information present in neural signals. As explained below, neural activity is naturally stratified into distinct frequency bands, each associated with different functional roles (Wang, 2010). Moreover, it turns out that the non-phase-locked activity and cross-frequency interactions that ERPs fail to capture play a key role in coordinating auditory and motor systems (Hyafil et al., 2015). Thus, neural oscillations have emerged as a promising new avenue to continue exploring the mechanisms underlying auditory-motor coupling.

Animal models reveal that, during self-produced sounds, inputs from motor cortices will act to suppress synchronized, low-frequency oscillations (<10 Hz) in auditory cortex, thereby shifting the local network toward desynchronization and transitioning firing patterns from phasic to tonic (Reznik & Mukamel, 2019). This desynchronized state is characterized by attenuated evoked local field potentials (LFPs)—which represent the sum of signals from within a radius of several hundred microns (Xing et al., 2009)—but a relative abundance of high-frequency oscillations (>20 Hz). Moreover, while the net population response is reduced, the auditory processing of weak stimuli that would otherwise be suppressed is sharpened (Curto et al., 2009; Marguet & Harris, 2011; Pachitariu et al., 2015). This occurs because motor inputs

hyperpolarize and stabilize the membrane potential of auditory pyramidal cells (Schneider et al., 2014), thereby sharpening frequency tuning curves, enhancing the signal-to-noise ratio (SNR) and ultimately improving tonal representation. Interestingly, in humans, auditory modulation has also been observed following M1 stimulation and actions that simply coincide with auditory stimuli (Regev et al., 2021). For instance, Morillon et al. (2014) found that pitch discrimination improved when participants simply engaged in silent finger tapping compared to passive listening, with the best performance occurring when the taps were synchronous with the sounds. Yet, since taps did not produce any sound, these findings suggest that mere temporal alignment between motor activity and auditory stimuli is sufficient to modulate auditory responses. Similarly in mice, sounds that temporally coincide with movement but are not *triggered* by movement also evoke reduced LFPs compared to identical sounds delivered during rest (Rummell et al., 2016; Zhou et al., 2014).

Of course, this is not to suggest that intentional and non-intentional auditory-motor coupling are indistinguishable. Auditory modulation has been shown to vary depending on whether individuals attribute self-generated sounds to themselves or an external source (Desantis et al., 2012). Moreover, a study examining putaminal LFPs in monkeys during a synchronization-continuation task identified an initial beta-band burst followed by a second beta increase during the continuation phase, thereby underscoring the nuanced role of intentionality in auditory-motor integration across primates (Rummell et al., 2016). Rather, findings that mere temporal alignment of movements and sounds induces auditory attenuation at the oscillatory and biophysical level highlight the fundamental nature of auditory-motor oscillations for hearing, as well as their exceptional sensitivity to the temporal dimension. Thus, ultimately, the intimate relationship between the biophysical and oscillatory aspects of this auditory streamlining mechanism provides the rationale for studying the dynamics underlying auditory-motor coupling.

1.2.3.2 Auditory-Motor Synchronization vs Auditory-Motor Coupling

At this point, it is important to distinguish between auditory-motor synchronization and auditory-motor coupling. In this thesis, synchronization refers to the purely behavioural interaction between movement and external stimuli, whereas coupling occurs in two key scenarios: (a) when two or more brain systems synchronize their rhythms through interaction, or (b) when one or several brain systems synchronize to a periodic stimulus. Study 1 will focus on the first case, whereas Study 2 will focus on the second case.

This being said, auditory-motor synchronization is still highly pertinent to this thesis and therefore merits some comment. For example, we note that synchronization of movements with a rhythmic stimulus involves complex timing processes and is a near-universal human trait (Honing, 2012) rarely observed in non-human species (Patel et al., 2009). In fact, humans typically achieve stable synchronization within just two to three stimulus repetitions and can sustain it over extended sequences (Thaut et al., 1998). Interestingly, behavioural studies show that motor responses to isochronous pulses often precede stimulus onset (Aschersleben, 2002), a phenomenon known as *negative mean asynchrony,* thereby providing the first clue on the role of predictive mechanisms in auditory-motor interactions. Notably, this anticipation tendency is generally absent during synchronization with more rhythmically or spectrotemporally complex stimuli (Patel et al., 2005), an idea that will be revisited toward the end of this thesis.

1.2.3.3 Auditory-Motor Frequency Bands

The most salient feature of neural oscillations is their spectral richness—that is, the fact that they can be grouped into distinct frequency bands, each associated with various functions (Buzsáki & Draguhn, 2004). The oscillations involved in auditory-motor coupling are no exception to this rule, and so it is vital to underline the basic contribution of each stratification. First, low-frequency neuronal oscillations,

particularly in the delta and theta bands, are believed to play a leading role in temporal predictions (Morillon et al., 2015) by providing a temporal window of integration through which auditory information is processed across long-range neural distances.

More relevant to this thesis, however, is the fact that alpha waves recorded over M1 are often phase-locked to periodic auditory stimuli (Sabate et al., 2011). Similarly, during auditory-cued tapping, significant alpha-band coherence has been observed across a widespread network involved in auditory-motor processes, including motor and premotor cortices, PPC, auditory cortex, thalamus, and cerebellum (Pollok et al., 2005). Furthermore, using a classical auditory oddball paradigm, Haig and Gordon (1998) demonstrated that participants' reaction times were influenced by pre-stimulus alpha phase coupling, aligning with earlier findings that alpha waves are associated with enhanced cortical information processing and working memory (Başar et al., 1997; Palva et al., 2005).

However, it could be argued that oscillations in the beta band play an even more critical role in sensory-motor integration at large. For example, Nijhuis et al. (2021) demonstrated that cortico-muscular synchronization occurs in the beta band during synchronization with both auditory and visual stimuli, as well as during imagined synchronization with visual stimuli. The study also found lower beta amplitudes during imagined movements compared to active synchronization, suggesting that imagined movements engage similar neural mechanisms but with reduced activation, an idea that is highly pertinent to Study 2 of this thesis. Furthermore, stability in beta oscillatory activity has been associated with more precise motor timing (Rosso et al., 2021), and passive listening to isochronous sounds has been shown to modulate beta-band power and phase coherence across cortico-subcortical areas, including M1, vPMC, SMA, and cerebellum (Fujioka et al., 2012). Crucially, beta oscillations play a key role in top-down modulation of sensory processing (Caras & Sanes, 2017), with coupled beta and delta waves reflecting

sensitivity to sensory cues in motor contexts (Saleh et al., 2010). Indeed, beta waves are prominent during movement preparation, diminish at movement onset (Jasper & Penfield, 1949), and increase when a movement is withheld (Pfurtscheller, 1981).

Beta oscillations can also entrain to auditory rhythms, like alpha oscillations, as seen in a study by Crasta et al. (2018), who demonstrated that listening to auditory rhythms before moving to them enhances entrainment and improves motor control in beta and gamma. Interestingly, Stegemöller and colleagues (2018) investigated task complexity and found that beta-band power increased during slow tapping rates while alpha-band power dominated during fast tapping rates. These authors additionally observed more power in both bands for music versus tones, and more power within the beta band in musicians (Stegemöller et al., 2018). Notably, this last set of findings suggests that the alpha and beta bands may in fact play equally important and complementary roles in auditory-motor coupling, an idea that aligns well with the results of this thesis.

Finally, gamma rhythms are associated with the processing of basic auditory information (Gurtubay et al., 2004; Steinschneider et al., 2008) and the selection of salient stimuli. These fast oscillations are particularly important for processing phonemes and rapid fluctuations in auditory signals, enabling the precise temporal encoding required for speech and sound discrimination. In the motor domain, fast gamma oscillations in the motor cortex are engaged during the active phase of voluntary motor actions (Cheyne et al., 2008). However, their variability relative to ongoing movements suggests gamma waves may not correspond directly to specific aspects of motor commands (Donoghue et al., 1998). Instead, these rhythms are hypothesized to serve a more global function, binding the activity of remote neuronal populations (Fries, 2009; Womelsdorf et al., 2007). This integrative role may support the formation and maintenance of motor plans (Donner et al., 2009; Pesaran et al., 2002) and could account for observed alterations in gamma-band power during tasks requiring competing motor responses (Gaetz et al., 2013).

1.2.3.4 How Auditory-Motor Oscillations Are Modelled

From a dynamical systems point of view, oscillations are defined as periodic temporal changes in a system's *state* parameters that lead to *stable states* in the brain's neural dynamics. These temporal changes in state naturally occur at both the cellular and neural levels and collectively play a critical role in various timing functions, such as the central representation of event timing (Large et al., 2015) and time perception (Gupta, 2014). Importantly, all brain oscillations are non-linear, and have been extensively modelled in mathematical and physical sciences for more than 50 years. Indeed, a wide array of mathematical models exists for describing oscillations, and the choice of model depends on the available data and desired predictions. Moreover, under certain assumptions, all oscillator models share a set of universal predictions and properties that are independent of particularities, rendering the oscillatory framework especially appealing for investigating behaviour and cognition. In addition, despite the existence of distinct principles governing brain structures (neurophysiological) and motor effector structures (biomechanical), both can be modelled together as coupled dynamical systems, providing a unifying framework for understanding their interaction (Large, 2008; Large et al., 2015).

1.2.3.5 Single-Neuron Oscillations

Brain oscillations are typically modelled across three levels of analysis, ranging from more microscopic and detailed to more abstract and global approaches. The first tier is the biophysical level, under which oscillations are described in terms of Hodgkin-Huxley and other similar equations that describe the electrical properties of single neurons. Indeed, individual neurons exhibit distinct frequency preferences due to their combination of low- and high-pass filtering properties, which technically qualifies (some of) them as resonant systems (Hutcheon & Yarom, 2000; Pike et al., 2000). The biophysical level has been instrumental in animal studies, advancing our understanding of neural pattern generation (Marder, 2000) and neural responses to external sound stimuli (Large & Crawford, 2002).

However, it would be erroneous to view neurons as mere passive integrators that only relay upstream information to other cells via action potentials. The diversity of ion channels across neuronal membranes, their regulated expression, and their opening kinetics endow neurons with sophisticated integration capabilities beyond simple activation thresholds and bistability (Damm et al., 2020). Accordingly, the neuron doctrine emphasizes the computational abilities of neurons by modelling them as bistable gates capable of processing information (Llinás, 1988). Moreover, single-neuron activity encompasses a wide range of patterns, including transient and sustained firing, subthreshold oscillations, and clock-like periodic firing. Thus, neuronal excitability exhibits a broad repertoire of possibilities (Connors & Regehr, 1996), with oscillations being just one example that emerges when spike timing becomes periodic (Kaneoke & Vitek, 1996; König, 1994).

“Oscillation” is therefore a key characteristic of neuronal activity, whether at the level of a single cell or across networks of multiple neurons. At the network level, oscillations reflect the activity of interconnected neurons, as exemplified by cortical organization, where six layers of diverse neuron types contribute to a complex interplay between elementary neuronal processes and cortical rhythms (Damm et al., 2020). However, even oscillations recorded using LFPs cannot be directly inferred from the stochastic spike trains of individual neurons (Jarvis & Mitra, 2001; Mureşan et al., 2008), which poses a challenge to understanding how "spike-to-spike" synchrony contributes to emergent cortical rhythms. That said, positive correlations have indeed been reported, such as in a study where primates viewed naturalistic scenes, where spikes and high-gamma LFPs were proven to originate from the same network (Ravignani et al., 2014). There is also fMRI evidence pointing to convergence of neuronal firing, field potentials, and fMRI activity in human auditory cortex (Mukamel et al., 2005). Thus, the growing accumulation of knowledge about brain activity and electrophysiology suggests a conjunct role of single neuron dynamics and the diversity of their connections in shaping network oscillations (Wang, 2010),

but further work is still needed to establish exactly how individual neuron properties give rise to large-scale network dynamics.

1.2.3.6 Population-Wide Oscillations

Large-scale network dynamics such as those explored in this thesis belong to the second tier, known as the oscillator level, under which mathematical simplifications are used to study population-level rhythms. At this level, rhythms observed in LFP and M/EEG recordings are defined as cyclical fluctuations in baseline neuronal activity that occur across the brain and exhibit 1/f frequency spectra with characteristic peaks across the frequency bands explained above. As alluded to at the start of this section, reciprocal excitatory and inhibitory connections form a widespread pattern in the brain and spinal cord, where action potentials from excitatory neurons activate inhibitory neurons, which, in turn, suppress the excitatory neurons (Hoppensteadt & Izhikevich, 1996). These reciprocal connections give rise to periodic activity that can be analysed using phase-response curves (PRC), which plot the normalized spike time shift against the perturbation time (Achuthan et al., 2011).

By understanding the PRCs of coupled neurons, important predictions about their phase differences have been made and formalized (Stiefel & Ermentrout, 2016). For instance, in a pair of identical coupled neurons, when the PRC shows a positive slope at zero crossing, the phase difference between the two neurons will progressively decrease to zero. This self-correcting mechanism forms the basis of population-wide oscillatory coupling, with deviations from synchrony being adjusted as the discharge of one neuron phase shifts the spikes of the other. This functional unit of periodic activity is formally referred to as a *neural oscillator*, and models such as the Wilson–Cowan model (Wilson & Cowan, 1973) have been formulated to explain how the interactions between excitatory and inhibitory subpopulations systematically converge and give rise to the characteristic behavioural states associated with said neural oscillators (Large, 2008; Large et al., 2015).

1.2.3.7 Canonical Formulations and Predictions

This leads us to the last tier of analysis, known as the canonical level, which involves simplified representations of oscillator-level models under specific assumptions about parameter values, achieved through normal-form analysis (Wiggins, 2003). As the name implies, canonical models are particularly valuable because they retain the essential dynamics of the original oscillator model while eliminating unnecessary complexities, thus serving as universal representations that capture the shared properties of all neural oscillator models. As a result, canonical models have been used to formulate the following generalized predictions about the behaviour and interaction of neural oscillators that apply to all exemplars, independent of context and specifics. However, while the following paragraphs provide a primary description of these predictions, the studies presented in this thesis will not directly assess these elements. Instead, the purpose of the next few subsections is to introduce the basic tenets of *neural resonance theory* (Large & Snyder, 2009), an important offshoot of dynamical systems theory (Izhikevich, 2006) that provides a particularly powerful framework for conceptualizing neural oscillations.

The first prediction of neural resonance theory is *spontaneous oscillatio*n, an endogenous periodicity that can arise or persist independently of external rhythmic stimuli. In the context of dynamical systems, spontaneous oscillation represents a *stable state* in which the system alternates between periods of rest and rhythmic activity, governed by intrinsic properties more so than external inputs (Large, 2008). Palmer and Krumhansl (1990) provided evidence of spontaneous oscillation in a perceptual task where participants evaluated the goodness-of-fit of auditory events presented within imagined metrical contexts. Specifically, participants were asked to imagine a repeating metrical structure and then judge how well stimuli aligned with said structure. Subjects consistently rated events as better fitting when they coincided with points of higher salience in the imagined metrical cycle, even in the

absence of an external rhythmic stimulus, suggesting that the brain may generate intrinsic oscillatory patterns to maintain internalized timing systems.

A key corollary of the spontaneous oscillation prediction is the existence of biologically preferred periods or eigenfrequencies (McAuley & Kidd, 1995), which are intrinsic time constants of neural systems that have also been corroborated by behavioural studies showing consistent periodic responses to rhythmic stimuli (Fraisse, 1978; McAuley et al., 2006; Parncutt, 1994). The capacity for spontaneous oscillation thus provides the foundation for endogenous periodicity, enabling the brain to anticipate, sustain, and integrate rhythmic patterns even in the absence of external cues. More importantly, this self-sustained oscillatory activity forms the basis for more complex processes, such as the ones assessed in this thesis.

The second prediction of neural resonance theory is *entrainment*, defined as the process through which two oscillators are drawn toward coupling via their interactions. Entrainment depends on two aspects of periodic signals: rate, or frequency of oscillation, and phase, defined as the instantaneous position within a given period. However, the word entrainment encompasses not only perfect period and phase synchronization but also the tendency toward that state, with perfect rate and phase synchrony being just one specific case often referred to as in-phase synchronization (Trost et al., 2017). Thus, studying auditory-motor coupling can be conceptualized as characterizing the frequency and phase anchoring between two or more oscillators, whereas auditory-motor synchronization would be defined by the phase and frequency matching between external auditory stimuli and movement. In both cases, frequency matching establishes appropriate time intervals while phase matching ensures timing accuracy. Following this interpretation, Study 1 of this thesis investigates intrinsic entrainment processes between auditory and motor regions of the brain via phase-to-phase functional connectivity metrics. Conversely, Study 2 assesses broadly time-locked desynchronization in the alpha band in response to

stimulus periodicities, and therefore does not examine *brain-stimulus* entrainment as is understood in the context of delta and theta synchronization to stimulus rates.

Due to the nonlinear nature of brain oscillations, neural resonance theory notably predicts a more general form of coupling, described as *mode-locked responses* of a neuron or neural circuit to exogenous or endogenous rhythms (Large et al., 2015). Mode-locking generalizes phase-locking by allowing a periodic stimulus to interact with intrinsic oscillatory dynamics at different rates, resulting in *k* cycles of the oscillation locking to *m* cycles of the stimulus, where *k* and *m* are integers. This phenomenon predicts neural responses at harmonics (2:1, 3:1), subharmonics (1:2, 1:3), integer ratios (3:2, 4:3), and other combinations of the frequencies present in the rhythmic stimulus. Therefore, non-linear neural entrainment specifically requires two key conditions (Large, 2008): 1) the restriction of frequency relationships between endogenous and external oscillators to harmonics, subharmonics, or integer ratios (mode-locking); and 2) phase synchronization in response to stimuli within the frequency range of neural oscillations (phase-locking).

As mentioned, Large and Palmer (Large & Palmer, 2002) demonstrated that nonlinear oscillators can successfully track tempo changes and use deviations from temporal expectations, embodied in their oscillatory activity, to discern structural interpretations such as musical phrasing and melody intended by performers. Moreover, some research lines have characterized optimal entrainment *itself* as inherently nonlinear, with increased tracking correlating with better performance up to a certain threshold, beyond which further increases in tracking are associated with poorer performance (Schmidt et al., 2023). These dynamics may also be band-specific, as increased cortical tracking in the delta band has been linked to improved speech comprehension, while increased tracking in the theta band has been associated with reduced comprehension. Technically, entrainment also extends to the muscular level, where auditory cues stabilize motor output by modulating muscular and kinematic variability across simple and complex movements For

instance, Thaut et al. (1991) observed reduced biceps and triceps co-contraction during a flexion-extension task when auditory cues matched the natural movement frequency as opposed to silence and, more recently, Yoles-Frenkel (2016) found reduced EMG and velocity profile variability during a finger-tapping task in response to auditory cues, particularly in the pre-tap period.

The third and last prediction of neural resonance theory is *higher-order resonance,* a concept intimately tied to the notion of mode-locking (Large et al., 2015). First, nonlinear oscillators exhibit filtering behaviour, responding most strongly to stimuli near their intrinsic frequency—although this property is subject to the third basic property of periodic signals besides frequency and phase: amplitude. Namely, at low stimulus amplitudes or strengths, high-frequency selectivity is maintained, but this selectivity deteriorates as stimulus amplitude increases due to nonlinear compression (Large, 2008). Crucially, however, nonlinear coupling can also give rise to oscillations at frequencies that are absent in the stimulus itself, the phenomenon properly known as *higher-order resonance*. While the strongest response usually occurs at the stimulus (carrier) frequency, oscillations also appear at harmonics, subharmonics, and more complex integer ratios of the stimulus frequency, giving rise to mode-locking (Large, 2008). Notably, this property exhibits the inverse relationship with amplitude: at low stimulus intensities, higher-order resonances are minimal, but their magnitude increases with stimulus strength. Thus, neural resonance theory predicts the emergence of metrical accents, even in the absence of a corresponding frequency in the stimulus, highlighting the capacity of neural systems to generate complex rhythmic structures and enrich temporal perception beyond the direct properties of the input (Large & Palmer, 2002).

In addition to these three predicted behaviours, neural oscillations have also been conceptualized to exhibit two fundamental albeit conflicting properties: the *maintenance tendency* and the *magnet effect*. The maintenance tendency refers to the steadiness of oscillations, reflecting the intrinsic stability and resistance to

external perturbations afforded by their self-sustaining nature. Conversely, the magnet effect describes the influence that one oscillator may exert on another bearing a different eigenfrequency to progressively draw it toward its own natural frequency. Thus, these conflicting dynamics highlight the delicate balance between preserving an oscillator's inherent rhythmic properties and allowing for adaptive synchronization in response to external or internal inputs (Damm et al., 2020).

1.2.3.8 Challenges Faced by the Oscillatory Model

Over the years, neural resonance theory and other dynamical systems models of oscillatory brain phenomena have received important criticisms. Namely, critics have argued that it is unclear whether oscillatory alignment results from 1) the summation of delayed, passive, transient evoked responses to rhythmic; 2) active, intrinsic brain oscillations that predictively align to rhythmic events; or 3) a combination of both mechanisms (Gnanateja et al., 2022).

In response, research in the auditory domain has sought to establish whether oscillations are an emergent property of the auditory system (evoked) or an inherent feature of the system (intrinsic). What is undisputed is that, when auditory stimuli are presented in a periodic pattern, ambient delta/theta oscillations entrain to the stimulus and induce fluctuations in beta and gamma rhythms, thereby reflecting an oscillatory hierarchy in the auditory cortex (Lakatos et al., 2005). Additionally, accumulating evidence does indeed suggest that this multi-level coupling may not be an epiphenomenon. For instance, an EEG study by Snyder and Large (2005) revealed *induced* beta and gamma power fluctuations coupled with periodic and metrical rhythms, revealing anticipatory responses to tones. Similarly, an MEG study by Fujioka et al. (2009) found both stimulus-driven and predictive responses in the beta and gamma bands. Further supporting these findings, Will and Berg (2007) reported delta rhythm synchronization, phasic responses in the theta band, and augmented phase synchronization throughout the beta/gamma range—all modulated by stimulus periodicity.

1.2.3.9 Better Evidence that Cortical Entrainment is Predictive

However, an important alternative explanation to the action of a neural oscillator was formalized in the *repetitive evoked response* hypothesis, which posits that a rhythmic stimulus can elicit rhythmic spikes in the brain region processing it without requiring any intrinsic oscillatory mechanisms (Shah, 2004). This proposition was originally countered by Nozaradan et al. (2011), who touted a sustained periodic EEG response tuned to the frequency of the beat when participants listened to musical samples. In a follow-up study, Fujioka et al. (2015) noted that beta-band oscillations were modulated depending on whether beats were perceived as accented, regardless of whether the accentuation was physically embedded in the stimulus or imagined. Tal et al. (2017) further investigated this phenomenon, and their results showed that neural responses in the auditory cortex were enhanced at the beat frequency and remained phase-locked to the timing of missing beats even when beats were omitted, thereby satisfying the prediction of *self-sustained oscillation* explained above. Moreover, in an MEG study involving a pitch distortion detection task, Doelling and Poeppel (2015) demonstrated that entrainment correlated with task performance and musicianship, and that cortical oscillations in non-musicians were phase-locked over a limited range of musical tempos, whereas musicians exhibited neural entrainment across the entire tested tempo range.

However appealing, it is important to realize that these accounts do not technically constitute stand-alone evidence for oscillatory entrainment (Damm et al., 2020). As an example, Novembre and Iannetti (2018) raised the possibility that these responses might simply be driven by auditory ERPs such as the biphasic negative-positive (N–P) vertex wave and the contingent negative variation (CNV), which are both modulated by non-periodic attention and expectation (Nobre & van Ede, 2018)—the CNV, in particular, being associated with timing error correction during auditory-motor synchronization (Jang et al., 2016). In response, David Poeppel's team recently set out to further investigate the distinction between neural entrainment and ERPs within

structured auditory streams. A first study highlighted the existence of a constant 4.5 Hz theta-band functional coupling rate between auditory and motor cortices that was impervious to stimulus presentation rates (Assaneo & Poeppel, 2018). Then, the team hypothesized that neural oscillators, compared to evoked models, would better predict and process rhythmic stimuli. Specifically, they analysed phase lags between neural responses and speech inputs as a function of stimulus (presentation) rate, predicting larger phase lags at higher stimulus rates for reactive evoked models versus a stable phase lag profile for predictive oscillatory models. Finally, in support of the oscillator model, MEG activity in participants' auditory cortex demonstrated a quickly adapting, constrained phase regime while listening to music at varying rates (Doelling et al., 2019). Thus, these two sets of findings respectively support the idea that both *intra-brain* and *brain-stimulus* non-linear oscillations can indeed *predict* rhythmic events, thereby legitimising the dynamical systems approach and its various offshoots.

1.3 The Present Investigation

To summarize, there are two key scenarios in which auditory-motor coupling takes place: 1) when two or more brain systems synchronize their rhythms through interaction, and 2) when one or several brain systems synchronize to a periodic stimulus. Study 1, contained in Chapter 2, will focus on the first case and will seek to determine the degree to which oscillations involved in auditory-motor connectivity are phase locked when the brain is at rest and not engaged in synchronizing with external stimulus periodicities. The overall rationale is that precise phase coupling has been suggested to be one of the key features underlying the auditory-motor's system capacity for fine temporal alignment (Morillon & Baillet, 2017), particularly when compared to the visual modality where spatial information is more relevant. To address this question, Study 1 examined functional connectivity between auditory and motor regions, as compared to visual and motor regions, using resting-state MEG. This approach allowed us to look at inter-regional connectivity in different

frequency bands with adequate spatial and temporal resolution (Baillet, 2017). Connectivity was assessed by deriving phase locking values (PLV; Lachaux et al., 1999) between predefined regions of interest within auditory/visual cortices with motor and premotor cortical regions. In addition, the recently developed phase transfer entropy metric (PTE; Lobier et al., 2014), was used to derive directionally specific connectivity estimates.

Study 2, contained in Chapter 3, will focus on the second scenario of brain-stimulus coupling, specifically as it pertains to alpha, a frequency band that does not entrain to stimulus presentation rates *per se* but whose amplitude fluctuations can align to stimulus periodicities. The overall rationale is that predictive timing is increasingly thought to involve the desynchronization of activity in this band, specifically before stimulus onsets. However, since most studies have been conducted in the context of visuomotor tasks (Thut et al., 2006), it is less clear whether this oscillatory phenomenon also pervades auditory-motor contexts, and to which degree it depends on musical training. Thus, Study 2 assessed these two gaps in the literature by examining whether alpha desynchronization measured with EEG is present during passive listening to a short melody that non-musician participants had previously learned how to play. The hypotheses of Study 2 involve the mu rhythm, a sub-component of the alpha band with a spectral centroid between 9-13 Hz (Pineda, 2005; Pineda et al., 2013). Notably, the mu band is known to become asynchronous before action execution and, in more attenuated form, during passive listening to sounds of familiar actions at effector-specific locations within contralateral sensorimotor cortex (Larionova et al., 2022). Therefore, mu suppression has been proposed as an index of motor preparation in the context of learned auditory-motor associations, an idea that was put to the test in Study 2

Finally, one noteworthy methodological contribution of Study 2 is that we developed a 3D time-frequency functional localizer on the active motor training data that helped

us identify channels, frequencies, and timepoints likely to exhibit mu suppression during passive listening to the previously learned melody.

1.3.1 Study 1's Hypotheses

Study 1 tested five specific hypotheses. First, that functional connectivity estimates in the auditory-motor domain would be larger than in the visuomotor modality. Second, that phase coupling between auditory and motor areas would be more predominant in right hemisphere, in line with prior auditory-motor literature (Palomar-García et al., 2016). Third, that the degree of auditory-motor coupling would scale with musical training as observed in rs-fMRI (Palomar-García et al., 2016) and EEG (Klein et al 2016). Fourth, that auditory-motor phase coupling estimates would be larger in the beta band than in other frequency bands (Morillon & Baillet, 2017). Fifth, that directed phase-based connectivity estimates would reveal different directions of information flow for each frequency band relevant to auditory-motor coupling, in line with the current notion of frequency band-specific information flow loops (R. Wang et al., 2019).

1.3.2 Study 2's Hypotheses

Study 2 tested three specific hypotheses. First, that our functional localizer would identify mu suppression occurring in the active condition within frequency, time, and channels that correspond with theoretical delineations of the mu rhythm (Fox et al., 2016; Pineda et al., 2013). Second, that the functional localizer would detect passive mu suppression occurring at the single-note level only during post-training exposure to the learned melody, in line with findings in musicians (Wu et al., 2016). Third, that suppression would be absent in other frequency bands, making it mu-band specific.

Chapter 2: Human Auditory-Motor Networks Show Frequency-Specific Phase-Based Coupling in Resting-State MEG

Bedford, O., Noly-Gandon, A., Ara, A., Wiesman, A. I., Albouy, P., Baillet, S., Penhune, V., & Zatorre, R. J. (2025). Human auditory-motor networks show frequency-specific phase-based coupling in resting-state MEG. *Human Brain Mapping*.

2.1 Front Page

Short title: *Auditory-motor phase-based coupling*

Authors: Oscar Bedford[1,2,3], Alix Noly-Gandon[1,2,3], Alberto Ara[1,2,3], Alex I. Wiesman[1], Philippe Albouy[2,3,4], Sylvain Baillet[1], Virginia Penhune[2,3,5], Robert J. Zatorre[1,2,3]

Affiliations: **1**. Montreal Neurological Institute, McGill University, Montréal, QC, Canada. **2**. International Laboratory for Brain, Music and Sound Research (BRAMS), Montréal, QC, Canada. **3**. Centre for Research on Brain, Language and Music (CRBLM), Montréal, QC, Canada. **4**. CERVO Brain Research Centre, School of Psychology, Université Laval, Québec City, QC, Canada. **5**. Department of Psychology, Concordia University, Montréal, QC, Canada

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements: This work was funded via research grants from *the Canadian Institutes of Health Research* and from the *Natural Sciences and Engineering Research Council of Canada* to RJZ, and by the *Healthy Brains for Healthy Lives* initiative of McGill University. RJZ is supported by the *Grand Prix Scientifique* FPARD-2021-6 from the *Fondation Pour l'Audition* (Paris, France) and is a senior fellow of the *Canadian Institute of Advanced Research*.

Conflict of Interest: None.

2.2 Abstract

Perception and production of music and speech rely on auditory-motor coupling, a mechanism which has been linked to temporally precise oscillatory coupling between auditory and motor regions of the human brain, particularly in the beta frequency band. Recently, brain imaging studies using MEG have also shown that accurate auditory temporal predictions specifically depend on phase coherence between auditory and motor cortical regions. However, it is not yet clear whether this tight oscillatory phase coupling is an intrinsic feature of the auditory-motor loop, or whether it is only elicited by task demands. Further, we do not know if phase synchrony is uniquely enhanced in the auditory-motor system compared to other sensorimotor modalities, or to which degree it is amplified by musical training. In order to resolve these questions, we measured the degree of phase locking between motor regions and auditory or visual areas in musicians and non-musicians using resting-state MEG. We derived phase locking values (PLVs) and phase transfer entropy (PTE) values from 90 healthy young participants. We observed significantly higher PLVs across all auditory-motor pairings compared to all visuo-motor pairings in all frequency bands. The pairing with the highest degree of phase synchrony was right primary auditory cortex with right ventral premotor cortex, a connection which has been highlighted in previous literature on auditory-motor coupling. Additionally, we observed that auditory-motor and visuo-motor PLVs were significantly higher across all structures in the right hemisphere, and we found the highest differences between auditory and visual PLVs in the theta, alpha, and beta frequency bands. Last, we found that the theta and beta bands exhibited a preference for a motor-to-auditory PTE direction and that the alpha and gamma bands exhibited the opposite preference for an auditory-to-motor PTE direction. Taken together, these findings confirm our hypotheses that motor phase synchrony is significantly enhanced in auditory compared to visual cortical regions at rest, that these differences are highest across the theta-beta spectrum of frequencies, and that there exist alternating information flow loops across auditory-motor structures as a function of frequency. In our view,

this supports the existence of an intrinsic, time-based coupling for low-latency integration of sounds and movements which involves synchronized phasic activity between primary auditory cortex with motor and premotor cortical areas.

2.3 Key Points

1) Human auditory-motor networks show stronger phase coupling than visual-motor networks at rest, particularly in the theta, alpha and beta frequency bands.

2) Auditory-motor phase-coupling is greatest in the right hemisphere and its strength varies consistently across motor and premotor regions.

3) Motor cortices send phase information to auditory cortex via a preferential top-down direction of information flow in the beta frequency band.

2.4 Introduction

Auditory-motor coupling is the neural system responsible for the precise temporal alignment between auditory and motor areas, and a key feature of human brain function that underlies the perception and production of speech and music (Poeppel & Assaneo, 2020; Patel & Iversen, 2014). One important aspect of this system is that it lends itself to the anticipation of upcoming events via phase locking of motor outputs to stimuli presentation rates, particularly in contexts with a high degree of periodicity (Repp & Su, 2013). Moreover, auditory-motor coupling displays a remarkably high degree of temporal sensitivity compared to other forms of sensory-motor coupling (Comstock et al., 2018) in the form of low-latency alignment of movements and sounds with minimal variability (Repp, 2003), a property which is particularly critical for both speech and music (Zatorre et al., 2007). Further, the degree of auditory-motor coupling as measured by resting-state fMRI has been shown to increase with musical training and is particularly enhanced across right-hemispheric structures (Palomar-García et al., 2017).

Recently, this system has been hypothesized to rely on the oscillatory coupling between the auditory and motor areas underlying the coordinated relationship between body movements and sound patterns in the environment (Lenc et al., 2021; Assaneo et al., 2019; Iversen & Balasubramaniam, 2016; Park et al., 2016; Park et al., 2015). Indeed, work using magnetoencephalography (MEG) has shown that accurate auditory temporal predictions that are independent of movement still depend on phase coupling between auditory and motor cortices, particularly within the beta band (Morillon & Baillet, 2017). Based on this and other work, it has been proposed that oscillatory influences from motor cortex modulate activity in auditory regions during perception (Morillon et al., 2019; Merchant et al., 2015; Patel & Iversen 2014; Arnal & Giraud, 2012). However, the specific neural mechanisms that enable the fine-tuned temporal synchronization between the auditory and motor system are not well understood. Moreover, it is not yet clear whether this temporal precision is unique to the auditory system, nor whether such mechanisms are only elicited in the context of specific task demands as opposed to being intrinsic to the system.

To address these questions, we examined functional connectivity between auditory and motor regions compared to visual and motor regions using resting-state MEG (rs-MEG), a method well suited to assess task-independent oscillatory mechanisms of whole-brain dynamics whilst preserving adequate spatial and temporal resolution (Baillet, 2017). We hypothesized that the low latency of auditory-motor coupling likely depends on an intrinsic oscillatory mechanism that makes use of continuous phase alignment across auditory and motor cortices, similar to the mechanism described in the task-based study by Morillon & Baillet (2017). Furthermore, if this oscillatory mechanism is an intrinsic feature of the auditory-motor system, we reasoned that it should also be observable at rest.

Anatomically, auditory and motor areas are interconnected via the dorsal auditory stream (Rauschecker, 1998), a set of mostly reciprocal pathways that is responsible

for sound localization in space (van der Heijden et al., 2019; Zatorre et al., 2002), as well as auditory motion processing, temporal processing, and sensorimotor functions (Rauschecker, 2018). Within the dorsal auditory stream, posterior superior temporal gyrus (pSTG) is linked to cortical motor regions via the arcuate fasciculus (AF), a white matter bundle containing both long and short fibres that interconnect the frontal, parietal, and temporal lobes (Catani & De Schotten, 2008) and which has been highlighted in many studies assessing auditory-motor coupling (Kornysheva & Schubotz, 2011). In particular, the AF links the ventral and dorsal portions of premotor cortex (vPMC and dPMC) with auditory cortex through direct connections (Petrides, 2013), as well as through a relay respectively situated at the inferior and superior parietal lobules (Rauschecker, 2011; Hoshi & Tanji, 2007).

Notably, PMC is consistently activated by passive music listening to familiar melodies after motor learning (Herholz et al., 2012), as well as anticipation (Leaver et al., 2009). In particular, our lab demonstrated that vPMC is recruited when subjects listen with anticipation and tap along to rhythms, whereas dPMC is engaged during movement synchronization and is responsive to higher-order features of rhythmic stimuli, such as metrical organization (Chen et al., 2008). Thus, both subdivisions of PMC are involved in planning aspects of auditory-motor coupling before relaying the information to M1 for movement effectuation. For this reason, the current study explores the intrinsic phase coupling of vPMC, dPMC, and M1, with primary auditory cortex (A1) and control region primary visual cortex (V1), respectively (Figure 1).

The choice to compare auditory-motor coupling to visuomotor coupling was informed by the fact that they both rely on dorsal stream processing (Rauschecker, 2018). For instance, bilateral damage to the posterior parietal cortex in humans can result in impairments of both auditory and visual spatial localization (Phan et al., 2000), indicating that this region is responsible for sensorimotor integration in both modalities (Sestieri et al., 2006). Moreover, regions such as the supplementary motor area are equally activated during separate auditory and visual

beat perception tasks (Araneda et al., 2017). However, other studies comparing motor coupling to auditory versus visual beats have reported that duration perception processing within early visual cortex operates independently and distinctly from auditory timing mechanisms (Zhou et al., 2014). Indeed, it is known that visual rhythm processing depends on additional computations in V1 involving the prediction of rhythmic onsets, which has led to the notion that visuomotor coupling likely requires auxiliary mechanisms for temporal processing (Zalta et al., 2020; Comstock et al., 2018). Thus, we believe that the visual system is an apt point of comparison because, being the other dominant sensory modality, it shares many privileged anatomical connections with the motor system whilst simultaneously displaying a less streamlined processing of temporal information than the auditory modality (Loeffler et al., 2018).

Another reason to compare auditory- and visuomotor coupling is that several of their respective structures are known to be functionally connected via the beta frequency band (Comstock et al., 2018). Processes such as sensory attenuation of self-generated sounds (Abbasi & Gross, 2020) or timing predictions in response to visual rhythms (Buchholz et al., 2019) are chiefly mediated by beta-band coupling between motor and auditory or visual cortices, respectively. In light of these discoveries, many have theorized that beta-band oscillations could be important for sensorimotor coupling across the brain at large (Morillon et al., 2019; Comstock et al., 2018; Merchant et al., 2015). This idea is supported by studies showing that the striatum utilizes the beta band as a means of coordinating other oscillatory frequency bands emanating from various different timing systems (Gu et al., 2015; Matell & Meck, 2004). However, the ubiquity of the beta band for motor processes has almost exclusively been shown in task-based paradigms, and thus the current study also aims to elucidate the resting frequency profile associated with auditory-motor coupling.

To examine intrinsic auditory-motor coupling in the human brain we used the Open MEG Archive (OMEGA), an open data repository of MEG and companion structural MRI data (Niso et al., 2016), which houses rs-MEG data for a large and well-characterized sample, including information on musical training. We then derived phase locking values (PLVs) between predefined regions of interest in the auditory and visual cortices with motor and premotor cortical regions. PLVs are an undirected measure of functional connectivity, which is why we also used a recently developed directed connectivity metric, called phase transfer entropy (PTE; Lobier et al., 2014).

We note that both of these measures require signals to share the same oscillatory frequency in order to determine the degree of phase-to-phase alignment over time (Cohen, 2014). The rationale for choosing frequency-bound metrics over cross-frequency coupling (CFC) methods is that rs-MEG networks show frequency-specific connectivity patterns within constituent brain structures, as well as across networks, and that frequencies observed in a given network tend to overlap with those observed during active tasks (Marzetti et al., 2019). Moreover, the current literature on resting state directed connectivity points to the existence of frequency band-specific information flow loops that the brain uses to functionally segregate long-range communication channels (Wang et al., 2019). Thus, there is much to be gleaned from assessing functional dynamics within discrete frequency bins.

We tested five hypotheses using the OMEGA database. First, we posited that functional connectivity in the auditory-motor domain should be stronger than in the visuomotor modality because auditory cortex shares many of the same connections with motor regions, but the literature indicates it is more efficient at temporal processing. Second, we hypothesized that phase coupling between auditory and motor areas would be more predominant in right hemisphere, in line with prior auditory-motor literature (Palomar-García et al., 2017). Third, we predicted that the degree of auditory-motor coupling would scale with musical training as observed in rs-fMRI (Palomar-García et al., 2017). Fourth, given the high prevalence of the beta

band in task-based motor paradigms, we also predicted that auditory-motor phase coupling would be stronger in the beta band than in other frequency bands. Fifth, we expected PTE values to reveal different directions of information flow for each frequency band relevant to auditory-motor coupling, in line with the notion of frequency band-specific information flow loops (Wang et al., 2019).

2.5 Materials & Methods

2.5.1 OMEGA database

The data were obtained from the Open MEG Archive (OMEGA), a centralized repository for multi-site MEG data aggregation which includes raw and processed data (Niso et al., 2016). The OMEGA repository was established by the McConnell Brain Imaging Centre and the Université de Montréal with support from the Québec Bioimaging Network. We used the 2016 OMEGA-BIDS release, which complies with the Brain Imaging Data Structure (BIDS) data storage system and contains over 300 rs-MEG scans from a total of 219 participants. Each participant's data includes an anatomical T1-weighted MRI volume and a screening questionnaire with demographic information such as age, sex, handedness, mental and physical health, linguistic abilities, and musical expertise.

2.5.2 Participants

2.5.2.1 Ethics Approval

The study protocol was approved by the McGill Human Research Ethics Board (REB) in accordance with the Declaration of Helsinki. The OMEGA database itself was approved by the REB of the Montreal Neurological Institute. All participants signed informed consent and agreed to have their anonymized data included in the OMEGA database.

2.5.2.2 Inclusion Criteria

We analysed the data obtained from right-handed participants between the ages of 18-47 years who reported no health issues and were not taking any medication that affects the central nervous system. Participants were required to have at least one high-quality eyes-open rs-MEG recording obtained prior to an experimental task, with a sampling rate of 2400 Hz, a minimum duration of 300 seconds, and a maximum duration of 600 seconds. Participants were also required to have one high-quality empty room MEG recording with a minimum duration of 120 seconds, obtained on the same date as the resting state recording. Additionally, they were required to have one high-quality T1-weighted anatomical MRI scan, as well as valid digitized scalp points and anatomical fiducial files. For participants with multiple resting state recordings obtained on different dates, the recording closest in time to the administration of the OMEGA screening questionnaire was selected.

2.5.2.3 Exclusion Criteria

Out of 219 potential participants, we excluded a total of 129 participants based on the following criteria. We excluded 62 participants who reported using psycho-pharmacological medication, whose demographic data indicated they were part of a patient group, or who reported a diagnosis of ADHD, chronic pain, neurological disease or general psychological illness. We further excluded 8 participants who reported being left-handed, and 1 whose data on handedness was missing. We then excluded 14 participants whose MEG data did not meet the inclusion criteria outlined above, as well as 19 subjects who did not have a T1-weighted anatomical MRI scan, and 7 subjects whose MRI scan presented quality issues. We excluded 12 subjects whose age was higher than 50 years. Finally, during data pre-processing we excluded a total of 6 subjects due to poor data quality and/or missing or inaccurate digitization data that interfered with MRI co-registration.

2.5.2.4 Study Sample

The final study sample consisted of 90 healthy participants (42 female), ranging in age from 19 to 47 years (mean = 27, SD = 6.7). To examine the effects of musicianship and musical training, an initial subset of 76 participants was selected based on the fact that they had valid responses to the Montreal Music History Questionnaire (MMHQ), which includes information on the number of years of musical practice, age of onset of practice, and hours of weekly practice (Coffey et al., 2011). These 76 of participants for whom we had valid MMHQ responses were divided into two subgroups based on their number of years of formal practice: musicians (5 or more years of formal musical or voice training) and non-musicians (0 years of formal musical or voice training). Six participants who reported 0-5 years of musical training were excluded from the subsample, resulting in a total of 29 musicians (mean years of training = 13 years, SD = 6.4) and 36 non-musicians for a grand total of 65 participants (11 female).

2.5.3 MEG Data Collection and Analysis

MEG data were collected at the McConnell Brain Imaging Centre in Montréal using a 275 first-order axial-gradiometer CTF system (Port Coquitlam, BC, Canada). Recordings lasted a minimum of 300 seconds, in accordance with the current minimum standard for neuroanatomical specificity (Wiesman et al., 2022) and were conducted with participants in the seated position as they fixated on a centrally presented crosshair. Participants were monitored during data acquisition via real-time audio-video feeds from inside the shielded room, and continuous head position was recorded for each session. Noise-cancellation was applied using CTF's software-based built-in third-order spatial gradient noise filters. Electrocardiography (ECG) and Electrooculography (EOG) measures were collected from each participant and used during data analysis for the modelling and removal of heartbeat and blink artifacts. Sixteen participants in our sample also wore an EEG cap during MEG data collection, but only their MEG data was utilized in this study.

2.5.3.1 Data Pre-Processing and Quality Control

MEG data preprocessing was conducted in Brainstorm (Tadel et al., 2011) following good-practice guidelines (Tadel et al., 2019). Noisy segments were identified and removed from the MEG signal recordings via visual inspection (Gross et al., 2013), and signals were filtered using a 60 Hz and harmonics notch filter, followed by a 0.3-80 Hz bandpass filter. To aid in the detection and removal of artifacts and bad channels, power spectrum density (PSD) plots were computed at successive stages of pre-processing for visual inspection. A total of 14 faulty channels were removed across 9 participants, with a maximum of three faulty channels removed for any given subject. Repetitive artifacts were removed using Brainstorm's built-in signal space projection (SSP) method (Tesche et al., 1995). Standard SSPs were used to remove heartbeats and eye blinks, and custom SSPs were used to remove saccades and other frequency-defined artifacts in a subset of participants for whom we had sufficient instances to warrant this targeted approach (n = 13). For custom SSPs, we utilized Brainstorm's standard parameters for EOG: an event was to be classed as a saccade only if it was restricted to activity in the 1.5-15Hz range, if this activity was 2 standard deviations above the amplitude threshold, and if the minimum duration between the current event and the next saccade event was longer than 800 ms.

2.5.3.2 Source Image Projection

T1-weighted MRI scans were pre-processed using the 'recon-all' pipeline in Freesurfer (Fischl, 2012) version 5.3, which includes segmentation and labelling of the cortical surface. The T1 images were then co-registered with the previously digitized head points and anatomical fiducials using the automatic registration function in Brainstorm (Tadel et al., 2011), followed by manual review and adjustments. Forward models were constructed for all subjects using the overlapping spheres approach (15,000 vertices, with current flows constrained normal to the cortical surface). This method fits one local sphere under each sensor

and is known to achieve reasonable accuracy relative to more complex boundary element methods (BEMs), due to the fact that magnetic fields in MEG are less sensitive to heterogeneity of tissue in the brain, skull and scalp compared to scalp potentials measured with EEG (Hämäläinen et al., 1993).

Neural sources were then estimated using a minimum norm source imaging kernel with non-scaled current density maps, as we were only interested in localizing sources on the cortex. The minimum norm solution is a cortically distributed source localization model that computes current source densities at equally spaced grid locations over the entire cortical surface, thereby affording a high level of spatial resolution (Tenke & Kayser, 2012). Dipole orientations were specified to be normal (perpendicular) with respect to each individual's cortex, a procedure known to boost the signal-to-noise ratio or SNR (Larson et al., 2014) because it optimally models the activity of pyramidal neuron macrocolumns (typically found in layers III and V), which are themselves perpendicular to the cortex and the drivers of the excitatory activity primarily detected by M/EEG sensors (Attal et al., 2007).

Similarly, we applied depth weighting in order to counteract the bias of non-scaled source estimates toward superficial sources (order [0,1] = 0.5; maximal amount = 10). In addition, we regularized the noise covariance by a factor of 0.1, and we used a strong L2 regularization parameter value ($\lambda = 0.3$) to constrain source estimates and thereby ensure good SNR ratios throughout. Noise covariance was estimated from empty-room recordings collected on the same day as the rs-MEG data, as is the gold standard for MEG (Mosher & Funke, 2020; Gross et al., 2013), and were included in the source imaging computation. Finally, we used the identity matrix in lieu of noise modelling and we applied baseline correction by subtracting the average value of each channel from the entire time window.

2.5.3.3 Regions of Interest

Five bilateral regions of interest (ROIs) were selected prior to computing the functional connectivity analyses: A1, V1, M1, vPMC, and dPMC (Figure 1). The Desikan-Killiany atlas (Desikan et al., 2006) was used to extract ROI templates for the A1 and V1 regions, whereas the Human Motor Area Template (HMAT; Mayka et al., 2006) was used to extract ROI templates for the M1, vPMC, and dPMC regions. The bilateral ROIs were then integrated into a custom hybrid atlas, which was manually generated in Brainstorm for each participant separately. Because all brains were defined in subject space, Brainstorm warped all five ROIs automatically onto each participant's brain upon selection. All participant's ROIs were visually inspected to ensure that the automatic warping was successful. We have included a multi-page supplementary document elucidating each step of the hybrid atlas creation in the GitHub repository listed below in subsection 2.5.

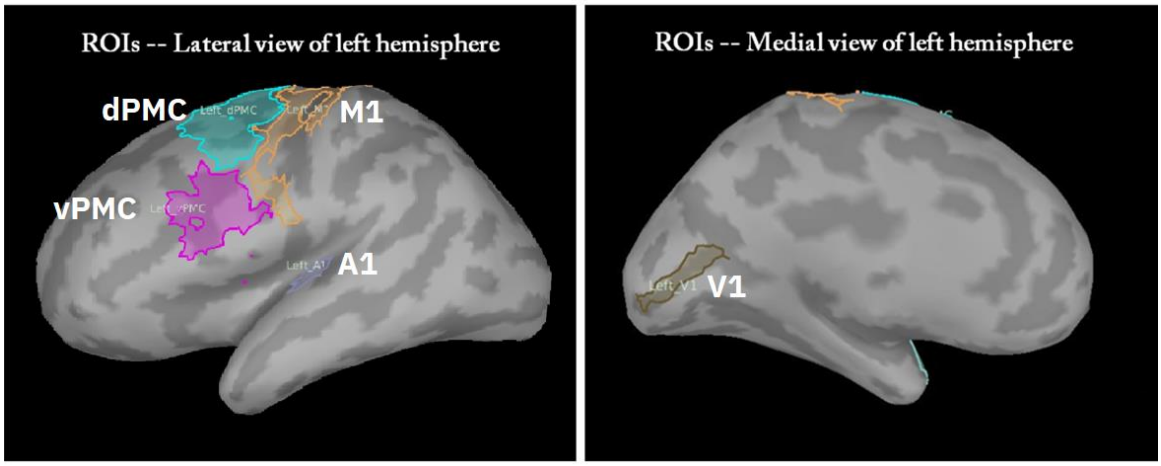


Figure 1. Example of the five ROIs displayed over an inflated brain (right hemisphere not shown) *Two bilateral ROIs extracted from the **Desikan-Killiany** atlas: **A1** (violet; left panel), **V1** (ochre; right panel). Three bilateral ROIs extracted from the **Human Motor Atlas Template** (HMAT): **M1** (orange; left panel), **dPMC** (cyan; left panel), **vPMC** (purple; left panel).*

2.5.3.4 Functional Connectivity Analyses

The strength of phase-to-phase coupling was evaluated using the phase-locking value (PLV), a metric which calculates the average phase angle differences between

pairs of sites over time in order to determine their degree of similarity (Lachaux et al., 1999). PLVs are reported on a scale from 0 to 1, where a value close to 1 indicates high phase coherence across regions. One of the main advantages of using PLVs over other methods is that they are thought to have a direct neurophysiological interpretation as a means by which distant neural populations can temporally coordinate periods of relative excitability (Cohen, 2014). Conversely, one potential pitfall of PLVs is volume conduction (Lachaux et al., 1999), also known as source leakage, where sources in the brain will generate large electromagnetic fields that are measured by more than one MEG sensor, thereby introducing spurious connections in the form of spatial autocorrelation at the sensor level (Nolte et al., 2004).

In this study we minimized the risk of volume conduction in various ways. First, we note the use of MEG, which is inherently more robust to volume conduction than EEG because magnetic fields are not diffused by head tissues, thereby leading to clearer signals (Hämäläinen et al., 1993). Moreover, the amplitude of a magnetic field decreases faster with distance than its associated electrical potential, meaning that volume conduction recorded by an MEG sensor will tend to be smaller than the true signal (Lachaux et al., 1999). In addition, the effects of volume conduction can be more effectively mitigated in MEG via various methods (Friston et al., 1997), with source localization (or source reconstruction) being one of the preferred approaches (Cohen, 2014).

Briefly, because source localization functions like a Laplacian spatial filter, it simultaneously increases spatial resolution and minimizes volume conduction by reducing the influence of distant sources (Cohen, 2014). Moreover, we used a minimum-norm model, which has good spatial resolution and is capable of correcting its own estimates based on the degree of correlated noise between sensor pairs (Engemann et al., 2015), thereby addressing the risk of volume conduction directly. Finally, we also used individual anatomies in order to model each person's brain and head geometry individually, a method which is known to boost source

localization and therefore indirectly mitigate the risk of volume conduction. Essentially, by individually constraining inverse estimates in order to obtain point-spread functions that are different enough across subjects to only overlap around true activations, as opposed to noise, which is thought to lead to improved group-level localization (Larson et al., 2014). Therefore, while we cannot absolutely rule out the possibility of volume conduction, we trust that our choice of methods corrected for this important risk to a reasonable degree, both directly as well as indirectly.

It is also important to note that PLVs are a non-directed measure and therefore cannot indicate the preferred direction of information flow between sites. For this reason, we also computed a secondary metric called phase transfer entropy (PTE), which is defined as a method for measuring the information flow between two time series (Lobier et al., 2014). PTE is an extension of the Transfer Entropy (TE) method, and its purpose is to infer the causal relationships between two non-linear and non-stationary time series based on their phase. Similar to Granger causality (Bressler & Seth, 2011), PTE compares the probability of future states of one phase, given the past states of both phases (Hillebrand et al., 2016). Besides being inherently more robust to volume conduction than non-directed metrics like PLV (Cohen, 2014), Lobier et al (2014) have confirmed that the PTE metric has four advantages for detecting directed brain connectivity: 1) PTE is reliable even in the combined presence of noise and sensor signal mixing which can result in reduced connectivity or artificial connectivity; 2) PTE can identify complex interactions; 3) PTE requires limited amounts of data and computation time; 4) PTE is effective for identifying information flow within an entire frequency band.

We first computed the PLVs for the six possible pairings of ROIs, and the time window was specified as being 300 seconds for all participants, as this was the shortest participant scan duration and the current minimum standard for neuroanatomical specificity (Wiesman et al., 2022). PLVs are sensitive to the duration of the time window across participants because phase angle clustering is directly affected by the

amount of data. Therefore, a fixed 5-minute window was used to ensure comparability across participants (Cohen, 2014). The frequency range of interest was 2-80 Hz and was divided into the following canonical frequency bands: delta (2-4 Hz), theta (5-7 Hz), alpha (9-12 Hz), beta (15-29 Hz), lower gamma (30-59 Hz) and higher gamma (60-80 Hz). The selected time-frequency decomposition method was the Hilbert transform, which allowed us to segregate and retain the phase information, whilst discarding the magnitude information from the calculation of phase angle timeseries. PLV vertex activities were aggregated using the 'Mean after' option in Brainstorm. In other words, within each ROI pair, one PLV for each vertex timeseries pair was calculated, and the results were averaged across all vertex pairings in order to obtain a single PLV at the end, as the PLV metric has been shown to be specifically sensitive to this approach in prior rs-MEG literature (Brkić et al., 2023). We then compared PLVs for the three auditory-motor connections to PLVs for the three visuo-motor connections for each subject, hemisphere, and frequency band. Namely, we compared PLVs for the A1-M1, A1-vPMC, and A1-dPMC connections to PLVs for the V1-M1, V1-vPMC and V1-dPMC connections, respectively. The PLV output measure was defined to be magnitude only in all cases, as we had no prior hypotheses about preferred phase-locking angles.

PTE values were calculated across the three auditory-motor pairings, and within each of the six canonical frequency bands. Vertex activities were aggregated over the entire duration of each individual's MEG recording using the 'PCA before' option in Brainstorm. In contrast to PLVs, PTE values are less susceptible to being altered by additional data points, as the directional information flow (e.g., whether signal A's activity predicts signal B's future activity) is generally stable, even as more data is added. Moreover, since PTE calculations benefit from a larger dataset to accurately capture signal interactions, our approach of using all available data maximizes the robustness and reliability of this measure (Lobier et al., 2014).

Principal component analysis (PCA) was first applied across all vertex timeseries within each ROI in a given pairing, effectively condensing all sources of phase information into a much more reduced set of orthogonal components. Then, the first component of each ROI, which represents the most significant source of variance, was used to compute a single PTE value for the ROI pairing in question. The exact amount of variance explained by this first component across all subjects, hemispheres and ROIs can be found in Supplementary Table 1. This method ensured that the most representative information across vertices was used in the PTE calculation. Despite this being a data-reducing approach, it was deemed appropriate for three reasons: 1) due to computational constraints, it is impractical to compute PTEs on every pair-wise combination of vertices from each pair of ROIs; 2) this approach has been shown to remain sensitive to very subtle effects in previous research (Müller et al., 2019); 3) any potential bias it produced would be consistent across our statistical contrast of interest, as is the case with the PLV metric.

2.5.4 Statistical Tests

Auditory-motor PLVs were compared to visuo-motor PLVs using a Generalized Linear Mixed Model (GLMM). This analytical approach was chosen in order to account for the non-normal distribution of the data points and to allow for the estimation of random effects associated with individual differences. Statistical analyses were conducted with the 'glmmTMB' package in RStudio version 2022.02.0, with the dependent variable parameterized as Beta-distributed and linked to parameter estimates in a log-odds ratio scale. We used a beta distribution for PLVs because they naturally exist on a scale ranging from 0 to 1 and may take any shape within those bounds (Cohen, 2014). Omnibus tests were computed from the model with F statistics, and Bonferroni-corrected t-tests were computed with paired-samples t-tests on the prediction scale. A total of three GLMM analyses were implemented across all frequency bands in order to investigate three distinct research questions.

The first analysis examined normative connectivity patterns in the full sample of 90 participants using a within-subjects design composed of three factors: Hemisphere (left; right), Modality (auditory; visual), and Motor region (M1; vPMC; dPMC). The second analysis compared connectivity between musicians (n = 29) and non-musicians (n = 36) using a four-way mixed-effects design composed of the same three within-subjects factors, as well as Musicianship (musician; non-musician) as a between-subjects factor. The third analysis explored the relationship between connectivity and years of formal musical training in the musician subsample (n=29) using a four-way mixed-effects design composed of the same three within-subjects factors as the first and second analyses, along with z-scored years of training as a between-subjects continuous regressor.

For these three models, as well as for all of the following models, the random effects term featured random intercepts and random slopes for all within-subjects fixed effects, and the grouping variable was set to subject ID. By including both random intercepts and random slopes, each model was able to account for individual differences in the relationship between the predictors and the dependent variable. Specifically, the random slopes allow the relationship between the within-subjects factors to vary across subjects, capturing how these effects might differ between individuals. Moreover, because the random effects term features subject ID as the grouping variable, the model effectively accounts for the repeated measures nature of the data, modelling changes in the dependent variable across the different conditions within each individual.

Because PLV strengths of the auditory and visual modalities were compared within each frequency band separately, we additionally performed cross-band comparison of the magnitudes obtained after calculating the difference between auditory and visual PLVs. We used a single GLMM analysis in the full sample of 90 participants, for which the dependent variable was defined as the value obtained after subtracting the visual PLV from the auditory PLV within a particular frequency band and motor region.

This new dependent variable was parameterized as being Gaussian-distributed because the central limit theorem states that the distribution of the difference between two independent and identically distributed random variables (i.e., the difference between two sets of PLVs) will approach a normal distribution, regardless of the underlying distribution of the original values (Cohen, 2014). For the model, we used a within-subjects design composed of two factors: Frequency band (delta; theta; alpha; beta; gamma1; gamma2) and Motor region (M1; vPMC; dPMC). Omnibus tests were computed from the model with F statistics, and Bonferroni-corrected t-tests were computed with paired-samples t-tests on the prediction scale. Note that we averaged the PLVs of each pairing across hemispheres before computing the cross-modality PLV differences.

Auditory-motor PTEs across all frequency bands were computed separately and compared using a single GLMM analysis in the full sample of 90 participants. The purpose of this analysis was to determine the direction of information flow between A1 and its three motor targets. The dependent variable for this analysis was the PTE value obtained from comparing the phases of the auditory signal at a given frequency band with respect to the phases of the motor signal at that same frequency band, for a given frequency band and motor region. Because PTEs are naturally scaled from 0 to ∞, the dependent variable was parametrized as following a gamma distribution.

For the model, we used a within-subjects design composed of two factors: Direction (auditory-to-motor; motor-to-auditory) and Motor region (M1; vPMC; dPMC). Omnibus tests were computed from the model with F statistics, and Bonferroni-corrected t-tests were computed with paired-samples t-tests on the response scale. Also, in order to validate the preferred auditory-motor directional edges we obtained from this GLMM, we additionally calculated a permutation statistic for the strongest edge for every auditory-motor pairing and frequency band against its own null distribution, closely paralleling the procedure described in the article that introduced

the PTE metric (Lobier et al., 2014; section "Noise and linear mixing"). The results of these null permutation tests and all associated scripts are available in the study's GitHub repository.

Goodness of fit estimates for all statistically significant models can be found in Supplementary Table 2.

2.5.5 Data Availability

We invite the reader to visit our GitHub repository link, which contains the Matlab scripts used for MEG preprocessing and data analysis, the RStudio scripts used for computing the various GLMM models, and all other supplementary tests and materials: https://github.com/OscarBedford/OMEGA_study

2.6 Results

2.6.1 Normative Connectivity

Auditory-motor PLVs were compared to visuo-motor PLVs in the entire sample of 90 participants in order to test the hypotheses that auditory-motor pairings would exhibit larger phase-coherence than visuo-motor pairings, that auditory-motor pairings would exhibit larger phase-coherence in right hemisphere, and that these patterns would be more pronounced in musicians.

2.6.1.1 Higher Auditory-Motor Connectivity and Consistent Increases in Auditory-Motor Connectivity Strengths

We obtained a statistically significant main effect of 'Modality' in all frequency bands, indicating larger auditory-motor PLVs than visuomotor PLVs (Tables 1&2, Figure 2, and supplementary figures 1-6) after Bonferroni correction in the delta ($F_{1,1055}$ = 744, $p < .001$), theta ($F_{1,1055}$ = 1930.73, $p < .001$), alpha ($F_{1,1055}$ = 1142, $p < .001$), beta ($F_{1,1055}$ = 4089.02, $p < .001$), lower gamma ($F_{1,1055}$ = 4740.6, $p < .001$), and higher gamma ($F_{1,1055}$ = 3909.75, $p < .001$) frequency bands. In addition, we obtained a statistically

significant interaction between the 'Modality' and 'Motor region' factors in the delta ($F_{2,1055} = 328.53$, $p < .001$), theta ($F_{2,1055} = 480.56$, $p < .001$) and beta ($F_{2,1055} = 675.18$, $p < .001$) frequency bands, as well as a statistically significant triple interaction between the 'Hemisphere', 'Modality' and 'Motor region' factors in the alpha ($F_{2,1055} = 3.66$, $p = 0.026$), lower gamma ($F_{2,1055} = 3.75$, $p = 0.024$) and higher gamma ($F_{2,1055} = 4.86$, $p = 0.008$) frequency bands.

Both the double and triple interactions pointed to a consistent and Bonferroni-corrected pattern of PLV increases only in the auditory-motor domain, which was the same across the entire spectrum of frequency bands. Namely, A1-M1 exhibited consistently higher PLVs than A1-dPMC, while A1-vPMC exhibited consistently higher PLVs than both A1-M1 and A1-dPMC (Tables 1&2, Figure 2, and supplementary figures 1-6). For the triple interactions, the contribution of the 'Hemisphere' factor is best explained by a lower-order interaction between the 'Hemisphere' and 'Modality' factors in lower gamma, as well as by a significant main effect of 'Hemisphere' in alpha and higher gamma.

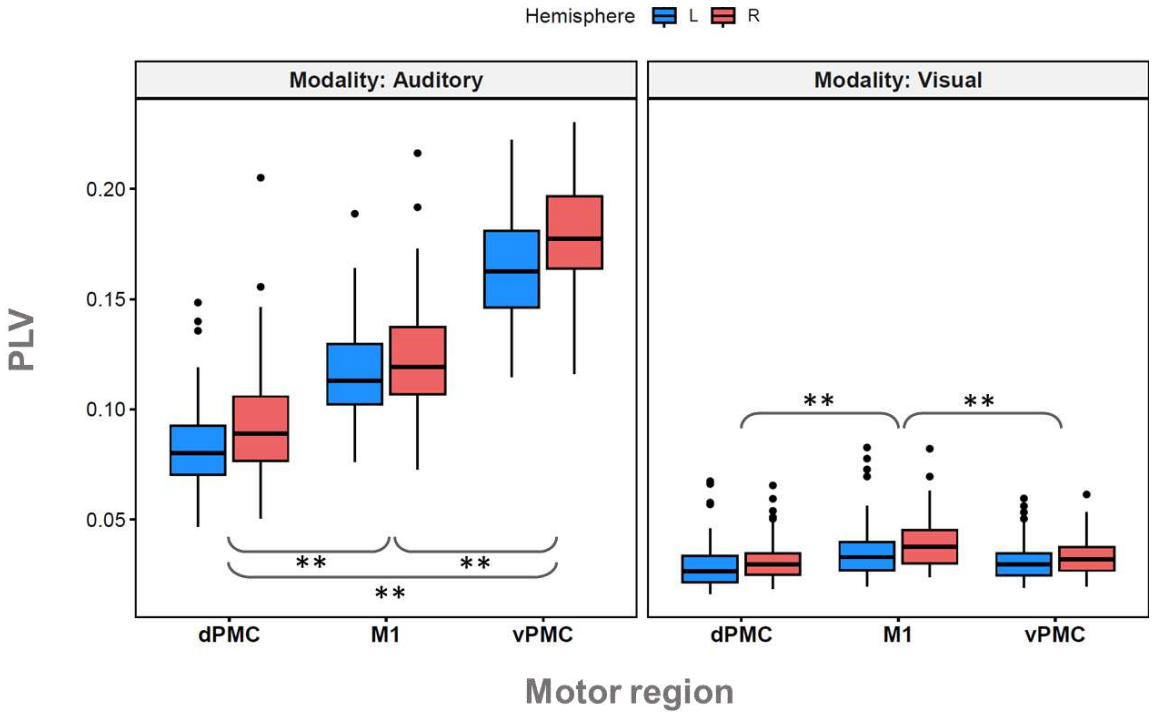


Figure 2. PLV results obtained in GLMM 1 (n=90) in the beta frequency band
***Interaction between 'Modality' and 'Motor region'**: tighter phase coupling for auditory-motor connections, and different coupling strengths for all three motor targets only in the auditory modality.*
***Main effect of 'Hemisphere'**: tighter phase coupling in right hemisphere across both modalities.*

Table 1. GLMM1 post-hoc tests: '*Modality*' by '*Motor region*' (*Auditory modality*)

Frequency band	Contrast	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value
delta	dPMC - M1	-0.2459	0.0125	1055	-19.658	<0.0001	**<0.0001**
	dPMC - vPMC	-0.6434	0.0183	1055	-35.249	<0.0001	**<0.0001**
	M1 - vPMC	-0.3975	0.0199	1055	-19.985	<0.0001	**<0.0001**
theta	dPMC - M1	-0.2363	0.011	1055	-21.389	<0.0001	**<0.0001**
	dPMC - vPMC	-0.6711	0.0143	1055	-46.900	<0.0001	**<0.0001**
	M1 - vPMC	-0.4348	0.0153	1055	-28.464	<0.0001	**<0.0001**
alpha	dPMC - M1	-0.276	0.0119	1055	-23.159	<0.0001	**<0.0001**
	dPMC - vPMC	-0.5962	0.0234	1055	-25.482	<0.0001	**<0.0001**
	M1 - vPMC	-0.3202	0.0245	1055	-13.064	<0.0001	**<0.0001**
beta	dPMC - M1	-0.347	0.0101	1055	-34.155	<0.0001	**<0.0001**
	dPMC - vPMC	-0.783	0.02	1055	-39.084	<0.0001	**<0.0001**
	M1 - vPMC	-0.436	0.0205	1055	-21.324	<0.0001	**<0.0001**
gamma1	dPMC - M1	-0.292	0.0107	1055	-27.240	<0.0001	**<0.0001**
	dPMC - vPMC	-0.8827	0.0154	1055	-57.325	<0.0001	**<0.0001**
	M1 - vPMC	-0.5907	0.0164	1055	-35.984	<0.0001	**<0.0001**
gamma2	dPMC - M1	-0.193	0.0095	1055	-20.358	<0.0001	**<0.0001**
	dPMC - vPMC	-0.688	0.0144	1055	-47.919	<0.0001	**<0.0001**
	M1 - vPMC	-0.495	0.0153	1055	-32.324	<0.0001	**<0.0001**

Table 2. GLMM1 post-hoc tests: '*Modality*' by '*Motor region*' (*Visual modality*)

Frequency band	Contrast	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value
delta	dPMC - M1	-0.0682	0.0148	1055	-4.612	<0.0001	**<0.0001**
	dPMC - vPMC	-0.1197	0.0245	1055	-4.892	<0.0001	**<0.0001**
	M1 - vPMC	-0.0515	0.0258	1055	-1.998	0.1379	0.827
theta	dPMC - M1	-0.0882	0.0146	1055	-6.046	<0.0001	**<0.0001**
	dPMC - vPMC	-0.0917	0.0203	1055	-4.521	<0.0001	**<0.0001**
	M1 - vPMC	-0.0035	0.021	1055	-0.168	>0.9999	>0.9999
alpha	dPMC - M1	-0.176	0.0148	1055	-11.883	<0.0001	**<0.0001**
	dPMC - vPMC	-0.1275	0.0299	1055	-4.269	<0.0001	**<0.0001**
	M1 - vPMC	0.0485	0.0307	1055	1.581	0.343	>0.9999
beta	dPMC - M1	-0.223	0.0154	1055	-14.442	<0.0001	**<0.0001**
	dPMC - vPMC	-0.071	0.026	1055	-2.734	0.019	0.115
	M1 - vPMC	0.152	0.0261	1055	5.825	<0.0001	**<0.0001**
gamma1	dPMC - M1	-0.2527	0.0162	1055	-15.623	<0.0001	**<0.0001**
	dPMC - vPMC	-0.2452	0.0199	1055	-12.307	<0.0001	**<0.0001**
	M1 - vPMC	0.0074	0.0203	1055	0.367	<0.0001	**<0.0001**
gamma2	dPMC - M1	-0.293	0.0135	1055	-21.63	<0.0001	**<0.0001**
	dPMC - vPMC	-0.184	0.02	1055	-9.177	<0.0001	**<0.0001**
	M1 - vPMC	0.108	0.0204	1055	5.309	<0.0001	**<0.0001**

2.6.1.2 Higher Right-Hemispheric Connectivity

We obtained a statistically significant main effect of 'Hemisphere', indicating overall larger PLVs in the right compared to the left hemisphere in 5 out of 6 frequency bands after Bonferroni correction; namely in the theta ($F_{1,1055} = 6.57$, $p = 0.011$; $t_{1055} = -3.235$; $p = 0.001$), alpha ($F_{1,1055} = 17.7$, $p < .001$; $t_{1055} = -4.014$; $p < .001$), beta ($F_{1,1055} = 27.76$, $p < .001$; $t_{1055} = -6.095$; $p < .001$), lower gamma ($F_{1,1055} = 15.66$, $p < .001$; $t_{1055} = -5.364$; $p < .001$), and higher gamma ($F_{1,1055} = 23.8$, $p < .001$; $t_{1055} = -5.552$; $p < .001$) frequency bands. In addition, we obtained a statistically significant interaction between the 'Hemisphere' and 'Modality' factors, indicating larger PLVs in the right compared to the left hemisphere after Bonferroni correction, exclusively in the auditory-motor modality, within the delta ($F_{2,1055} = 11.2$, $p < .001$; $t_{auditory(1055)} = -2.357$; $p = 0.037$; $t_{visual(1055)} = 1.409$; $p = 0.318$) and lower gamma bands ($F_{2,1055} = 7.17$, $p = 0.008$; $t_{auditory(1055)} = -5.902$; $p < 0.001$; $t_{visual(1055)} = 1.862$; $p = 0.126$).

2.6.1.3 No Effects of Musicianship or Years of Musical Training on Connectivity

The 'Musicianship' and 'Years of training' factors, respectively included in a second (n=65) and third (n=29) GLMM analysis, were not associated with any statistically significant main effects or interactions.

2.6.2 Cross-Band 'Auditory – Visual' Connectivity Differences

We performed a cross-band comparison of the magnitudes obtained from calculating the difference between auditory and visual PLVs at every motor region in order to determine whether there were any frequency bands and motor regions wherein the two modalities differed the most. The GLMM analysis we conducted led to a statistically significant interaction between the 'Motor region' and 'Frequency band' factors after averaging all PLVs across hemispheres ($F_{10,1583} = 17.5$, $p < .001$).

Bonferroni-corrected t-tests indicated that, within dPMC, the alpha frequency band exhibited a significantly larger PLV difference across modalities than the other 5 bands, namely delta ($t_{1583} = 7.36$, $p < .001$), theta ($t_{1583} = 4.82$, $p = 0.001$), beta ($t_{1583} = 5.2$, $p < .001$), lower gamma ($t_{1583} = 6.98$, $p < .001$) and higher gamma ($t_{1583} = 4.05$, $p = 0.036$) frequency bands (Table 3a). Similarly, Bonferroni-corrected t-tests revealed that, within M1, the alpha frequency band exhibited a significantly larger PLV difference across modalities than the other 5 bands, namely delta ($t_{1583} = 7.37$, $p < .001$), theta ($t_{1583} = 5.72$, $p < .001$), beta ($t_{1583} = 4.95$, $p = 0.001$), lower gamma ($t_{1583} = 9.44$, $p < .001$) and higher gamma ($t_{1583} = 8.84$, $p < .001$) frequency bands (Table 3b). Finally, Bonferroni-corrected t-tests revealed that, within vPMC, the alpha frequency band exhibited a significantly larger PLV difference across modalities than 3 other bands, namely delta ($t_{1583} = 6.92$, $p < .001$), lower gamma ($t_{1583} = 6.30$, $p < .001$) and higher gamma ($t_{1583} = 6.38$, $p < .001$) frequency bands (Table 3c).

Thus, the significant interaction between 'Frequency band' and 'Motor region' was driven by a significant post-hoc difference between the alpha band and every other frequency band within dPMC and M1, compared to the alpha band and every other band *except* the beta and theta bands within vPMC (Table 3ab versus Table 3c). Overall, these results indicate that the alpha frequency band exhibited the largest difference between the auditory and visual modalities across all three motor targets, followed closely by the theta and beta frequency bands particularly within vPMC (Figure 3).

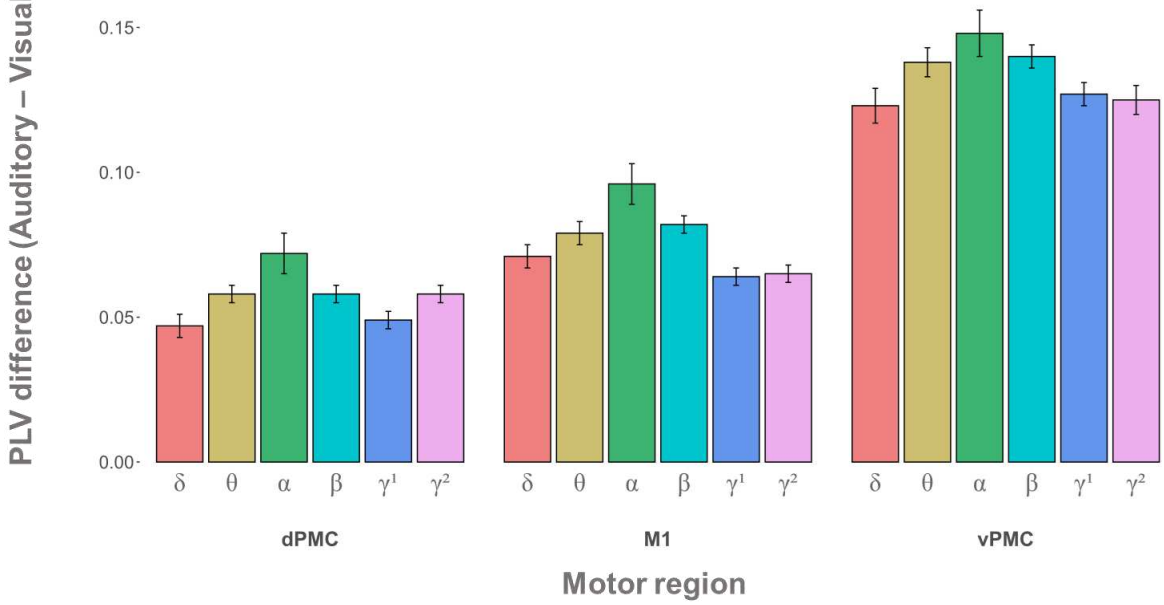


Figure 3. PLV differences across modalities in GLMM 2 (n=90)
***Interaction between 'Frequency band' and 'Motor region'**: all differences are positive, indicating larger PLVs in the auditory modality. Among these, the alpha, beta and theta frequency bands exhibit the largest PLV magnitude differences across all three motor targets.*

Table 3a. GLMM2 post-hoc tests: '*Frequency band*' by '*Motor region*' *(dPMC)*

Motor region	Contrast	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value
dPMC	alpha - beta	0.015	0.003	1583	5.205	<0.0001	**<0.0001**
	alpha - delta	0.025	0.003	1583	7.363	<0.0001	**<0.0001**
	alpha - gamma1	0.024	0.003	1583	6.977	<0.0001	**<0.0001**
	alpha - gamma2	0.014	0.004	1583	4.052	0.001	**0.036**
	alpha - theta	0.014	0.003	1583	4.822	<0.0001	**0.001**
dPMC	beta - delta	0.01	0.004	1583	2.409	0.242	>0.9999
	beta - gamma1	0.009	0.004	1583	2.127	0.503	>0.9999
	beta - gamma2	-0.001	0.004	1583	-0.117	>0.9999	>0.9999
	beta - theta	0	0.004	1583	-0.094	>0.9999	>0.9999
dPMC	delta - gamma1	-0.001	0.005	1583	-0.249	>0.9999	>0.9999
	delta - gamma2	-0.011	0.005	1583	-2.268	0.352	>0.9999
	delta - theta	-0.01	0.004	1583	-2.438	0.223	>0.9999
dPMC	gamma1 - gamma2	-0.009	0.005	1583	-2.017	0.659	>0.9999
	gamma1 - theta	-0.009	0.004	1583	-2.163	0.461	>0.9999
	gamma2 - theta	0	0.004	1583	0.031	>0.9999	>0.9999

Table 3b. GLMM2 post-hoc tests: '*Frequency band*' by '*Motor region*' (*M1*)

Motor region	Contrast	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value
M1	alpha - beta	0.014	0.003	1583	4.951	<0.0001	**0.001**
	alpha - delta	0.025	0.003	1583	7.374	<0.0001	**<0.0001**
	alpha - gamma1	0.032	0.003	1583	9.443	<0.0001	**<0.0001**
	alpha - gamma2	0.031	0.004	1583	8.836	<0.0001	**<0.0001**
	alpha - theta	0.017	0.003	1583	5.716	<0.0001	**<0.0001**
M1	beta - delta	0.011	0.004	1583	2.591	0.145	>0.9999
	beta - gamma1	0.018	0.004	1583	4.280	<0.0001	**0.013**
	beta - gamma2	0.017	0.004	1583	3.940	0.001	0.057
	beta - theta	0.003	0.004	1583	0.774	>0.9999	>0.9999
M1	delta - gamma1	0.007	0.005	1583	1.557	>0.9999	>0.9999
	delta - gamma2	0.006	0.005	1583	1.299	>0.9999	>0.9999
	delta - theta	-0.008	0.004	1583	-1.830	>0.9999	>0.9999
M1	gamma1 - gamma2	-0.001	0.005	1583	-0.23	>0.9999	>0.9999
	gamma1 - theta	-0.015	0.004	1583	-3.483	0.008	0.344
	gamma2 - theta	-0.014	0.004	1583	-3.168	0.023	>0.9999

Table 3c. GLMM2 post-hoc tests: '*Frequency band*' by '*Motor region*' (*vPMC*)

Motor region	Contrast	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value
vPMC	alpha - beta	0.008	0.003	1583	2.824	0.072	>0.9999
	alpha - delta	0.025	0.004	1583	6.916	<0.0001	**<0.0001**
	alpha - gamma1	0.021	0.003	1583	6.303	<0.0001	**<0.0001**
	alpha - gamma2	0.023	0.004	1583	6.379	<0.0001	**<0.0001**
	alpha - theta	0.01	0.003	1583	3.287	0.016	0.699
vPMC	beta - delta	0.017	0.004	1583	3.845	0.002	0.085
	beta - gamma1	0.013	0.004	1583	3.182	0.022	>0.9999
	beta - gamma2	0.015	0.004	1583	3.458	0.008	0.377
	beta - theta	0.002	0.004	1583	0.485	>0.9999	>0.9999
vPMC	delta - gamma1	-0.003	0.005	1583	-0.711	>0.9999	>0.9999
	delta - gamma2	-0.002	0.005	1583	-0.304	>0.9999	>0.9999
	delta - theta	-0.015	0.005	1583	-3.322	0.014	0.618
vPMC	gamma1 - gamma2	0.002	0.005	1583	0.391	>0.9999	>0.9999
	gamma1 - theta	-0.012	0.004	1583	-2.656	0.12	>0.9999
	gamma2 - theta	-0.013	0.005	1583	-2.952	0.048	>0.9999

2.6.3 Directed Connectivity within Auditory-Motor Regions

In order to determine the direction of information flow between A1 and its three motor targets, we computed auditory-motor PTE values across all frequency bands separately. Within a given band, we compared PTE values using a GLMM analysis, which revealed a statistically significant interaction between the 'Direction' and 'Motor region' factors in the right-hemisphere theta frequency band ($F_{2,527} = 3.41$, $p = 0.034$), bilaterally in the alpha frequency band (Left: $F_{2,527} = 12.09$, $p < .001$; Right: $F_{2,527} = 8.7$, $p < .001$), and bilaterally in the beta frequency band (Left: $F_{2,527} = 6.11$, $p = 0.002$; Right: $F_{2,527} = 13.34$, $p < .001$). We also obtained a statistically significant main effect of 'Direction', revealing a globally dominant motor-to-sensory direction in the left-hemisphere theta frequency band that survived Bonferroni correction (Table 4a; $F_{1,527} = 5.38$, $p = 0.02$), as well as a globally dominant sensory-to-motor direction in the left-hemisphere higher gamma band (Table 4g; $F_{1,527} = 14.89$, $p < .001$) and right-hemisphere gamma band (Table 4h; $F_{1,527} = 7.33$, $p = 0.007$) that also survived corrections for multiple comparisons. The delta and lower gamma bands were not associated with any significant 'Direction' by 'Motor region' interaction, nor any significant main effect of 'Direction'.

In the right-hemisphere theta band (Table 4b), Bonferroni-corrected t-tests revealed that only the direction contrast in dPMC was statistically significant (t_{527} = 3.79, p = 0.001), indicating that the direction of phase information was predominantly top-down, as it was strongest from Right-dPMC to Right-A1, and therefore consistent with the predominantly top-down direction in the left-hemisphere theta band (Figure 4; column 1). In the left-hemisphere alpha frequency band (Table 4c), the direction contrasts in dPMC (t_{527} = -5.598, p < .001), M1 (t_{527} = -3.053, p = 0.007), and vPMC (t_{527} = -4.137, p < .001) all survived Bonferroni correction. Similarly, in the right-hemisphere alpha frequency band (Table 4d), the direction contrasts in dPMC (t_{527} = -3.978, p < .001) and vPMC (t_{527} = -3.399, p = 0.002) also survived correction for multiple comparisons. These results in the alpha frequency band indicate that the direction of phase information in this band was predominantly bottom-up, as in all cases it was strongest from bilateral A1 to bilateral motor regions (Figure 4; column 2).

In the left-hemisphere beta frequency band (Table 4e), Bonferroni-corrected t-tests revealed that only the direction contrast in M1 (t_{527} = 2.435, p = 0.046) was statistically significant. In the right-hemisphere beta frequency band (Table 4f), Bonferroni-corrected t-tests revealed that the direction contrasts in dPMC (t_{527} = 2.658, p = 0.024) and M1 (t_{527} = 4.250, p < .001) were also statistically significant. These results in the beta frequency band indicate that the direction of phase information in this band was predominantly top-down, as in all cases it was strongest from bilateral motor regions to bilateral A1 (Figure 4; column 3). Overall, these results indicate that the general direction of phase information in our data went predominantly from auditory to motor nodes in the theta and beta bands, and predominantly from motor to auditory nodes in the alpha band and, to a lesser extent, in the higher gamma bands (Figure 4; column 4).

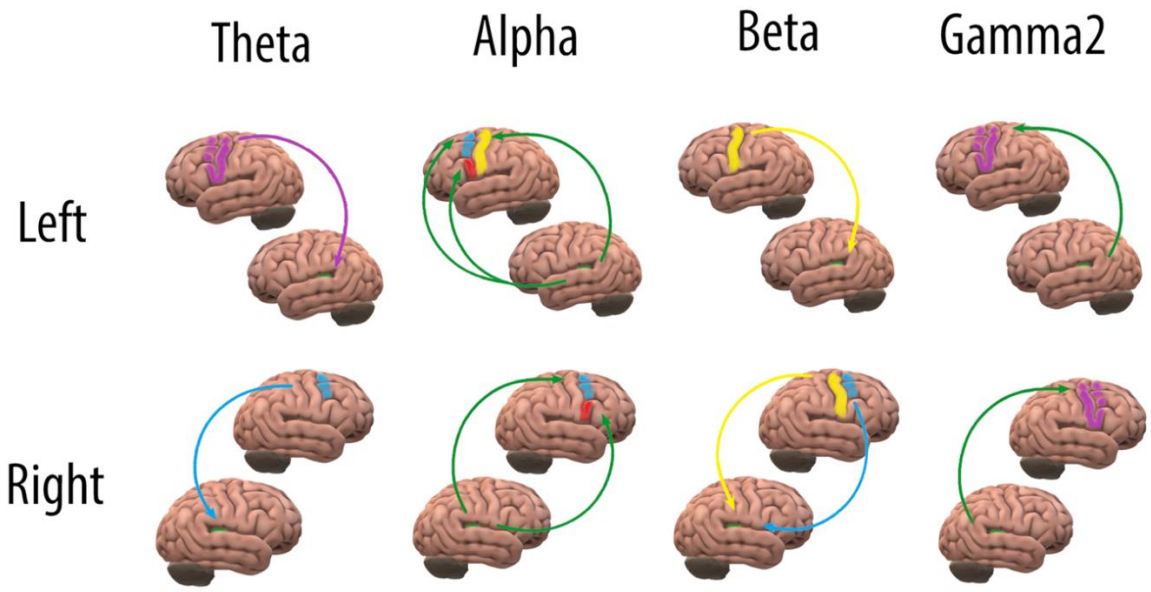


Figure 4. Predominant auditory-motor PTE values across the frequency spectrum in GLMM 3 (n=90)
Interaction between 'Direction' and 'Motor region'*: In the right theta frequency band, the direction of phase information is strongest from dPMC to A1. In the left alpha frequency band, the direction of information is strongest from A1 to all three motor nodes, and in the right alpha frequency band it is strongest from A1 to dPMC and vPMC. In the left beta frequency band, the direction of information is strongest from M1 to A1, and in the right beta frequency band it is strongest from M1 and dPMC to A1.*
Main effect of 'Direction'*: In the left theta frequency band, the direction of phase information is strongest from the average activity across motor regions to A1. In bilateral gamma2, the direction of phase information is strongest from A1 to the average activity across motor regions.*

Table 4. GLMM3 post-hoc tests: '*Direction*' by '*Motor region*' interaction, and '*Direction*' main effect

	Frequency band	Hemisphere	Contrast	Motor region	Estimate	SE	df	t-ratio	p-value	Bonferroni p-value	
4a	theta	left	M2S - S2M	AVG	0.0037	0.0016	527	2.332	0.02	0.02	4a
4b	theta	right	M2S - S2M	dPMC	0.006	0.00160	527	3.789	<.0001	0.001	
				M1	0.0032	0.002	527	1.598	0.1110	0.332	4b
				vPMC	0.0037	0.0020	527	1.893	0.0590	0.177	
4c	alpha	left	M2S - S2M	dPMC	-0.0259	0.0046	527	-5.598	<.0001	<.0001	
				M1	-0.0147	0.0048	527	-3.053	0.002	0.007	4c
				vPMC	-0.021	0.005	527	-4.137	<.0001	<.0001	
4d	alpha	right	M2S - S2M	dPMC	-0.0186	0.0047	527	-3.978	<.0001	<.0001	
				M1	-0.0093	0.0049	527	-1.912	0.056	0.169	4d
				vPMC	-0.0187	0.0055	527	-3.399	0.001	0.002	
4e	beta	left	M2S - S2M	dPMC	0.0002	0.0016	527	0.113	0.91	>0.9999	
				M1	0.0052	0.0022	527	2.435	0.015	0.046	4e
				vPMC	-0.0012	0.0021	527	-0.607	0.544	>0.9999	
4f	beta	right	M2S - S2M	dPMC	0.005	0.0019	527	2.658	0.008	0.024	
				M1	0.0094	0.0022	527	4.250	<.0001	<.0001	4f
				vPMC	-0.0004	0.0024	527	-0.147	0.883	>0.9999	
4g	gamma2	left	M2S - S2M	AVG	-0.001	0.0003	527	-3.832	<.0001	<.0001	4g
4h	gamma2	right	M2S - S2M	AVG	-0.0007	0.0003	527	-2.689	0.007	0.007	4h

Finally, the results of the PTE null permutation tests (not reported) confirmed that the PTE values we obtained for each preferred direction were, respectively for each individual directional edge, statistically significantly higher than a gaussian distribution of 1000 phase-shuffled PTE values. Moreover, most of the opposing, non-preferred directions within a given edge were also statistically significantly higher than their respective gaussian distribution of phase-shuffled PTE values. The only exceptions were the non-preferred top-down direction in the left and right high gamma band, whose respective PTE values had a lower value than their associated distribution of phase-shuffled counterparts, and therefore did not survive the null permutation test. The results of all these null permutation tests are available in the form of annotated histograms in the study's GitHub repository.

2.7 Discussion

2.7.1 Summary

The goal of this study was to use rs-MEG data to assess the resting connectivity patterns between auditory-motor cortical regions across a broad spectrum of frequency bands, and to compare these patterns to resting visuomotor cortical connectivity (Figure 1). Phase-based metrics were used because they specifically allowed us to determine the degree of temporal synchronization between pairs of regions of interest, which were defined independently using anatomical atlases. We obtained significantly higher phase-locking values (PLVs) across auditory-motor regions compared to visuomotor regions (Figure 2 and supplementary figures 1-6), indicating that tight phase locking constitutes an inherent property of cortico-cortical functional connectivity between primary auditory cortex and cortical motor areas at rest.

We also found a unique pattern of consistent increases in PLV strength only in the auditory-motor modality (Figure 2 and supplementary figures 1-6), which we relate to the underlying anatomy linking A1 to the motor cortices, as well as to the functional

differences between motor regions. Moreover, we observed an overall right-hemispheric predominance (Figure 2 and supplementary figures 1-6) which we consider in the larger context of anatomical and functional studies showing higher indices of interconnectivity and efficiency in this hemisphere, particularly for A1. When comparing frequency bands, we observed that the alpha band had the biggest modality difference in PLV strength across all three motor targets, followed closely by the beta and theta bands (Figure 3), which we interpret as further evidence of an intrinsic and modality-specific mechanism enabling the low-latency temporal alignment of sounds and movements across the theta-beta frequency spectrum.

This conclusion was supported by our directed connectivity analysis using phase transfer entropy (PTE) values, which revealed that auditory-motor information flow in the alpha frequency band (and to a lesser extent in the higher gamma band) preferentially followed an auditory-to-motor direction, whereas in the theta and beta frequency bands it preferentially followed a motor-to-auditory direction (Figure 4). Notably, this pattern of directed connectivity results is well-aligned with the notions of segregated information flow loops and cross-frequency coupling, as well as with the common claim that auditory-motor coupling relies on top-down beta-band modulations.

2.7.2 Normative Connectivity

2.7.2.1 Higher Auditory-Motor Connectivity and Consistent Increases in Auditory-Motor Connectivity Strengths

We observed a significantly higher degree of average phase-locked connectivity between auditory-motor regions compared to visuomotor regions (Figure 2 and supplementary figures 1-6), as originally predicted. Because this greater temporal synchronization was observed at rest, we can interpret it to be an inherent property of cortico-cortical connections between primary auditory cortex with motor and premotor areas. As such, we believe that this coupling may underlie the higher degree

of temporal alignment between sounds and movements routinely observed in auditory-motor compared to visuomotor paradigms (Hove et al., 2013; Kato & Konishi, 2006; Repp & Penel, 2004).

It is true that bodily entrainment to *moving* visual stimuli, as opposed to visual flashes, can sometimes yield variability and latency estimates that are comparable to those seen in the auditory modality (Hove et al., 2010). However, given our resting state findings, we maintain that the auditory modality must retain privileged access to the motor system and possibly to timing systems in general even when compared to moving visual stimuli, an idea which is in line with the following studies. First, we note on a broader level, that visual timing seems to require the additional engagement of timing centres, and yet still produces behavioural outcomes that are, at best, equivalent to those of the auditory timing system (Jäncke et al., 2000). For instance, visual rhythm perception induces increased beta activity at event onsets in V1 (Comstock & Balasubramaniam, 2017) which, as Comstock and colleagues point out, could potentially be the result of "compensating for a weaker connection to the motor timing system" (2018). Indeed, this visual activity may in fact reflect "additional processing of visual information required to interface with the motor system" more effectively (Comstock et al., 2018), such as the trajectory parameters of moving visual stimuli. In contrast, the auditory modality's superiority in timing tasks correlates with greater activation within motor structures, such as the SMA and premotor cortex (Jäncke et al., 2000). Additionally, when compared to moving visual stimuli, auditory stimuli elicit optimal behavioural outcomes which correlate with greater activation in regions like the putamen (Hove et al., 2013).

Additionally, we note that our reported difference between auditory- and visuomotor coupling strengths is unlikely to be due to differences in SNRs between the A1 and V1 ROIs in our data. This is because V1's source dipoles are oriented tangentially to the scalp, and this orientation directs the resulting magnetic fields to be optimally and preferentially captured by the closest sensors (Ahlfors et al., 2010), instead of those

on the more radially oriented occipital poles, despite the latter being physically closer to said sensors. Moreover, the exact opposite situation is true for A1, as its source dipoles are oriented radially to the scalp, and therefore its closest sensors will be significantly less sensitive to its magnetic fields, with some being nearly insensitive (Ahlfors et al., 2010). Thus, we believe it is reasonable to expect similar SNRs for A1 and V1, especially since this is what has been reported in prior literature (Goldenholz et al., 2009).

Furthermore, unlike regions that are both medial and deep, there is evidence to suggest that more superficial medial sources like V1 have point spread values that are equivalently low as those in lateral cortical regions (Muthukumaraswamy & Singh, 2013), as they are ultimately not too far from their closest sensors. In addition, we note that: 1) our choice of a cortically distributed source model (Tenke & Kayser, 2012), in conjunction with the use of individual anatomies (Larson et al., 2014), was driven by the requirement for a high degree of spatial specificity with which to specifically address low SNRs; 2) our selection of depth-weighting parameter values was optimized to counter the fact that minimum-norm estimates have a bias for superficial currents (Lin et al., 2006), and therefore to improve SNR in medial and deep regions; 3) our decision to use a strong L2 regularization parameter value was precisely informed by our preference for a smoother solution over the risk of obtaining a noisy one (Lapalme et al., 2006). Therefore, we have no reason to believe that our V1 ROI suffered from lower SNRs than the motor ROIs in our set, nor A1 in particular, and we trust that the methods deployed to mitigate the risk of low SNRs adequately protected our connectivity estimates throughout.

Within the auditory-motor but not the visuomotor modality we also observed a consistent pattern of variation in PLV strengths across motor regions, whereby connectivity between A1 to vPMC was the greatest, followed by A1-M1, and lastly by A1-dPMC (Figure 2 and supplemental figures 1-6). In our opinion, the anatomical pathways of the arcuate fasciculus (AF) may be partly responsible for these

differences. The AF is known to directly connect posterior auditory regions with both the vPMC (Schubotz et al., 2010; Catani & Mesulam, 2008) and the dPMC (Petrides, 2013). In addition, there exists an indirect AF pathway connecting vPMC and pSTG (Hoshi & Tanji, 2007) which consists of two segments: an anterior segment linking vPMC and other frontal areas with the inferior parietal lobule, and a posterior segment linking the inferior parietal lobule with pSTG (Catani & Mesulam, 2008; Catani et al., 2005).

The dPMC is similarly connected to the auditory cortex by additional indirect pathways which relay at the superior parietal lobule (Hoshi & Tanji, 2007) via the superior longitudinal fasciculus (Makris et al., 2005; Kamali et al., 2014). Finally, although there are no known anatomical tracts connecting A1 to M1 directly, a substantial number of fibres connect both vPMC and dPMC with M1 (Petrides & Pandya, 2006; Dum & Strick, 2005). Thus, while there is not necessarily a one-to-one relationship between anatomical connectivity and the functional interactions we describe below, we believe that this multitude of anatomical pathways connecting auditory, motor and premotor cortical areas is nonetheless likely to be relevant. Therefore, future work relating the microstructural properties (myelination, fiber organization, white-matter volume…) of these links to the pattern of results reported here may help to clarify the differences in functional coupling strength we obtained.

On a related note, we believe that the differences between auditory- and visuomotor coupling strengths here reported are unlikely to be due to differences in anatomical path lengths between these two modalities. The major anatomical link between the visual and motor cortices occurs via the dorsal visual stream pathway, which is similar to the dorsal auditory stream because it goes from occipital to frontal areas of the cortex via a relay at parietal regions (Glickstein, 2000). However, there is evidence that visual input can reach M1 at relatively short latencies and quickly modulate corticospinal excitability in subjects at rest (Strigaro et al., 2015). Moreover, there exists a long associative bundle called the inferior fronto-occipital fascicle that

directly connects the occipital cortex and other posterior areas to the frontal lobe (Sarubbo et al., 2013).

This structure is primarily known for transmitting information from frontal to occipital cortex for the purposes of top-down control, but it is also suspected to contain a direct afferent pathway originating at occipital cortex which is capable of transmitting visual information to motor regions in a rapid manner (Martino et al., 2010). Thus, long and short path lengths are similarly present in both the auditory- and visuomotor modalities, and rapid and reciprocal signal transmission between nodes has been reported in both systems, suggesting that the differences observed in the current study are not anatomical but rather functional in nature.

This is a plausible interpretation because functional connectivity does not necessarily exhibit a one-to-one correspondence with white matter fibre tracts and relays. A functional connection generally results from information flowing through multiple, often polysynaptic paths, even in the presence of a direct anatomical connection between neural elements (Avena-Koenigsberger et al., 2018). Therefore, one should not always expect shorter anatomical paths to be more highly correlated with functional connectivity measures or vice versa. In fact, a recent study combined structural and functional data to infer sequences of nodes that were structurally connected and temporally synchronized and found that such pathways actually exhibit a greater correlation with functional connectivity than the topologically shortest ones (Griffa et al., 2017). Thus, it is essential to assess the connectivity profiles of the three auditory-motor pairings in relation to the distinct cognitive processes they support.

In that respect, our pattern of results is consistent with the different functional specializations of cortical motor regions previously described in the auditory-motor literature (Zatorre et al., 2007). For instance, connectivity between auditory and vPMC has been linked to the processing of so-called direct auditory-motor transformations (Baumann et al., 2007; Lahav et al., 2007; Bangert et al., 2006), which involve a one-

to-one matching of auditory features to motor acts. As evidence supporting this idea, there is the fact that both A1 and vPMC are active during passive exposure to music for which one has an associated motor program (Lahav et al., 2007), for example. Furthermore, substantial evidence points to the involvement of the vPMC (Alho et al., 2014) and nearby region pars opercularis (Burton et al., 2000) in auditory-motor mappings relevant for speech. Indeed, vPMC has been found to mediate the mapping of speech sounds such as phonemes onto their articulatory representations (Devlin & Watkins, 2007) and has also been shown to modulate neural activity in secondary auditory cortex as a function of increasing syllable articulation rate (Paus et al., 1996).

Moreover, Ruspantini and colleagues (2011) used TMS-induced disruptions to prove a critical involvement of the vPMC in visuomotor coupling, while Kornysheva and Schubotz (2011) used the same method to establish its involvement in auditory-motor coupling. In both studies, the effect of vPMC disruption was specific to externally paced timing, as opposed to self-paced timing, suggesting that the vPMC is a critical node for modality-independent externally paced motor timing (Kornysheva & Schubotz, 2011). That said, it is important to note that although some individual areas such as vPMC may overlap onto both modalities, this does not negate the fact that the auditory and visual circuits which share vPMC are functionally distinct at a network level.

The functional connectivity between A1 and M1, the pairing which exhibited an intermediate degree of phase-locking in our data, has been similarly well established in studies assessing the cortical excitability of M1 in response to speech. For instance, several studies (Watkins et al., 2003; Campbell et al., 2001; MacSweeney et al., 2000) applied TMS to the M1 lip area in order to elicit motor-evoked potentials and found enhanced amplitudes during viewing and listening to speech, compared to control conditions.

Lastly, A1-dPMC has been associated with so-called indirect auditory-motor transformations, which represent motor information instructed by sensory cues in a

many-to-many mapping rather than by direct one-to one mapping of sensory properties (Lega et al., 2016; Chen et al., 2012; Chen et al., 2008). Therefore, the comparatively high degree of functional connectivity observed between A1 and vPMC in prior resting state studies, as well as our own data, could be due to the fact that this connection is involved in one-to-one auditory-motor processes which are externally paced and crucial for executing recurrent actions, such as speech articulation and musical practice. Conversely, the fact that the A1-dPMC link is typically involved in less assiduous higher order processes and new learning could explain why this pairing displays the lowest index of resting functional connectivity in our data. Finally, due to the somatotopic nature of M1, we suspect that our intermediate A1-M1 connection may have been stronger if we had only included the lip and hand subsections, as these recruit the main effectors of the recurrent auditory-motor programs the vPMC supports.

We suggest two ways of testing our claim that tight phase locking is an inherent property of cortico-cortical functional connectivity between primary auditory cortex and cortical motor areas at rest. First, we would expect individual differences in PLV strength between auditory and motor areas to be positively related to performance on auditory-motor tasks, as was the case in a recent study showing that performance on a motor task could be predicted by beta-band functional connectivity between M1 and left STG in the preceding rest period (Sugata et al., 2020). Thus, future inquiries could extend the scope of prior literature by including all the cortical areas and frequency bands used in the present study in order to comprehensively compare the effect of individual differences at rest to behavioural performance.

Second, we would expect that stimulating either node within a given auditory-motor pairing with single-pulse TMS would transiently enhance the consistency of phase locking at the naturally preferred frequency and thus drive up the PLV for that connection. This prediction would be consistent with Kawasaki and colleagues (2014), who showed that the consistency of endogenous phase locking

across stimulation trials is transiently enhanced after single-pulse stimulation, presumably via phase resetting (Glim et al., 2019; Sauseng et al., 2007).

Moreover, single pulses are assumed to act on intrinsic neural systems, thereby maximally affecting those frequencies which already arise naturally within a particular region (Rosanova et al., 2009), meaning that it would be possible to probe the strength of the putative mechanism for ongoing auditory-motor synchronization observed in the current study without needing to entrain oscillators to a particular eigenfrequency via repetitive TMS (Thut et al., 2011). Moreover, phase coupling is known to be involved in long range communication processes (Guggisberg et al., 2015), and TMS has proven to have more sustained effects on phase information, despite showcasing more transient effects on signal power (Thut et al., 2011). Thus, the literature suggests it would be plausible to stimulate a motor node with single pulses and to expect the effects on phase coupling to be detectable, to spread a long distance, and to last long enough to be measured with rs-MEG.

2.7.2.2 Higher Right-Hemispheric Connectivity

We obtained a pattern of heightened connectivity in the right hemisphere and exclusively within the auditory-motor modality for the delta and lower gamma bands. Thus, in two out of six frequency bands we verified our hypothesis that only phase coupling between auditory and motor areas would be predominant in the right hemisphere. In addition, we obtained an overall pattern of heightened connectivity in the right hemisphere across the theta-gamma frequency spectrum (Figure 2 and supplementary figures 1-6). At present, we do not have a theory that can explain why our hypothesis was only fully validated in the delta and lower gamma bands specifically, especially since these are the only two bands that did not reveal any preferred information flow direction between auditory and motor nodes (Figure 4). Therefore, we have chosen to focus the interpretation of these results on the right-hemispheric predominance across auditory-motor connectivity estimates, as this is

the common finding that pervades all frequency bands in our data, as well as the literature (Palomar-García et al., 2017).

From this perspective, our finding is consistent with the theory that hemispheric differences are related to network properties, an idea supported by graph-theoretic metrics of anatomical connectivity showing higher indices of interconnectivity and efficiency in right hemisphere specifically (Iturria-Medina et al., 2011). Measures of functional connectivity density have similarly shown greater short- and long-range connectivity in the right versus left superior temporal cortex (Tomasi & Volkow, 2012). Moreover, a recent DTI study investigated the propensity for left and right A1 to communicate with other brain areas by quantifying the centrality of the auditory network across a spectrum of communication mechanisms, from shortest path communication to diffusive spreading, and also found that the right A1 is better integrated in the connectome, facilitating more efficient communication with other areas (Mišić et al., 2018). Similarly, repetitive TMS to the right A1, but not left A1, results in distributed changes to fMRI connectivity profiles in auditory and motor-related networks which are visible even in the absence of an overt task (Andoh et al., 2015).

It should be noted that there are several auditory-motor studies that point to a predominance of left-hemisphere nodes in the context of specific tasks, as opposed to resting state. For instance, Pollok and colleagues (2008) used TMS to assess functional connectivity between dPMC and M1 in the context of finger-tap timing and found that left dPMC is more interconnected with left and right M1 than right dPMC. Similarly, the aforementioned Morillon & Baillet study (2017) found that the right associative auditory cortex was most strongly connected to the left – not right – somatomotor cortex in their beat-tracking task. That said, we are not aware of any evidence for stronger left-hemispheric effects between auditory and motor systems in resting state data. Therefore, we believe that the right-hemispheric predominance in our auditory-motor data is most consistent with the anatomical and functional

studies showing higher indices of interconnectivity and efficiency in this hemisphere, particularly in A1.

2.7.2.3 No Effects of Musicianship or Years of Musical Training on Connectivity

We observed no effect of musicianship on the strength of functional connectivity in the beta band, nor any of the other frequency bands, either within the auditory-motor pairs as predicted nor within the visuomotor pairs. This observation is at odds with what we hypothesized, as well as with several M/EEG task-based and resting state functional connectivity studies, which indicate that musicians show increased phase synchrony over distributed cortical areas, predominantly in the gamma band, as compared to nonmusicians (Bhattacharya & Petsche, 2005; Snyder & Large, 2005; Sokolov et al., 2004). Klein and colleagues (2016) additionally showed that musicians at rest exhibit increased intra- and interhemispheric phase coherence within the theta and alpha frequency bands and between brain regions which are typically involved in music perception and production, such as auditory, sensorimotor, and prefrontal cortices. Moreover, in that study, mean coherence within this network was positively related to musical skill and the total number of training hours, indicating that the differences were specifically due to musical training (Klein et al., 2006).

Given these consistent results in the literature, we do not believe that our null result abrogates the relationship between auditory-motor phase coherence and musicianship. More likely there is some other aspect of the measures used or the population sampled here that accounts for the discrepancy. For example, we were unable to examine the role of age of start of training, which is known to be important in promoting neural plasticity (Penhune, 2021), and which therefore should be examined more closely in future studies. Similarly, our sample did not provide the opportunity to examine the possible role of expertise on different musical instruments. Different instruments make distinct motor demands on the player; for instance, piano vs French horn require bimanual vs unimanual control, respectively,

while different combinations of effectors (lips/tongue, fingers, hands, feet) are used for wind, keyboard or percussion instruments. Thus, more detailed future studies along these lines may yet clarify the role of musicianship in auditory-motor coupling regions at rest.

2.7.3 Cross-Band 'Auditory – Visual' Connectivity Differences

The largest differences in phase coupling strength between auditory-motor pairings and their corresponding visuomotor pairings were observed in the alpha band across all three motor targets (Figure 3). Moreover, the largest cross-modality difference was observed in the alpha band with respect to vPMC and was closely followed by similarly large cross-modality differences in the theta and beta bands, indicating that these three groups of frequencies may be of particular relevance to the intrinsic mechanism for low-latency auditory-motor coupling outlined above.

We note that this idea is supported by our directed connectivity analyses discussed below, which highlight that auditory-motor information flow in the alpha band (and to a lesser extent in the higher gamma band) preferentially follows an auditory-to-motor direction, whereas in the theta and beta bands it preferentially follows a motor-to-auditory direction (Figure 4). Moreover, whole-brain rs-MEG studies have emphasized the particular contribution of alpha and beta oscillatory signals for the generation of the canonical resting-state networks (Brookes et al., 2011; De Pasquale et al., 2010), a fact that conforms well with findings that low-frequency oscillations coordinate long-range communication (Von Stein et al., 2000). In addition, the task-based literature has shown that the theta, alpha and beta bands all play an important role in auditory-motor coupling specifically. For instance, in a simultaneous EEG-fMRI study using a dual auditory vs visual attention task, attention to the auditory domain yielded a positive correlation of theta-band amplitudes and BOLD activity in right A1, right M1 and right rolandic operculum, thereby implicating both auditory and motor systems in auditory attention via the theta band (Wang et al., 2016).

As for the alpha band, a growing body of evidence points to the existence of an independent alpha auditory rhythm distinct from other known generators (Weisz et al., 2011; Lehtelä et al., 1997; Tiihonen et al., 1991), and one auditory-motor study found that active syllable discrimination produced more robust sensorimotor alpha event-related desynchronization than passive listening to syllables or discriminating between tones, indicating more sensorimotor processing in the active condition (Bowers et al., 2013). Pre-stimulus alpha event-related desynchronization over sensorimotor cortices has also been observed in expert pianists during passive listening to trained melodies, but not untrained melodies, indicating that action representations can be used as a predictive model in order to guide listening, and that these action representations implicate the alpha band (Wu et al., 2016).

Regarding the beta band, several studies have highlighted its role in the representation of various time durations in the brain (Bartolo & Merchant, 2015; Cirelli et al., 2014; Kononowicz & van Rijn, 2015), a feature which is especially relevant to auditory-motor coupling. Moreover, the beta band has been implicated in numerous functional connectivity studies of motor learning which specifically utilize phase-based methods (Tropini et al., 2011; Deeny et al., 2009). For instance, Morillon & Baillet (2017) used MEG during a task which involved either passively listening to a melody or superimposing a beat by mentally "accentuating" every other note and found that PLVs corresponding to the melody rate were strongest in right auditory association cortex whereas PLVs corresponding to the imagined beat were strongest in left primary motor cortex.

It is important to note that the theta, alpha, and beta bands have also been related to numerous visuomotor processes in the context of task-based paradigms. For instance, the aforementioned study using a dual visual vs auditory attention task additionally reported a supramodal positive correlation between theta-band amplitudes and BOLD activity in right vPMC (Wang et al., 2016). Similarly, other visuomotor studies have found that pre-stimulus alpha power dictates apparent

motion perception (Sanders et al., 2014), and one study comparing MEG activity in visuomotor tasks to interspersed rest periods reported task-related increases in beta-band phase coherence between areas involved in movement, motor planning, and vision (Bardouille & Boe, 2012).

Thus, while the literature on these three bands indicates that heightened activity generally underpins both auditory- and visuomotor function in task-based contexts, our results in these bands indicate that primary auditory cortex is characterized by much stronger phase coupling with cortical motor and premotor areas than primary visual cortex, particularly within vPMC. Therefore, we believe that our cross-modal differences support the notion of an intrinsic and modality-specific mechanism for low-latency temporal alignment of sounds and movements which, at a minimum, involves primary auditory cortex and its oscillatory-based coupling with cortical motor and premotor areas across the theta-beta frequency spectrum.

2.7.4 Directed Connectivity within Auditory-Motor Regions

Our PTE analyses revealed that the direction of phase information was strongest from motor nodes to our primary auditory cortex node in right hemisphere within the theta band, from bilateral auditory cortex nodes to bilateral motor nodes in the alpha band, and from bilateral motor to bilateral auditory cortex nodes in the beta band. Our analyses additionally revealed that the direction of phase information was strongest from the average activity collapsed across left motor nodes to our left primary auditory cortex node within the theta band, as well as bilaterally from our auditory cortex nodes to the average activity collapsed across motor nodes within the higher gamma band. (Figure 4).

Each of these directionally opposite information flows was validated using its own null permutation test and corresponds to what is sometimes referred to in the literature as a (frequency band-specific) information flow loop (Wang et al., 2019), a putative mechanism for segregating long-range communication channels in order to

prevent crosstalk. Thus, our pattern of alternating information flow directionalities across the theta-gamma range lends additional evidence to the prevalence of information flow loops in resting state contexts (Wang et al., 2019; Dauwan et al., 2019; Engels et al., 2017; Hillebrand et al., 2016).

The fact that the preferred directionality of phase information in our data alternated across contiguous frequency bands is also well aligned with the concept of cross-frequency coupling (CFC), and with the notion of alpha-gamma and theta-beta communication channels being key players in auditory processing. Of all types of CFC, phase-amplitude coupling (PAC) between the phase of a low frequency and the amplitude of a high frequency has been highlighted as an ideal candidate for long-range communication, especially at rest (Canolty & Knight, 2010). For instance, one whole-brain PAC study found that resting alpha phase dynamics are coupled to resting power bursts in gamma, and that this phenomenon is located predominantly and almost exclusively within A1 sources (Florin & Baillet, 2015), in line with our own results.

On the other hand, frontal lobe to temporal lobe theta phase synchronization has been touted as the foundation for music-induced pleasant emotions (Ara & Marco-Pallarés, 2020), and theta-beta band PAC information appears to capture music emotions much more robustly than theta-gamma PAC information does, as there appears to be consistent desynchronization between the theta and beta frequency bands during the processing of music-induced emotions (Xu et al., 2023). This might be related to the fact that theta power in right inferior frontal cortex and beta oscillations in right A1 have been associated with correct detection of pitch changes (Florin et al., 2017). Similarly, excessive beta activity in basal ganglia is linked to motor impairment in Parkinson's disease and has been shown to be suppressed via PAC to theta (and delta) band phase activity in response to music, leading to directly proportional improvements in clinical measures such as gait (Jin et al., 2022).

Finally, the finding of a motor-to-auditory direction within the beta band is particularly noteworthy because it matches the findings of at least two recent M/EEG task-based auditory-motor studies demonstrating a motor-to-auditory direction of phase information within this frequency band (Morillon & Baillet, 2017; Nicolaou et al., 2017). Moreover, one recent MEG study combined resting-state periods with task periods where participants had to alternate between attention to visual and auditory stimuli, and reported a preferred resting motor-to-auditory directionality of phase information in the theta and upper alpha bands, an auditory-to-motor directionality of phase information in the lower alpha band, and a tendency toward a motor-to-auditory directionality of phase information in the beta band; a pattern which closely matches our own phase-based results (Hanna et al., 2023).

More broadly, our resting-state directed connectivity finding in the beta band is well aligned with the literature on the modulatory role of beta-band oscillations (Abbasi & Gross, 2020) for top-down predictions of temporal structure (Arnal & Giraud, 2012) originating in motor structures like sensorimotor and premotor cortices (Fujioka et al., 2015), and specifically targeting cortical auditory structures (Samiee et al., 2022). Lastly, the preferred motor-to-auditory direction found in our data is also concordant with the idea that beta activity is a vehicle for top-down signalling during sensory processing in general (Chao et al., 2018; Michalareas et al., 2016; Bastos et al., 2015; Bressler & Richter, 2015).

2.8 Conclusion

The current study provides evidence that tight phase coupling is a unique and intrinsic property of the resting functional connectivity between the primary auditory cortex and cortical motor areas. We believe that this intrinsic phase coupling is important because it may underlie the low-latency temporal alignment of movements with sounds observed in auditory-motor paradigms. This idea is supported by a right-hemispheric predominance which is consistent with the literature, and by the fact that the degree of phase coupling between auditory cortex

and the cortical motor system was region-specific in a manner which is broadly consistent with the anatomical and functional differences associated with each motor area.

Another point of support is the fact that the biggest differences in resting phase coupling strength between the auditory and visual modalities were found in the theta-beta frequency spectrum, despite the fact that this collection of bands is similarly activated by both auditory- and visuomotor task-based paradigms. Moreover, each band within this spectrum was associated with a distinct direction of information flow and was therefore aligned with the notions of information flow loops and cross-frequency coupling. Finally, the finding of a preferred motor-to-auditory direction in the beta frequency band is consistent with the auditory-motor task-based articles that inspired our study. Thus, our results fully validate the idea that resting auditory-motor connectivity displays consistent properties which hold important clues about the system's functionality.

2.9 References

Abbasi, O., & Gross, J. (2020). Beta-band oscillations play an essential role in motor–auditory interactions. *Human Brain Mapping*, *41*(3), 656–665.

Ahlfors, S. P., Han, J., Lin, F. H., Witzel, T., Belliveau, J. W., Hämäläinen, M. S., & Halgren, E. (2010). Cancellation of EEG and MEG signals generated by extended and distributed sources. *Human Brain Mapping*, *31*(1), 140–149.

Alho, J., Lin, F. H., Sato, M., Tiitinen, H., Sams, M., & Jääskeläinen, I. P. (2014). Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Frontiers in Psychology*, *5*, 394.

Andoh, J., Matsushita, R., & Zatorre, R. J. (2015). Asymmetric interhemispheric transfer in the auditory network: evidence from TMS, resting-state fMRI, and diffusion imaging. *Journal of Neuroscience*, *35*(43), 14602–14611.

Ara, A., & Marco-Pallarés, J. (2020). Fronto-temporal theta phase-synchronization underlies music-evoked pleasantness. *NeuroImage*, *212*, 116665.

Araneda, R., Renier, L., Ebner-Karestinos, D., Dricot, L., & Volder, A. G. (2017). Hearing, feeling or seeing a beat recruits a supramodal network in the auditory dorsal stream. *European Journal of Neuroscience*, *45*(11), 1439–1450.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience*, *22*(4), 627–632.

Attal, Y., Bhattacharjee, M., Yelnik, J., Cottereau, B., Lefèvre, J., Okada, Y., & Baillet, S. (2007). Modeling and detecting deep brain activity with MEG & EEG. *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4937–4940.

Avena-Koenigsberger, A., Mišić, B., & Sporns, O. (2018). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, *19*(1), 17–33.

Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, *20*(3), 327–339.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, *30*(3), 917-926.

Bardouille, T., & Boe, S. (2012). State-related changes in MEG functional connectivity reveal the task-positive sensorimotor network. *PloS One*, *7*(10), 48682.

Bartolo, R., & Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *Journal of Neuroscience*, *35*(11), 4635–4640.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390–401.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain Research*, 65–78.

Bhattacharya, J., & Petsche, H. (2005). Phase synchrony analysis of EEG during music perception reveals changes in functional connectivity due to musical expertise. *Signal Processing*, *85*(11), 2161–2177.

Bowers, A., Saltuklaroglu, T., Harkrider, A., & Cuellar, M. (2013). Suppression of the μ rhythm during speech and non-speech discrimination revealed by independent component analysis: Implications for sensorimotor integration in speech processing. *PloS One*, *8*(8), 72024.

Bressler, S. L., & Richter, C. G. (2015). Interareal oscillatory synchronization in top-down neocortical processing. *Current Opinion in Neurobiology*, *31*, 62–66.

Bressler, S. L., & Seth, A. K. (2011). Wiener–Granger causality: a well established methodology. *Neuroimage*, *58*(2), 323–329.

Brkić, D., Sommariva, S., Schuler, A. L., Pascarella, A., Belardinelli, P., Isabella, S. L., & Pellegrino, G. (2023). The impact of ROI extraction method for MEG connectivity estimation: practical recommendations for the study of resting state data. *NeuroImage*, *284*, 120424.

Brookes, M. J., Woolrich, M., Luckhoo, H., Price, D., Hale, J. R., Stephenson, M. C., & Morris, P. G. (2011). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proceedings of the National Academy of Sciences*, *108*(40), 16783–16788.

Buchholz, V. N., David, N., Sengelmann, M., & Engel, A. K. (2019). Belief of agency changes dynamics in sensorimotor networks. *Scientific Reports*, *9*(1).

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: an fMRI investigation. *Journal of Cognitive Neuroscience*, *12*(4), 679–690.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., & David, A. S. (2001). Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning. *Cognitive Brain Research*, *12*(2), 233–243.

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, *14*(11), 506–515.

Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *57*(1), 8–16.

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex*, *44*(8), 953-961.

Catani, M., & Schotten, M. T. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132.

Chao, Z. C., Takaura, K., Wang, L., Fujii, N., & Dehaene, S. (2018). Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron*, *100*(5), 1252–1266.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*(12), 2844–2854.

Chen, J. L., Rae, C., & Watkins, K. E. (2012). Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage*, *59*(2), 1200–1208.

Cirelli, L. K., Bosnyak, D., Manning, F. C., Spinelli, C., Marie, C., Fujioka, T., & Trainor, L. J. (2014). Beat-induced fluctuations in auditory cortical beta-band activity: using EEG to measure age-related changes. *Frontiers in Psychology*, *5*, 742.

Coffey, E. B. J., Herholz, S. C., Scala, S., & Zatorre, R. J. (2011, June). Montreal Music History Questionnaire: a tool for the assessment of music-related experience in music cognition research. *The Neurosciences and Music IV: Learning and Memory, Conference*.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

Comstock, D., & Balasubramaniam, R. (2017). Beta-band response synchronizes and predicts rhythmic flashing visual stimuli. *47th Annual Society for Neuroscience Meeting*.

Comstock, D. C., Hove, M. J., & Balasubramaniam, R. (2018). Sensorimotor synchronization with auditory and visual modalities: Behavioural and neural differences. *Frontiers in Computational Neuroscience*, *12*, 53.

Dauwan, M., Hoff, J. I., Vriens, E. M., Hillebrand, A., Stam, C. J., & Sommer, I. E. (2019). Aberrant resting-state oscillatory brain activity in Parkinson's disease patients with visual hallucinations: An MEG source-space study. *NeuroImage: Clinical*, *22*, 101752.

Deeny, S. P., Haufler, A. J., Saffer, M., & Hatfield, B. D. (2009). Electroencephalographic coherence during visuomotor performance: a comparison of cortico-cortical communication in experts and novices. *Journal of Motor Behavior*, *41*(2), 106–116.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968–980.

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS. *Brain*, *130*(3), 610–622.

Donoghue, T., Schaworonkow, N., & Voytek, B. (2022). Methodological considerations for studying neural oscillations. *European Journal of Neuroscience*, *55*(11–12), 3502–3527.

Dum, R. P., & Strick, P. L. (2005). Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *Journal of Neuroscience*, *25*(6), 1375–1386.

Engels, M. M. A., Yu, M., Stam, C. J., Gouw, A. A., Flier, W. M., Scheltens, P. H., & Hillebrand, A. (2017). Directional information flow in patients with Alzheimer's disease. A source-space resting-state MEG study. *NeuroImage: Clinical*, *15*, 673–681.

Engemann, D., Strohmeier, D., Larson, E., & Gramfort, A. (2015). Mind the noise covariance when localizing brain sources with M/EEG. In *2015 International Workshop on Pattern Recognition in NeuroImaging* (pp. 9–12). IEEE.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774–781.

Florin, E., & Baillet, S. (2015). The brain's resting-state activity is shaped by synchronized cross-frequency coupling of neural oscillations. *Neuroimage*, *111*, 26–35.

Florin, E., Vuvan, D., Peretz, I., & Baillet, S. (2017). Pre-target neural oscillations predict variability in the detection of small pitch changes. *PloS One*, *12*(5), 0177836.

Friston, K. J., Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1994). *Functional and effective connectivity in neuroimaging: a synthesis*. *2*(1-2), 56–78.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, *35*(45), 15187–15198.

Glickstein, M. (2000). How are visual areas of the brain connected to motor areas for the sensory guidance of movement? *Trends in Neurosciences*, *23*(12), 613–617.

Glim, S., Okazaki, Y. O., Nakagawa, Y., Mizuno, Y., Hanakawa, T., & Kitajo, K. (2019). Phase-amplitude coupling of neural oscillations can be effectively probed with concurrent TMS-EEG. *Neural Plasticity*.

Goldenholz, D. M., Ahlfors, S. P., Hämäläinen, M. S., Sharon, D., Ishitobi, M., Vaina, L. M., & Stufflebeam, S. M. (2009). Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human Brain Mapping*, *30*(4), 1077–1086.

Griffa, A., Ricaud, B., Benzi, K., Bresson, X., Daducci, A., Vandergheynst, P., & Hagmann, P. (2017). Transient networks of spatio-temporal connectivity map communication pathways in brain functional systems. *NeuroImage*, *155*, 490-502.

Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., & Schoffelen, J. M. (2013). Good practice for conducting and reporting MEG research. *Neuroimage*, *65*, 349–363.

Gu, B. M., Rijn, H., & Meck, W. H. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neuroscience & Biobehavioural Reviews*, *48*, 160–185.

Guggisberg, A. G., Rizk, S., Ptak, R., Pietro, M., Saj, A., Lazeyras, F., & Pignat, J. M. (2015). Two intrinsic coupling types for resting-state integration in the human brain. *Brain Topography*, *28*, 318–329.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413.

Hanna, J., Kim, C., Rampp, S., Buchfelder, M., & Müller-Voggel, N. (2024). Decreasing alpha flow releases task-specific processing paths. *Imaging Neuroscience*, *2*, 1–24.

Heijden, K., Rauschecker, J. P., Gelder, B., & Formisano, E. (2019). Cortical mechanisms of spatial hearing. *Nature Reviews Neuroscience*, *20*(10), 609-623.

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, *24*(6), 1382–1397.

Hillebrand, A., Tewarie, P., Dellen, E., Yu, M., Carbo, E. W., Douw, L., & Stam, C. J. (2016). Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proceedings of the National Academy of Sciences*, *113*(14), 3867–3872.

Hoshi, E., & Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties. *Current Opinion in Neurobiology*, *17*(2), 234–242.

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *Neuroimage*, *67*, 313–321.

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6).

Iturria-Medina, Y., Pérez Fernández, A., Morris, D. M., Canales-Rodríguez, E. J., Haroon, H. A., García Pentón, L., & Melie-García, L. (2011). Brain hemispheric structural efficiency and interconnectivity rightward asymmetry in human and nonhuman primates. *Cerebral Cortex*, *21*(1), 56–67.

Iversen, J. R., & Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioural Sciences*, *8*, 175–180.

Jäncke, L., Loose, R., Lutz, K., Specht, K., & Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cognitive Brain Research*, *10*(1–2), 51–66.

Jin, L., Shi, W., Zhang, C., & Yeh, C. H. (2022). Frequency nesting interactions in the subthalamic nucleus correlate with the step phases for Parkinson's disease. *Frontiers in Physiology*, *13*, 890753.

Kamali, A., Sair, H. I., Radmanesh, A., & Hasan, K. M. (2014). Decoding the superior parietal lobule connections of the superior longitudinal fasciculus/arcuate fasciculus in the human brain. *Neuroscience*, *277*, 577–583.

Kato, M., & Konishi, Y. (2006). Auditory dominance in the error correction process: a synchronized tapping study. *Brain Research*, *1084*(1), 115–122.

Kawasaki, M., Uno, Y., Mori, J., Kobata, K., & Kitajo, K. (2014). Transcranial magnetic stimulation-induced global propagation of transient phase resetting associated with directional information flow. *Frontiers in Human Neuroscience*, *8*, 173.

Klein, C., Liem, F., Hänggi, J., Elmer, S., & Jäncke, L. (2016). The "silent" imprint of musical training. *Human Brain Mapping*, *37*(2), 536–546.

Kononowicz, T. W., & Rijn, H. (2015). Single trial beta oscillations index time estimation. *Neuropsychologia*, *75*, 381–389.

Kornysheva, K., & Schubotz, R. I. (2011). Impairment of auditory-motor timing and compensatory reorganization after ventral premotor cortex stimulation. *PLoS One*, *6*(6), 21421.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, *8*(4), 194–208.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, *27*(2), 308–314.

Lapalme, E., Lina, J. M., & Mattout, J. (2006). Data-driven parceling and entropic inference in MEG. *NeuroImage*, *30*(1), 160–171.

Larson, E., Maddox, R. K., & Lee, A. K. (2014). Improving spatial localization in MEG inverse imaging by leveraging intersubject anatomical differences. *Frontiers in Neuroscience*, *8*, 96205.

Leaver, A. M., Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, *29*(8), 2477–2485.

Lega, C., Stephan, M. A., Zatorre, R. J., & Penhune, V. (2016). Testing the role of dorsal premotor cortex in auditory-motor association learning using transcranical magnetic stimulation (TMS. *PLoS One*, *11*(9), 0163380.

Lehtelä, L., Salmelin, R., & Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neuroscience Letters*, *222*(2), 111–114.

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society B*, *376*(1835), 20200325.

Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, *27*(1), 1–13.

Lobier, M., Siebenhühner, F., Palva, S., & Palva, J. M. (2014). Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *Neuroimage*, *85*, 853–872.

Loeffler, J., Cañal-Bruland, R., Schroeger, A., Tolentino-Castro, J. W., & Raab, M. (2018). Interrelations between temporal and spatial cognition: The role of modality-specific processing. *Frontiers in Psychology*, *9*, 2609.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., & Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise: an event-related fMRI study. *Neuroreport*, *11*(8), 1729–1733.

Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness, V. S., & Pandya, D. N. (2005). Segmentation of Subcomponents within the Superior Longitudinal Fascicle in Humans: A Quantitative, In Vivo, DT-MRI Study. *Cerebral Cortex*, *15*(6), 854–869.

Martino, J., Brogna, C., Robles, S. G., Vergani, F., & Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex*, *46*(5), 691–699.

Marzetti, L., Basti, A., Chella, F., D'Andrea, A., Syrjälä, J., & Pizzella, V. (2019). Brain functional connectivity through phase coupling of neuronal oscillations: a perspective from magnetoencephalography. *Frontiers in Neuroscience*, *13*, 964.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research*, *21*(2), 139–170.

Mayka, M. A., Corcos, D. M., Leurgans, S. E., & Vaillancourt, D. E. (2006). Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *Neuroimage*, *31*(4), 1453-1474.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140093.

Michalareas, G., Vezoli, J., Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron*, *89*(2), 384–397.

Mišić, B., Betzel, R. F., Griffa, A., Reus, M. A., He, Y., Zuo, X. N., & Zatorre, R. J. (2018). Network-based asymmetry of the human auditory system. *Cerebral Cortex*, *28*(7), 2655–2664.

Morillon, B., Arnal, L. H., Schroeder, C. E., & Keitel, A. (2019). Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neuroscience & Biobehavioural Reviews*, *107*, 136–142.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42), 8913-8921.

Mosher, J. C., & Funke, M. E. (2020). Towards best practices in clinical magnetoencephalography: patient preparation and data acquisition. *Journal of Clinical Neurophysiology*, *37*(6), 498–507.

Müller, F., Niso, G., Samiee, S., Ptito, M., Baillet, S., & Kupers, R. (2019). A thalamocortical pathway for fast rerouting of tactile information to occipital cortex in congenital blindness. *Nature Communications*, *10*(1), 5154.

Muthukumaraswamy, S. D., & Singh, K. D. (2013). Visual gamma oscillations: the effects of stimulus type, visual field coverage and stimulus motion on MEG and EEG recordings. *Neuroimage*, *69*, 223–230.

Nicolaou, N., Malik, A., Daly, I., Weaver, J., Hwang, F., Kirke, A., & Nasuto, S. J. (2017). Directed motor-auditory EEG connectivity is modulated by music tempo. *Frontiers in Human Neuroscience*, *11*, 502.

Niso, G., Rogers, C., Moreau, J. T., Chen, L. Y., Madjar, C., Das, S., & Baillet, S. (2016). OMEGA: the open MEG archive. *Neuroimage*, *124*, 1182–1187.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., & Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, *115*(10), 2292–2307.

Palomar-García, M. Á., Zatorre, R. J., Ventura-Campos, N., Bueichekú, E., & Ávila, C. (2017). Modulation of functional connectivity in auditory–motor networks in musicians compared with nonmusicians. *Cerebral Cortex*, *27*(5), 2768–2778.

Park, H., Ince, R. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, *25*(12), 1649–1653.

Park, H., Kayser, C., Thut, G., & Gross, J. (2016). Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *Elife*, *5*, 14521.

Pasquale, F., Della Penna, S., Snyder, A. Z., Lewis, C., Mantini, D., Marzetti, L., & Corbetta, M. (2010). Temporal dynamics of spontaneous MEG activity in brain networks. *Proceedings of the National Academy of Sciences*, *107*(13), 6040-6045.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.

Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K. J., & Evans, A. C. (1996). Modulation of cerebral blood flow in the human auditory cortex during speech: Role of motor-to-sensory discharges. *European Journal of Neuroscience*, *8*(11), 2236–2246.

Penhune, V. B. (2021). Understanding sensitive period effects in musical training. In *Sensitive Periods of Brain Development and Preventive Interventions* (pp. 167-188). Springer International Publishing.

Pernier, J., Perrin, F., & Bertrand, O. (1988). Scalp current density fields: concept and properties. *Electroencephalography and Clinical Neurophysiology*, *69*(4), 385-389.

Petrides, M. (2013). *Neuroanatomy of language regions of the human brain*. Academic Press.

Petrides, M., & Pandya, D. N. (2006). Efferent association pathways originating in the caudal prefrontal cortex in the macaque monkey. *Journal of Comparative Neurology*, *498*(2), 227–251.

Phan, M. L., Schendel, K. L., Recanzone, G. H., & Robertson, L. C. (2000). Auditory and visual spatial localization deficits following bilateral parietal lobe lesions in a patient with Balint's syndrome. *Journal of Cognitive Neuroscience*, *12*(4), 583-600.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature reviews neuroscience*, *21*(6), 322–334.

Pollok, B., Rothkegel, H., Schnitzler, A., Paulus, W., & Lang, N. (2008). The effect of rTMS over left and right dorsolateral premotor cortex on movement timing of either hand. *European Journal of Neuroscience*, *27*(3), 757–764.

Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology*, *8*(4), 516–521.

Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, *271*(1–2), 16–25.

Rauschecker, J. P. (2018). Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex*, *98*, 262–268.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, *35*(4), 355–370.

Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, *68*, 252–270.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012. *Psychonomic Bulletin & Review*, *20*, 403–452.

Rosanova, M., Casali, A., Bellina, V., Resta, F., Mariotti, M., & Massimini, M. (2009). Natural frequencies of human corticothalamic circuits. *Journal of Neuroscience*, *29*(24), 7679–7685.

Ruspantini, I., Mäki, H., Korhonen, R., D'Ausilio, A., & Ilmoniemi, R. J. (2011). The functional role of the ventral premotor cortex in a visually paced finger tapping task: a TMS study. *Behavioural Brain Research*, *220*(2), 325–330.

Samiee, S., Vuvan, D., Florin, E., Albouy, P., Peretz, I., & Baillet, S. (2022). Cross-frequency brain network dynamics support pitch change detection. *Journal of Neuroscience*, *42*(18), 3823–3835.

Sanders, L. L. O., Auksztulewicz, R., Hohlefeld, F. U., Busch, N. A., & Sterzer, P. (2014). The influence of spontaneous brain oscillations on apparent motion perception. *Neuroimage*, *102*, 241–248.

Sarubbo, S., Benedictis, A., Maldonado, I. L., Basso, G., & Duffau, H. (2013). Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain Structure and Function*, *218*, 21–37.

Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., & Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? *A Critical Discussion*. *Neuroscience*, *146*(4), 1435–1444.

Schubotz, R. I., Anwander, A., Knösche, T. R., Cramon, D. Y., & Tittgemeyer, M. (2010). Anatomical and functional parcellation of the human lateral premotor cortex. *Neuroimage*, *50*(2), 396–408.

Sestieri, C., Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., Olivetti Belardinelli, M., & Romani, G. L. (2006). What" versus "where" in the audiovisual domain: An fMRI study. *NeuroImage*, *33*(2), 672–680.

Siems, M., & Siegel, M. (2020). Dissociated neuronal phase-and amplitude-coupling patterns in the human brain. *NeuroImage*, *209*, 116538.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126.

Sokolov, A., Pavlova, M., Lutzenberger, W., & Birbaumer, N. (2004). Reciprocal modulation of neuromagnetic induced gamma activity by attention in the human visual and auditory cortex. *Neuroimage*, *22*(2), 521–529.

Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, *38*(3), 301–313.

Strigaro, G., Ruge, D., Chen, J. C., Marshall, L., Desikan, M., Cantello, R., & Rothwell, J. C. (2015). Interaction between visual and motor cortex: a transcranial magnetic stimulation study. *The Journal of Physiology*, *593*(10), 2365–2377.

Sugata, H., Yagi, K., Yazawa, S., Nagase, Y., Tsuruta, K., Ikeda, T., & Kawakami, K. (2020). Role of beta-band resting-state functional connectivity as a predictor of motor learning ability. *NeuroImage*, *210*, 116562.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 1–13.

Tadel, F., Bock, E., Niso, G., Mosher, J. C., Cousineau, M., Pantazis, D., & Baillet, S. (2019). MEG/EEG group analysis with Brainstorm. *Frontiers in Neuroscience*, *76*.

Tenke, C. E., & Kayser, J. (2012). Generator localization by current source density (CSD): implications of volume conduction and field closure at intracranial and scalp resolutions. *Clinical Neurophysiology*, *123*(12), 2328–2345.

Tesche, C. D., Uusitalo, M. A., Ilmoniemi, R. J., Huotilainen, M., Kajola, M., & Salonen, O. (1995). Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. *Electroencephalography and Clinical Neurophysiology*, *95*(3), 189–200.

Thut, G., Schyns, P. G., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology*, *2*, 170.

Tiihonen, J., Hari, R., Kajola, M., Karhu, J., Ahlfors, S., & Tissari, S. (1991). Magnetoencephalographic 10-Hz rhythm from the human auditory cortex. *Neuroscience Letters*, *129*(2), 303–305.

Tomasi, D., & Volkow, N. D. (2012). Laterality patterns of brain functional connectivity: gender effects. *Cerebral Cortex*, *22*(6), 1455–1462.

Tropini, G., Chiang, J., Wang, Z. J., Ty, E., & McKeown, M. J. (2011). Altered directional connectivity in Parkinson's disease during performance of a visually guided task. *Neuroimage*, *56*(4), 2144–2156.

Wang, R., Ge, S., Zommara, N. M., Ravienna, K., Espinoza, T., & Iramina, K. (2019). Consistency and dynamical changes of directional information flow in different brain states: A comparison of working memory and resting-state using EEG. *NeuroImage*, *203*, 116188.

Wang, W., Viswanathan, S., Lee, T., & Grafton, S. T. (2016). Coupling between theta oscillations and cognitive control network during cross-modal visual and auditory attention: supramodal vs modality-specific mechanisms. *PLoS One*, *11*(7), 0158465.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989-994.

Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in Psychology*, *2*, 73.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data–a practical approach. *Journal of Neuroscience Methods*, *250*, 34–46.

Wiesman, A. I., Silva Castanheira, J., & Baillet, S. (2022). Stability of spectral estimates in resting-state magnetoencephalography: Recommendations for minimal data duration with neuroanatomical specificity. *Neuroimage*, *247*, 118823.

Wu, C. C., Hamm, J. P., Lim, V. K., & Kirk, I. J. (2016). Mu rhythm suppression demonstrates action representation in pianists during passive listening of piano melodies. *Experimental Brain Research*, *234*, 2133–2139.

Xu, J., Hu, L., Qiao, R., Hu, Y., & Tian, Y. (2023). Music-emotion EEG coupling effects based on representational similarity. *Journal of Neuroscience Methods*, *398*, 109959.

Zalta, A., Petkoski, S., & Morillon, B. (2020). Natural rhythms of periodic temporal attention. *Nature Communications*, *11*(1), 1051.

Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is "where" in the human auditory cortex? *Nature Neuroscience*, *5*(9), 905–909.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558.

Zhou, B., Yang, S., Mao, L., & Han, S. (2014). Visual feature processing in the early visual cortex affects duration perception. *Journal of Experimental Psychology: General*, *143*(5), 1893.

Synapse Between the Two Studies

The previous chapter explored cortical auditory-motor functional connectivity dynamics across the frequency spectrum using rs-MEG and unveiled critical findings. Among these, we highlight a remarkable strength of intrinsic phase coupling between auditory and motor areas as compared to visual and motor areas, a pattern of increasing connectivity strengths across these areas, and intrinsic directions of information flow that are consistent with those observed in task-based studies. This notable alignment with task-based paradigms provides a bridge to Chapter 3, which introduces an EEG experiment that instead focuses on synchronization to external stimuli. Specifically, within a frequency band that does not entrain to stimulus periodicities as such but whose amplitude dynamics do align with stimulus periodicities, and which already played a prominent role in the previous chapter: the alpha band. Indeed, while most accounts in this context have focused on entraining frequencies like delta and theta, as well as time-locked processes relating to the beta band, desynchronization in the alpha band has become increasingly present in task-based sensory-motor coupling paradigms. In particular, the upcoming chapter will assess the degree to which learned auditory-motor associations in a musical context result in anticipatory mu-band suppression at the single-note level in non-musicians.

Chapter 3: Mu Suppression Reveals Auditory-Motor Predictions After Short Motor Training in Non-Musicians

Bedford, O., Ara, A., Albouy, P., Zatorre, R. J. & Penhune, V. (in preparation). *Mu Suppression Reveals Auditory-Motor Predictions After Short Motor Training in Non-Musicians*.

3.1 Front Page

Authors: Oscar Bedford[1,2,3], Alberto Ara[1,2,3], Philippe Albouy[2,3,4], Robert J. Zatorre[1,2,3], Virginia Penhune[2,3,5]

Affiliations: **1**. Montreal Neurological Institute, McGill University, Montréal, QC, Canada. **2**. International Laboratory for Brain, Music and Sound Research (BRAMS), Montréal, QC, Canada. **3**. Centre for Research on Brain, Language and Music (CRBLM), Montréal, QC, Canada. **4**. CERVO Brain Research Centre, School of Psychology, Université Laval, Québec City, QC, Canada. **5**. Department of Psychology, Concordia University, Montréal, QC, Canada

3.2 Abstract

Bidirectional coupling between the auditory and motor systems is an important brain mechanism that supports speech and music processes. Prior literature has shown the presence of motor activity during passive listening to known melodies, and our lab recently demonstrated with TMS that this activity is anticipatory, occurs in non-musicians, and can be elicited at the single note level after a single motor training session. However, the associated oscillatory dynamics remain unclear. Previous work using EEG has linked mu band (9-13Hz) suppression to auditory-motor coupling in musicians, but never in non-musicians, nor at the single note level. To this end, we recruited 24 non-musicians who underwent motor training of a simple melody, which was both preceded and followed by passive listening to the trained and untrained melodies. Based on an independent functional localizer, we demonstrated significant mu suppression immediately preceding each tone during passive listening to the trained compared to untrained melodies after learning. The effect was maximal over dorsal motor regions contralateral to the hand used. We found no significant effects in other time windows or frequency bands, including the beta band, indicating specificity of the response. These findings provide evidence that motor activity is predictive for learned auditory-motor sequences heard in passive listening contexts at the individual note level, and that it is independent of musical training. Moreover, our findings align with predictive coding and with the common coding theory, both of which posit that the motor system makes predictions about how current actions will produce specific sensory events.

3.3 Introduction

Auditory-motor coupling is the neural mechanism underlying the coordinated relationship between movements and sounds (Iversen & Balasubramaniam, 2016) which can be highly temporally precise (Repp, 2005). Information flow in this system is thought to be bidirectional, with motor regions sending top-down predictions to auditory centres to streamline perception (Patel & Iversen, 2014). Evidence in favour

of this idea can be found in studies like the one put forth by Lahav and colleagues (2007), where passive listening to previously learned melodies, but not unlearned melodies, was associated with activity in motor cortex and vPMC in non-musicians. However, an important limitation of fMRI activation-based studies such as the one by Lahav et al. (2007) is that they lack the temporal resolution to determine whether motor activity is predictive or reactive.

From a cognitive level, the common coding theory (Prinz, 1990) explains the bidirectional nature of the auditory-motor coupling system by positing that the planning of an action and the perception of the sensory consequences of said action are represented contiguously in the brain (Hommel, 2015). Thus, according to this theory, the brain operates on an internal model of the relationship between the body and the environment, containing interacting inverse and forward components (Wolpert et al., 1994). Historically, the embodied cognition framework only referred to inverse models to explain action-based effects on perception, thereby ascribing a passive role to the motor system. However, more current models propose that available sensory predictions of planned actions are projected onto the internal representation of the auditory stimulus (Halász & Cunnington, 2012). It follows that one of the roles of the motor system must be to generate forward models capable of making predictions about perception. The open question is whether the motor activity recorded by Lahav et al. (2007) and others reflects the workings of this forward modelling mechanism.

One candidate index for forward modelling that can be captured with time-sensitive methods like EEG is an oscillatory fluctuation pattern in the mu rhythm, a sub-component of the alpha band that has a spectral centroid between 9-13 Hz (Pineda, 2005) and is thought to involve large groups of pyramidal neurons in M1 (Niedermeyer, 1997). While normal mu activity reflects synchronized neural patterns which are prominent when the body is at rest, activity over M1 during self-movement is known to become asynchronous before action execution. This leads to reduced

amplitudes in the mu band, a fluctuation otherwise known as *mu suppression* (Pineda, 2005). Crucially, mu suppression also occurs during passive listening to sounds of familiar actions, lateralizing to the contralateral hemisphere at effector-specific locations within sensorimotor cortex (Pineda et al., 2013). Specifically, Wu and colleagues (2016) observed mu suppression in expert pianists undergoing passive listening of a previously practiced musical piece, but not untrained melodies. Importantly, mu suppression anticipated the presentation of each melody, indicating that this activity may indeed reflect forward modelling processes.

However, the study by Wu et al. (2016) could not resolve whether the observed mu suppression was due to extensive musical training or is an inherent property of human auditory-motor coupling. Moreover, this study did not utilize the proper timescale to establish whether mu suppression precedes note onsets, which is critical for determining whether this measure is truly an index of forward modelling (Novembre & Keller, 2014). In contrast, Stephan and colleagues (2018) used single-pulse TMS on non-musicians to demonstrate that anticipation of upcoming tones during passive listening to a practiced melody reduces the threshold for motor activation, which is evidence that the corresponding fingers are preactivated, and which indicates predictive motor preparation at the effector level. Thus, their finding provided solid evidence that the motor system of non-musicians is indeed capable of anticipating previously embodied sounds on a note-by-note basis, even when motoric output is neither required nor executed. However, Stefan et al. (2018) were not able to determine the correspondence between their TMS result and the mu suppression observed by others using EEG.

In light of these gaps in the literature, the current study formulated two aims. First, to assess the fast temporal dynamics of motor preparation as indexed by mu suppression, at the single-note level, within a non-musician population, in a brief learning context. Second, to create a data-driven functional localizer that makes use of the sensor-level EEG data recorded during the motor training condition in order to

better establish the presence of mu suppression in passive listening conditions. In addition, we outlined three distinct hypotheses. First, that our functional localizer would identify mu suppression occurring in the active condition within frequency, time, and channels that correspond with theoretical delineations of the mu rhythm (Fox et al., 2016; Pineda et al., 2013). Second, that the functional localizer coordinates would detect passive mu suppression occurring at the single-note level only during post-training exposure to the learned melody, in line with findings in musicians (Wu et al., 2016). Third, that suppression would be absent in other frequency bands, making it mu-band specific.

3.4 Methods

3.4.1 Participants

We recruited 27, healthy, right-handed non-musicians. Three participants were excluded due to poor performance (n=1) or poor EEG data quality (n=2). Thus, the final sample consisted of 24 participants (13 females) between the ages of 18 and 35 years (mean: 22.6 ± 4.6 years). Consistent with previous studies, non-musicians were defined as those with less than 3 years of lifetime musical training and/or experience and reported no active music-making within the 3 years prior to the study (Slater & Kraus, 2016), based on information from the Goldsmiths Musical Sophistication index (Müllensiefen et al., 2014). Exclusion criteria included being left-handed and having a history of psychiatric or neurological disorders. All participants gave written informed consent prior to the start of the experimental session. The study protocol conformed to the principles of the Declaration of Helsinki (World Medical Association, 2001) and was approved by the McGill Human Research Ethics Board.

3.4.2 Task and Stimuli

The goal of this experiment was to assess the degree of mu suppression during passive listening to learned melodies. EEG was recorded across the experiment,

which featured three conditions summarized in Figure 1 below: Pre-training passive listening (PRE); Melody playback training (TRAIN) and Post-training passive listening (POST). In the PRE and POST conditions participants heard two types of melodies: 1) the trained melody (Trained); and 2) untrained melodies (Untrained). In the TRAIN condition participants either listened to and played back to blocks containing repetitions of the Trained melody (Regular) or catch blocks containing repetitions of the reversed Trained melody (Reverse).

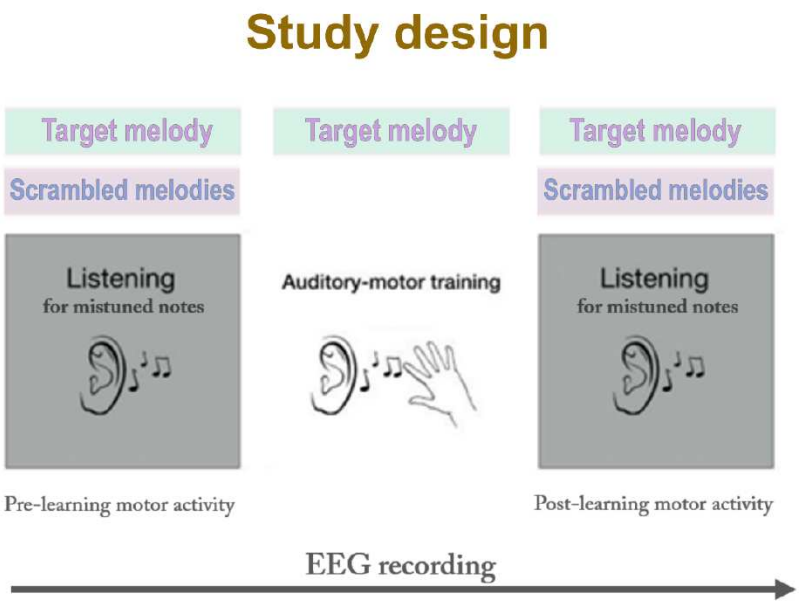


Figure 1. Study design

***PRE (left column)**: subjects underwent one passive listening block of the Trained melody (Target melody in the figure) and one passive listening block of the Untrained melodies (Scrambled melodies in the figure) while listening for mistuned notes. **TRAIN (middle column)**: subjects underwent auditory motor training of the Trained melody (Target melody in the figure) over blocks where it was presented normally (Regular) and blocks in which it was presented in reverse order (Reverse). **POST (right column)**: identical to PRE except for the number of mistuned notes.*

Forty-one unique 8-note melodies were created using the four notes in a C-major arpeggio which corresponded to the four digits of the right hand when played back on the piano-like keyboard (C4 (259 Hz) = index, E4 (329 Hz) = middle; G4 (389 Hz) = ring; C5 (531 Hz) = pinkie). Melodies were 12 seconds long, featuring a fixed tone duration of 600ms and an inter- stimulus interval of 700ms for a total onset-to-onset asynchrony of 1500ms. Melodies were constructed such that each note was

presented twice and no notes were repeated sequentially (Stephan et al., 2018). The Trained melody is represented in Figure 2 below:

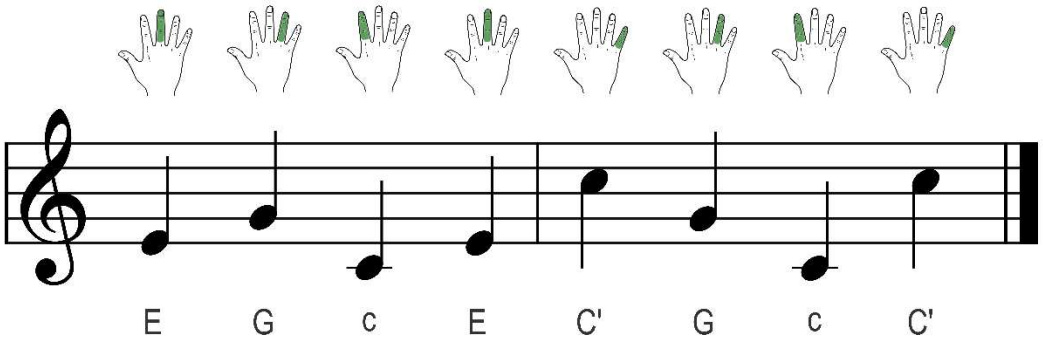


Figure 2. Trained melody sequence
The Trained melody is based on a C-major arpeggio. Each pitch is associated to one of the fingers of the right hand and repeats twice. The sequence has isochronous timing and undefined meter.

All sounds were delivered via E-A-RTONE 3A foam-tipped insert-type earbuds (E-A-R® Auditory Systems). Sounds were synthesized with Adobe® Audition v.3.0 (Adobe Systems Incorporated, 2007), all musical tones were rendered using a piano timber, and every melody presentation was cued by a single woodblock sound.

3.4.3 PRE and POST passive listening conditions

During the PRE and POST passive listening conditions, participants listened to one block of 40 repetitions of the Trained melody and one block of 40 unique Untrained melodies (12.7 minutes per block). As an attentional control, seven catch trials were introduced in each condition in which the last note of the melody was mistuned to a C#5 note (554 Hz) instead of the natural C5 note (531 Hz). Participants were required to silently count the mistuned notes and report them at the end of each block. The presentation order of the Trained and Untrained blocks was the same across passive listening conditions, but it was counter-balanced across participants. Inter-trial intervals (ITIs) for melodies presented in the passive listening blocks varied randomly between a 3.3–3.7 second range (0.1 second step size) in order to prevent the anticipation of melody onsets.

3.4.4 Melody Playback Training (TRAIN)

The TRAIN condition contained 15 blocks (10 trials each) of practice on the Trained melody, and 2 blocks (10 trials each) of practice on the Reverse melody for a total of 17 blocks. Reverse blocks corresponded to the 3rd and 15th block positions within the TRAIN condition, while all other blocks were Regular. The ITI for all blocks was fixed at 3.5 seconds. On each trial, participants heard the melody through headphones and were asked to reproduce the notes by pressing the four keys on a purpose-built piano-like keyboard. Notes produced by the participants' key presses were played back through the headphones one octave in pitch below the target melody presentation to provide discernible auditory feedback. Participants were instructed to synchronize their responses as accurately as possible to the melodies, and no further instructions or performance feedback was provided. Keypresses activated a switch on the keyboard that sent TTL pulses from a parallel port connected to a Beaglebone® Black microcomputer with a Bela real-time module (Thangaraj, 2016), based on the system introduced by Zappi & McPherson (2014). Pulses were received by the recording computer with minimal delay and constant jitter and timings were used as individual event markers for EEG.

3.4.5 Behavioural Data Analysis

Individual performance in the motor training task (TRAIN) was evaluated based on the percentage of correct key presses and the mean absolute timing difference between key presses and target tone onsets within each of the 17 training blocks. A key press was considered accurate if it matched the presented tone within a ±300ms window and was executed without overlapping with other key presses. Reaction times were calculated only for correct key presses. If missing values for average reaction times occurred within a training block, they were interpolated using the block average from all other participants.

3.4.6 EEG Data Collection

EEG recordings were collected using a 64-channel active electrode cap (actiCAP, Brain Products GmbH, Gilching, Germany), a unipolar EEG amplifier (BrainAmp, Brain Products GmbH, Gilching, Germany), and designated recording software (BrainVision Recorder, Version 1.20.0801, Brain Products GmbH, Gilching, Germany, 2019), all from the same manufacturer. The sampling rate was set to 1 kHz, the ground electrode was left at the default FPz location, and the reference electrode was placed on the right canthus in order to monitor eye blinks. Cap electrodes were located at 60 standard positions according to the 10-10 system (Chatrian et al., 1988; Fp1/2, AFz, AF3/4, AF7/8, Fz, F7/8, F5/6, F3/4, F1/2, FCz, FT7/8, FC5/6, FC3/4, FC1/2, Cz, T7/8, C5/C6, C3/4, C1/2, CPz, TP7/8, CP5/6, CP3/4, CP1/2, Pz, P7/8, P5/6, P3/4, P2/1, POz, PO7/8, PO3/4, Oz, O1/2).

Electrodes corresponding to channels FT9, FT10 were respectively placed on the left outer canthus and right cheek and used to capture horizontal and vertical electrooculography recordings with which to monitor saccadic eye movements and blinks during online recordings, as well as to remove ocular artifacts during offline data preprocessing. Electrodes corresponding to channels TP9 and TP10 were respectively placed on the left and right mastoid locations to be used as re-reference sites for offline data preprocessing. During the EEG recording, participants were instructed to fix their gaze onto an onscreen fixation cross. Impedance levels were assessed at the end of each block in the PRE and POST conditions, as well as halfway through the TRAIN condition, and continuously kept below 10 kΩ. During passive listening, participants were instructed to remain as still as possible, and their right hand was monitored to ensure that no movement contaminated the EEG recording.

3.4.7 EEG Pre-Processing

EEG data was pre-processed using a combination of Brainstorm (Tadel et al., 2011) and MATLAB R2021a version 9.10.0 (The MathWorks Inc, 2022) custom scripts, and

consisted of standard preprocessing and artifact removal steps. First, EEG signals were bandpass-filtered between 1-100 Hz, notch-filtered at 60 Hz and associated harmonics, and re-referenced to bilateral mastoids. Noisy segments and channels were identified and removed from the EEG signal recordings via visual inspection (Gross et al., 2013). To aid in the detection and removal of artifacts and bad channels, power spectrum density (PSD) plots were computed at successive stages of pre-processing. A total of 29 faulty channels were removed across 11 participants, with a maximum of 8 faulty channels removed for any given subject (mean: 2.64 ± 2.11 channels). Repetitive eye blink artifacts were removed for all subjects using Brainstorm's built-in signal space projection (SSP) method (Tesche et al., 1995), and saccade artifacts were additionally removed for a total of 3 subjects.

3.4.8 EEG Data Analysis

3.4.8.1 Time-Frequency Analysis

After preprocessing, data were downsampled to 200 Hz and epoched from -1400 to 200 milliseconds relative to tone onsets to account for the approximately ± 200 milliseconds of edge artifacts introduced by convolution. Individual EEG epochs were convolved over a linearly scaled frequency range of 1-50 Hz using a family of Gaussian-tapered scaling Morlet wavelets, with the mother wavelet centred at 1 Hz and an initial full-width half maximum (FWHM) of 3s. This approach yielded time-frequency magnitude estimates. Time-frequency plots were generated from individual trial epochs and subsequently averaged within each subject and condition. This approach was chosen because it yields 'total magnitude' estimates, which capture both evoked and induced components (Cohen, 2014), as we had no prior hypothesis regarding whether mu-suppression would be phase-locked or non-phase-locked to individual tone onsets. All resulting time-frequency plots were normalized via average baseline correction using Brainstorm's method (Tadel et al., 2011), which converts raw magnitude estimates to z-scores (formula: $x\_std = (x - \mu) / \sigma$). The baseline segment from -1400 to 100 milliseconds

relative to tone onset was selected to encompass the entire interval between successive tones.

3.4.8.2 3D Functional Localizer

We created a mask using a 3D functional localizer of active mu suppression based on correctly performed trials from the TRAIN condition compared to rest. The function of this mask was to help isolate channels, frequencies, and timepoints likely to exhibit mu suppression during passive listening to the Trained melody. For each subject, we selected the two most accurate blocks from the last four Regular blocks of the TRAIN condition. Initially, 160 time-frequency plots (10 melodies * 8 notes * 2 blocks) were extracted, which were subsequently filtered to remove plots associated with poor performance, such as incorrect key presses or excessive key press latency. Time-frequency plots were then generated for each successful trial and averaged to produce one average time-frequency plot per participant, resulting in a total of 24 plots for the sample. Each plot contained data from 64 channels, across 50 frequencies, and 321 timepoints (each representing a 5ms interval).

Resting data were extracted from the intervals between melody presentations across the two PRE passive listening blocks. Because between-melody periods ranged from 3.3s to 3.7s it was possible to extract two resting trials per period while maintaining a gap of at least -150ms before the onset of the next melody. For the shorter intervals (3.3s, 3.4s, and 3.5s), the first resting window overlapped with activity related to the offset of the previous tone by either -1500ms, -1450ms, or -1400ms with respect to current tone onset. However, these cases accounted for only 6 out of the 160 resting trials (3.75%), so any potential contamination was mitigated through averaging. Moreover, we did not analyse activity prior to -1000ms relative to tone onset, as we had no hypotheses regarding tone offset periods. For each subject, 160 time-frequency plots were generated for each epoch and subsequently averaged to produce one time-frequency plot per participant, resulting

in 24 plots for the full sample. Each plot contained data from 64 channels, across 50 frequencies, and 321 timepoints (5ms interval per timepoint).

The 24 resting time-frequency plots were contrasted with the 24 active motor training plots using a cluster-based permutation t-test (one-tailed), with the alternative hypothesis being that active datapoints would have significantly lower baseline-corrected activation values than rest datapoints. This test was implemented using FieldTrip's native 'ft_freqstatistics' function within Brainstorm (Oostenveld et al., 2011; Tadel et al., 2011). To control for multiple comparisons, the cluster permutation test was configured to run 1000 Monte Carlo permutations, safely above the recommended minimum of 800 for EEG data (Candia-Rivera & Valenza, 2022; Pernet et al., 2015). The significance level for both the t-test and the cluster permutation procedure was set at the conventional alpha threshold of 5% ($p = 0.05$). No averaging was performed across dimensions. Regions of interest (ROIs) were allowed to vary independently in time and frequency, as long as clusters retained contiguity, which was defined as one time-frequency pixel connected to a neighbouring channel.

The resulting group-average t-map was subsequently 'thresholded' to retain only pixels with an associated p-value of less than 0.05. This 'thresholded' t-map was then applied as a binary mask to each subject's passive listening time-frequency plots. Namely, the mask was divided into 5 ROIs (details in subsection 3.5.2.1) and data were aggregated by averaging the nonzero z-score values within each ROI's constituent pixels, resulting in a single value per ROI, channel, condition, and subject. Finally, the scalar values for each ROI were then averaged across subjects, producing a group-averaged value for each channel and condition.

3.4.9 Statistical Analysis

We used a Generalized Linear Mixed Model (GLMM) approach for all statistical analyses because it allows for the estimation of random effects associated with

individual differences within the grouping variable (Stroup, 2012). Statistical analyses were conducted with the 'glmmTMB' package (Brooks et al., 2017) in RStudio version 2022.02.0 (Posit team, 2022).

3.4.9.1 Behavioural Performance (TRAIN)

We conducted two statistical analyses to ascertain that subjects had successfully learned the target melody sequence. The first was designed to test for learning across the 15 Regular blocks and consisted of a one-factor within-subjects design across two separate GLMM analyses: one featuring 'Accuracy' and the other featuring 'Reaction Time' as the dependent variable. In both of these GLMMs, we implemented a single factor within-subjects design with 'Trial' as the within-subjects fixed factor. This factor featured 15 levels, each consisting of one of the 15 Regular blocks. The same 'Trial' factor was included as part of the random effects structure in both GLMMs, which additionally allowed for random slopes and intercepts to be calculated, and 'subject' was used as grouping variable (1 + Trial || subject).

The second statistical analysis was designed to test for sequence-specific learning and consisted of a two-factor within-subjects design across two separate GLMM analyses: one featuring 'Accuracy' and the other featuring 'Reaction Time' as the dependent variable. For Regular blocks, scores representing the beginning and end of the motor training phase were respectively derived from the average of the 2nd and 14th blocks of the TRAIN condition within each participant, as these blocks preceded the two Reverse blocks. For the Reverse blocks, the average of each block (3rd and 15th blocks of the TRAIN condition) was used to represent each participant's scores. In both of these GLMMs, we implemented a 2x2 within-subjects design with 'Block' (Regular vs Reverse) and 'Time' (beginning vs end of training) as the within-subjects fixed factors (Block * Time). The same two factors were included as part of the random effects structure in both GLMMs, which additionally allowed for random slopes and intercepts to be calculated, and 'subject' was used as grouping variable (1 + Trial || subject).

Across both statistical analyses, accuracy data followed a bimodal, non-normal distribution. Therefore, we converted the percentages to proportions and used a Beta-family distribution with a logit link function to better parameterize its shape. Conversely, across both statistical analyses, Reaction Time data aligned well with the theoretical gamma distribution indicated for the analysis of reaction times (McGill & Gibbon, 1965). Therefore, we selected a gamma distribution with an inverse link function to parameterize its shape. Across all four GLMMs, omnibus tests were computed from the model with F statistics, and FDR-corrected t-statistics were computed with paired-samples t-tests on the prediction scale. All resulting metrics were back-transformed to parameter estimates in the response scale.

3.4.9.2 EEG Activity during Passive Listening (PRE and POST)

In order to determine the degree of frequency band-specific suppression across all four passive listening blocks, the group-averaged z-score values across channels and conditions was chosen as the dependent variable. All values followed a normal distribution, as expected for averaged data, so the dependent variable was parameterized as gaussian. Omnibus tests were computed from the model with F statistics, and FDR-corrected t-statistics were computed with paired-samples t-tests directly on the response scale. A total of five GLMM designs with identical structure were implemented across each of the five ROIs. Namely, we implemented a 2x2 within-subjects design with 'Condition' (Learn vs Rand) and 'Time' (Pre vs Post) as fixed effect factors (Condition*Time). Note that the Condition factor's level names Learn and Rand respectively correspond to the Trained and Untrained melodies described above. The same two factors were included as part of the random effects structure, which additionally allowed for random slopes and intercepts, and 'channel' was used as the grouping variable (1 + Condition + Time || channel).

3.5 Results

3.5.1 Behavioural Performance

3.5.1.1 Accuracy

In GLMM1, which tested for a main effect of learning, we obtained a significant main effect of 'Trial' ($F(14,329 = 46.022$; p-value < 0.0001), indicating globally higher accuracy scores as training progressed (Figure 3). Specifically, FDR-corrected post-hoc tests revealed that all 6 blocks at the end of training, excluding the second catch block, displayed significantly higher accuracies than all 6 blocks at the start of training, excluding the first catch block (Supplementary Table 1).

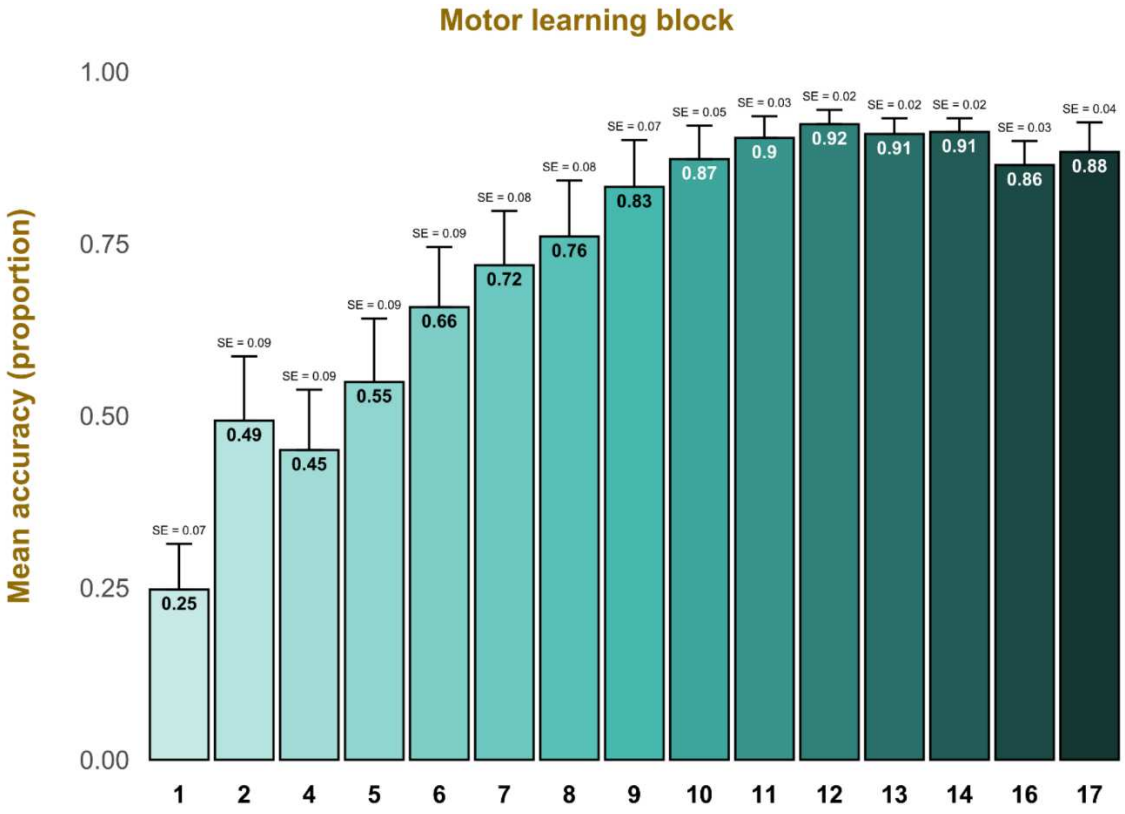


Figure 3. GLMM1: Group-average training accuracy as a function of block (Regular only)
Overall accuracy scores improved throughout the TRAIN condition, while variability (SE) diminished. Blocks 3 and 15 are excluded because they correspond to the Reverse blocks of the TRAIN condition.

In GLMM2, which tested for sequence-specific learning (Figure 4), we obtained significant main effects for both fixed effect factors 'Block' ($F(1,88) = 64.842$;

p-value < 0.0001) and 'Time' (F(1,88) = 35.125; p-value < 0.0001), as well as the global interaction term for these two factors (F(1,88) = 14.376; p-value = 0.0003).

FDR-corrected post hoc contrasts (Table 1) revealed significantly higher accuracies within the Interaction term for Regular block #14 as compared to Regular block #2 (t(88) = 5.350; corrected p-value < 0.0001), as well as Reverse block 1 (t(88) = 17.550; corrected p-value < 0.0001) and Reverse block 2 (t(88) = 12.410; corrected p-value < 0.0001) blocks. In other words, the end of training displayed significantly higher accuracies than the other three conditions, indicating that the participants learned the specific target melody sequence appropriately. Moreover, Reverse block 1 was associated with significantly lower accuracy scores than Regular block #2 (t(88) = -4.900; corrected p-value < 0.0001) as well as Reverse block 2 (t(88) = -3.720; corrected p-value = 0.0002). This indicated that participants also learned the finger-to-tone mapping throughout the experiment and were able to apply it when the melody was reversed a second time. In addition, Regular block #2 was associated with significantly higher accuracy scores than Reverse block 2 (t(88) = 2.570; corrected p-value = 0.0118).

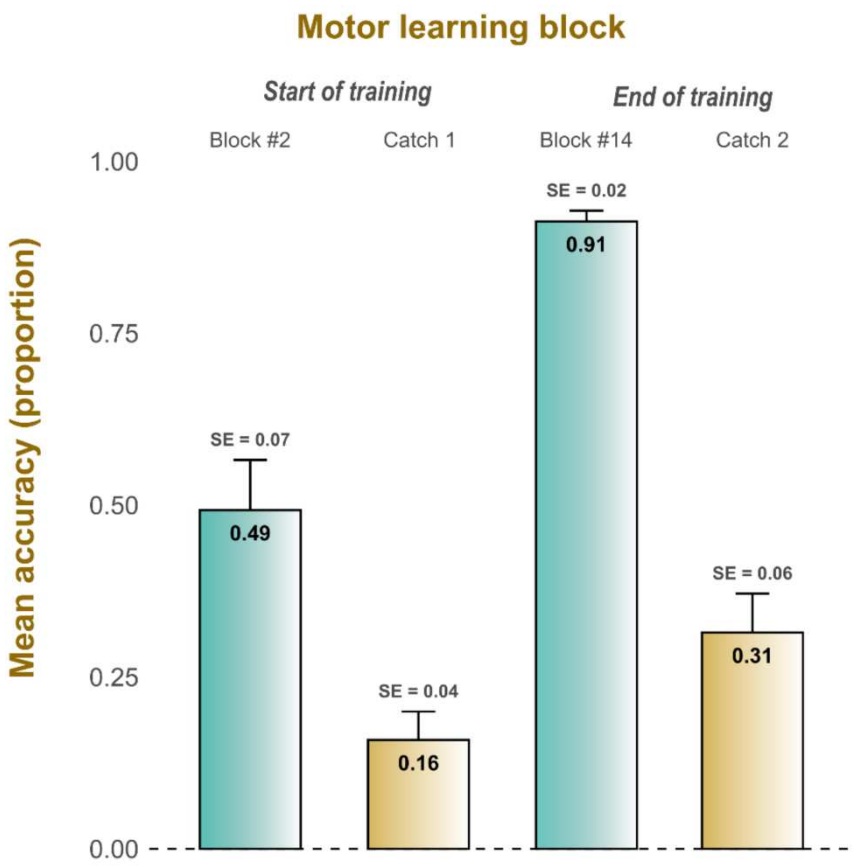


Figure 4. GLMM2: Sequence-specific learning

*Accuracy scores for **Regular blocks #2 and #14** (green) indicate motor learning from beginning to end of the TRAIN condition. Accuracy scores for **Reverse blocks 1 and 2** (Catch blocks in the figure; yellow) are significantly lower than for Regular blocks, indicating sequence-specific learning. The difference between Reverse blocks 1 and 2 is significant and indicates independent sound-to-finger mapping learning.*

Table 1. GLMM2 post-hoc tests: '*Block*' by '*Time*' (*DV = accuracy scores*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Catch End - Reg End	-0.65	0.05	88	-12.410	<.0001	**<.0001**
Catch End - Catch Start	0.17	0.05	88	3.720	0.0004	**0.0005**
Catch End - Reg Start	-0.21	0.08	88	-2.570	0.0118	**0.0118**
Reg End - Catch Start	0.82	0.05	88	17.550	<.0001	**<.0001**
Reg End - Reg Start	0.44	0.08	88	5.350	<.0001	**<.0001**
Catch Start - Reg Start	-0.38	0.08	88	-4.900	<.0001	**<.0001**

3.5.1.2 Reaction Times

In GLMM3, which used the same structure as the prior GLMM1 to test for a main effect of learning, we obtained a significant main effect of 'Trial' ($F(14,329) = 18.391$; p-value < 0.0001), indicating globally lower latencies as training progressed. Specifically, FDR-corrected post-hoc tests revealed that all 6 finishing blocks, excluding the second catch block, displayed significantly lower latencies than all 6 starting blocks, excluding the first catch block (Supplementary Table 2).

Group-averaged reaction times in the second Regular block at the start of motor training was 0.14s (SE = 0.01s), whereas group-averaged reaction times in the fourteenth Regular block at the end of training was 0.11s (SE = 0.01s). The group-averaged reaction times in the first Reverse block was 0.18s (SE = 0.01s) and the group-averaged reaction times in the second Reverse block was 0.14s (SE = 0.01s).

In GLMM4, which compared the scores for these four conditions, we obtained significant main effects for both fixed effect factors 'Block' ($F(1,88) = 20.672$; p-value < 0.0001) and 'Time' ($F(1,88) = 36.075$; p-value < 0.0001). The global interaction term for these two factors was not statistically significant ($F(1,88) = 1.650$; p-value = 0.4528).

FDR-corrected post hoc contrasts (Table 2) revealed significantly lower latencies within the Interaction term for Regular block #14 as compared to Regular block #2 ($t(88) = -4.150$; corrected p-value = 0.0004), Reverse block 1 ($t(88) = -6.720$; corrected p-value < 0.0001), and Reverse block 2 ($t(88) = -3.510$; corrected p-value = 0.0016). In other words, the end of training displayed significantly lower latencies than the other three conditions, indicating that participants learned the specific target melody sequence appropriately. Moreover, Reverse block 1 was associated with significantly higher latency scores than Regular block #2 ($t(88) = 6.720$; corrected p-value < 0.0001) and Reverse block 2 ($t(88) = 4.030$; corrected p-value = 0.0004). This indicated that participants also learned the finger-to-tone mapping throughout

the experiment and were able to apply it when the melody was reversed a second time. There was no statistical difference between Regular block #2 and Reverse block 2 (t(88) = 0.660; corrected p-value = 0.5140).

Table 2. GLMM4 post-hoc tests: '*Block*' by '*Time*' (*DV = latency scores*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Catch End - Reg End	0.03	0.01	88	3.510	0.0011	0.0016
Catch End - Catch Start	-0.04	0.01	88	-4.030	0.0002	0.0004
Catch End - Reg Start	-0.01	0.01	88	-1	0.514	0.514
Reg End - Catch Start	-0.07	0.01	88	-6.720	<.0001	<.0001
Reg End - Reg Start	-0.03	0.01	88	-4.150	0.0002	0.0004
Catch Start - Reg Start	0.03	0.01	88	3.290	0.0017	0.002

3.5.2 Brain Activity

3.5.2.1 3D Functional Localizer

The cluster permutation analysis yielded a functional localizer t-map containing 51 channels, 50 frequency bins, and 321 timepoints (each representing a 5ms window). Channels FT9, FT10, TP9, and TP10 had been used as external electrodes and did not show any activity, so they were excluded along with non-surviving channels due to absence of pertinent activity, leaving a total of 47 channels for analysis (Figure 5).

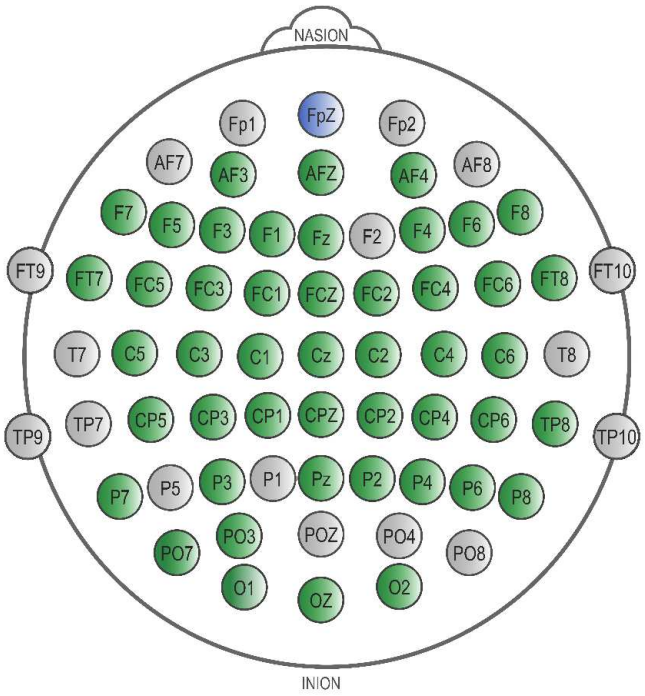


Figure 5. Topography of the functional localizer

Green *indicates channels that survived the cluster-permutation test (n=47).* ***Grey*** *indicates the 17 channels that either did not survive the cluster-permutation test (n=13) or that did not show any activity (n=4).* ***Blue*** *indicates the ground channel.*

Two significant 3D clusters survived the cluster-based permutation test (p = 0.000999, 1000 permutations) when comparing active and resting conditions. The first cluster (cluster-level statistic = -232759.2563, size = 83,729 pixels) spanned all 47 channels, covering the frequency range of 15-50 Hz (channel PO7) and the time window from -550 milliseconds to 175 milliseconds (channel P7). This cluster was manually divided into two 3D regions of interest (ROIs), spectrally corresponding to the beta (47 channels; 16-30 Hz) and gamma (47 channels; 31-50 Hz) frequency bands (Figure 6 – example in channel C3). The second cluster (cluster-level statistic = -140321, size = 46,200 pixels) spanned all 47 channels, frequencies from 1-15 Hz (channel F5), and timepoints from -1100 milliseconds to 0 milliseconds (channel PO3). This cluster was manually subdivided into three distinct 3D ROIs, spectrally corresponding to the delta-theta (47 channels; 1-8 Hz), early mu (45 channels; 9-15 Hz), and late mu (31 channels; 9-15 Hz) frequency bands (Figure 6 – example in channel C3).

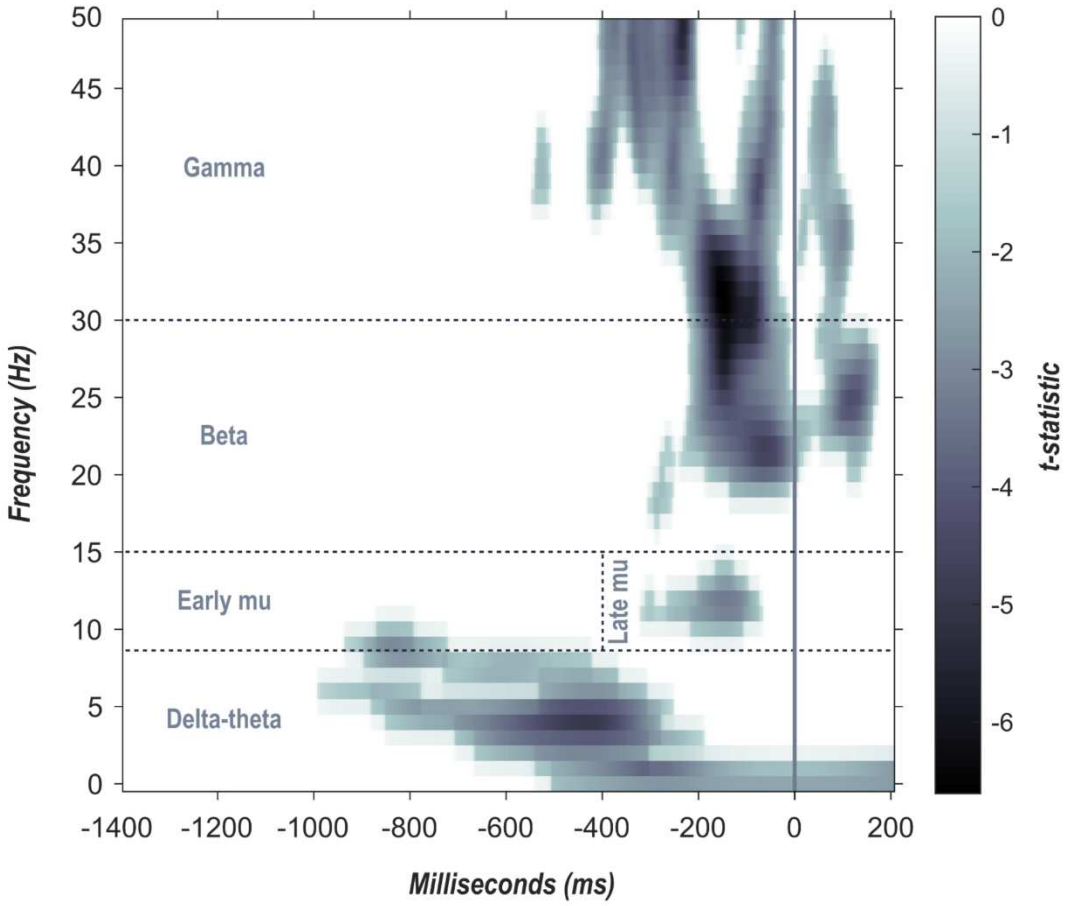


Figure 6. Spectrotemporal distribution of clusters and partitions into ROIs (channel C3)
*Non-white activity indicates pixels that 1) were associated with significantly lower baseline-corrected values in group-averaged active data as compared to rest data; and 2) survived standard thresholding ($p < 0.05$). Surviving pixels were masked equally onto the passive data in binary fashion, regardless of t-statistic value. The two natural clusters include all activity above 15Hz (**Cluster 1**) and all activity below 15Hz (**Cluster 2**). Cluster 1 was partitioned into **Beta** (15-30Hz) and **Gamma** (31-50Hz) ROIs. Cluster 2 was partitioned into **Delta-theta** (1-8Hz), **Early mu** (9-15Hz) and **Late mu** (9-15Hz) ROIs. The temporal distribution of these ROIs varied slightly in each channel, but spectral delineations did not.*

The late mu ROI was interpreted as capturing motor preparation before tone onset, based on its distinct, isolated location within cluster 2. Moreover, it appeared in a reduced subset of channels, was left-lateralized, and showed a decreasing number of activated pixels from the left-most channels toward the midline and the first row of right-hemisphere channels (Figure 7).

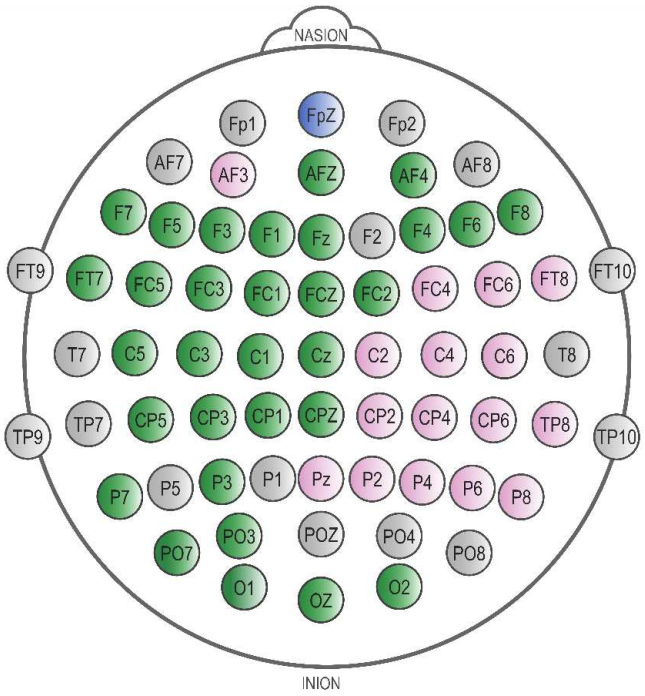


Figure 7. Topography of the Late Mu ROI

Green *indicates channels that survived the cluster-permutation test and show late mu activity (n=31).* ***Pink*** *indicates the surviving channels that did not show any late mu activity but did show activity in other ROIs (n=16).* ***Grey*** *indicates the 17 channels that either did not survive the cluster-permutation test (n=13) or that did not show any activity (n=4).* ***Blue*** *indicates the ground channel.*

3.5.2.2 EEG Activity during Passive Listening blocks

Late mu

In the late mu ROI, we obtained significant main effects for both fixed effect factors 'Condition' ($F(1,116) = 11.176$; p-value = 0.0011) and 'Time' ($F(1,116) = 23.679$; p-value < 0.0001). The interaction term for these two fixed effect factors was not statistically significant ($F(1,116) = 0.236$; p-value = 0.6283). FDR-corrected post hoc tests revealed significantly lower activity within 'Condition' for Learn as compared to Rand ($t(116) = -3.340$; corrected p-value = 0.0011), as well as significantly lower activity within 'Time' for Post as compared to Pre ($t(116) = -4.870$; corrected p-value < 0.0001).

FDR-corrected post hoc contrasts (Table 3) additionally revealed significantly lower activity in the interaction between 'Learn_Post' with 'Learn_Pre' ($t(116) = -3.100$;

corrected p-value = 0.0048), 'Rand_Pre' (t(116) = -5.750; corrected p-value < 0.0001), and 'Rand_Post' (t(116) = -2.130; corrected p-value = 0.0420). In other words, the 'Learn_Post' passive listening condition displayed significantly more mu suppression than the other three conditions. As well, 'Rand_Pre' exhibited significantly more activity than 'Rand_Post' (t(116) = -3.780; corrected p-value = 0.0006) and 'Learn_Pre' (t(116) = -2.790; corrected p-value = 0.0093). Finally, 'Learn_Pre' and 'Rand_Post' were not significantly different from each other (-t(116) = -0.830; corrected p-value = 0.4080). Taken together, these results indicate that the effect that was hypothesized to occur only in the mu band was indeed present within the late mu ROI, as passive listening to the target melody after motor training revealed significantly more mu suppression than in the other three passive listening blocks (Figure 8).

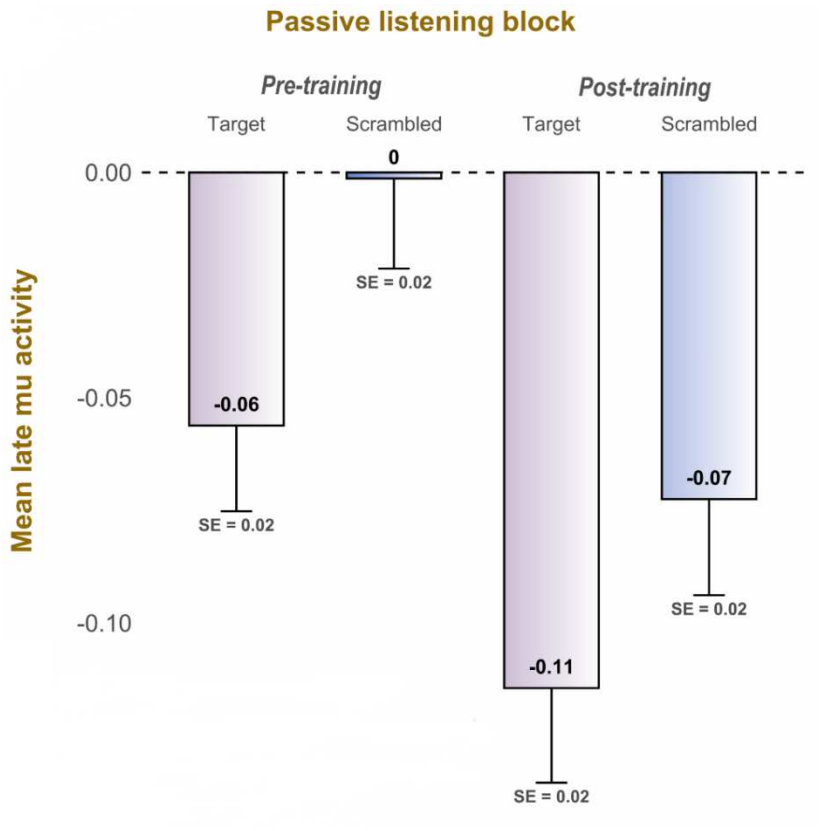


Figure 8. GLMM5: Channel-averaged late mu suppression across passive listening blocks

Z-scores representing the difference between magnitude and baseline in each pixel of the late mu ROI were averaged across subjects within each passive listening condition and for each channel separately. Figure 8 shows the means and variation (SE) after then averaging across channels within each condition. The Training POST melody block (Post-training Target in the Figure) displayed significantly more late mu suppression than the other three passive listening blocks.

Table 3. Late mu ROI's GLMM post-hoc tests: '*Condition*' by '*Time*' (*DV = mean activity*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Learn Post - Rand Post	-0.04	0.02	116	-2.130	0.035	0.042
Learn Post - Learn Pre	-0.06	0.02	116	-3.100	0.0024	0.0048
Learn Post - Rand Pre	-0.11	0.02	116	-5.750	<.0001	<.0001
Rand Post - Learn Pre	-0.02	0.02	116	-1	0.408	0.408
Rand Post - Rand Pre	-0.07	0.02	116	-3.780	0.0002	0.0006
Learn Pre - Rand Pre	-0.05	0.02	116	-2.790	0.0062	0.0093

Early mu

In the early mu ROI, we obtained significant main effects for both fixed effect factors 'Condition' (F(1,172) = 13.867; p-value = 0.0003) and 'Time' (F(1,172) = 11.086;

p-value = 0.0011), as well as the interaction term for these two factors (F(1,172) = 22.092; p-value < 0.0001).

FDR-corrected post hoc contrasts (Table 4) revealed significantly higher activity within the Interaction term for 'Learn_Post' only as compared to 'Rand_Pre' (t(172) = 4.970; corrected p-value < 0.0001). In other words, the 'Learn_Post' passive listening condition was not statistically different than the 'Learn_Pre' (t(172) = -0.560; corrected p-value = 0.7601) or 'Rand_Post' (t(172) = -0.480; corrected p-value = 0.7601) conditions. As well, 'Rand_Pre' exhibited significantly less activity than 'Rand_Post' (t(172) = 5.590; corrected p-value < 0.0001) and 'Learn_Pre' (t(172) = 5.930; corrected p-value < 0.0001). Finally, 'Learn_Pre' and and 'Rand_Post' were not significantly different from each other (t(172) = -0.100; corrected p-value = 0.9235). Taken together, these results indicate that the effect that was hypothesized to occur only in the mu band was not present within the early mu ROI, as passive listening to the target melody after motor training revealed enhancement, as opposed to suppression, which was additionally greater than in the other three passive listening blocks (Figure 9 – top right).

Table 4. Early mu ROI's GLMM post-hoc tests: '*Condition*' by '*Time*' (*DV = mean activity*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Learn Post - Rand Post	0	0.01	172	0	0.6334	0.7601
Learn Post - Learn Pre	-0.01	0.01	172	-1	0.5775	0.7601
Learn Post - Rand Pre	0.05	0.01	172	4.970	<.0001	**<.0001**
Rand Post - Learn Pre	0	0.01	172	0	0.9235	0.9235
Rand Post - Rand Pre	0.06	0.01	172	5.590	<.0001	**<.0001**
Learn Pre - Rand Pre	0.06	0.01	172	5.930	<.0001	**<.0001**

Beta

In the beta ROI we obtained significant main effects for both fixed effect factors 'Condition' (F(1,180) = 17.916; p-value < 0.0001) and 'Time' (F(1,180) = 35.136; p-value < 0.0001), as well as the interaction term for these two factors (F(1,180) = 20.165; p-value < 0.0001). FDR-corrected post hoc contrasts (Table 5) revealed significantly lower activity within the Interaction term for 'Learn_Post' only

as compared to 'Rand_Pre' (t(180) = -7.220; corrected p-value < 0.0001). In other words, the 'Learn_Post' passive listening condition did not display significantly more beta suppression than 'Learn_Pre' (t(180) = -1.540; corrected p-value = 0.1811) or 'Rand_Post' (t(180) = -0.050; corrected p-value = 0.9569). As well, 'Rand_Pre' exhibited significantly more activity than 'Rand_Post' (t(180) = -7.420; corrected p-value < 0.0001) and 'Learn_Pre' (t(180) = -6.160; corrected p-value < 0.0001). Finally, 'Learn_Pre' and 'Rand_Post' were not significantly different from each other (t(180) = -1.440; corrected p-value = 0.1811). Taken together, these results indicate that the effect that was hypothesized to occur only in the mu band was not present within the beta ROI, as passive listening to the target melody after motor training revealed the same beta suppression as listening to the target melody before motor training, as well as to scrambled melodies after training (Figure 9 – bottom left).

Table 5. Beta ROI's GLMM post-hoc tests: '*Condition*' by '*Time*' (*DV = mean activity*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Learn Post - Rand Post	0	0.02	180	0	0.9569	0.9569
Learn Post - Learn Pre	-0.03	0.02	180	-1.540	0.1241	0.1811
Learn Post - Rand Pre	-0.13	0.02	180	-7.220	<.0001	**<.0001**
Rand Post - Learn Pre	-0.03	0.02	180	-1.440	0.1509	0.1811
Rand Post - Rand Pre	-0.13	0.02	180	-7.420	<.0001	**<.0001**
Learn Pre - Rand Pre	-0.1	0.02	180	-6.160	<.0001	**<.0001**

Delta-theta

In the delta-theta ROI we obtained significant main effects for both fixed effect factors 'Condition' (F(1,180) = 79.650; p-value < 0.0001) and 'Time' (F(1,180) = 105.401; p-value < 0.0001). The interaction term for these two fixed effect factors was not statistically significant (F(1,180) = 2.966; p-value = 0.0867). FDR-corrected post hoc tests revealed significantly higher activity within 'Condition' for Learn as compared to Rand (t(180) = 8.920; corrected p-value < 0.0001), as well as significantly higher activity within 'Time' for Post as compared to Pre (t(180) = 10.270; corrected p-value < 0.0001). Taken together, these results indicate that the effect that was hypothesized to occur only in the mu band was not present

within the delta-theta ROI, as passive listening was associated with enhanced activity with respect to baseline in all four blocks, as opposed to suppressed activity (Figure 9 – top left).

Gamma

In the gamma ROI we obtained significant main effects for both fixed effect factors 'Condition' ($F(1,180) = 123.668$; p-value < 0.0001) and 'Time' ($F(1,180) = 18.415$; p-value < 0.0001), as well as the interaction term for these two factors ($F(1,180) = 6.543$; p-value < 0.0114). FDR-corrected post hoc contrasts (Table 6) revealed significantly lower activity within the Interaction term for 'Learn_Post' only as compared to 'Rand_Pre' ($t(180) = -11.170$; corrected p-value < 0.0001) and 'Rand_Post' ($t(180) = -6.630$; corrected p-value < 0.0001). In other words, the 'Learn_Post' passive listening condition did not display significantly more gamma suppression than 'Learn_Pre' ($t(180) = -1.230$; corrected p-value = 0.2218). As well, 'Rand_Pre' exhibited significantly more activity than 'Rand_Post' ($t(180) = -4.840$; corrected p-value < 0.0001) and 'Learn_Pre' ($t(180) = -10.020$; corrected p-value < 0.0001). Finally, 'Rand_Post' was associated with significantly more activity than 'Learn_Pre' ($t(180) = -1.230$; corrected p-value = 0.2218). Taken together, these results indicate that the effect that was hypothesized to occur only in the mu band was not present within the gamma ROI, as passive listening to the target melody after motor training revealed the same gamma suppression as listening to the target melody before motor training (Figure 9 – bottom right).

Table 6. Gamma ROI's GLMM post-hoc tests: '*Condition*' by '*Time*' (*DV = mean activity*)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Learn Post - Rand Post	-0.07	0.01	180	-6.630	<.0001	**<.0001**
Learn Post - Learn Pre	-0.01	0.01	180	-1.230	0.2218	0.2218
Learn Post - Rand Pre	-0.12	0.01	180	-11.170	<.0001	**<.0001**
Rand Post - Learn Pre	0.06	0.01	180	5.480	<.0001	**<.0001**
Rand Post - Rand Pre	-0.05	0.01	180	-4.840	<.0001	**<.0001**
Learn Pre - Rand Pre	-0.11	0.01	180	-10.020	<.0001	**<.0001**

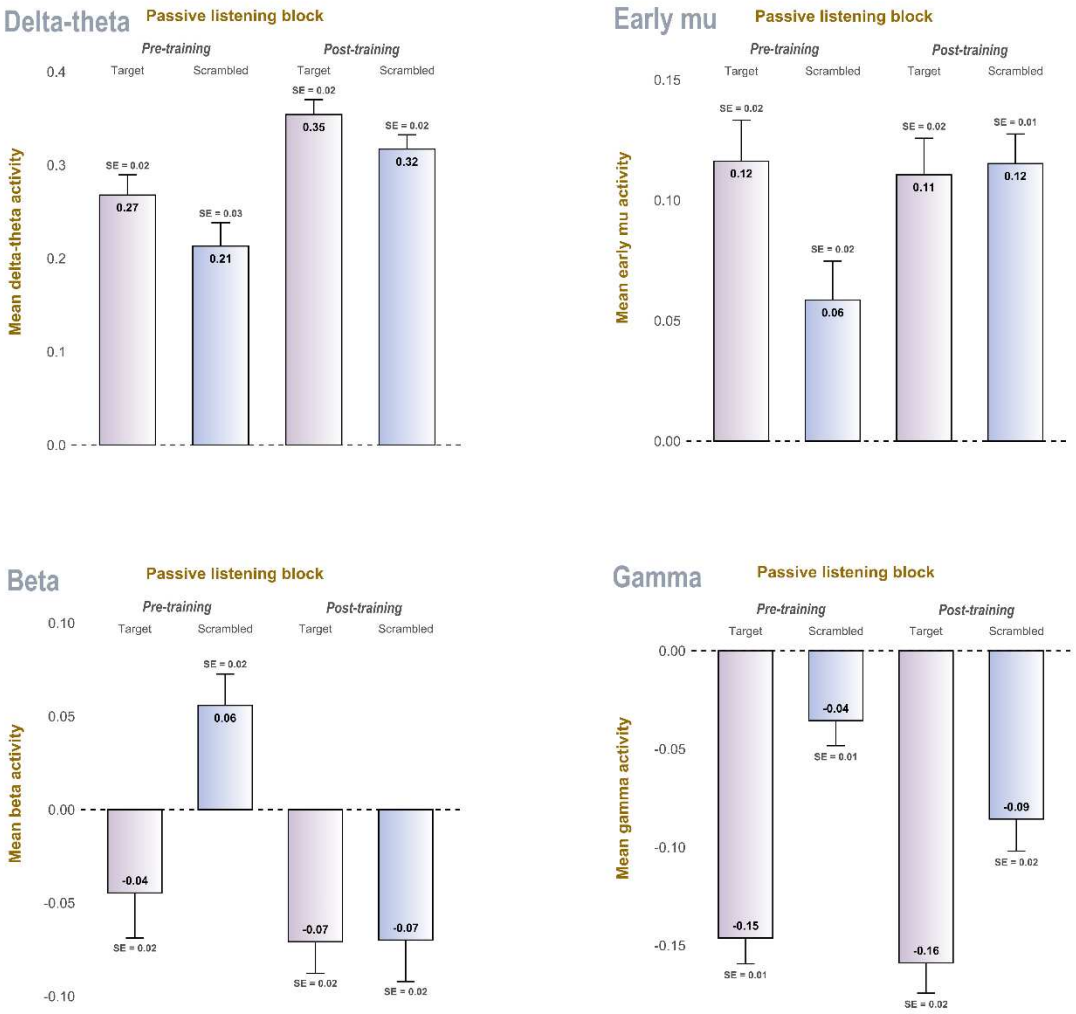


Figure 9. GLMM5: Channel-averaged suppression across passive listening blocks in control ROIs
Z-scores representing the difference between magnitude and baseline in each pixel of each ROI were averaged across subjects, within each passive listening condition and for each channel separately. Figure 8 shows the means and variation (SE) after then averaging across channels within each condition, for each of the four control ROIs. The Training POST block (Post-training Target in the Figure) did not display significantly more suppression than the other three passive listening blocks in any ROI. ***Top-left****: Delta-theta ROI;* ***Top-right****: Early mu ROI;* ***Bottom-left****: Beta ROI;* ***Bottom-right****: Gamma ROI.*

3.6 Discussion

3.6.1 Summary

The goal of this study was to compare anticipatory mu suppression occurring at the single-note level in non-musicians during passive listening to learned versus non-

learned melodies. We hypothesized that the coordinates of active mu suppression occurring during motor training of the target melody would correspond to the coordinates of passive mu suppression occurring during subsequent passive exposure to that same melody, but not to other melodies or moments of passive exposure. Moreover, we expected suppression effects during passive listening to only be significantly different across listening blocks in the mu band. To test these ideas, we developed a functional localizer to identify mu suppression across timepoints, and channels. These coordinates were then used to demonstrate that mu suppression was greater during passive listening to the learned melody compared to not-learned melodies post training, and not pre training. Moreover, no significant suppression effects were found in other frequency bands identified from the localizer. Therefore, our findings support the notion that mu suppression can be observed and measured at the single-note level in non-musicians, and that this effect is frequency-band specific. Overall, this finding aligns well with the current literature on the proactive role of covert, predictive motor processing of time-sensitive regularities in supporting auditory perception.

3.6.2 Behavioural Results

Our behavioural results demonstrate that subjects successfully learned the target melody sequence. First, because performance improved across blocks, and second because improvements were sequence-specific (greater for learned than Reverse blocks). These findings align with Stephan et al. (2018), who reported end-of-training accuracies well above chance and shortened latencies of around 100ms. In our data, accuracy improved more prominently than reaction times, consistent with the findings that accuracy in motor training tasks develops before speed in humans (Hikosaka et al., 2002) and monkeys (Hikosaka et al., 1995). Additionally, we obtained evidence that participants were able to learn the tone-to-finger mapping independently of the melody sequence, as performance in the second Reverse block

improved compared to the first. However, this kind of generalized learning was much weaker than the learning of the repeated sequence.

3.6.3 3D Functional Localizer

Before proceeding to the EEG results, it is important to note that, according to the cluster permutation framework, a 3D ROI can only be described if it is validated through cross-validation or within a condition separate from the one that generated it (Maris & Oostenveld, 2007). This criterion is only met by the late mu ROI identified by our functional localizer, because it was the only 3D ROI that was subsequently validated in the passive listening data based upon a prior experimental hypothesis. Accordingly, given our choice of methodology, all other ROIs *except* late mu should be thought of as control regions, which is why the upcoming discussion will focus exclusively on describing and interpreting the findings in late mu.

As for the localizer itself, we predicted that it would identify active mu suppression across credible frequencies, channels, and timepoints; a hypothesis that we deem to have been confirmed by the coordinates obtained for the late mu ROI. Spectrally, this ROI is a credible candidate of mu suppression because it spans 10-15Hz, thereby matching the upper portion of the alpha band typically ascribed to the mu rhythm (Pfurtscheller & Lopes da Silva, 1999; Pineda, 2005).

Spatially, the late mu ROI is also a credible candidate of mu suppression because of its apparent clustering in left central channels around M1, which is thought to be generator of mu rhythms (Pfurtscheller & Neuper, 1997), as supported by evidence from intracranial EEG recordings (Jasper & Penfield, 1949), MEG source-space analysis (Salmelin & Hari, 1994), and cortex-muscle coherence (Salenius et al., 1997). This being said, it is important to note that sensor-level EEG has a limited capacity to pinpoint exact neural sources and generators, especially when compared to source-resolved MEG (Baillet, 2017). This issue is compounded by the fact that electrical signals are more likely to spread to neighbouring channels than magnetic

signals (Hämäläinen et al., 1993; Lachaux et al., 1999), which ultimately means that we cannot at present conclude that the source of our late mu ROI is indeed M1, although it is reasonable to conclude that it originates either in M1 and/or surrounding dorsal motor structures.

Temporally, the -400 to 0ms timeframe of the late mu ROI would not appear to align with the -2000ms onset reported in many classical mu suppression studies (Pineda, 2005). However, Fox et al. (2016) note in their meta-analysis that not enough studies have assessed the chronometry of mu suppression to establish its bounds, and that suppression occurs before, during, and even after observation of an action, depending on the study. More importantly, the inter-onset interval in our data was 1500ms and therefore incompatible with the longer -2000ms onset of suppression found in other studies. This is because our study assessed the degree of mu suppression occurring before every note, a scale that is much faster than that of most music-based studies, which tend to assess the general degree of mu suppression effects occurring before every melody (Wu et al., 2016, 2017). Thus, given the absence of a theoretical time-window for finger-level mu-suppression in a musical context, we instead base our comparison on the temporal structure of other measures of effector-level motor preparation.

In this respect, we posit that our late mu ROI has a credible temporal structure because its onset overlaps with that of the late component of the Lateralized-Readiness Potential or LRP, a classical ERP that indexes motor preparation. More precisely, the LRP reflects the activation of motor cortical areas in anticipation of movement, and is an ERP that originates primarily in M1, lateralizing to the hemisphere contralateral to the effector preparing for movement. Moreover, the LRP consists of a gradual negative wave that begins several hundred milliseconds before movement onset and can be divided into an early component, which reflects the planning or selection of the motor response, and a late component, which captures the immediate preparation for movement execution. Notably, this late component

begins at around -400ms prior to movement onset (Shibasaki & Hallett, 2006; Trevena & Miller, 2002), closely matching the timing we observed for late mu suppression.

Similarly, the mid-point of our late mu ROI's period coincides with the Motor Evoked Potentials (MEPs) elicited at -200ms prior to tone onset in Stephan et al. (2018), further solidifying our claim. MEPs are involuntary muscle twitches that can be elicited via brain stimulation to effector-specific sites in contralateral M1 and are typically measured with electromyography (Pascual-Leone et al., 1998). These potentials are said to index the degree of motor excitability, which is itself interpreted to be a measure of the degree of motor preparation in the system (Lepage et al., 2008). In the study by Stephan et al. (2008), the authors used single-pulse TMS on non-musicians to trigger said MEPs, which were accordingly used to determine the degree of motor preparation to upcoming tones during passive listening to practiced and non-practiced melodies. As with mu suppression in our own data, MEPs were significantly larger to the practiced melody, particularly -200ms before tone onsets, in an effector-specific manner, indicating predictive motor preparation at the finger level. Thus, our late mu ROI aligns especially well with Stephan et al. (2018), suggesting that the MEPs in that study may be related to mu-suppression in ours, at least indirectly, as they both share a common M1 generator and very similar timing.

This being said, it should be noted that Lepage et al. (2008) found no correlation between MEPs and mu suppression in their combined TMS-EEG study, despite finding the expected increase in MEP amplitude and the expected suppression in the mu band across action observation, imagination, and execution. The authors therefore indicate that these components may reflect different processes taking place within the matching of actions to observations, and point to the issue of inter-subject variability as a potential contributor to this discrepancy (Lepage et al., 2008). At any rate, we contend that the temporal overlap between our late mu ROI with the LRPs and MEPs measured in similar studies globally supports the timing coordinates we obtained for mu suppression, and that these connections merit future investigation.

3.6.4 Late Mu Suppression during Passive Listening to the Learned Melody

Based on the coordinates provided by the functional localizer, we obtained significantly greater mu suppression estimates during exposure to the target melody after training compared to the other three passive listening conditions. Therefore, this result supports our second hypothesis that motor training would lead to greater mu suppression during passive exposure to the trained melody. Moreover, condition-specific suppression was not found in any of the other four control ROIs, supporting our third hypothesis that the predicted effect would be mu-specific.

We note that the convergence between active and passive mu suppression in our data is consistent with the literature showing a high degree similarity between action and perception, particularly in M1 (Fadiga et al., 2005). Moreover, Arnstein et al. (2011) found similar BOLD increases during both action observation and execution in areas connected to M1, such as dPMC and vPMC, inferior parietal lobe, and posterior S1, (Pineda et al., 2013). In addition, we note that the greater post-training mu suppression while listening to the target melody, compared to the untrained melodies, aligns with many neuroimaging studies reporting that only passive listening to learned melodies elicits activity in cortical and subcortical motor areas (Baumann et al., 2007; Lahav et al., 2007; Lappe et al., 2008, Herholz et al., 2016).

Within this niche, our finding converges especially well with the mu suppression reported for musicians in Wu et al. (2016), but notably diverges from the lack of mu suppression in non-musicians reported by the same authors (Wu et al., 2017). Several methodological differences may account for this discrepancy. First, while Wu et al. (2017) did report mu suppression prior to effectuated movements (Audiomotor trials), they did not use these data to localize suppression during passive listening (Audio-T, Audio-W, Audio-U). Our success in masking with our late mu ROI suggests that leveraging the larger mu suppression in active conditions may be essential for isolating passive mu suppression in non-musicians, as this passive component may be fainter or noisier than in musicians. Moreover, given their sample size

of 13 subjects, Wu et al. (2017) may have lacked statistical power to detect the comparably subtler mu suppression effect in passive listening. Another key difference is that motor training and passive exposure to learned sequences occurred in separate sessions, unlike in our study, suggesting that mu suppression effects in non-musicians may be short-lived and only detectable shortly after motor training. Therefore, we argue that future research should examine the possibility that mu suppression is dependent on time of recording by measuring it at progressively later moments after training. In addition, the minimal amount of motor training required to elicit passive mu suppression and the potential role of sleep for consolidation all merit investigation. We also believe that the precise influence of musical training on mu band enhancement/suppression should now be explored within the context of the same study, not only across studies.

More globally, our mu suppression effect points to sequence-specific activity and suggests top-down predictive timing processes (Kornysheva & Diedrichsen, 2014; Schubotz & von Cramon, 2002). Specifically, our finding supports the idea that the process of predicting *which* event will occur next demands additional resources from cortical motor areas than simply predicting *when* it will occur (Novembre & Keller, 2014; Stupacher et al., 2013). Given the specific involvement of vPMC during passive listening to known melodies (Lahav et al., 2007) and its putative role in matching sounds to movements in a one-to-one fashion (Zatorre et al., 2007), it is likely that mu suppression is contingent on vPMC recruitment. The involvement of dPMC may also be necessary for mu suppression, particularly during the motor learning stage, given its hypothesized role in orchestrating complex movement plans (Hoshi & Tanji, 2007). However, our experiment does not support the idea of dPMC involvement as robustly as vPMC's, given that our motor task only required simple one-to-one matching, as opposed to the many-to-many matching of sounds to movements that the dPMC is specifically thought to enable (Zatorre et al., 2007).

In addition, our finding provides support to the current common coding theory's thesis that forward models are generated in a predictive fashion, when available, to aid perception (Prinz, 1990). Specifically, the temporal structure of our observed mu suppression suggests that motor cortices are activated not solely in response to auditory inputs, via inverse modelling, but also in predictive anticipation of the sensory consequences of the next action at the finger level, which certainly requires forward modelling. Under this framework, forward models exist because they help the brain anticipate and interpret incoming sensory events more efficiently, a notion that is central to predictive coding formulations (Friston et al., 2010), notably predictive timing (Arnal & Giraud, 2012). In particular, forward models are thought to be critical in fast-paced, dynamic environments, where real-time sensory feedback alone may be insufficient for optimal performance (Welniarz et al., 2021).

Finally, our finding in non-musicians provides support to the notion that motor predictions during auditory-motor perception are an intrinsic process of the human brain (Patel & Iversen, 2014). This is supported by early spectral analyses showing that mu suppression occurs in most adults (Pfurtscheller & Aranibar, 1979), as well as functional connectivity studies indicating that the auditory and motor systems exhibit a strong functional connection in the brain that is independent of musical training (Bedford et al., in press). In this context, learned auditory-motor associations may specifically engage predictive processes that differ from those that occur naturally during beat perception tasks (Iversen & Patel, 2008). Based on our data, we propose that the former are more complex and require desynchronization of motor alpha activity, whereas the latter probably only require ongoing beta-mediated or beta-dominant motor activity (Fujioka et al., 2015).

3.6.5 Methodological considerations

One challenge in detecting mu suppression coming from M1 is distinguishing it from concurrent occipital alpha activity, which is linked to both visual stimulation and attention, and is independent of movement (Hobson & Bishop, 2017). We note

several reasons that support our claim that the mu band (upper alpha) activity we obtained is decidedly not occipital alpha. First, because participants were instructed to maintain their gaze on a fixation cross equally throughout the entire experiment, including the motor training part. Additionally, EOG signals were recorded to detect any abnormal saccadic movements indicative of occipital alpha, and none were observed. Second, because the experiment engaged participant attention equally across all passive listening blocks through the distractor task of counting mistuned notes. Third, because the presence of occipital channels in the 3D mask does not imply that the activity in those electrodes reflects occipital alpha. Since the data was limited to sensor-space, it is more likely that mu suppression preceding movements spread from central to posterior channels, especially since mu suppression is much greater before actual movements than imagined ones (McFarland et al., 2000).

Another important methodological consideration is that our study does not permit the direct exploration of how motor training impacts the magnitude of mu suppression on an individual basis, as our statistical approach was not designed to assess brain-behaviour relationships. This is because we chose to average passive listening data across subjects in order to gain insight into the distribution of channel values across conditions. Consequently, the results cannot at present be extrapolated to the general population either, although we have no reason to believe our sample differs from the wider population in any meaningful way. Similarly, our deliberate choice to not include the beta band as part of our hypothesis set circumscribes our results to the upper alpha band. Notably, the beta band is thought to contain a secondary component or at the very least a subharmonic of the mu rhythm (Hari, 2006). However, the neural generators of this secondary beta component do not necessarily overlap with those of the main mu band component we explored in this experiment (Larionova et al., 2022). Therefore, given the aforementioned spatial resolution limits of sensor-level EEG, we decided to avoid the risk of conflating the two mu complex components and focus only on mu proper. Finally, we note that the functional localizer developed in this study could inform

present-day speech models aiming to distinguish mu suppression during speech versus non-speech auditory tasks, where it is methodologically challenging to differentiate sensorimotor activity related to linguistic processing from attention and working memory (Cuellar et al., 2012).

3.7 Conclusion

This study presented the first evidence of mu suppression during passive listening to a trained melody in non-musicians, demonstrating that this effect stems from an auditory-motor mechanism that is independent of prior musical training. Furthermore, the observation of mu suppression within a continuous training session suggests that this phenomenon can emerge rapidly during learning. Notably, mu suppression was detected at the single-note level, showing for the first time that individual units of a learned melody sequence can trigger mu suppression, not just the entire sequence itself. Ultimately, these findings underscore the versatility of the human auditory-motor system and support the idea that motor activity can support passive auditory perception by predicting upcoming consequences of actions.

3.8 References

actiCap (64 channels) [Apparatus]. (2019). Gilching, Germany: Brain Products GmbH.

Adobe Systems Incorporated. (2007). *Adobe Systems Incorporated. (2007). Adobe Audition (Version 3.0) [Computer software]. Adobe Systems Incorporated*. [Computer software].

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., & Gazzola, V. (2011). μ-Suppression during Action Observation and Execution Correlates with BOLD in Dorsal

Premotor, Inferior Parietal, and SI Cortices. *The Journal of Neuroscience*, *31*(40), 14243–14249.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain Research*, *1161*, 65–78.

Bedford, O., Noly-Gandon, A., Ara, A., Wiesman, A. I., Albouy, P., Baillet, S., Penhune, V., & Zatorre, R. J. (2025). Human auditory-motor networks show frequency-specific phase-based coupling in resting-state MEG. *Human Brain Mapping*.

BrainAmp [Apparatus]. (2019). Gilching, Germany: Brain Products GmbH.

BrainVision Recorder, Version 1.20.0801, Brain Products GmbH, Gilching, Germany. (2019). [Computer software].

Brooks, M. E., Kristensen, K., Benthem, K. J. van, Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400.

Candia-Rivera, D., & Valenza, G. (2022). Cluster permutation analysis for EEG series based on non-parametric Wilcoxon–Mann–Whitney statistical tests. *SoftwareX*, *19*, 101170.

Chatrian, G.-E., Lettich, E., & Nelson, P. L. (1988). Modified Nomenclature for the "10%" Electrode System1. *Journal of Clinical Neurophysiology*, *5*(2), 183.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

Cuellar, M., Bowers, A., Harkrider, A. W., Wilson, M., & Saltuklaroglu, T. (2012). Mu suppression as an index of sensorimotor contributions to speech processing:

Evidence from continuous EEG signals. *International Journal of Psychophysiology*, *85*(2), 242–248.

Fadiga, L., Craighero, L., & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Current opinion in neurobiology*, *15*(2), 213-218.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., Ferrari, P. F., & van IJzendoorn, M. H. (2016). Assessing Human Mirror Activity With EEG Mu Rhythm: A Meta-Analysis. *Psychological Bulletin*, *142*(3), 291–313.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *The Journal of Neuroscience*, *35*(45), 15187–15198.

Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., & Oostenveld, R. (2013). Good practice for conducting and reporting MEG research. *Neuroimage*, *65*, 349–363.

Halász, V., & Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sciences*, *2*(2), 130–146.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, *65*(2), 413.

Hari, R. (2006). Action–perception connection and the cortical mu rhythm. In C. Neuper & W. Klimesch (Eds.), *Progress in Brain Research*, *159*, 253-260. Elsevier.

Herholz, S. C., Coffey, E. B., Pantev, C., & Zatorre, R. J. (2015). Dissociation of neural networks for predisposition and for training-related plasticity in auditory-motor learning. *Cerebral Cortex, 26*(7), 3125-3134.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, *12*(2), 217–222.

Hikosaka, O., Rand, M. K., Miyachi, S., & Miyashita, K. (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. *Journal of Neurophysiology*, *74*(4), 1652–1661.

Hobson, H. M., & Bishop, D. V. M. (2017). The interpretation of mu suppression as an index of mirror neuron activity: past, present and future. *Royal Society Open Science*, *4*(3), 160662.

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology*, *6*, 1318.

Iversen, J. R., & Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioural Sciences*, *8*, 175–180.

Iversen, J., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. *Proceedings of the 10th International Conference on Music Perception and Cognition*.

Jasper, H., & Penfield, W. (1949). Electrocorticograms in man: Effect of voluntary movement upon the electrical activity of the precentral gyrus. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *183*(1–2), 163–174.

Kornysheva, K., & Diedrichsen, J. (2014). Human premotor areas parse sequences into their spatial and temporal features. *eLife*, *3*, e03043.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human brain mapping*, *8*(4), 194-208.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *The Journal of Neuroscience*, *27*(2), 308–314.

Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical Plasticity Induced by Short-Term Unimodal and Multimodal Musical Training. *The Journal of Neuroscience*, *28*(39), 9632–9639.

Lepage, J.-F., Saint-Amour, D., & Théoret, H. (2008). EEG and neuronavigated single-pulse TMS in the study of the observation/execution matching system: Are both techniques measuring the same process? *Journal of Neuroscience Methods*, *175*(1), 17–24.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Mathias, B., Palmer, C., Perrin, F., & Tillmann, B. (2015). Sensorimotor Learning Enhances Expectations During Auditory Perception. *Cerebral Cortex*, *25*(8), 2238–2254.

McFarland, D. J., Miner, L. A., Vaughan, T. M., & Wolpaw, J. R. (2000). Mu and Beta Rhythm Topographies During Motor Imagery and Actual Movements. *Brain Topography*, *12*(3), 177–186.

McGill, W. J., & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, *2*(1), 1–18.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS One*, *9*(2), e89642.

Niedermeyer, E. (1997). Alpha rhythms as physiological and abnormal phenomena. *International Journal of Psychophysiology*, *26*(1), 31–49.

Novembre, G., & Keller, P. E. (2014). A conceptual review on action-perception coupling in the musicians' brain: what is it good for? *Frontiers in Human Neuroscience*, *8*, 603.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*.

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, *250*, 85–93.

Pfurtscheller, G., & Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalography and Clinical Neurophysiology*, *46*(2), 138–146.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857.

Pfurtscheller, G., & Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neuroscience Letters*, *239*(2), 65–68.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing." *Brain Research Reviews*, *50*(1), 57–68.

Pineda, J. A., Grichanik, M., Williams, V., Trieu, M., Chang, H., & Keysers, C. (2013). EEG sensorimotor correlates of translating sounds into actions. *Frontiers in Neuroscience*, *7*, 203.

Posit team. (2022). *RStudio: Integrated Development Environment for R*. *Posit Software, PBC, Boston, MA*. [Computer software].

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In O. Neumann & W. Prinz (Eds.), *Relationships Between Perception and Action: Current Approaches* (pp. 167–201). Springer.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.

Salenius, S., Portin, K., Kajola, M., Salmelin, R., & Hari, R. (1997). Cortical Control of Human Motoneuron Firing During Isometric Contraction. *Journal of Neurophysiology*, *77*(6), 3401–3405.

Salmelin, R., & Hari, R. (1994). Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience*, *60*(2), 537–550.

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting Perceptual Events Activates Corresponding Motor Schemes in Lateral Premotor Cortex: An fMRI Study. *NeuroImage*, *15*(4), 787–796.

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential?. *Clinical neurophysiology*, *117*(11), 2341-2356.

Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Cognitive Processing*, *17*(1), 79–87.

Stephan, M. A., Lega, C., & Penhune, V. B. (2018). Auditory prediction cues motor preparation in the absence of movements. *NeuroImage*, *174*, 288–296.

Stroup, W. W. (2012). *Generalized linear mixed models: modern concepts, methods and applications*. CRC press.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, *82*(2), 127–136.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A User-Friendly Application for MEG/EEG Analysis. *Computational Intelligence and Neuroscience*, *2011*, 1–13.

Tesche, C. D., Uusitalo, M. A., Ilmoniemi, R. J., Huotilainen, M., Kajola, M., & Salonen, O. (1995). Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. *Electroencephalography and Clinical Neurophysiology*, *95*(3), 189–200.

Thangaraj, K. T. D. P. (2016). A New Algorithm For Beaglebone Black. *Global Journal of Advanced Engineering Technologies and Sciences*, *3*(1), Article 1.

The MathWorks Inc. (2022). *MATLAB version: 9.10.0 (R2021a), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com* [Computer software]. The MathWorks Inc.

Trevena, J. A., & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Consciousness and cognition*, *11*(2), 162-190.

Welniarz, Q., Worbe, Y., & Gallea, C. (2021). The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency. *Frontiers in Systems Neuroscience*, *15*, 644059.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. (1994). Forward dynamic models in human motor control: Psychophysical evidence. *Advances in Neural Information Processing Systems*, *7*.

World Medical Association. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, *79*(4), 373–374.

Wu, C. C., Hamm, J. P., Lim, V. K., & Kirk, I. J. (2016). Mu rhythm suppression demonstrates action representation in pianists during passive listening of piano melodies. *Experimental Brain Research*, *234*(8), 2133–2139.

Wu, C. C., Hamm, J. P., Lim, V. K., & Kirk, I. J. (2017). Musical training increases functional connectivity, but does not enhance mu suppression. *Neuropsychologia*, *104*, 223–233.

Chapter 4: General Discussion

4.1 Recapitulation of Results

4.1.1 Study 1's Main Results

Study 1 used rs-MEG to test whether the intrinsic functional connectivity patterns between cortical auditory and motor areas exhibit larger ipsilateral phase-to-phase coupling estimates than between visual and motor areas. Based on prior studies, this effect was expected to be more pronounced in musicians and most prominent in right-hemispheric connections, particularly in the beta band. Moreover, phase-based effective connectivity measurements were expected to show alternating directions of information flow across frequency bands, consistent with recent findings.

This study generated five key results. First, functional connectivity estimates between auditory and motor areas were significantly larger than between visual and motor areas across the frequency spectrum, not just in the beta band. Second, auditory-motor connections demonstrated great consistency in the degree of functional connectivity strength across the frequency spectrum, with the A1-vPMC link being systematically largest. Third, right-hemispheric connectivity was associated with greater PLVs across both the auditory- and visuomotor modalities, not just the former. Fourth, cross-modal differences in connectivity strength were most pronounced in the alpha band, followed closely by the beta and theta bands, particularly in vPMC. Fifth, effective connectivity revealed a preferred auditory-to-motor direction in the alpha and upper gamma bands, versus a preferred motor-to-auditory direction in the theta and beta bands. No differences in functional connectivity were observed between musicians and non-musicians, nor did the strength of auditory-motor connections scale with years of musical training.

4.1.2 Study 2's Main Results

Study 2 utilized EEG in a non-musician sample to determine whether predictive auditory-motor representations associated with learned melody sequences have an intrinsic oscillatory correlate, specifically in the form of mu suppression over the sensorimotor cortex contralateral to the dominant hand. Based on prior motor learning studies in musicians, it was hypothesized that this effect would emerge after a single motor training session and therefore be measurable during passive exposure to the learned melody. However, this preparatory motor effect was also expected to occur at the single note/finger level, a novel hypothesis in the literature. Given the reported challenges of detecting passive mu suppression in non-musicians, a group-level, cluster-based 3D functional localizer was developed using independent data from the active motor training condition to test the hypotheses.

This study produced four key findings. First, the functional localizer successfully identified mu suppression in the active condition across frequency, time, and channels, in accordance with theoretical delineations. Second, the localizer captured passive mu suppression occurring at the single-note level most prominently during post-training exposure to the learned melody. Third, suppression was mu-band specific and absent in the delta-theta, beta, and gamma bands. Fourth, mu suppression was observed only 400–100 milliseconds before tone onset, with no suppression detected in an earlier component within the same band emerging halfway between inter-tone intervals.

4.2 Interpretations

4.2.1 Juxtaposing the Findings

The two sets of findings presented in this thesis differ in ways that distinctly shape their scope and interpretability. Thus, it is essential to enumerate these differences and their related strengths and weaknesses, as well as the limits around what is to be extracted from each of the two studies.

Study 1 explores effects occurring during idling states untied to experimental conditions, within a large but mixed sample that was derived from a well-curated publicly accessible database (Niso et al., 2016), and in the context of intrinsic *intra-brain* oscillatory coupling. These features render this study's results less behaviourally controlled, yet readily generalizable to a broader population and to functional connectivity processes relevant to non-musical domains such as speech. Moreover, Study 1's results are based on source-resolved magnetic signals, span both hemispheres, all frequency bands, and multiple cortical regions, and refer to oscillatory phase synchronization effects. This means that this study's findings can be classified as spectrally broad yet fine, spatially broad yet fine, and highly time locked. Consequently, a large amount of high-resolution spatiotemporal and spectral information can be extracted about the precise temporal alignment between defined auditory and motor cortical regions across the frequency spectrum, and much can be inferred about how this information relates to intrinsic, population-wide mechanisms. Finally, given the limits associated with resting state results, the findings would be well-complemented by future studies using different paradigms, including active states.

Conversely, Study 2 is anchored to statistical comparisons across task conditions, was conducted in our lab using a small but homogenous sample and focuses on the highly specific context of *brain-stimulus* interactions subserving a keyboard-based musical task. These features render this study more behaviourally controlled, but its findings can only be extrapolated to non-musicians and music-related oscillatory processes, with potentially less relevance for speech. Moreover, Study 2's results are based on electrical signals resolved to sensor space, they are tied to the activity of a narrow sub-band within left dorsal sensorimotor areas and refer to desynchronization effects across a mixture of evoked and induced components. This means that this study's findings can be classified as spectrally narrow and fine, spatially narrow yet coarse, and more loosely time locked. Accordingly, a large amount of robust, high-resolution spectrotemporal information can be extracted

about the time-frequency bounds of anticipatory upper alpha desynchronization and, therefore, much can be inferred about the predictive role of motor alpha in music production. Moreover, given the coordinates of mu suppression in our study, it is quite likely that our findings in non-musicians point to the same mechanism that other studies have reported in musicians, although a direct head-to-head comparison between the two groups would be needed to definitively establish this correspondence.

However, the exact temporal nature of anticipatory mu suppression's relationship to stimulus onsets—whether phase-locked or just time-locked—cannot be determined from this study's results. Furthermore, while the topography in this study broadly supports the notion that the upper alpha desynchronization specifically emanates from an M1 generator (Pineda et al., 2013) and not from neighbouring premotor cortices, spatially resolved replication, for example using MEG, is needed to establish that this is indeed the case. Finally, while informative, the statistical choice of using channels instead of subjects as the grouping variable is uncommon in the literature and does not allow for brain-behaviour relationships to be fully explored.

4.2.2 Interpreting the Three Key Findings Across Each Study

4.2.2.1 Functional Dorsal/Ventral Distinctions in Premotor Cortex

In our view, there are three key findings across each study presented in this thesis. The first relates to the **consistently distinct levels of auditory-motor coupling** strength observed in all bands across the three cortical motor targets in Study 1. This result integrates particularly well into the functional differences typically ascribed to dPMC and vPMC, where the former is considered to code 'supramodal' action plans or sequences, while the latter is thought to code 'surface properties' of those behaviours (Schubotz & Von Cramon, 2003). This distinction is even more relevant in the context of task-based auditory-motor coupling studies, some of which were conducted in our lab. Within this niche, vPMC has been specifically tied to direct

auditory-motor transformations that match sounds to movements in a one-to-one manner, while dPMC has been associated with indirect, more abstract, many-to-many transformations (Hoshi & Tanji, 2007). Notably, these concepts are buttressed by findings that vPMC is more active upon exposure to music for which one has a specific one-to-one learned motor programme (Lahav et al., 2007), whereas more dorsal motor recruitment is seen when keyboard learning involves different fingerings for more complex melodies requiring a many-to-many mapping (Herholz et al., 2016). Moreover, a prior rs-fMRI study from our lab specifically highlighted the strength of the right A1-vPMC connection (Palomar-García et al., 2016), thereby rendering the analogous result obtained in Study 1 a partial replication of this finding.

Our claim, therefore, is that the distinctive strength of the A1-vPMC connection in Study 1 is indeed related to the fact that these two regions are functionally tied to the execution of ubiquitous movements involving simple, one-to-one transformations of actions to sounds. Crucially, this interpretation matches vPMC's unique somatotopic organization, where oro-facial and oro-laryngeal movement representations predominate and support the recurrent movements of speech (Rizzolatti & Arbib, 1998). Interestingly, upper limb and hand-related representations also dominate vPMC's body map and are thought to support fine motor skills (Rizzolatti et al., 2002), such as those that would be implicated in playing musical instruments. Equally important is the fact that vPMC is characterized by a higher density of multi-sensory receptive fields compared to other motor and premotor regions (Bremmer et al., 2001), effectively rendering this region a proper integration hub. Indeed, primate studies indicate that certain cells in vPMC (and IPS) respond to tactile stimulation of a given part of the animal's body when paired with an auditory stimulus, but only if the latter is presented near the body part (Schlack et al., 2005). Moreover, when the body part moves, the auditory receptive field also shifts, in line with vPMC's characteristic egocentric frame of reference (Graziano & Cooke, 2006).

Behavioural studies in humans have similarly shown that reaction times to a tactile stimulus administered to the hand are faster if concurrent task-irrelevant auditory stimuli are presented near the hand as opposed to farther in space (Bassolino et al., 2010) and that this facilitation is nullified following inhibitory brain stimulation to vPMC (and IPS) but not control sites (Serino et al., 2011). Musicians have also been shown to exhibit a higher intercorrelation of grey matter thickness between A1 and vPMC than non-musicians, which is thought to be due to frequent co-activations throughout the lifespan (Bermudez et al., 2009). Thus, both efferent and afferent streams of information involving vPMC, its grey matter intercorrelations with A1, and the high volume of tracts supporting the A1-vPMC link collectively support our claim that this functional connection exhibits the largest estimate of coupling in our data because it is essential to everyday functions such as speech or musical production.

In contrast, dPMC appears to be better placed for translating auditory information into motor instructions to conduct complex action plans (Hoshi & Tanji, 2004), an idea consistent with its potential role of transposing allocentric and limb coordinate systems. Rostral dPMC has particularly been tied to higher-order parameters of movement (Picard & Strick, 1996), like selecting actions that are conditionally tied to a specific sound (Bermudez & Zatorre, 2005). Critically, there is causal evidence of this putative relationship between dPMC and conditional movements, as dPMC disruptions—but not vPMC disruptions—impair such behaviours (Kurata & Hoffman, 1994), as well as the ability to coordinate movement across time (Davare et al., 2006). Alternative explanations of dPMC function have similarly focused on its potential role in organizing movement sequences (Janata & Grafton, 2003), or more generally in predicting sequences of sounds and movements alike (Schubotz & Von Cramon, 2003). Therefore, the fact that the A1-dPMC link was weakest in our study dovetails well with prior evidence that this functional connection is used on an *ad-hoc* basis when auditory-motor demands surpass vPMC's simple matching capabilities, such as during the acquisition of new auditory-motor associations during learning—especially if sequencing and novel mappings are involved. Thus, we

submit the hypothesis that dPMC activity was instrumental in orchestrating the learning processes that gave rise to the mu suppression effects observed in Study 2.

4.2.2.2 Top-Down Motor Oscillations: A Default Organizational Principle

The second key finding of this thesis corresponds to the **intrinsic, preferential motor-to-auditory direction** found in the beta and theta bands in Study 1. Critically, this result aligns with accounts of motor beta encoding temporal predictions in both speech (Park et al., 2015) and music (Morillon & Baillet, 2017), where precise timing is essential. In speech, beta-band predictions have notably been elicited in multimodal studies. For example, Schepers et al. (2013) found greater beta desynchronization in the superior temporal sulcus after auditory onsets in audiovisual versus audio-only conditions, clearly indicating that viewing lip movements enhances auditory predictability via this band. In music, beta-driven motor signals weave the precise alignment of instrument-related actions to the beat, such as when striking a piano key on time (Grahn & Brett, 2007). Supporting this idea, studies of passive listening to temporally regular tone sequences reveal that random tone omissions disrupt beta desynchronization and *rebound* patterns, inducing abnormally long beta enhancements that persist until the next tone onset (Fujioka et al., 2009, 2012).

Beta bursts are known to align with the phase of slower theta oscillations in auditory cortex through cross-frequency coupling (Hyafil et al., 2015; Samiee et al., 2022). Therefore, the concomitant motor-to-auditory preference we observed in the *theta* band supports a multi-band predictive coding framework, where these two bands would work together to encode temporal predictions across nested timescales. As evidence, theta oscillations in speech have been shown to follow slower temporal structures, such as phrases, while beta oscillations encode faster dynamics, such as syllables (Giraud & Poeppel, 2012). Accordingly, theta oscillations are said to reflect the perception of melodic phrases or musical passages (Large & Snyder, 2009), while

beta activity is known to keep track of metrical accents (Fujioka et al., 2015; Large et al., 2015).

Relatedly, our finding in the beta band and its opposite counterpart in the gamma band specifically lend support to the theoretical *beta-gamma hierarchical feedback-feedforward loop* for predictive timing outlined by Arnal & Giraud (2012). Namely, beta oscillations are characteristically found in deep 5/6 cortical layers (Roopun et al., 2006) and are thought to carry predictive information whenever these layers of higher-order sites discharge onto more superficial layers of downstream regions (Buschman & Miller, 2007). Mechanistically, beta oscillations achieve this goal by specifically upregulating the activity of *inhibitory* interneurons (Barone & Rossiter, 2021), which *suppresses* cortical excitability at lower levels. This is supported by the fact that beta activity from motor regions dampens neural responses in auditory regions to irrelevant stimuli, effectively enhancing the SNR of task-relevant sounds (Caras & Sanes, 2017).

On the other hand, the preferred auditory-to-motor direction we found in upper gamma fits well with the fact that gamma is prominently generated in superficial layers 2/3 (Roopun et al., 2008) and ostensibly codes for sensory prediction errors whenever information is sent from these layers within lower-order regions to layer IV of the next hierarchical stage (Douglas & Martin, 2004). That said, only the microcircuitry of this bidirectional loop has been proven experimentally, while the related functional links to predictive timing remain hypothetical, for now. Thus, in our view, the biggest value of Study 1's beta finding is that it provides validation to an emerging body of multi-disciplinary work linking this band to predictive timing processes, thus fulfilling one of the main aims of this thesis.

4.2.2.3 Mu-Band Dynamics Reflect Auditory-Motor Associations

The key finding of Study 2 is the fact that **mu suppression reflects learned auditory-motor associations** when the next note in a melody sequence is predictive, even

amongst people with limited musical experience. This finding appears to overlap with similar music-based studies looking at the dynamics of mu suppression in musicians (Wu et al., 2016), but it notably differs from the beat-tracking literature mostly underscoring the singular role of beta in such predictive processes. In our view, this discrepancy rests on the fact that, while beta can track predictable beat structures within both familiar and unfamiliar stimuli, the presence of mu suppression indexes motor preparation processes with respect to familiar stimuli for which one has a learned auditory-motor associations—even if such preparation occurs covertly or is never transformed into an action. As such, the magnitude of suppression in these contexts is likely to reflect the strength of this association. Moreover, this would explain why mu suppression has not been detected in prior pulse and beat-tracking literature, or in prior studies not explicitly focusing on predictive processes during playing or listening to known melodies, where the demands for preparing a specific movement to generate a specific sound are either non-existent, vastly mitigated by expertise, or perhaps simply eclipsed by beta enhancement.

We propose that the magnitude of mu suppression may *additionally* be contingent on the type of melodic stimulus we used, which is complex with respect to the simplest analogous case of an isochronous stream of pulses, as it demands ongoing dual predictions about both *what* and *when* events will occur next. In such cases, vPMC activity may be necessary and scale with spectral complexity and/or number of sound-effector pairings, in turn leading to greater mu suppression. Furthermore, melodic stimuli processing is thought to engage motor and premotor cortices more extensively than beat tracking, most especially when the former involves *sequencing* (Novembre & Keller, 2014). Thus, the degree of sequence complexity in our data may have also resulted in proportional dPMC activation that amplified suppression. Therefore, while our specific finding indicates that the presence of mu suppression is tied to previously learned auditory-motor associations and that the amount of this suppression is most likely reflective of association strength, magnitude of suppression may also be affected by ongoing spectral prediction demands and/or the

orchestration of various sound-effector pairings, as well as motor sequencing complexity—three possibilities that merit further investigation.

Finally, we note that in our behavioural data there is a complete absence of the anticipation tendency typically seen in beat alignment studies using simple isochronous timings (Aschersleben, 2002). As mentioned in the General Introduction, this observation has a direct tie to the literature, which predicts that the anticipation tendency will quickly dissipate during synchronization to non-isochronous or metrical stimuli (Patel et al., 2005), as well as to spectrotemporally rich stimuli such as music (Iversen & Patel, 2008). However, the fact that mu suppression was consistently anticipatory with respect to tone onsets—both in the active and passive conditions of the experiment—is at odds with our finding that keypresses always occurred after tone onset during training. Thus, our study directly points to an important dissociation between the predictive nature of biomechanical versus neural oscillators, thereby justifying the distinction made at the start of this thesis between auditory-motor synchronization and auditory-motor coupling. Indeed, the positive mean asynchronies we observed—similarly reported by others in contexts where the anticipation tendency is extinguished—clearly do not imply the absence of anticipatory oscillatory processes. Therefore, while biomechanical and neural oscillators are evidently related and have been integrated into dynamical systems models, our study also points to their timing divergence.

4.2.3 Interpreting the Convergent Findings

In the present thesis, two findings reported in one study have a mirror in the other. Recognizing and interpreting these links is as important as analysing the three key findings of each study independently. Therefore, we will take one last step to examine these connections before distilling and integrating all that has been covered in this manuscript into a unitary framework.

It is presently accepted that auditory-motor coupling is a bi-directional system (Chen et al., 2009; Patel & Iversen, 2014), and both studies in this thesis corroborate this idea by highlighting the proactive role of motor cortices in resting and task-based auditory-motor interactions. This is because, together, the two studies uncovered intrinsic mechanisms for top-down signalling from motor to auditory cortices and provided evidence that this motor activity can predict auditory consequences of actions. Moreover, the fact that the findings address key gaps in our understanding of intra-band dynamics is, itself, demonstrative of the fact that band-specific analyses remain relevant amidst the growing emphasis on cross-frequency coupling methods. Indeed, Study 1 provided evidence that top-down beta influences on A1 are not only task-dependent, even though they appear in countless task-based studies, whereas Study 2 highlighted the role of alpha in the predictive processing of sounds—a function more commonly attributed to motor beta.

Moreover, in Study 1, the alpha band was found to exhibit the largest phase-to-phase coupling differences between auditory- and visuomotor domains. This finding could be explained away by the fact that alpha-band predominance is a feature of resting states, but it is important to note that this has been specifically tied to *occipital* alpha (Mahjoory et al., 2020). Thus, it could be argued that the difference between auditory- and visuomotor phase coupling should be *minimal*, or even reversed, if only occipital alpha were at play during rest, given that V1 lies at its centre. Instead, the fact that this difference is *maximal* suggests that auditory-motor coupling involves at least one other kind of alpha, a notion consistent with recent studies revealing multiple distinct alpha generators with different functional and anatomical profiles (Bollimunta et al., 2011; Haegens et al., 2015; Keitel & Gross, 2016). We note that no other directional preference was observed for the A1-M1 link in other frequency bands besides the *bottom-up* one in alpha and the *top-down* one in beta, which suggests that bidirectional communication between these two regions would require the alpha-beta axis. Study 2 converges on this idea, as it showed that sensorimotor alpha, specifically the mu sub-band, is implicated in both active and passive

auditory-motor coupling, thus making it even more likely that the intrinsic functional connectivity along this band in Study 1 has concrete significance for auditory-motor processes. Therefore, we propose the existence of an intrinsic functional alpha-beta pathway that supports fast, reciprocal transmissions between *primary* auditory and motor cortices.

4.3 Distillation and Integration

4.3.1 Overall Significance and One Last Meta-Finding

Together, the two studies presented in this thesis provide support to the idea that oscillatory coupling is fundamental to auditory-motor processes. As highlighted, auditory-motor coupling involves both *intra-brain* and *brain-stimulus* oscillatory interactions. In this context, the breadth of results from our studies demonstrates that both phase-to-phase interactions across auditory and motor regions and motor time locking of power fluctuation to stimuli periodicities in high frequency bands are both intrinsic to the human brain, irrespective of musical training. This is especially relevant in the current context, where task-based and musician-focused accounts still dwarf the limited number of studies on the intrinsic mechanisms. Crucially, the findings also emphasize the importance of precise time-locked processes in auditory-motor interactions, thereby justifying the use of methods with high temporal resolution, such as M/EEG, as a complement to techniques with high spatial resolution, such as fMRI. As mentioned in the beginning, auditory-motor investigations cannot continue to progress without integrating oscillatory time-based information, and the novel findings in this manuscript are a testament to this fact.

Finally, our global or meta-finding is simply this: cortical auditory-motor coupling exhibits an inherently *hierarchical* nature along multiple dimensions; an idea that is consistent with the wider literature. Notably, this arrangement provides a complementary picture of inter-regional interactions along the dorsal auditory stream, presenting several particularities that warrant independent, careful study.

For example, along the spatial dimension, there appears to be a ranking of connections between auditory and motor cortices whose relation to the underlying anatomy is not straightforward and merits further investigations. Then, along the spectral dimension, there appears to be a clear, intrinsic division of labour within the higher frequencies between what can be described as mostly *feedforward* bands, such as alpha and gamma, and a mostly *feedback* band in beta. In addition, this division implies *alternation* along contiguous frequency bands, which suggests that the auditory-motor system intrinsically prevents functional leakage from occurring. Finally, along the time dimension, there appears to be a clear assignment of these regions and frequency bands where alpha and beta bidirectionally dominate the motor preparation stage along cortical auditory-motor areas, as seen in Study 1, and low frequency oscillations delta and theta provide the means for oscillatory tracking and *entrainment* to the produced sound within auditory centres (Gross et al., 2013; Stefanics et al., 2010).

4.3.2 Integration with the Frameworks, Models, and Neuroanatomy

In order to integrate the results of this manuscript with all the material reviewed in the Introduction, we now outline an original explanation of how mu suppression intertwines with the three canonical stages of the beta cycle in the context of either playing or passively listening to a known, spectrally rich melody. Namely, beta *enhancement* during movement preparation, beta *suppression* at movement onset, and beta *rebound* upon movement cessation. The *enhancement* stage of our explanation rests on the combined findings of this thesis, the results in the wider literature, and several important concepts from the models. Conversely, the parts about the *suppression* and *rebound* stages are not directly supported by our findings and are therefore more speculative. However, we believe that the large number of anatomical descriptions, high-level theories, and findings discussed over this thesis have enabled us to draw a reasonable picture and put together a compelling web of nested hypotheses about what might be taking place in these two stages.

Our explanation revolves around the central idea that, in these auditory-motor situations, the shift from baseline alpha to suppression in M1 represents the transition between ceasing to integrate feedforward sensory afferents from A1 and kickstarting motor preparation. This notion is concretely based on the fact that the preferentially *bottom-up* A1-M1 connection alpha found in Study 1 offers a direct interpretational bridge to Study 2, which instead showed that alpha suppression over M1 is anticipatory and therefore driven by *top-down* influences. Indeed, this discrepancy suggests a task-induced switch from the intrinsically bottom-up alpha pattern in favour of a top-down beta pattern at movement *preparation* onset, when beta enhancement is known to abound. More precisely, we propose that the call to prepare a timely keypress involves the transient suppression of afferent auditory alpha information received at M1, specifically in the form of mu suppression, to accommodate top-down, beta-coded information from vPMC, dPMC and SMA cortices to M1, as well as to cerebellum (Kotz & Schwartze, 2010; Wolpert et al., 1998) and A1 (Tian & Poeppel, 2012), either directly or through the M1 relay. In our view, these secondary motor cortices would respectively inform the downstream areas about *what*, *how*, and *when* dimensions of the upcoming movement, in line with the functionality of these motor areas.

The chief stream of beta information travelling from secondary motor areas to M1 would include the motor command itself, which, if allowed to be triggered, will travel down to motor neurons of the spinal cord, and produce a movement. In addition, we theorize that the pre/motor streams received by A1 and cerebellum would instead involve *efference copies* of the motor command, in line with the broad definition of action simulations posited by auditory-motor coupling models like Active Sensing (Schroeder et al., 2010). Note that the global explanation we have thus far provided is highly consistent with findings that alpha power is *inversely* related to both the BOLD signal (Laufs et al., 2003) and cortical excitability (Romei et al., 2007), whereas alpha-band *suppression* is synonymous with the release of cortical inhibition, the

facilitation of motor execution and, critically, inter-regional communication (Fries et al., 2015).

In sum, our claim for this stage of the beta cycle is that predictive alignment to *what* and *when* aspects of realistic musical sequences: 1) involves inhibition of A1 sensory efferents to M1; 2) elicits an increase in M1 responsiveness to predictive action-planning information from secondary motor areas; and 3) either causes or is caused by a transient, timely suppression of mu activity—most likely the former.

From a cognitive perspective, particularly within the *common coding* theory (Maes et al., 2014), this explanation accounts for the transition from the default state of inverse modelling, which involves decoding how ongoing sensory inputs were produced by specific movements (Wolpert et al., 1995), to forward modelling, which involves predicting the sensory consequences of upcoming actions (Halász & Cunnington, 2012). Crucially, given that the function of forward models in auditory-motor coupling is still being solidified in the literature (Maes et al., 2014), we note that our explanation provides testable hypotheses about the specific moment and site of implementation of such forward models. Namely, that the period before mu suppression occurs involves the generation of inverse models via alpha-band signalling from A1 to motor and premotor cortices, whereas the period after mu suppression occurs involves forward models being sent via beta-coded signals from premotor cortices to M1, and from pre/motor cortices to A1 and cerebellum.

Moving on from this first beta enhancement stage would place us squarely on the motor execution and subsequent cessation side of the equation, where the literature reviewed throughout this manuscript enables us to formulate a purely theoretical continuation of the rest of the beta cycle. As mentioned, at movement onset, M1 will send the motor command down the spinal tract, and motor beta will desynchronize. Here, we propose that this desynchronization in M1 is meant to inhibit the incoming flow of top-down motor planning information from premotor areas to accommodate fast, gamma-based cerebellar error correction signals (Bastos et al., 2012) that can

shift or interrupt the movement as it is happening. To substantiate this idea, we recall the findings discussed in the General Introduction that gamma oscillations in the motor cortex are engaged during the active phase of voluntary motor actions (Cheyne et al., 2008) and that competing motor responses generate alterations in this band (Gaetz et al., 2013).

Moreover, as mentioned, online perturbations in auditory feedback consistently induce compensatory reflexes that correlate with increases in cerebellar activity during speech production (Tourville et al., 2008), and recent models such as the *forward model theory* (Welniarz et al., 2021) are increasingly placing cerebellum at the centre of perception of agency (Barone & Rossiter, 2021). More importantly, the cerebellar error correction signals we hypothesize would depend on the efference copy that was received by cerebellum during the motor preparation stage, an idea that is directly aligned with the formulations of the aforementioned *internal cerebellar forward model on motor planning and motor control* (Wolpert et al., 1998) whereby M1 sends efference copies of actions to the cerebellum for the generation of motor-to-somatosensory predictions, as well as similar evidence in the auditory-motor domain that self-initiated sounds suppress auditory N100 ERPs in controls but not patients with cerebellar lesions (Knolle et al., 2012).

Shortly after, at movement cessation, motor beta will start to rebound, and we propose that rebound onset is what will now prompt cortical layers 5/6 in M1 to begin ramping up the discharge of beta waves downstream to A1. Or, more accurately, to the inhibitory *interneurons* in auditory cortex that are known to suppress activity in superficial layers of A1 (Barone & Rossiter, 2021). On the one hand, this beta-dominant discharge from M1 would result in an overall dampening of all cortical activity in A1, a fact that is well-documented for sounds that are self-produced in humans (Martikainen, 2004), monkeys (Eliades & Wang, 2008), and mice (Nelson et al., 2013; Schneider et al., 2014).

On the other hand, we predict that this would also lead to a *relative* enhancement of beta activity in A1, a sharpening of neuronal tuning curves, and an increase in global SNR—all of which is also known to occur across various animal models (Curto et al., 2009; Marguet & Harris, 2011; Pachitariu et al., 2015) and has been hypothesized to occur in humans as well (Caras & Sanes, 2017). As explained at various points, the purpose of this mechanism is to *streamline* the upcoming processing of the sensory consequences of the action that just took place, an idea that finds its home in the *common coding* theory. However, the dynamical systems perspective also offers up a conceptualization of this process as being an optimization of neural population entrainment in A1 to the upcoming self-generated stimulus (Stefanics et al., 2010). Under this view, the optimization is achieved through either delta or theta phase resetting (Mégevand et al., 2018), depending on the presentation rate of the stimulus stream. Additionally, it may involve delta-beta or theta-beta mode-locking (Arnal et al., 2015), which would further enhance the temporal synchronization of neural activity to the upcoming sensory event.

Should there be a spectrotemporal match within A1 between the efference copy received from pre/motor areas (Tian & Poeppel, 2012) and the subsequently produced sound, the predictive coding framework dictates that the cycle would loop back around onto the default alpha-mediated auditory-to-motor flow state without additional computations (Friston, 2005; Rao & Ballard, 1999). In this case, A1 would relay spectrotemporal codes to vPMC, M1, and dPMC via alpha, as suggested by the findings in Study 1. Conversely, in case of a spectrotemporal mismatch between the efference copy and the ground truth of the self-produced sound, Arnal & Giraud (2012) and others claim that A1 cortical layers 2/3 (Roopun et al., 2008) will *additionally* send prediction errors via fast gamma waves (Bastos et al., 2012) to the next cortical dorsal auditory stream region's layer IV (Douglas & Martin, 2004). In this scenario, the area in question would likely include planum temporale, where we propose that the error signal will be re-routed for timely integration, in line with the

redirective function ascribed to this region (Griffiths & Warren, 2002; Hickok & Saberi, 2012).

At the cortical level, this information would continue travelling further up the dorsal auditory stream from layers 2/3 of the sender region (Roopun et al., 2008) to layer IV of the receiver region (Douglas & Martin, 2004), once again, as predicted in Arnal & Giraud (2012). However, we hypothesize that the gamma-bound error signal will also be routed down the direct corticothalamic pathway from PT to MGB (Hickok & Saberi, 2012)—or even sooner from A1 to MGB (Winer, 1992)—after which it might be relayed from thalamus to cerebellum and basal ganglia. Notably, this idea rests on the aforementioned claim made by others that the cerebellum receives information about incoming stimuli directly from auditory centres, encodes the temporal relationships between them, and sends this information upwards to motor cortices (Kotz & Schwartze, 2010).

As for basal ganglia, justification for their involvement can be found in the newly revised ASAP hypothesis (Cannon & Patel, 2021), which proposes that dorsal striatum in particular will sequence neural activity to support rhythmic anticipation and, more importantly, motor alignment. According to the authors, the dorsal striatum achieves this goal by organizing the efference copies specifically generated in SMA, thus rendering this model a close match to our idea that pre/motor cortices send movement efference copies to cerebellum during motor preparation for error-correction. Finally, after prompt processing in these two subcortical areas, we propose that striatal-cerebellar outputs might ascend to synergistically innervate thalamic aspects VA and VL, where the interplay of excitatory and inhibitory discharges are known to fine-tune the next movement cycle (Hintzen et al., 2018; Lanciego et al., 2012), ostensibly in a way that would account for the gamma-coded prediction error (Bastos et al., 2012) in due time.

4.4 Future Directions

Starting with Study 1, the first future step would be to expand on our focused exploration of key auditory- and visuomotor pairings by undertaking a network-level analysis of auditory-motor and visuomotor systems that is informed by our findings. In particular, this could involve visual cortical regions for higher-order processing than V1 that may have more direct connections to motor and premotor areas, as doing so would help to re-test our claim that auditory-motor coupling involves widespread phase-to-phase coupling to a larger degree than visuomotor coupling. A good candidate area would be V5/MT, given its established role in motion perception and its known connectivity with dorsal stream regions such as dPMC and PPC (Grosbras et al., 2012; Friston & Büchel, 2000). Similarly, given the past and presently growing attention being given to SMA, with its links to timing via connections with basal ganglia (Schwartze & Kotz, 2024), we suggest adding this area to the expanded motor ROI set. The inclusion of PPC itself would also be interesting, given its pivotal role in integrating multisensory information (Whitlock, 2017) and coordinating motor functions (Andersen & Cui, 2009), with its specific role in visuospatial processing (Bai et al., 2021) adding another anchor point to an expanded comparison between auditory-motor and visuomotor systems.

Exploring the connections between and across motor regions would also provide insights into how these regions engage each other to support processes such as motor preparation. More precisely, given the nested hypotheses outlined in the previous section, we predict that resting premotor and SMA cortices predominantly send *top-down* beta information to M1, and that M1 conversely preferentially sends alpha information back up to these secondary motor cortices. Under this network approach, it would also be highly informative to include all potential contralateral connections between ROIs, and to place special focus on the link going from right A1 to left M1, given that this connection has emerged before in the literature (Klein et al., 2016) and a similar pairing was highlighted in Morillon & Baillet (2017) between

right associative auditory cortex and left sensorimotor cortex. Finally, since our study explores the resting state of the brain and is therefore not tied to any stimuli or experimental conditions, it would be useful for future research to attempt to replicate our findings. On that note, while Study 1 used the original release of the OMEGA repository (Niso et al., 2016), the latest release features hundreds of new scans, so the use of this version of the dataset for replication should be prioritized by others.

Regarding Study 2, the first future step will be promptly taken by ourselves, and it involves rendering the data and statistical models more amenable to exploring *brain-behaviour* interactions. Presently, results indicate that the subject-averaged collection of channels displays significantly more suppression, exclusively in the mu band, to the learned melody after training than in any of the other conditions. This is a good preliminary result that is informative of sample-average effects, but since it does not account for individual differences among the participants, presently it cannot be extrapolated beyond the confines of the study. More importantly, performing new statistics at the subject level will allow us to deploy more complex, mixed-effects models where *between-subjects* variables, such as motor learning slope, can be correlated to each person's degree of mu suppression. Another improvement would be to formulate and evaluate a specific hypothesis about fluctuations in the beta band, given the intimate relation between this band and predictive processing discussed throughout this thesis. Indeed, as mentioned in Chapter 3, most consider beta to be a subcomponent or at the very least a subharmonic of the mu complex altogether, although it seems that its neural generators do not necessarily overlap with those of the upper alpha component (Larionova et al., 2022). At any rate, this being the state of the art certainly warrants the eventual inclusion of beta into our passive mu suppression paradigm, provided a methodology such as MEG is used to distinguishing the distinct sources.

On this note, it is critical that future studies with higher spatial resolution help us pinpoint the precise generator of the upper alpha suppression we report to ensure

that it is indeed sensorimotor mu suppression, especially given that we were the first to report that suppression occurs at such fast rates. Finally, it would be highly illuminating for future studies to evaluate our hypothesis that the degree of mu suppression is also affected by task complexity, specifically spectral/effector and sequencing requirements. This could be done by systematically varying the melodic structure in terms of the number and type of melodic and rhythmic transitions, which in turn would allow us to gauge the amount of concurrent vPMC and dPMC involvement. If proven to play a role, it would be fascinating to determine whether these relationships are positive or negative, as well as linear or non-linear.

General Conclusion

The present thesis investigated the intimate relationship between the auditory-motor system and neural oscillations, the now proverbial "clocks" of the brain (Buzsáki & Draguhn, 2004). Chapter 1 provided a review of the literature surrounding conceptual and implementation-level auditory-motor frameworks and theories, a description of the key cortical and subcortical neuroanatomy subserving the system, and a cursory account of how neural oscillations have been studied, modelled, challenged, and vindicated over the last five decades. Chapter 2 sought to determine the degree to which common oscillatory patterns of auditory-motor activity are task-dependent or instead reflect intrinsic phase-based connectivity between auditory and motor regions, as compared to visual and motor regions. Results demonstrated that resting functional connectivity patterns across cortical auditory and motor areas are permeated by larger, multi-band phase-to-phase coupling indices than visual and motor areas, thereby supporting the notion that precisely time-locked auditory-motor connections are inherent to this system. The second objective of this study was to pinpoint the degree of similarity between auditory-motor resting state networks and phase-based connectivity patterns associated with auditory-motor tasks. In this regard, effective connectivity revealed significant overlap with task-related patterns, such as a predominant motor-to-auditory direction of information flow in the beta band.

Chapter 3 was designed to assess whether auditory-motor co-activations observed in fMRI activation-based and stimulation studies are predictive of auditory events. The data revealed that non-musicians do develop anticipatory adaptations to learned melodies over short training periods in the form of mu-band suppression, indicating that predictive processes involve the upper alpha band. This study also examined whether the presence or absence of concurrent movements induces differences in the oscillatory processing of sounds associated with a learned motor program, and

results revealed common coordinates for both active and passive mu band suppression, thereby supporting claims in the literature of common mechanisms underlying both action and motor imagery.

Finally, the current Chapter 4 unified the theoretical, anatomical and empirical threads covered in the other three chapters into a cohesive framework that, as is to be expected and celebrated in science, produced a number of hypotheses that is at least one order of magnitude above the number of hypotheses these seven years of intense PhD work have managed to resolve. Thus, overall, the current manuscript's highest aim has been to become a faithful, comprehensive snapshot of the past, present, and future inroads into what makes all our auditory-motor oscillatory "clocks" tick in time to the music.

General Bibliography

Achuthan, S., Butera, R. J., & Canavier, C. C. (2011). Synaptic and intrinsic determinants of the phase resetting curve for weak coupling. *Journal of Computational Neuroscience*, *30*(2), 373–390.

Akkal, D., Dum, R. P., & Strick, P. L. (2007). Supplementary Motor Area and Presupplementary Motor Area: Targets of Basal Ganglia and Cerebellar Output. *The Journal of Neuroscience*, *27*(40), 10659–10673.

Albouy, P., Weiss, A., Baillet, S., & Zatorre, R. J. (2017). Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. *Neuron*, *94*(1), 193-206.e5.

Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K., & Procyk, E. (2013). The Location of Feedback-Related Activity in the Midcingulate Cortex Is Predicted by Local Morphology. *The Journal of Neuroscience*, *33*(5), 2217–2228.

Amiez, C., Sallet, J., Procyk, E., & Petrides, M. (2012). Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration. *NeuroImage*, *63*(3), 1078–1090.

Andersen, L. M., & Dalal, S. S. (2021). The cerebellar clock: Predicting and timing somatosensory touch. *NeuroImage*, *238*, 118202.

Andoh, J., Matsushita, R., & Zatorre, R. J. (2015). Asymmetric Interhemispheric Transfer in the Auditory Network: Evidence from TMS, Resting-State fMRI, and Diffusion Imaging. *The Journal of Neuroscience*, *35*(43), 14602–14611.

Arnal, L. H. (2012). Predicting "When" Using the Motor System's Beta-Band Oscillations. *Frontiers in Human Neuroscience*, *6*.

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta–Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral Cortex*, *25*(9), 3077–3085.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Aschersleben, G. (2002). Temporal Control of Movements in Sensorimotor Synchronization. *Brain and Cognition*, *48*(1), 66–79.

Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, *4*(2), eaao3842.

Baer, L. H., Park, M. T. M., Bailey, J. A., Chakravarty, M. M., Li, K. Z. H., & Penhune, V. B. (2015). Regional cerebellar volumes are related to early musical training and finger tapping performance. *NeuroImage*, *109*, 130–139.

Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, *20*(3), 327–339.

Barone, J., & Rossiter, H. E. (2021). Understanding the Role of Sensorimotor Beta Oscillations. *Frontiers in Systems Neuroscience*, *15*, 655886.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645.

Başar, E., Schürmann, M., Başar-Eroglu, C., & Karakaş, S. (1997). Alpha oscillations in brain functioning: an integrative theory. *International Journal of Psychophysiology*, *26*(1–3), 5–29.

Bassolino, M., Serino, A., Ubaldi, S., & Làdavas, E. (2010). Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia*, *48*(3), 803–811.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711.

Belin, P., & Zatorre, R. J. (2000). 'What', 'where' and 'how' in auditory cortex. *Nature Neuroscience*, *3*(10), 965–966.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical Correlates of Musicianship as Revealed by Cortical Thickness and Voxel-Based Morphometry. *Cerebral Cortex*, *19*(7), 1583–1596.

Bermudez, P., & Zatorre, R. J. (2005). Conditional Associative Memory for Musical Stimuli in Nonmusicians: Implications for Absolute Pitch. *The Journal of Neuroscience*, *25*(34), 7718–7723.

Bollimunta, A., Mo, J., Schroeder, C. E., & Ding, M. (2011). Neuronal Mechanisms and Attentional Modulation of Corticothalamic Alpha Oscillations. *The Journal of Neuroscience*, *31*(13), 4935–4943.

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., Zilles, K., & Fink, G. R. (2001). Polymodal Motion Processing in Posterior Parietal and Premotor Cortex. *Neuron*, *29*(1), 287–296.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth.

Buschman, T. J., & Miller, E. K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science*, *315*(5820), 1860–1862.

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926–1929.

Cannon, J. J., & Patel, A. D. (2021). How Beat Perception Co-opts Motor Neurophysiology. *Trends in Cognitive Sciences*, *25*(2), 137–150.

Caras, M. L., & Sanes, D. H. (2017). Top-down modulation of sensory cortex gates perceptual learning. *Proceedings of the National Academy of Sciences*, *114*(37), 9972–9977.

Chainay, H., Krainik, A., Tanguy, M.-L., Gerardin, E., Le Bihan, D., & Lehéricy, S. (2004). Foot, face and hand representation in the human supplementary motor area: *NeuroReport*, *15*(5), 765–769.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to Musical Rhythms Recruits Motor Regions of the Brain. *Cerebral Cortex*, *18*(12), 2844–2854.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The Role of Auditory and Premotor Cortex in Sensorimotor Transformations. *Annals of the New York Academy of Sciences*, *1169*(1), 15–34.

Cheyne, D., Bells, S., Ferrari, P., Gaetz, W., & Bostan, A. C. (2008). Self-paced movements induce high-frequency gamma oscillations in primary motor cortex. *NeuroImage*, *42*(1), 332–342.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Comstock, D. C., Hove, M. J., & Balasubramaniam, R. (2018). Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences. *Frontiers in Computational Neuroscience*, *12*, 53.

Connors, B. W., & Regehr, W. G. (1996). Neuronal firing: Does function follow form? *Current Biology*, *6*(12), 1560–1562.

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "What" and "When" in the Auditory System: Temporal Predictability Enhances Repetition Suppression. *The Journal of Neuroscience*, *31*(50), 18590–18597.

Crasta, J. E., Thaut, M. H., Anderson, C. W., Davies, P. L., & Gavin, W. J. (2018). Auditory priming improves neural synchronization in auditory-motor entrainment. *Neuropsychologia*, *117*, 102–112.

Curto, C., Sakata, S., Marguet, S., Itskov, V., & Harris, K. D. (2009). A Simple Model of Cortical Dynamics Explains Variability and State Dependence of Sensory Responses in Urethane-Anesthetized Auditory Cortex. *The Journal of Neuroscience*, *29*(34), 10600–10612.

Da Costa, S., Van Der Zwaag, W., Marques, J. P., Frackowiak, R. S. J., Clarke, S., & Saenz, M. (2011). Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. *The Journal of Neuroscience*, *31*(40), 14067–14075.

Damm, L., Varoqui, D., De Cock, V. C., Dalla Bella, S., & Bardy, B. (2020). Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. *Neuroscience & Biobehavioral Reviews*, *112*, 553–584.

Daroff, R. B., & Aminoff, M. J. (2014). *Encyclopedia of the neurological sciences*. Academic press.

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, *24*(3), 955–958.

Davare, M., Andres, M., Cosnard, G., Thonnard, J.-L., & Olivier, E. (2006). Dissociating the Role of Ventral and Dorsal Premotor Cortex in Precision Grasping. *The Journal of Neuroscience*, *26*(8), 2260–2268.

de Leon, A. S., & Das, J. M. (2024). Neuroanatomy, Dentate Nucleus. In *StatPearls*. StatPearls Publishing.

Dehaene, S., & Cohen, L. (2007). Cultural Recycling of Cortical Maps. *Neuron*, *56*(2), 384–398.

Desantis, A., Weiss, C., Schütz-Bosbach, S., & Waszak, F. (2012). Believing and Perceiving: Authorship Belief Modulates Sensory Attenuation. *PLoS ONE*, *7*(5), e37959.

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences*, *116*(20), 10113–10121.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45).

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of Choice-Predictive Activity in Human Motor Cortex during Perceptual Decision Making. *Current Biology*, *19*(18), 1581–1585.

Donoghue, J. P., Sanes, J. N., Hatsopoulos, N. G., & Gaál, G. (1998). Neural Discharge and Local Field Potential Oscillations in Primate Motor Cortex During Voluntary Movements. *Journal of Neurophysiology*, *79*(1), 159–173.

Douglas, R. J., & Martin, K. A. C. (2004). Neuronal Circuits of the Neocortex. *Annual Review of Neuroscience*, *27*(1), 419–451.

Dum, R. P., & Strick, P. L. (2004). Motor Areas in the Frontal Lobe: The Anatomical Substrate for the Central Control of Movement. In *Motor Cortex in Voluntary Movements*. CRC Press.

Dum, R. P., & Strick, P. L. (2005). Frontal Lobe Inputs to the Digit Representations of the Motor Areas on the Lateral Surface of the Hemisphere. *The Journal of Neuroscience*, *25*(6), 1375–1386.

Eggermont, J. J. (2010). The Auditory Cortex: The Final Frontier. In R. Meddis, E. A. Lopez-Poveda, R. R. Fay, & A. N. Popper (Eds.), *Computational Models of the Auditory System* (pp. 97–127). Springer US.

Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*(7198), 1102–1106.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716.

Evarts, E. V. (1968). Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of Neurophysiology*, *31*(1), 14–27.

Farina, D., & Negro, F. (2015). Common Synaptic Input to Motor Neurons, Motor Unit Synchronization, and Force Control. *Exercise and Sport Sciences Reviews*, *43*(1), 23–33.

Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage*, *75*, 27–35.

Foster, N. E. V., & Zatorre, R. J. (2010). A Role for the Intraparietal Sulcus in Transforming Musical Pitch Information. *Cerebral Cortex*, *20*(6), 1350–1359.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., Ferrari, P. F., & van IJzendoorn, M. H. (2016). Assessing Human Mirror Activity With EEG Mu Rhythm: A Meta-Analysis. *Psychological Bulletin*, *142*(3), 291–313.

Fraisse, P. (1978). Time and Rhythm Perception. In *Perceptual Coding* (pp. 203–254). Elsevier.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480.

Fries, P. (2009). Neuronal Gamma-Band Synchronization as a Fundamental Process in Cortical Computation. *Annual Review of Neuroscience*, *32*(1), 209–224.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *The Journal of Neuroscience*, *35*(45), 15187–15198.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and Gamma Rhythms in Human Auditory Cortex during Musical Beat Processing. *Annals of the New York Academy of Sciences*, *1169*(1), 89–92.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *The Journal of Neuroscience*, *32*(5), 1791–1802.

Gaetz, W., Liu, C., Zhu, H., Bloy, L., & Roberts, T. P. L. (2013). Evidence for a motor gamma-band network governing response interference. *NeuroImage*, *74*, 245–253.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*(3), 361–377.

Gallese, V., & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3–4), 455–479.

Gazzola, V., & Keysers, C. (2009). The Observation and Execution of Actions Share Motor and Somatosensory Voxels in all Tested Subjects: Single-Subject Analyses of Unsmoothed fMRI Data. *Cerebral Cortex*, *19*(6), 1239–1255.

Georgopoulos, A., Kalaska, J., Caminiti, R., & Massey, J. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *The Journal of Neuroscience*, *2*(11), 1527–1537.

Ghitza, O., & Greenberg, S. (2009). On the Possible Role of Brain Rhythms in Speech Perception: Intelligibility of Time-Compressed Speech with Periodic and Aperiodic Insertions of Silence. *Phonetica*, *66*(1–2), 113–126.

Gierhan, S. M. E. (2013). Connections for auditory language in the human brain. *Brain and Language*, *127*(2), 205–221.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517.

Gnanateja, G. N., Devaraju, D. S., Heyne, M., Quique, Y. M., Sitek, K. R., Tardif, M. C., Tessmer, R., & Dial, H. R. (2022). On the Role of Neural Oscillations Across

Timescales in Speech and Music Processing. *Frontiers in Computational Neuroscience*, *16*.

Godschalk, M., Mitz, A. R., Duin, B. V., & Burga, H. V. D. (1995). Somatotopy of monkey premotor cortex examined with microstimulation. *Neuroscience Research*, *23*(3), 269–279.

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage*, *52*(3), 862–874.

Grahn, J. A. (2009). The Role of the Basal Ganglia in Beat Perception: Neuroimaging and Neuropsychological Investigations. *Annals of the New York Academy of Sciences*, *1169*(1), 35–45.

Grahn, J. A. (2012). Neural Mechanisms of Rhythm Perception: Current Findings and Future Perspectives. *Topics in Cognitive Science*, *4*(4), 585–606.

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Graziano, M. S. A., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*(6), 845–859.

Graziano, M. S. A., Taylor, C. S. R., & Moore, T. (2002a). Complex Movements Evoked by Microstimulation of Precentral Cortex. *Neuron*, *34*(5), 841–851.

Graziano, M. S. A., Taylor, C. S. R., Moore, T., & Cooke, D. F. (2002b). The Cortical Control of Movement Revisited. *Neuron*, *36*(3), 349–362.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348–353.

Gupta, D. S. (2014). Processing of sub- and supra-second intervals in the primate brain results from the calibration of neuronal oscillators via sensory, motor, and feedback processes. *Frontiers in Psychology*, *5*.

Gurtubay, I. G., Alegre, M., Labarga, A., Malanda, A., & Artieda, J. (2004). Gamma band responses to target and non-target auditory stimuli in humans. *Neuroscience Letters*, *367*(1), 6–9.

Haber, S. N. (2016). Corticostriatal circuitry. *Dialogues in Clinical Neuroscience*, *18*(1), 7–21.

Haegens, S., Barczak, A., Musacchia, G., Lipton, M. L., Mehta, A. D., Lakatos, P., & Schroeder, C. E. (2015). Laminar Profile and Physiology of the α Rhythm in Primary Visual, Auditory, and Somatosensory Regions of Neocortex. *The Journal of Neuroscience*, *35*(42), 14341–14352.

Haig, A. R., & Gordon, E. (1998). Prestimulus EEG alpha phase synchronicity influences N100 amplitude and reaction time. *Psychophysiology*, *35*(5), 591–595.

Halász, V., & Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sciences*, *2*(2), 130–146.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281–1292.

He, Y., Li, Y., Pu, Z., Chen, M., Gao, Y., Chen, L., Ruan, Y., Pan, X., Zhou, Y., Ge, Y., Zhou, J., Zheng, W., Huang, Z., Li, Z., & Chen, J.-F. (2020). Striatopallidal Pathway Distinctly Modulates Goal-Directed Valuation and Acquisition of Instrumental Behavior via Striatopallidal Output Projections. *Cerebral Cortex*, *30*(3), 1366–1381.

Herholz, S. C., Coffey, E. B. J., Pantev, C., & Zatorre, R. J. (2016). Dissociation of Neural Networks for Predisposition and for Training-Related Plasticity in Auditory-Motor Learning. *Cerebral Cortex*, *26*(7), 3125–3134.

Herrojo Ruiz, M., Maess, B., Altenmüller, E., Curio, G., & Nikulin, V. V. (2017). Cingulate and cerebellar beta oscillations are engaged in the acquisition of auditory-motor sequences. *Human Brain Mapping*, *38*(10), 5161–5179.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673–682.

Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the Human Planum Temporale Supports Sensory-Motor Integration for Speech Processing. *Journal of Neurophysiology*, *101*(5), 2725–2732.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Hickok, G., & Saberi, K. (2012). Redefining the Functional Organization of the Planum Temporale Region: Space, Objects, and Sensory–Motor Integration. In D. Poeppel, T. Overath, A. N. Popper, & R. R. Fay (Eds.), *The Human Auditory Cortex* (pp. 333–350). Springer.

Highben, Z., & Palmer, C. (2004). Effects of Auditory and Motor Mental Practice in Memorized Piano Performance. *Bulletin of the Council for Research in Music Education*, 58–67.

Hintzen, A., Pelzer, E. A., & Tittgemeyer, M. (2018). Thalamic interactions of cerebellum and basal ganglia. *Brain Structure and Function*, *223*(2), 569–587.

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology*, *6*, 1318.

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*(1), 85–91.

Hoppensteadt, F. C., & Izhikevich, E. M. (1996). Synaptic organizations and dynamical properties of weakly connected neural oscillators. *Biological Cybernetics*, *75*(2), 117–127.

Hoshi, E., & Tanji, J. (2004). Functional specialization in dorsal and ventral premotor areas. In *Progress in Brain Research* (Vol. 143, pp. 507–511). Elsevier.

Hoshi, E., & Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties. *Current Opinion in Neurobiology*, *17*(2), 234–242.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor Adaptation in Speech Production. *Science*, *279*(5354), 1213–1216.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the Auditory Cortex during Speech: An MEG Study. *Journal of Cognitive Neuroscience*, *14*(8), 1125–1138.

Husain, M., & Nachev, P. (2007). Space and the parietal cortex. *Trends in Cognitive Sciences*, *11*(1), 30–36.

Hutcheon, B., & Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends in Neurosciences*, *23*(5), 216–222.

Hyafil, A., Giraud, A.-L., Fontolan, L., & Gutkin, B. (2015). Neural Cross-Frequency Coupling: Connecting Architectures, Mechanisms, and Functions. *Trends in Neurosciences*, *38*(11), 725–740.

Iacoboni, M., Woods, R. P., & Mazziotta, J. C. (1998). Bimodal (auditory and visual) left frontoparietal circuitry for sensorimotor integration and sensorimotor learning. *Brain*, *121*(11), 2135–2143.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*(4), 304–313.

Iversen, J., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. *Proceedings of the 10th International Conference on Music Perception and Cognition*.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-Down Control of Rhythm Perception Modulates Early Auditory Responses. *Annals of the New York Academy of Sciences*, *1169*(1), 58–73.

Izhikevich, E. M. (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*(7), 682–687.

Jang, J., Jones, M., Milne, E., Wilson, D., & Lee, K.-H. (2016). Contingent negative variation (CNV) associated with sensorimotor timing error correction. *NeuroImage*, *127*, 58–66.

Jarvis, M. R., & Mitra, P. P. (2001). Sampling Properties of the Spectrum and Coherency of Sequences of Action Potentials. *Neural Computation*, *13*(4), 717–749.

Jasper, H., & Penfield, W. (1949). Electrocorticograms in man: Effect of voluntary movement upon the electrical activity of the precentral gyrus. *Archiv Für Psychiatrie Und Nervenkrankheiten*, *183*(1–2), 163–174.

Jo, H.-G., Wittmann, M., Hinterberger, T., & Schmidt, S. (2014). The readiness potential reflects intentional binding. *Frontiers in Human Neuroscience*, *8*.

Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., Smith, S. M., Higham, D. J., & Matthews, P. M. (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences*, *101*(36), 13335–13340.

Kaneoke, Y., & Vitek, J. L. (1996). Burst and oscillation as disparate neuronal properties. *Journal of Neuroscience Methods*, *68*(2), 211–223.

Keitel, A., & Gross, J. (2016). Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLOS Biology*, *14*(6), e1002498.

Keller, P. E., & Koch, I. (2008). Action Planning in Sequential Skills: Relations to Music Performance. *Quarterly Journal of Experimental Psychology*, *61*(2), 275–291.

Knierim, J. (2016). Spinal reflexes and descending motor pathways. *Neuroscience Online*.

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The Cerebellum Generates Motor-to-Auditory Predictions: ERP Lesion Evidence. *Journal of Cognitive Neuroscience*, *24*(3), 698–706.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, *297*(5582), 846–848.

König, P. (1994). A method for the quantification of synchrony and oscillatory properties of neuronal activity. *Journal of Neuroscience Methods*, *54*(1), 31–37.

Kornysheva, K., & Diedrichsen, J. (2014). Human premotor areas parse sequences into their spatial and temporal features. *eLife*, *3*, e03043.

Kosslyn, S. M. (1994). *Image and brain*. MIT Press.

Kotz, S. A., Brown, R. M., & Schwartze, M. (2016). Cortico-striatal circuits and the timing of action and perception. *Current Opinion in Behavioral Sciences*, *8*, 42–45.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*(9), 392–399.

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience*, *25*(3), 401–420.

Kurata, K., & Hoffman, D. S. (1994). Differential effects of muscimol microinjection into dorsal and ventral aspects of the premotor cortex of monkeys. *Journal of Neurophysiology*, *71*(3), 1151–1164.

Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, *8*(4), 194–208.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. *The Journal of Neuroscience*, *27*(2), 308–314.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus

Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911.

Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional Neuroanatomy of the Basal Ganglia. *Cold Spring Harbor Perspectives in Medicine*, *2*(12), a009621.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, *19*(4), 527–566.

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *The Psychology of Time*, 189–231.

Large, E. W., & Crawford, J. D. (2002). Auditory Temporal Computation: Interval Selectivity Based on Post-Inhibitory Rebound. *Journal of Computational Neuroscience*, *13*(2), 125–142.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural Networks for Beat Perception in Musical Rhythm. *Frontiers in Systems Neuroscience*, *9*.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*(1), 1–37.

Large, E. W., & Snyder, J. S. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences*, *1169*(1), 46–57.

Larionova, E. V., Garakh, Zh. V., & Zaytseva, Yu. S. (2022). The Mu Rhythm in Current Research: Theoretical and Methodological Aspects. *Neuroscience and Behavioral Physiology*, *52*(7), 999–1016.

Leaver, A. M., & Rauschecker, J. P. (2016). Functional Topography of Human Auditory Cortex. *Journal of Neuroscience*, *36*(4), 1416–1428.

Lega, C., Stephan, M. A., Zatorre, R. J., & Penhune, V. (2016). Testing the Role of Dorsal Premotor Cortex in Auditory-Motor Association Learning Using Transcranical Magnetic Stimulation (TMS). *PLOS ONE*, *11*(9), e0163380.

Lewis, J. W., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *The Journal of Comparative Neurology*, *428*(1), 112–137.

Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of Supplementary Motor Areas in Auditory Processing and Auditory Imagery. *Trends in Neurosciences*, *39*(8), 527–542.

Liu, H., & Larson, C. R. (2007). Effects of perturbation magnitude and voice F level on the pitch-shift reflex. *The Journal of the Acoustical Society of America*, *122*(6), 3671–3677.

Llinás, R. R. (1988). The Intrinsic Electrophysiological Properties of Mammalian Neurons: Insights into Central Nervous System Function. *Science*, *242*(4886), 1654–1664.

Lobier, M., Siebenhühner, F., Palva, S., & Palva, J. M. (2014). Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *Neuroimage*, *85*, 853–872.

Luo, J., Hage, S. R., & Moss, C. F. (2018). The Lombard Effect: From Acoustics to Neural Mechanisms. *Trends in Neurosciences*, *41*(12), 938–949.

MacEvoy, S. P., Hanks, T. D., & Paradiso, M. A. (2008). Macaque V1 Activity During Natural Vision: Effects of Natural Scenes and Saccades. *Journal of Neurophysiology*, *99*(2), 460–472.

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, *4*.

Mahjoory, K., Schoffelen, J.-M., Keitel, A., & Gross, J. (2020). The frequency gradient of human resting-state brain oscillations follows cortical hierarchies. *eLife*, *9*, e53715.

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing Expectancy Violations during Music Performance and Perception: An ERP Study. *Journal of Cognitive Neuroscience*, *22*(10), 2401–2413.

Marder, E. (2000). Motor pattern generation. *Current Opinion in Neurobiology*, *10*(6), 691–698.

Marguet, S. L., & Harris, K. D. (2011). State-Dependent Representation of Amplitude-Modulated Noise Stimuli in Rat Auditory Cortex. *The Journal of Neuroscience*, *31*(17), 6414–6420.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190.

Martikainen, M. H. (2004). Suppressed Responses to Self-triggered Sounds in the Human Auditory Cortex. *Cerebral Cortex*, *15*(3), 299–302.

Matano, S. (2001). Brief communication: Proportions of the ventral half of the cerebellar dentate nucleus in humans and great apes. *American Journal of Physical Anthropology*, *114*(2), 163–165.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367.

McAuley, J. D., & Kidd, G. R. (1995). Temporally directed attending in the discrimination of tempo: Further support for an entrainment model. *The Journal of the Acoustical Society of America*, *97*(5_Supplement), 3278–3278.

McIntosh, G. C., Brown, S. H., Rice, R. R., & Thaut, M. H. (1997). Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *62*(1), 22–26.

Mégevand, P., Mercier, M. R., Groppe, D. M., Golumbic, E. Z., Mesgarani, N., Beauchamp, M. S., Schroeder, C. E., & Mehta, A. D. (2018). *Phase resetting in human auditory cortex to visual speech*.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140093.

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*.

Milardi, D., Quartarone, A., Bramanti, A., Anastasi, G., Bertino, S., Basile, G. A., Buonasera, P., Pilone, G., Celeste, G., Rizzo, G., Bruschetta, D., & Cacciola, A. (2019). The Cortico-Basal Ganglia-Cerebellar Network: Past, Present and Future Perspectives. *Frontiers in Systems Neuroscience*, *13*, 61.

Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, *8*.

Moran, D. W., & Schwartz, A. B. (1999). Motor Cortical Representation of Speed and Direction During Reaching. *Journal of Neurophysiology*, *82*(5), 2676–2692.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42).

Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology*, *31*, 230–238.

Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A.-L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences*, *107*(43), 18688–18693.

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, *5*(1), 5255.

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., & Malach, R. (2005). Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex. *Science*, *309*(5736), 951–954.

Mureşan, R. C., Jurjuţ, O. F., Moca, V. V., Singer, W., & Nikolić, D. (2008). The Oscillation Score: An Efficient Method for Estimating Oscillation Strength in Neuronal Activity. *Journal of Neurophysiology*, *99*(3), 1333–1353.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, *9*(11), 856–869.

Nambu, A., Tokuno, H., & Takada, M. (2002). Functional significance of the cortico–subthalamo–pallidal 'hyperdirect' pathway. *Neuroscience Research*, *43*(2), 111–117.

Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., & Mooney, R. (2013). A Circuit for Motor Cortical Modulation of Auditory Cortical Activity. *The Journal of Neuroscience*, *33*(36), 14342–14353.

Nijhuis, P., Keller, P. E., Nozaradan, S., & Varlet, M. (2021). Dynamic modulation of cortico-muscular coupling during real and imagined sensorimotor synchronisation. *NeuroImage*, *238*, 118209.

Ninokura, Y., Mushiake, H., & Tanji, J. (2004). Integration of Temporal Order and Object Information in the Monkey Lateral Prefrontal Cortex. *Journal of Neurophysiology*, *91*(1), 555–560.

Niso, G., Rogers, C., Moreau, J. T., Chen, L.-Y., Madjar, C., Das, S., Bock, E., Tadel, F., Evans, A. C., Jolicoeur, P., & Baillet, S. (2016). OMEGA: The Open MEG Archive. *NeuroImage*, *124*, 1182–1187.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48.

Novembre, G., & Iannetti, G. D. (2018). Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences*, *115*(47).

Novembre, G., & Keller, P. E. (2014). A conceptual review on action-perception coupling in the musicians' brain: what is it good for? *Frontiers in Human Neuroscience*, *8*, 603.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *Journal of Neuroscience*, *31*(28), 10234–10240.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *Journal of Neuroscience*, *32*(49), 17572–17581.

Oane, I., Barborica, A., & Mindruta, I. R. (2023). Cingulate Cortex: Anatomy, Structural and Functional Connectivity. *Journal of Clinical Neurophysiology*, *40*(6), 482–490.

Pachitariu, M., Lyamzin, D. R., Sahani, M., & Lesica, N. A. (2015). State-Dependent Population Coding in Primary Auditory Cortex. *The Journal of Neuroscience*, *35*(5), 2058–2073.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 728–741.

Palomar-García, M.-Á., Zatorre, R. J., Ventura-Campos, N., Bueichekú, E., & Ávila, C. (2016). Modulation of Functional Connectivity in Auditory–Motor Networks in Musicians Compared with Nonmusicians. *Cerebral Cortex*, bhw120.

Palva, S., Linkenkaer-Hansen, K., Näätänen, R., & Palva, J. M. (2005). Early Neural Correlates of Conscious Somatosensory Perception. *The Journal of Neuroscience*, *25*(21), 5248–5258.

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, *25*(12), 1649–1653.

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception*, *11*(4), 409–464.

Parsons, L. M., Sergent, J., Hodges, D. A., & Fox, P. T. (2005). The brain basis of piano performance. *Neuropsychologia*, *43*(2), 199–215.

Patel, A. D. (2007). *Music, Language, and the Brain*. Oxford University Press.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology*, *19*(10), 827–830.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*(2), 226–238.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, *60*(4), 389–443.

Pesaran, B., Pezaris, J. S., Sahani, M., Mitra, P. P., & Andersen, R. A. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature Neuroscience*, *5*(8), 805–811.

Petrides, M., & Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *Journal of Comparative Neurology*, *273*(1), 52–66.

Pfordresher, P. Q., Mantell, J. T., Brown, S., Zivadinov, R., & Cox, J. L. (2014). Brain responses to altered auditory feedback during musical keyboard production: An fMRI study. *Brain Research*, *1556*, 28–37.

Pfurtscheller, G. (1981). Central beta rhythm during sensorimotor activities in man. *Electroencephalography and Clinical Neurophysiology*, *51*(3), 253–264.

Picard, N., & Strick, P. L. (1996). Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation. *Cerebral Cortex*, *6*(3), 342–353.

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, *11*(6), 663–672.

Pike, F. G., Goddard, R. S., Suckling, J. M., Ganter, P., Kasthuri, N., & Paulsen, O. (2000). Distinct frequency preferences of different types of rat hippocampal neurones in response to oscillatory input currents. *The Journal of Physiology*, *529*(1), 205–213.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing." *Brain Research Reviews*, *50*(1), 57–68.

Pineda, J. A., Grichanik, M., Williams, V., Trieu, M., Chang, H., & Keysers, C. (2013). EEG sensorimotor correlates of translating sounds into actions. *Frontiers in Neuroscience*, *7*, 203.

Pollok, B., Gross, J., Müller, K., Aschersleben, G., & Schnitzler, A. (2005). The cerebral oscillatory network associated with auditorily paced finger movements. *NeuroImage*, *24*(3), 646–655.

Praamstra, P., Kourtis, D., Fei Kwok, H., & Oostenveld, R. (2006). Neurophysiology of Implicit Timing in Serial Choice Reaction-Time Performance. *The Journal of Neuroscience*, *26*(20), 5448–5455.

Pranjić, M., Braun Janzen, T., Vukšić, N., & Thaut, M. (2024). From Sound to Movement: Mapping the Neural Mechanisms of Auditory–Motor Entrainment and Synchronization. *Brain Sciences*, *14*(11), 1063.

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In O. Neumann & W. Prinz (Eds.), *Relationships Between Perception and Action: Current Approaches* (pp. 167–201). Springer.

Procyk, E., Wilson, C. R. E., Stoll, F. M., Faraut, M. C. M., Petrides, M., & Amiez, C. (2014). Midcingulate Motor Map and Feedback Detection: Converging Data from Humans and Monkeys. *Cerebral Cortex*.

Pulvermüller, F., Huss, M., Kherif, F., Moscoso Del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, *103*(20), 7865–7870.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). Circuits within the Cerebellum. In *Neuroscience. 2nd edition*. Sinauer Associates.

Rahimi-Balaei, M., Ghiyamihoor, F., Rad, A. A., Ashtari, N., Toback, M., Bergen, H., & Marzban, H. (2023). The Embryology and Anatomy of the Cerebellum. In H. Marzban (Ed.), *Development of the Cerebellum from Molecular Aspects to Diseases* (pp. 33–44). Springer International Publishing.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.

Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology*, *8*(4), 516–521.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*(6), 718–724.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800–11806.

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology*, *5*.

Regev, M., Halpern, A. R., Owen, A. M., Patel, A. D., & Zatorre, R. J. (2021). Mapping Specific Mental Content during Musical Imagery. *Cerebral Cortex*, *31*(8), 3622–3640.

Repp, B. H., & Knoblich, G. (2004). Perceiving Action Identity: How Pianists Recognize Their Own Performances. *Psychological Science*, *15*(9), 604–609.

Reznik, D., & Mukamel, R. (2019). Motor output, neural states and auditory perception. *Neuroscience & Biobehavioral Reviews*, *96*, 116–126.

Reznik, D., Ossmy, O., & Mukamel, R. (2015). Enhanced Auditory Evoked Activity to Self-Generated Sounds Is Mediated by Primary and Supplementary Motor Cortices. *The Journal of Neuroscience*, *35*(5), 2173–2180.

Rijntjes, M., Dettmers, C., Büchel, C., Kiebel, S., Frackowiak, R. S. J., & Weiller, C. (1999). A Blueprint for Movement: Functional and Anatomical Representations in the Human Motor System. *The Journal of Neuroscience*, *19*(18), 8043–8048.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, *12*(2), 149–154.

Rocha, G. S., Freire, M. A. M., Britto, A. M., Paiva, K. M., Oliveira, R. F., Fonseca, I. A. T., Araújo, D. P., Oliveira, L. C., Guzen, F. P., Morais, P. L. A. G., & Cavalcanti, J. R. L. P. (2023). Basal ganglia for beginners: the basic concepts you need to know and their role in movement control. *Frontiers in Systems Neuroscience*, *17*, 1242929.

Rohenkohl, G., & Nobre, A. C. (2011). Alpha Oscillations Related to Anticipatory Attention Follow Temporal Expectations. *The Journal of Neuroscience*, *31*(40), 14076–14084.

Rolls, E. T. (2019). The cingulate cortex and limbic systems for emotion, action, and memory. *Brain Structure and Function*, *224*(9), 3001–3018.

Roopun, A. K., Kramer, M. A., Carracedo, L. M., Kaiser, M., Davies, C. H., Traub, R. D., Kopell, N. J., & Whittington, M. A. (2008). Period concatenation underlies interactions between gamma and beta rhythms in neocortex. *Frontiers in Cellular Neuroscience*, *2*, 1.

Roopun, A. K., Middleton, S. J., Cunningham, M. O., LeBeau, F. E. N., Bibbig, A., Whittington, M. A., & Traub, R. D. (2006). A beta2-frequency (20–30 Hz) oscillation in nonsynaptic networks of somatosensory cortex. *Proceedings of the National Academy of Sciences*, *103*(42), 15646–15650.

Rosso, M., Leman, M., & Moumdjian, L. (2021). Neural Entrainment Meets Behavior: The Stability Index as a Neural Outcome Measure of Auditory-Motor Coupling. *Frontiers in Human Neuroscience*, *15*.

Ruiz, M. H., Jabusch, H.-C., & Altenmüller, E. (2009). Detecting Wrong Notes in Advance: Neuronal Correlates of Error Monitoring in Pianists. *Cerebral Cortex*, *19*(11), 2625–2639.

Ruiz, M. H., Strübing, F., Jabusch, H.-C., & Altenmüller, E. (2011). EEG oscillatory patterns are associated with error prediction during music performance and are altered in musician's dystonia. *NeuroImage*, *55*(4), 1791–1803.

Rummell, B. P., Klee, J. L., & Sigurdsson, T. (2016). Attenuation of Responses to Self-Generated Sounds in Auditory Cortical Neurons. *The Journal of Neuroscience*, *36*(47), 12010–12026.

Sabate, M., Llanos, C., Enriquez, E., Gonzalez, B., & Rodriguez, M. (2011). Fast modulation of alpha activity during visual processing and motor control. *Neuroscience*, *189*, 236–249.

Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., & Hatsopoulos, N. G. (2010). Fast and Slow Oscillations in Human Primary Motor Cortex Predict Oncoming Behaviorally Relevant Cues. *Neuron*, *65*(4), 461–471.

Samiee, S., Vuvan, D., Florin, E., Albouy, P., Peretz, I., & Baillet, S. (2022). Cross-Frequency Brain Network Dynamics Support Pitch Change Detection. *The Journal of Neuroscience*, *42*(18), 3823–3835.

Schellekens, W., Bakker, C., Ramsey, N. F., & Petridou, N. (2022). Moving in on human motor cortex. Characterizing the relationship between body parts with non-rigid population response fields. *PLOS Computational Biology*, *18*(4), e1009955.

Schepers, I. M., Schneider, T. R., Hipp, J. F., Engel, A. K., & Senkowski, D. (2013). Noise alters beta-band activity in superior temporal cortex during audiovisual speech processing. *NeuroImage*, *70*, 101–112.

Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K.-P., & Bremmer, F. (2005). Multisensory Space Representations in the Macaque Ventral Intraparietal Area. *Journal of Neuroscience*, *25*(18).

Schmidt, F., Chen, Y., Keitel, A., Rösch, S., Hannemann, R., Serman, M., Hauswald, A., & Weisz, N. (2023). Neural speech tracking shifts from the syllabic to the modulation rate of speech as intelligibility decreases. *Psychophysiology*, *60*(11), e14362.

Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, *513*(7517), 189–194.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176.

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218.

Schubotz, R. I., & Von Cramon, D. Y. (2003). Functional–anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. *NeuroImage*, *20*, S120–S131.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

Segado, M., Hollinger, A., Thibodeau, J., Penhune, V., & Zatorre, R. J. (2018). Partially overlapping brain networks for singing and cello playing. *Frontiers in Neuroscience*, *12*, 351.

Seitz, R. J., Stephan, K. M., & Binkofski, F. (2000). Control of action as mediated by the human frontal lobe. *Experimental Brain Research*, *133*(1), 71–80.

Seo, J. P., & Jang, S. H. (2013). Different Characteristics of the Corticospinal Tract According to the Cerebral Origin: DTI Study. *American Journal of Neuroradiology*, *34*(7), 1359–1363.

Serino, A., Canzoneri, E., & Avenanti, A. (2011). Fronto-parietal Areas Necessary for a Multisensory Representation of Peripersonal Space in Humans: An rTMS Study. *Journal of Cognitive Neuroscience*, *23*(10), 2956–2967.

Shah, A. S. (2004). Neural Dynamics and the Fundamental Mechanisms of Event-related Brain Potentials. *Cerebral Cortex*, *14*(5), 476–483.

Sharma, N., & Baron, J.-C. (2013). Does motor imagery share neural networks with executed movement: a multivariate fMRI analysis. *Frontiers in Human Neuroscience*, *7*.

Simon, S. R., Meunier, M., Piettre, L., Berardi, A. M., Segebarth, C. M., & Boussaoud, D. (2002). Spatial Attention and Memory Versus Motor Preparation: Premotor Cortex Involvement as Revealed by fMRI. *Journal of Neurophysiology*, *88*(4), 2047–2057.

Singh, I. (2006). *Textbook of human neuroanatomy*. Jaypee Brothers Publishers.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6), 482–489.

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase Entrainment of Human Delta Oscillations Can Mediate the Effects of Expectation on Reaction Speed. *The Journal of Neuroscience*, *30*(41), 13578–13585.

Stegemöller, E. L., Izbicki, P., & Hibbing, P. (2018). The influence of moving with music on motor cortical activity. *Neuroscience Letters*, *683*, 27–32.

Steinschneider, M., Fishman, Y. I., & Arezzo, J. C. (2008). Spectrotemporal Analysis of Evoked and Induced Electroencephalographic Responses in Primary Auditory Cortex (A1) of the Awake Monkey. *Cerebral Cortex*, *18*(3), 610–625.

Stephan, M. A., Lega, C., & Penhune, V. B. (2018). Auditory prediction cues motor preparation in the absence of movements. *NeuroImage*, *174*, 288–296.

Stiefel, K. M., & Ermentrout, G. B. (2016). Neurons as oscillators. *Journal of Neurophysiology*, *116*(6), 2950–2960.

Styns, F., Van Noorden, L., Moelants, D., & Leman, M. (2007). Walking on music. *Human Movement Science*, *26*(5), 769–785.

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Zion Golumbic, E. (2017). Neural Entrainment to the Beat: The "Missing-Pulse" Phenomenon. *The Journal of Neuroscience*, *37*(26), 6331–6341.

Tanji, J. (2001). Sequential Organization of Multiple Movements: Involvement of Cortical Motor Areas. *Annual Review of Neuroscience*, *24*(1), 631–651.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *The Journal of Neuroscience*, *31*(10), 3805–3812.

Thaut, M. H., Miller, R. A., & Schauer, L. M. (1998). Multiple synchronization strategies in rhythmic sensorimotor tasks: phase vs period correction. *Biological Cybernetics*, *79*(3), 241–250.

Thaut, M., & Hoemberg, V. (Eds.). (2016). *Handbook of neurologic music therapy* (First published in paperback). Oxford University Press.

Thaut, M., Schleiffers, S., & Davis, W. (1991). Analysis of EMG Activity in Biceps and Triceps Muscle in an Upper Extremity Gross Motor Task under the Influence of Auditory Rhythm. *Journal of Music Therapy*, *28*(2), 64–88.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band Electroencephalographic Activity over Occipital Cortex Indexes Visuospatial Attention Bias and Predicts Visual Target Detection. *The Journal of Neuroscience*, *26*(37), 9494–9502.

Tian, X., & Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, *6*.

Todd, N. P. M., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, *39*(3), 1429–1443.

Trost, W., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia*, *96*, 96–110.

Van Vugt, F. T., & Ostry, D. J. (2018). The Structure and Acquisition of Sensorimotor Maps. *Journal of Cognitive Neuroscience*, *30*(3), 290–306.

Vogt, B. A. (2016). Midcingulate cortex: Structure, connections, homologies, functions and diseases. *Journal of Chemical Neuroanatomy*, *74*, 28–46.

Voogd, J., Shinoda, Y., Ruigrok, T. J. H., & Sugihara, I. (2013). Cerebellar Nuclei and the Inferior Olivary Nuclei: Organization and Connections. In M. Manto, J. D. Schmahmann, F. Rossi, D. L. Gruol, & N. Koibuchi (Eds.), *Handbook of the Cerebellum and Cerebellar Disorders* (pp. 377–436). Springer Netherlands.

Wang, R., Ge, S., Zommara, N. M., Ravienna, K., Espinoza, T., & Iramina, K. (2019). Consistency and dynamical changes of directional information flow in different brain states: A comparison of working memory and resting-state using EEG. *NeuroImage*, *203*, 116188.

Wang, X.-J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, *90*(3), 1195–1268.

Warren, J. E., Wise, R. J. S., & Warren, J. D. (2005). Sounds do-able: auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*(12), 636–643.

Welniarz, Q., Worbe, Y., & Gallea, C. (2021). The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency. *Frontiers in Systems Neuroscience*, *15*, 644059.

Wiggins, S. (2003). *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Second Edition). Springer New York.

Will, U., & Berg, E. (2007). Brain wave synchronization and entrainment to periodic acoustic stimuli. *Neuroscience Letters*, *424*(1), 55–60.

Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, *13*(2), 55–80.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.

Winer, J. A. (1992). The Functional Architecture of the Medial Geniculate Body and the Primary Auditory Cortex. In D. B. Webster, A. N. Popper, & R. R. Fay (Eds.), *The Mammalian Auditory Pathway: Neuroanatomy* (pp. 222–409). Springer.

Wise, S. P., Boussaoud, D., Johnson, P. B., & Caminiti, R. (1997). PREMOTOR AND PARIETAL CORTEX: Corticocortical Connectivity and Combinatorial Computations. *Annual Review of Neuroscience*, *20*(1), 25–42.

Wollman, I., Penhune, V., Segado, M., Carpentier, T., & Zatorre, R. J. (2018). Neural network retuning and neural predictors of learning success associated with cello training. *Proceedings of the National Academy of Sciences*, *115*(26).

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration. *Science*, *269*(5232), 1880–1882.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347.

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science*, *316*(5831), 1609–1612.

Wu, C. C., Hamm, J. P., Lim, V. K., & Kirk, I. J. (2016). Mu rhythm suppression demonstrates action representation in pianists during passive listening of piano melodies. *Experimental Brain Research*, *234*(8), 2133–2139.

Xing, D., Yeh, C.-I., & Shapley, R. M. (2009). Spatial Spread of the Local Field Potential and its Laminar Variation in Visual Cortex. *The Journal of Neuroscience*, *29*(37), 11540–11549.

Yoles-Frenkel, M., Avron, M., & Prut, Y. (2016). Impact of Auditory Context on Executed Motor Actions. *Frontiers in Integrative Neuroscience*, *10*.

Zacks, J. M. (2008). Neuroimaging Studies of Mental Rotation: A Meta-analysis and Review. *Journal of Cognitive Neuroscience*, *20*(1), 1–19.

Zarate, J. M., & Zatorre, R. J. (2005). Neural Substrates Governing Audiovocal Integration for Vocal Pitch Regulation in Singing. *Annals of the New York Academy of Sciences*, *1060*(1), 404–408.

Zatorre, R. J. (2024). *From perception to pleasure: the neuroscience of music and why we love it*. Oxford University Press.

Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is “where” in the human auditory cortex? *Nature Neuroscience*, *5*(9), 905–909.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, *107*(13), 5768–5773.

Zhou, M., Liang, F., Xiong, X. R., Li, L., Li, H., Xiao, Z., Tao, H. W., & Zhang, L. I. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nature Neuroscience*, *17*(6), 841–850.

List of Abbreviations

A1	Primary Auditory Cortex
aCC	Anterior Cingulate Cortex
AF	Arcuate Fasciculus
ASAP	Action Simulation for Auditory Prediction (hypothesis)
BEM	Boundary Element Method
BG	Basal Ganglia
BIDS	Brain Imaging Data Structure
BOLD	Blood Oxygen Level-Dependent (signal)
CC	Cingulate Cortex
CFC	Cross-Frequency Coupling
CNV	Contingent Negative Variation
CST	Corticospinal Tract
DAT	Dynamic Attending Theory
dPMC	Dorsal Premotor Cortex
DTI	Diffusion Tensor Imaging
EEG	Electroencephalography
EOG	Electrooculography

ERP	Event-Related Potential
F	F-statistic
FDR	False Discovery Rate
fMRI	Functional Magnetic Resonance Imaging
FWHM	Full-Width Half Maximum
GLMM	Generalized Linear Mixed Model
GPe	Globus Pallidus Externus
GPi	Globus Pallidus Internus
HMAT	Human Motor Area Template
Hz	Hertz (unit)
IPL	Inferior Parietal Lobule
IPS	Intraparietal Sulcus
ITI	Inter-Trial Interval
kΩ	kilo-ohms (unit)
LFP	Local Field Potential
LRP	Lateralized Readiness Potential
M1	Primary Motor Cortex
mCC	Midcingulate Cortex
MEG	Magnetoencephalography

MEP	Motor Evoked Potential	
MGN	Medial Geniculate Nucleus	
MRI	Magnetic Resonance Imaging	
ms	Milliseconds	
N–P	Negative-Positive (vertex wave)	
OMEGA	Open MEG Archive	
PAC	Phase-Amplitude Coupling	
PCA	Principal Component Analysis	
pCC	Posterior Cingulate Cortex	
PLV	Phase Locking Value	
POST	Passive listening block after training	(Study 2)
PPC	Posterior Parietal Cortex	
PRC	Phase-Response Curve	
PRE	Passive listening block before training	(Study 2)
PSD	Power Spectrum Density	
PT	Planum Temporale	
PTE	Phase Transfer Entropy	
REB	Research Ethics Board	
ROI	Region of Interest	

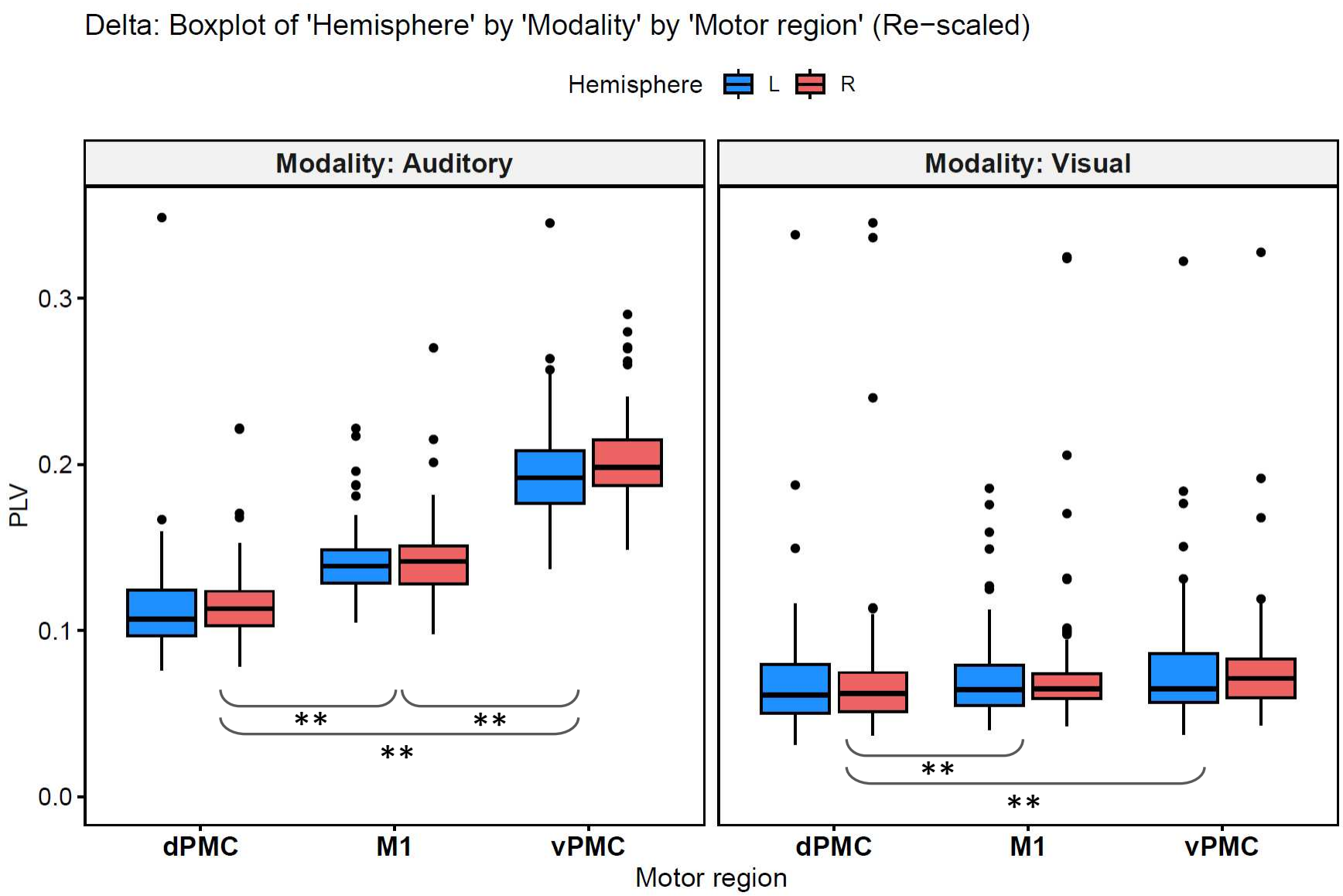
rs-fMRI	Resting-State Functional Magnetic Resonance Imaging
rs-MEG	Resting-State Magnetoencephalography
S2	Secondary Somatosensory Cortex
SE	Standard Error
SLF	Superior Longitudinal Fasciculus
SLF-tp	Temporoparietal Tract of the Superior Longitudinal Fasciculus
SMA	Supplementary Motor Area
SNc	Substantia Nigra Pars Compacta
SNR	Signal-to-Noise Ratio
SNr	Substantia Nigra Pars Reticulata
SPL	Superior Parietal Lobule
Spt	Sylvian–parietal–temporal
SSP	Signal Space Projection
STG	Superior Temporal Gyrus
STN	Subthalamic Nucleus
t	t-statistic
TE	Transfer Entropy
t-map	t-statistic map
TMS	Transcranial Magnetic Stimulation

TRAIN	Melody playback training condition	(Study 2)
TTL	Transistor-Transistor Logic (used for sending event markers)	
V1	Primary Visual Cortex	
V5/MT	Visual area 5/middle temporal visual area	
VA	Ventral Anterior Nucleus	(thalamus)
VL	Ventrolateral Nucleus	(thalamus)
vPMC	Ventral Premotor Cortex	

A

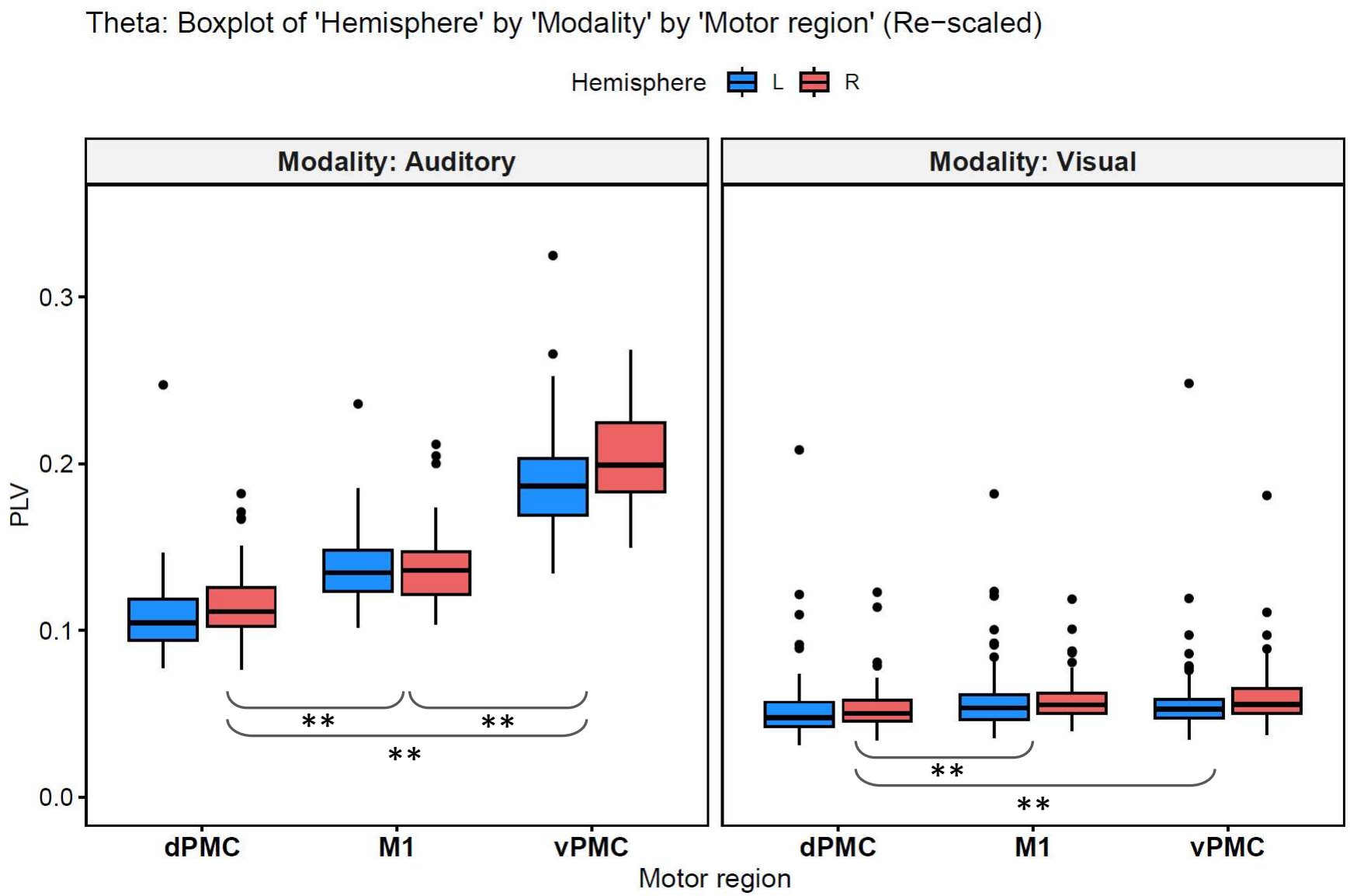
Appendix A
Supplementary Material for Study 1

Supplementary Figure 1. PLV results obtained in GLMM 1 (n=90) in the delta frequency band

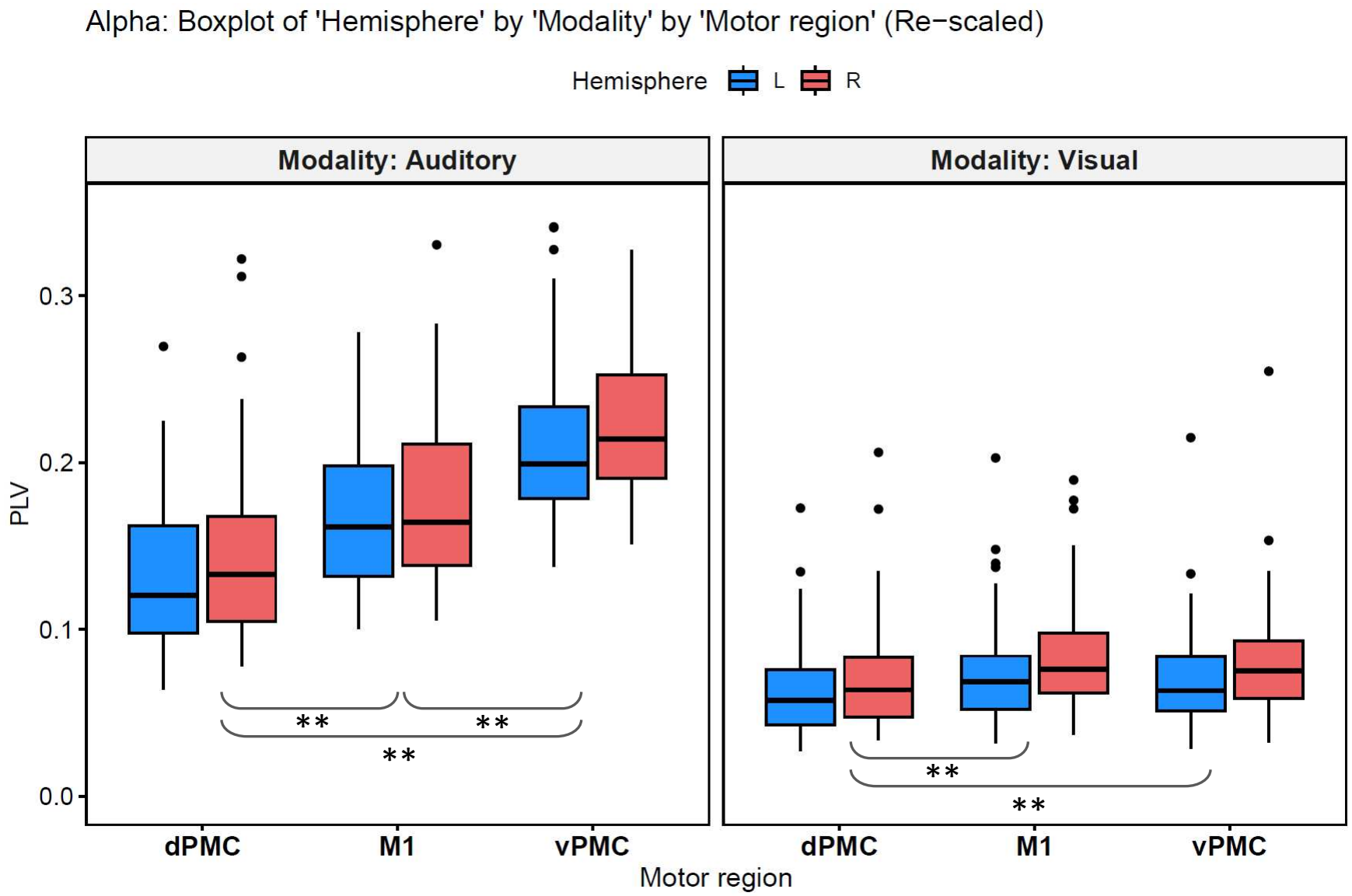


Please note: outliers above 0.35 have been clipped from this plot

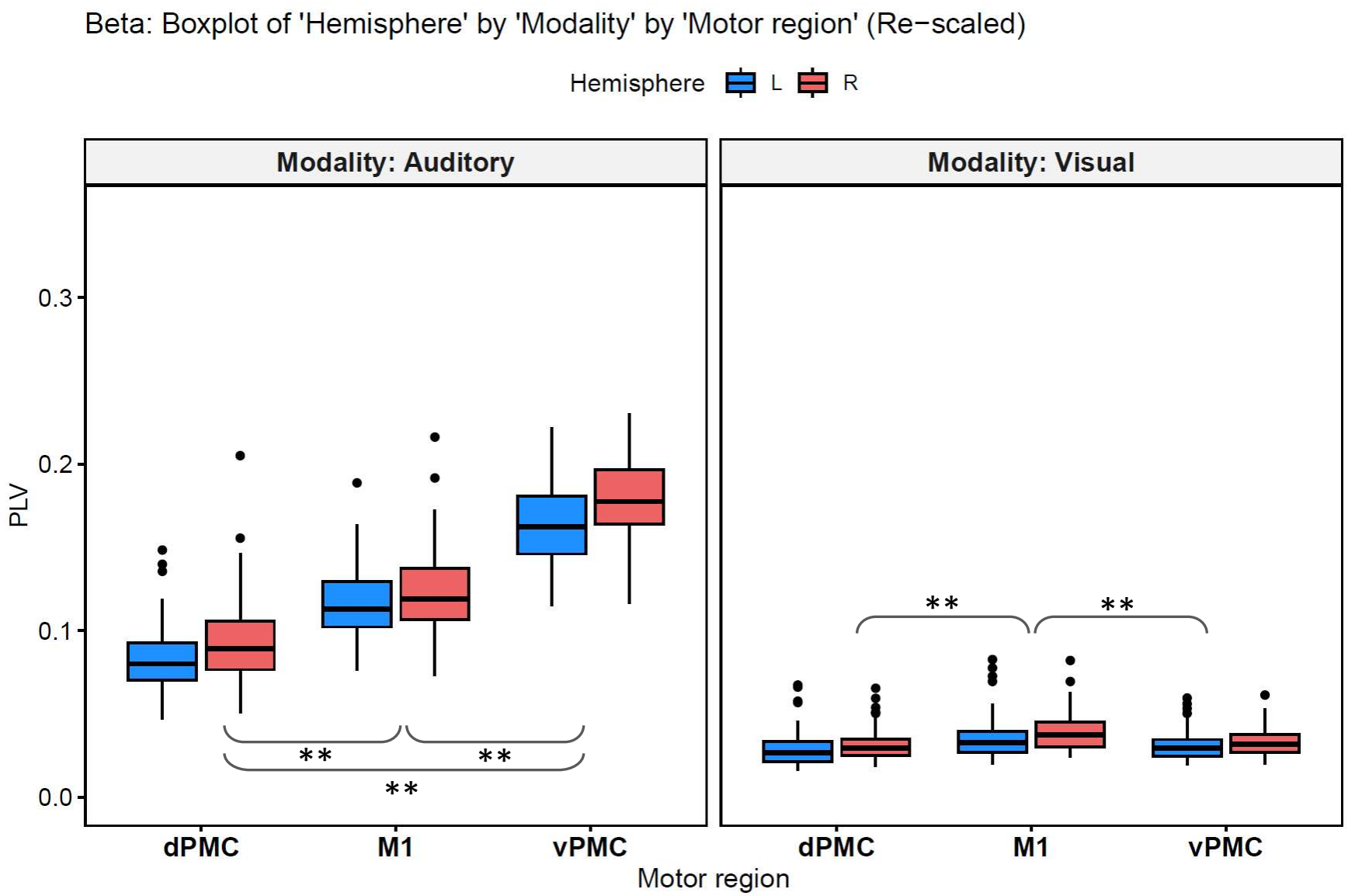
Supplementary Figure 2. PLV results obtained in GLMM 1 (n=90) in the theta frequency band



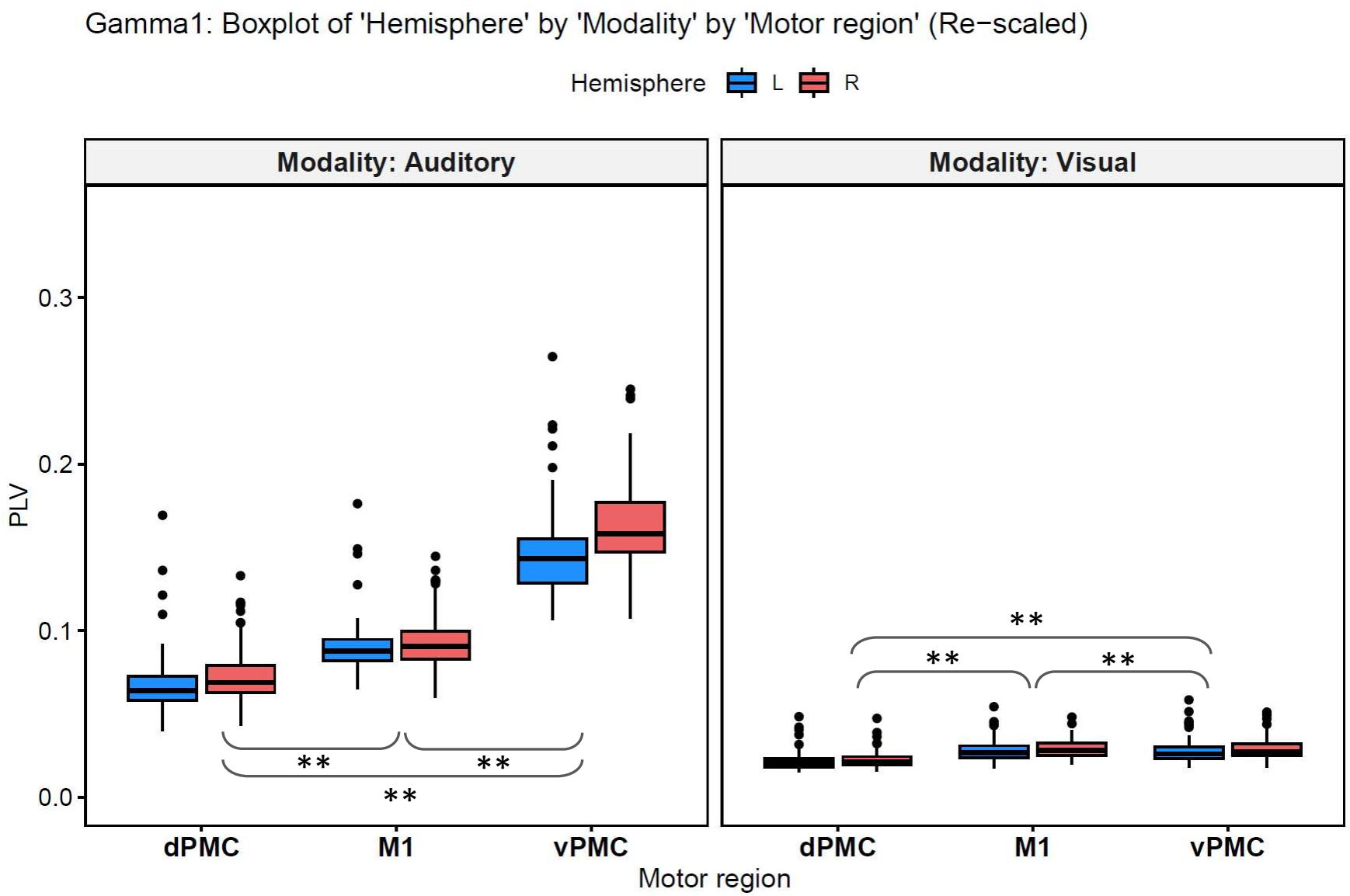
Supplementary Figure 3. PLV results obtained in GLMM 1 (n=90) in the alpha frequency band



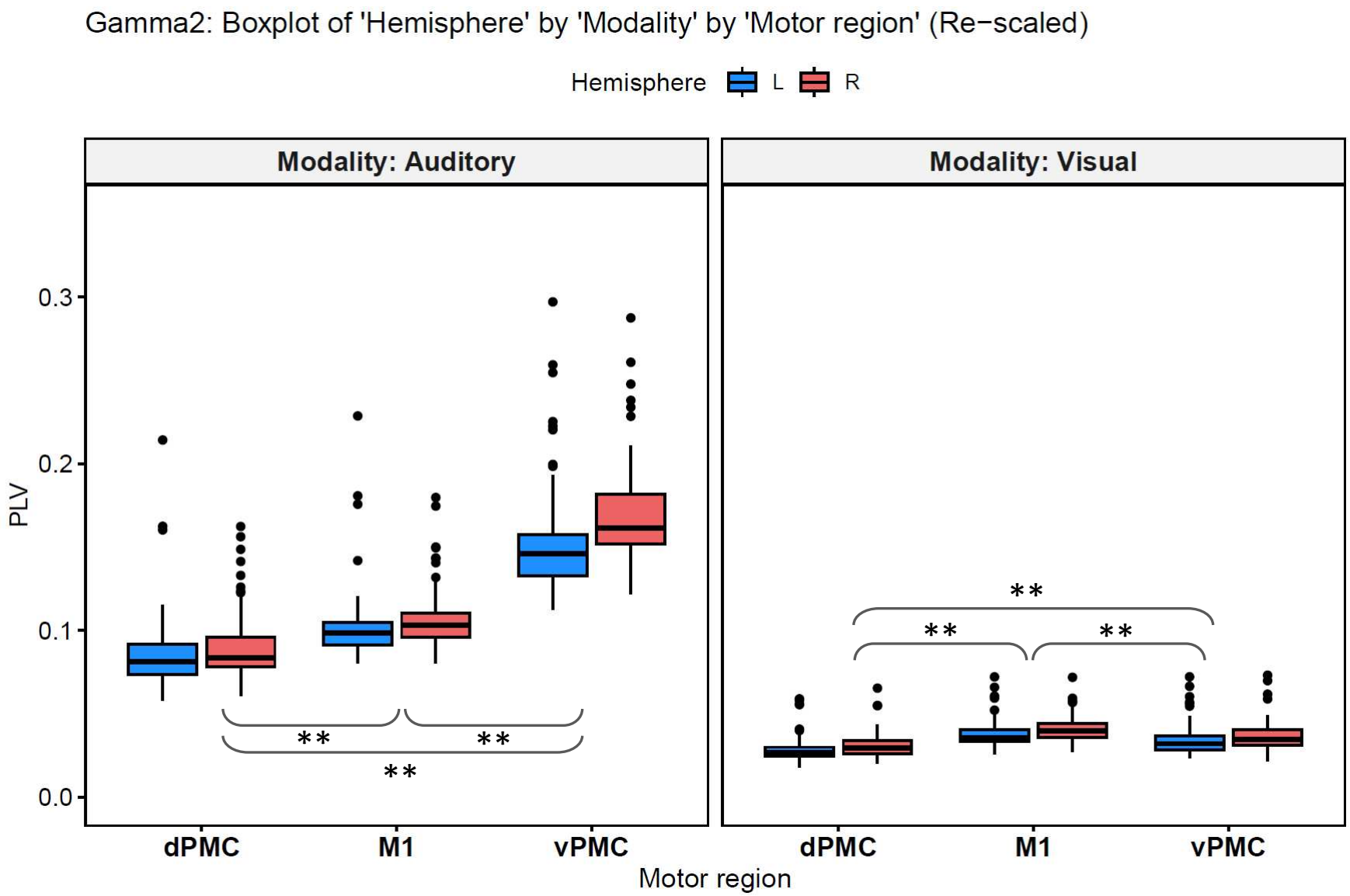
Supplementary Figure 4. PLV results obtained in GLMM 1 (n=90) in the beta frequency band



Supplementary Figure 5. PLV results obtained in GLMM 1 (n=90) in the gamma1 frequency band



Supplementary Figure 6. PLV results obtained in GLMM 1 (n=90) in the gamma2 frequency band



Sup. Table 1. Percentage of variance explained by each ROI's first Principal Component

ID	PC1-A1	PC1-dPMC	PC1-M1	PC1-vPMC
sub-0001/LEFT	57%	26%	23%	34%
sub-0002/LEFT	48%	37%	32%	36%
sub-0003/LEFT	50%	31%	35%	38%
sub-0004/LEFT	48%	39%	26%	38%
sub-0005/LEFT	60%	30%	26%	38%
sub-0007/LEFT	57%	31%	23%	38%
sub-0008/LEFT	48%	24%	30%	49%
sub-0011/LEFT	59%	36%	27%	36%
sub-0012/LEFT	52%	46%	38%	39%
sub-0013/LEFT	46%	35%	36%	55%
sub-0014/LEFT	50%	34%	35%	30%
sub-0015/LEFT	53%	42%	34%	36%
sub-0016/LEFT	61%	51%	63%	84%
sub-0018/LEFT	53%	44%	41%	42%
sub-0020/LEFT	67%	37%	23%	33%
sub-0021/LEFT	53%	30%	47%	47%
sub-0023/LEFT	100%	100%	100%	100%
sub-0024/LEFT	56%	33%	22%	32%
sub-0025/LEFT	59%	38%	42%	49%
sub-0026/LEFT	54%	30%	29%	42%
sub-0028/LEFT	50%	27%	20%	36%
sub-0029/LEFT	49%	28%	25%	36%
sub-0030/LEFT	60%	33%	21%	40%
sub-0031/LEFT	55%	31%	30%	53%
sub-0032/LEFT	48%	29%	35%	30%
sub-0033/LEFT	59%	28%	24%	35%
sub-0035/LEFT	48%	36%	35%	30%
sub-0036/LEFT	54%	57%	34%	40%
sub-0037/LEFT	57%	39%	39%	49%
sub-0040/LEFT	59%	36%	25%	41%
sub-0045/LEFT	47%	42%	38%	36%
sub-0046/LEFT	47%	31%	31%	42%
sub-0047/LEFT	52%	30%	24%	33%
sub-0048/LEFT	52%	28%	23%	29%
sub-0049/LEFT	63%	27%	25%	32%
sub-0050/LEFT	57%	30%	27%	32%
sub-0052/LEFT	63%	39%	38%	45%
sub-0055/LEFT	58%	67%	84%	29%
sub-0056/LEFT	56%	32%	24%	32%
sub-0058/LEFT	46%	56%	74%	33%
sub-0063/LEFT	59%	30%	24%	47%
sub-0064/LEFT	52%	32%	44%	43%
sub-0065/LEFT	61%	35%	24%	38%
sub-0069/LEFT	52%	31%	39%	30%
sub-0070/LEFT	59%	28%	26%	35%
sub-0074/LEFT	48%	31%	42%	32%
sub-0076/LEFT	55%	38%	31%	33%
sub-0077/LEFT	54%	28%	20%	38%
sub-0082/LEFT	66%	59%	50%	55%
sub-0083/LEFT	93%	74%	73%	70%
sub-0084/LEFT	49%	29%	22%	32%
sub-0086/LEFT	63%	51%	31%	40%
sub-0087/LEFT	50%	65%	29%	73%
sub-0090/LEFT	60%	35%	26%	35%
sub-0091/LEFT	46%	47%	36%	59%
sub-0093/LEFT	74%	73%	67%	68%
sub-0094/LEFT	54%	28%	25%	33%
sub-0097/LEFT	46%	29%	25%	36%
sub-0099/LEFT	44%	30%	24%	32%
sub-0101/LEFT	55%	27%	27%	32%
sub-0102/LEFT	60%	27%	22%	39%
sub-0103/LEFT	100%	99%	100%	100%
sub-0104/LEFT	65%	31%	29%	31%
sub-0105/LEFT	48%	29%	29%	39%
sub-0106/LEFT	61%	31%	27%	31%
sub-0134/LEFT	54%	32%	23%	45%
sub-0145/LEFT	47%	33%	28%	36%
sub-0146/LEFT	50%	39%	28%	31%
sub-0148/LEFT	59%	74%	57%	69%
sub-0149/LEFT	46%	27%	21%	30%
sub-0150/LEFT	53%	28%	26%	36%
sub-0151/LEFT	60%	32%	37%	48%
sub-0152/LEFT	55%	25%	26%	31%
sub-0154/LEFT	62%	29%	24%	37%
sub-0156/LEFT	57%	33%	23%	39%
sub-0157/LEFT	54%	30%	23%	38%
sub-0166/LEFT	50%	27%	20%	31%
sub-0169/LEFT	67%	33%	29%	37%
sub-0171/LEFT	52%	36%	20%	36%
sub-0172/LEFT	51%	50%	53%	31%
sub-0174/LEFT	79%	41%	67%	49%
sub-0175/LEFT	57%	33%	24%	30%
sub-0176/LEFT	99%	100%	100%	100%
sub-0179/LEFT	45%	24%	26%	32%
sub-0181/LEFT	51%	32%	26%	30%
sub-0184/LEFT	45%	61%	38%	33%
sub-0191/LEFT	100%	100%	100%	100%
sub-0195/LEFT	48%	25%	21%	43%
sub-0208/LEFT	54%	34%	40%	37%
sub-0210/LEFT	55%	25%	25%	31%

ID	PC1-A1	PC1-dPMC	PC1-M1	PC1-vPMC
sub-0001/RIGHT	55%	25%	22%	32%
sub-0002/RIGHT	53%	33%	30%	36%
sub-0003/RIGHT	47%	39%	45%	39%
sub-0004/RIGHT	49%	35%	29%	28%
sub-0005/RIGHT	49%	32%	30%	34%
sub-0007/RIGHT	48%	29%	27%	32%
sub-0008/RIGHT	52%	34%	31%	54%
sub-0011/RIGHT	59%	31%	34%	30%
sub-0012/RIGHT	52%	46%	38%	39%
sub-0013/RIGHT	57%	35%	28%	41%
sub-0014/RIGHT	62%	38%	32%	35%
sub-0015/RIGHT	63%	37%	37%	40%
sub-0016/RIGHT	83%	57%	68%	84%
sub-0018/RIGHT	53%	37%	37%	30%
sub-0020/RIGHT	46%	31%	25%	33%
sub-0021/RIGHT	52%	40%	25%	40%
sub-0023/RIGHT	100%	100%	100%	100%
sub-0024/RIGHT	56%	26%	23%	35%
sub-0025/RIGHT	55%	36%	32%	48%
sub-0026/RIGHT	47%	29%	26%	38%
sub-0028/RIGHT	59%	27%	24%	34%
sub-0029/RIGHT	59%	26%	24%	28%
sub-0030/RIGHT	55%	37%	36%	49%
sub-0031/RIGHT	68%	34%	24%	46%
sub-0032/RIGHT	45%	31%	40%	34%
sub-0033/RIGHT	51%	33%	23%	37%
sub-0035/RIGHT	49%	33%	29%	30%
sub-0036/RIGHT	66%	33%	49%	37%
sub-0037/RIGHT	54%	40%	38%	40%
sub-0040/RIGHT	59%	41%	24%	35%
sub-0045/RIGHT	49%	38%	41%	37%
sub-0046/RIGHT	64%	34%	34%	38%
sub-0047/RIGHT	52%	26%	21%	30%
sub-0048/RIGHT	46%	27%	24%	34%
sub-0049/RIGHT	50%	29%	30%	33%
sub-0050/RIGHT	58%	27%	21%	34%
sub-0052/RIGHT	76%	54%	40%	60%
sub-0055/RIGHT	56%	59%	82%	32%
sub-0056/RIGHT	59%	27%	25%	29%
sub-0058/RIGHT	52%	39%	63%	30%
sub-0063/RIGHT	59%	27%	20%	40%
sub-0064/RIGHT	48%	28%	26%	44%
sub-0065/RIGHT	59%	36%	27%	34%
sub-0069/RIGHT	47%	33%	38%	38%
sub-0070/RIGHT	49%	31%	30%	29%
sub-0074/RIGHT	67%	36%	47%	29%
sub-0076/RIGHT	50%	42%	35%	38%
sub-0077/RIGHT	52%	31%	25%	34%
sub-0082/RIGHT	74%	33%	31%	32%
sub-0083/RIGHT	95%	83%	80%	56%
sub-0084/RIGHT	50%	29%	20%	30%
sub-0086/RIGHT	57%	43%	38%	39%
sub-0087/RIGHT	52%	35%	33%	38%
sub-0090/RIGHT	58%	27%	41%	39%
sub-0091/RIGHT	57%	35%	35%	53%
sub-0093/RIGHT	90%	61%	76%	67%
sub-0094/RIGHT	52%	27%	29%	32%
sub-0097/RIGHT	54%	32%	24%	28%
sub-0099/RIGHT	54%	29%	23%	34%
sub-0101/RIGHT	53%	26%	25%	29%
sub-0102/RIGHT	60%	29%	22%	28%
sub-0103/RIGHT	100%	100%	100%	100%
sub-0104/RIGHT	67%	30%	33%	36%
sub-0105/RIGHT	75%	27%	40%	39%
sub-0106/RIGHT	56%	33%	27%	35%
sub-0134/RIGHT	50%	29%	23%	33%
sub-0145/RIGHT	47%	34%	24%	33%
sub-0146/RIGHT	60%	61%	37%	56%
sub-0148/RIGHT	53%	50%	23%	49%
sub-0149/RIGHT	52%	26%	22%	34%
sub-0150/RIGHT	49%	32%	26%	34%
sub-0151/RIGHT	65%	27%	41%	46%
sub-0152/RIGHT	49%	30%	28%	32%
sub-0154/RIGHT	55%	27%	26%	50%
sub-0156/RIGHT	46%	32%	29%	35%
sub-0157/RIGHT	58%	29%	25%	29%
sub-0166/RIGHT	49%	33%	23%	32%
sub-0169/RIGHT	51%	35%	36%	37%
sub-0171/RIGHT	52%	34%	28%	39%
sub-0172/RIGHT	51%	36%	47%	32%
sub-0174/RIGHT	65%	40%	34%	39%
sub-0175/RIGHT	53%	27%	21%	35%
sub-0176/RIGHT	100%	100%	100%	100%
sub-0179/RIGHT	46%	27%	24%	30%
sub-0181/RIGHT	57%	31%	23%	31%
sub-0184/RIGHT	64%	43%	27%	57%
sub-0191/RIGHT	100%	100%	100%	100%
sub-0195/RIGHT	46%	29%	28%	36%
sub-0208/RIGHT	57%	40%	39%	42%
sub-0210/RIGHT	53%	29%	21%	32%

Sup. Table 2. Goodness of fit measures for all GLMM models and all frequency bands

Model	AIC	BIC	logLik	deviance	Precision (φ)
GLMM1_Delta	-5668.2	-5543.6	2859.1	-5718.2	1150
GLMM1_Theta	-6246.2	-6121.6	3148.1	-6296.2	1210
GLMM1_Alpha	-5423.1	-5298.4	2736.5	-5473.1	1100
GLMM1_Beta	-6803.2	-6678.6	3426.6	-6853.2	1640
GLMM1_Gamma1	-7368.8	-7244.2	3709.4	-7418.8	2060
GLMM1_Gamma2	-7118.3	-6993.7	3584.2	-7168.3	2270

Model	AIC	BIC	logLik	deviance	Dispersion (σ^2)
GLMM2_PLV_Differences	-8345.5	-8146.1	4209.7	-8419.5	7.76E-05

Model	AIC	BIC	logLik	deviance	Dispersion (σ^2)
GLMM3_Delta_Left	-2136.9	-2081.2	1081.5	-2162.9	3.96E-05
GLMM3_Delta_Right	-2024.2	-1968.4	1025.1	-2050.2	9.18E-05
GLMM3_Theta_Left	-3075.5	-3019.7	1550.7	-3101.5	4.62E-05
GLMM3_Theta_Right	-3000.7	-2944.9	1513.3	-3026.7	6.37E-05
GLMM3_Alpha_Left	-2567	-2511.2	1296.5	-2593	3.18E-04
GLMM3_Alpha_Right	-2510.2	-2454.4	1268.1	-2536.2	1.83E-04
GLMM3_Beta_Left	-3444	-3388.2	1735	-3470	1.44E-05
GLMM3_Beta_Right	-3316.4	-3260.6	1671.2	-3342.4	3.22E-05
GLMM3_Gamma1_Left	-4152.7	-4096.9	2089.4	-4178.7	2.47E-05
GLMM3_Gamma1_Right	-4090.6	-4034.8	2058.3	-4116.6	2.18E-05
GLMM3_Gamma2_Left	-4517.9	-4462.1	2272	-4543.9	5.86E-05
GLMM3_Gamma2_Right	-4528.3	-4472.5	2277.2	-4554.3	4.27E-05

B

Appendix B
Supplementary Material for Study 2

Supplementary Table 1. Post-hoc contrasts for GLMM1 (Accuracy ~ Trial)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Trial_1-Trial_10	-0.76	0.03	329	-23.380	<.0001	<.0001
Trial_1-Trial_11	-0.77	0.03	329	-23.680	<.0001	<.0001
Trial_1-Trial_12	-0.79	0.03	329	-22.850	<.0001	<.0001
Trial_1-Trial_13	-0.78	0.03	329	-23.140	<.0001	<.0001
Trial_1-Trial_14	-0.78	0.03	329	-23.790	<.0001	<.0001
Trial_1-Trial_16	-0.74	0.04	329	-19.380	<.0001	<.0001
Trial_1-Trial_17	-0.77	0.03	329	-22.080	<.0001	<.0001
Trial_1-Trial_2	-0.32	0.07	329	-4.820	<.0001	<.0001
Trial_1-Trial_4	-0.22	0.06	329	-3.520	0.001	0.0021
Trial_1-Trial_5	-0.39	0.06	329	-6.700	<.0001	<.0001
Trial_1-Trial_6	-0.55	0.05	329	-10.530	<.0001	<.0001
Trial_1-Trial_7	-0.61	0.03	329	-17.770	<.0001	<.0001
Trial_1-Trial_8	-0.7	0.03	329	-20.200	<.0001	<.0001
Trial_1-Trial_9	-0.72	0.03	329	-22.250	<.0001	<.0001
Trial_10-Trial_11	-0.01	0.03	329	-0.58	0.6086	0.6588
Trial_10-Trial_12	-0.04	0.02	329	-1.450	0.201	0.2706
Trial_10-Trial_13	-0.02	0.03	329	-0.7	0.5387	0.5954
Trial_10-Trial_14	-0.02	0.03	329	-0.75	0.5134	0.5796
Trial_10-Trial_16	0.02	0.04	329	0.59	0.6086	0.6588
Trial_10-Trial_17	-0.01	0.03	329	-0.36	0.7685	0.8234
Trial_10-Trial_2	0.44	0.08	329	5.570	<.0001	<.0001
Trial_10-Trial_4	0.53	0.08	329	7.080	<.0001	<.0001
Trial_10-Trial_5	0.37	0.07	329	5.140	<.0001	<.0001
Trial_10-Trial_6	0.2	0.06	329	3.270	0.0022	0.0041
Trial_10-Trial_7	0.15	0.04	329	3.330	0.0018	0.0034

Trial_10-Trial_8	0.06	0.04	329	1.590	0.1605	0.2234
Trial_10-Trial_9	0.03	0.03	329	1.010	0.3886	0.4665
Trial_11-Trial_12	-0.02	0.02	329	-1.000	0.3905	0.4665
Trial_11-Trial_13	0	0.02	329	-0.15	0.8928	0.9014
Trial_11-Trial_14	0	0.02	329	-0.18	0.8823	0.9014
Trial_11-Trial_16	0.04	0.04	329	1.010	0.3886	0.4665
Trial_11-Trial_17	0	0.03	329	0.16	0.8911	0.9014
Trial_11-Trial_2	0.45	0.08	329	5.640	<.0001	<.0001
Trial_11-Trial_4	0.55	0.08	329	7.130	<.0001	<.0001
Trial_11-Trial_5	0.38	0.07	329	5.210	<.0001	<.0001
Trial_11-Trial_6	0.22	0.06	329	3.430	0.0014	0.0028
Trial_11-Trial_7	0.16	0.05	329	3.550	0.0009	0.0019
Trial_11-Trial_8	0.07	0.04	329	1.980	0.0711	0.1051
Trial_11-Trial_9	0.05	0.03	329	1.500	0.183	0.2495
Trial_12-Trial_13	0.02	0.02	329	0.81	0.4745	0.5415
Trial_12-Trial_14	0.02	0.02	329	0.85	0.464	0.539
Trial_12-Trial_16	0.06	0.04	329	1.570	0.1617	0.2234
Trial_12-Trial_17	0.02	0.03	329	0.98	0.391	0.4665
Trial_12-Trial_2	0.47	0.08	329	5.680	<.0001	<.0001
Trial_12-Trial_4	0.57	0.08	329	7.130	<.0001	<.0001
Trial_12-Trial_5	0.4	0.08	329	5.250	<.0001	<.0001
Trial_12-Trial_6	0.24	0.07	329	3.600	0.0008	0.0018
Trial_12-Trial_7	0.18	0.05	329	3.760	0.0005	0.0012
Trial_12-Trial_8	0.09	0.04	329	2.460	0.0237	0.0383
Trial_12-Trial_9	0.07	0.03	329	2.120	0.0538	0.0843
Trial_13-Trial_14	0	0.02	329	-0.03	0.9797	0.9797
Trial_13-Trial_16	0.04	0.04	329	1.080	0.3576	0.4579

Trial_13-Trial_17	0.01	0.03	329	0.28	0.8189	0.858
Trial_13-Trial_2	0.46	0.08	329	5.630	<.0001	<.0001
Trial_13-Trial_4	0.55	0.08	329	7.110	<.0001	<.0001
Trial_13-Trial_5	0.38	0.07	329	5.200	<.0001	<.0001
Trial_13-Trial_6	0.22	0.06	329	3.440	0.0013	0.0026
Trial_13-Trial_7	0.17	0.05	329	3.550	0.0009	0.0019
Trial_13-Trial_8	0.08	0.04	329	2.030	0.0643	0.0964
Trial_13-Trial_9	0.05	0.03	329	1.570	0.1617	0.2234
Trial_14-Trial_16	0.04	0.04	329	1.120	0.3401	0.4409
Trial_14-Trial_17	0.01	0.03	329	0.31	0.8001	0.8486
Trial_14-Trial_2	0.46	0.08	329	5.650	<.0001	<.0001
Trial_14-Trial_4	0.55	0.08	329	7.150	<.0001	<.0001
Trial_14-Trial_5	0.39	0.07	329	5.220	<.0001	<.0001
Trial_14-Trial_6	0.22	0.06	329	3.470	0.0012	0.0025
Trial_14-Trial_7	0.17	0.05	329	3.600	0.0008	0.0018
Trial_14-Trial_8	0.08	0.04	329	2.080	0.0577	0.0878
Trial_14-Trial_9	0.05	0.03	329	1.630	0.1494	0.2149
Trial_16-Trial_17	-0.03	0.04	329	-0.83	0.4671	0.539
Trial_16-Trial_2	0.42	0.08	329	5.260	<.0001	<.0001
Trial_16-Trial_4	0.51	0.08	329	6.740	<.0001	<.0001
Trial_16-Trial_5	0.34	0.07	329	4.790	<.0001	<.0001
Trial_16-Trial_6	0.18	0.06	329	2.830	0.0085	0.0149
Trial_16-Trial_7	0.13	0.05	329	2.620	0.0154	0.0261
Trial_16-Trial_8	0.04	0.04	329	0.86	0.4627	0.539
Trial_16-Trial_9	0.01	0.04	329	0.26	0.8253	0.858
Trial_17-Trial_2	0.45	0.08	329	5.550	<.0001	<.0001
Trial_17-Trial_4	0.54	0.08	329	7.030	<.0001	<.0001

Trial_17-Trial_5	0.38	0.07	329	5.110	<.0001	<.0001
Trial_17-Trial_6	0.21	0.06	329	3.320	0.0019	0.0036
Trial_17-Trial_7	0.16	0.05	329	3.350	0.0017	0.0033
Trial_17-Trial_8	0.07	0.04	329	1.770	0.1141	0.1664
Trial_17-Trial_9	0.04	0.03	329	1.260	0.2791	0.3663
Trial_2-Trial_4	0.1	0.08	329	1.250	0.2791	0.3663
Trial_2-Trial_5	-0.07	0.07	329	-0.98	0.391	0.4665
Trial_2-Trial_6	-0.23	0.07	329	-3.170	0.003	0.0054
Trial_2-Trial_7	-0.29	0.07	329	-4.420	<.0001	<.0001
Trial_2-Trial_8	-0.38	0.07	329	-5.140	<.0001	<.0001
Trial_2-Trial_9	-0.4	0.08	329	-5.390	<.0001	<.0001
Trial_4-Trial_5	-0.17	0.07	329	-2.350	0.0307	0.0488
Trial_4-Trial_6	-0.33	0.07	329	-4.570	<.0001	<.0001
Trial_4-Trial_7	-0.39	0.06	329	-6.060	<.0001	<.0001
Trial_4-Trial_8	-0.48	0.07	329	-6.680	<.0001	<.0001
Trial_4-Trial_9	-0.5	0.07	329	-6.920	<.0001	<.0001
Trial_5-Trial_6	-0.16	0.07	329	-2.450	0.0237	0.0383
Trial_5-Trial_7	-0.22	0.06	329	-3.830	0.0004	0.001
Trial_5-Trial_8	-0.31	0.07	329	-4.660	<.0001	<.0001
Trial_5-Trial_9	-0.33	0.07	329	-4.940	<.0001	<.0001
Trial_6-Trial_7	-0.06	0.05	329	-1.070	0.3621	0.4581
Trial_6-Trial_8	-0.15	0.06	329	-2.460	0.0237	0.0383
Trial_6-Trial_9	-0.17	0.06	329	-2.880	0.0076	0.0135
Trial_7-Trial_8	-0.09	0.04	329	-2.110	0.0556	0.0859
Trial_7-Trial_9	-0.12	0.04	329	-2.750	0.0108	0.0186
Trial_8-Trial_9	-0.03	0.04	329	-0.71	0.5354	0.5954

Legend for Supplementary Table 1

Contrast	Specific comparison between two levels of ‘Accuracy’
Estimate	Estimated difference between two levels (response scale)
SE	Standard error of the estimate
df	Degrees of freedom
t-ratio	Ratio of the contrast estimate to its standard error (SE)
p-value	Probability of observing the data under a true null hypothesis
FDR p-value	p-value after FDR correction for multiple comparisons

Supplementary Table 2. Post-hoc contrasts for GLMM3 (Latency ~ Trial)

Contrast	Estimate	SE	df	t-ratio	p-value	FDR p-value
Trial_1 - Trial_10	0.04	0.01	329	6.680	<.0001	<.0001
Trial_1 - Trial_11	0.04	0.01	329	6.950	<.0001	<.0001
Trial_1 - Trial_12	0.04	0.01	329	7.450	<.0001	<.0001
Trial_1 - Trial_13	0.04	0.01	329	7.420	<.0001	<.0001
Trial_1 - Trial_14	0.04	0.01	329	6.900	<.0001	<.0001
Trial_1 - Trial_16	0.04	0.01	329	7.140	<.0001	<.0001
Trial_1 - Trial_17	0.05	0.01	329	7.510	<.0001	<.0001
Trial_1 - Trial_2	0.01	0.01	329	1.550	0.1787	0.2548
Trial_1 - Trial_4	0.02	0.01	329	2.810	0.0122	0.0285
Trial_1 - Trial_5	0.02	0.01	329	3.020	0.0071	0.018
Trial_1 - Trial_6	0.03	0.01	329	5.160	<.0001	<.0001
Trial_1 - Trial_7	0.03	0.01	329	4.410	0.0001	0.0005
Trial_1 - Trial_8	0.03	0.01	329	4.280	0.0001	0.0005
Trial_1 - Trial_9	0.04	0.01	329	6.360	<.0001	<.0001
Trial_10 - Trial_11	<0.01	<0.01	329	0.54	0.6654	0.7354
Trial_10 - Trial_12	0.01	<0.01	329	1.590	0.1709	0.2548
Trial_10 - Trial_13	0.01	<0.01	329	1.520	0.182	0.2548
Trial_10 - Trial_14	<0.01	<0.01	329	0.43	0.7208	0.7802
Trial_10 - Trial_16	0.01	<0.01	329	1.610	0.1707	0.2548
Trial_10 - Trial_17	0.01	<0.01	329	2.020	0.0766	0.1319
Trial_10 - Trial_2	-0.03	0.01	329	-4.280	0.0001	0.0005
Trial_10 - Trial_4	-0.02	0.01	329	-2.370	0.0342	0.0641
Trial_10 - Trial_5	-0.02	0.01	329	-3.400	0.0023	0.0069
Trial_10 - Trial_6	-0.01	<0.01	329	-2.510	0.0255	0.0515
Trial_10 - Trial_7	-0.01	<0.01	329	-2.600	0.0203	0.0426

Trial_10 - Trial_8	-0.01	0.01	329	-1.360	0.242	0.3343
Trial_10 - Trial_9	<0.01	<0.01	329	-0.12	0.9228	0.9373
Trial_11 - Trial_12	<0.01	<0.01	329	1.070	0.3549	0.4384
Trial_11 - Trial_13	<0.01	<0.01	329	0.99	0.3898	0.4704
Trial_11 - Trial_14	<0.01	<0.01	329	-0.1	0.9284	0.9373
Trial_11 - Trial_16	<0.01	<0.01	329	1.140	0.3259	0.4173
Trial_11 - Trial_17	0.01	<0.01	329	1.520	0.182	0.2548
Trial_11 - Trial_2	-0.03	0.01	329	-4.570	<.0001	<.0001
Trial_11 - Trial_4	-0.02	0.01	329	-2.660	0.0183	0.0402
Trial_11 - Trial_5	-0.02	0.01	329	-3.750	0.0007	0.0024
Trial_11 - Trial_6	-0.01	<0.01	329	-3.000	0.0072	0.018
Trial_11 - Trial_7	-0.01	<0.01	329	-3.030	0.007	0.018
Trial_11 - Trial_8	-0.01	0.01	329	-1.710	0.1484	0.2473
Trial_11 - Trial_9	<0.01	<0.01	329	-0.62	0.6286	0.7334
Trial_12 - Trial_13	<0.01	<0.01	329	-0.08	0.9389	0.9389
Trial_12 - Trial_14	<0.01	<0.01	329	-1.170	0.317	0.4109
Trial_12 - Trial_16	<0.01	<0.01	329	0.21	0.8622	0.8876
Trial_12 - Trial_17	<0.01	<0.01	329	0.52	0.6654	0.7354
Trial_12 - Trial_2	-0.03	0.01	329	-5.120	<.0001	<.0001
Trial_12 - Trial_4	-0.02	0.01	329	-3.210	0.0042	0.0116
Trial_12 - Trial_5	-0.03	0.01	329	-4.410	0.0001	0.0005
Trial_12 - Trial_6	-0.02	<0.01	329	-3.950	0.0004	0.0016
Trial_12 - Trial_7	-0.02	<0.01	329	-3.840	0.0006	0.0022
Trial_12 - Trial_8	-0.01	0.01	329	-2.380	0.0342	0.0641
Trial_12 - Trial_9	-0.01	<0.01	329	-1.590	0.1709	0.2548
Trial_13 - Trial_14	<0.01	<0.01	329	-1.090	0.3457	0.4321
Trial_13 - Trial_16	<0.01	<0.01	329	0.27	0.8193	0.8517

Trial_13 - Trial_17	<0.01	<0.01	329	0.6	0.6363	0.7342
Trial_13 - Trial_2	-0.03	0.01	329	-5.080	<.0001	<.0001
Trial_13 - Trial_4	-0.02	0.01	329	-3.180	0.0044	0.0118
Trial_13 - Trial_5	-0.03	0.01	329	-4.360	0.0001	0.0005
Trial_13 - Trial_6	-0.02	<0.01	329	-3.880	0.0005	0.0019
Trial_13 - Trial_7	-0.02	<0.01	329	-3.780	0.0007	0.0024
Trial_13 - Trial_8	-0.01	0.01	329	-2.330	0.0377	0.0694
Trial_13 - Trial_9	-0.01	<0.01	329	-1.520	0.182	0.2548
Trial_14 - Trial_16	0.01	<0.01	329	1.230	0.292	0.3881
Trial_14 - Trial_17	0.01	<0.01	329	1.620	0.1698	0.2548
Trial_14 - Trial_2	-0.03	0.01	329	-4.520	0.0001	0.0005
Trial_14 - Trial_4	-0.02	0.01	329	-2.610	0.0203	0.0426
Trial_14 - Trial_5	-0.02	0.01	329	-3.680	0.0009	0.0029
Trial_14 - Trial_6	-0.01	<0.01	329	-2.910	0.0092	0.022
Trial_14 - Trial_7	-0.01	<0.01	329	-2.950	0.0083	0.0203
Trial_14 - Trial_8	-0.01	0.01	329	-1.640	0.1671	0.2548
Trial_14 - Trial_9	<0.01	<0.01	329	-0.52	0.6654	0.7354
Trial_16 - Trial_17	<0.01	<0.01	329	0.27	0.8193	0.8517
Trial_16 - Trial_2	-0.03	0.01	329	-5.000	<.0001	<.0001
Trial_16 - Trial_4	-0.02	0.01	329	-3.200	0.0042	0.0116
Trial_16 - Trial_5	-0.03	0.01	329	-4.290	0.0001	0.0005
Trial_16 - Trial_6	-0.02	<0.01	329	-3.730	0.0008	0.0026
Trial_16 - Trial_7	-0.02	0.01	329	-3.700	0.0008	0.0026
Trial_16 - Trial_8	-0.02	0.01	329	-2.390	0.0342	0.0641
Trial_16 - Trial_9	-0.01	<0.01	329	-1.620	0.1698	0.2548
Trial_17 - Trial_2	-0.04	0.01	329	-5.290	<.0001	<.0001
Trial_17 - Trial_4	-0.02	0.01	329	-3.440	0.0021	0.0065

Trial_17 - Trial_5	-0.03	0.01	329	-4.620	<.0001	<.0001
Trial_17 - Trial_6	-0.02	<0.01	329	-4.220	0.0001	0.0005
Trial_17 - Trial_7	-0.02	0.01	329	-4.100	0.0002	0.0008
Trial_17 - Trial_8	-0.02	0.01	329	-2.650	0.0184	0.0402
Trial_17 - Trial_9	-0.01	<0.01	329	-2.000	0.0793	0.1343
Trial_2 - Trial_4	0.01	0.01	329	1.330	0.2502	0.3412
Trial_2 - Trial_5	0.01	0.01	329	1.190	0.3086	0.405
Trial_2 - Trial_6	0.02	0.01	329	2.750	0.0145	0.0331
Trial_2 - Trial_7	0.02	0.01	329	2.280	0.042	0.076
Trial_2 - Trial_8	0.02	0.01	329	2.550	0.023	0.0474
Trial_2 - Trial_9	0.03	0.01	329	4.100	0.0002	0.0008
Trial_4 - Trial_5	<0.01	0.01	329	-0.33	0.7856	0.8332
Trial_4 - Trial_6	0.01	0.01	329	0.9	0.4427	0.5282
Trial_4 - Trial_7	<0.01	0.01	329	0.57	0.6506	0.7354
Trial_4 - Trial_8	0.01	0.01	329	1.040	0.3662	0.4471
Trial_4 - Trial_9	0.02	0.01	329	2.250	0.0445	0.0792
Trial_5 - Trial_6	0.01	0.01	329	1.580	0.1717	0.2548
Trial_5 - Trial_7	0.01	0.01	329	1.120	0.3322	0.4203
Trial_5 - Trial_8	0.01	0.01	329	1.570	0.1732	0.2548
Trial_5 - Trial_9	0.02	0.01	329	3.200	0.0042	0.0116
Trial_6 - Trial_7	<0.01	<0.01	329	-0.43	0.7208	0.7802
Trial_6 - Trial_8	<0.01	0.01	329	0.38	0.7568	0.8109
Trial_6 - Trial_9	0.01	<0.01	329	2.230	0.0458	0.0802
Trial_7 - Trial_8	<0.01	0.01	329	0.67	0.5904	0.6965
Trial_7 - Trial_9	0.01	0.01	329	2.390	0.0342	0.0641
Trial_8 - Trial_9	0.01	0.01	329	1.240	0.292	0.3881

Legend for Supplementary Table 2

Contrast	Specific comparison between two levels of ‘Latency’
Estimate	Estimated difference between two levels (response scale)
SE	Standard error of the estimate
df	Degrees of freedom
t-ratio	Ratio of the contrast estimate to its standard error (SE)
p-value	Probability of observing the data under a true null hypothesis
FDR p-value	p-value after FDR correction for multiple comparisons