

# White matter structural bases for predictive tapping synchronization

Pamela García-Saldivar <sup>a</sup>, Cynthia de León<sup>a</sup>, Luis Concha <sup>a,b,\*</sup>, Hugo Merchant <sup>a,\*</sup>

<sup>a</sup>*Institute of Neurobiology, Universidad Nacional Autónoma de México. Blvd. Juriquilla 3001, Querétaro, Querétaro, México.*

<sup>b</sup>*International Laboratory for Brain, Music and Sound (BRAMS), Montreal, Québec, Canada.*

---

## Abstract

We determined the intersubject association between rhythmic entrainment abilities of human subjects during a synchronization continuation tapping task (SCT) and the macro and microstructural properties of their superficial (SWM) and deep (dWM) white matter. Diffusion-weighted images were obtained from 32 subjects who also performed the SCT with auditory or visual metronomes and five tempos ranging from 550 to 950 ms. We developed a method to determine the fiber density of U-fibers running tangentially to the cortex. Notably, the right audiomotor system showed individual differences in the density of U-fibers that were correlated with the degree of predictive entrainment across subjects. These correlations were selective for the synchronization epoch with auditory metronomes and were specific for tempos around 1.5 Hz. In addition, there was a significant association between predictive rhythmic entrainment and the density and bundle diameter of the corpus callosum (CC), forming a chronotopic map where behavioural correlations of short and long intervals were found with the anterior and posterior portions of the CC. Finally, the fiber bundle cross-section of the arcuate fasciculus, the CC, and the Superior Longitudinal Fasciculus showed a significant correlation with the mean asynchronies of the auditory SCT. These findings suggest that the structural properties of the SWM and dWM in the audiomotor system support the predictive abilities of subjects during rhythmic tapping, where the density of cortical U-fibers are linked to the preferred tapping tempo, and the bundle properties of CC define an interval selective topography that has an anterior posterior gradient.

*Keywords: audiomotor system, tap synchronization, superficial white matter, corpus callosum*

---

## 1. Introduction

Moving in synchrony with a beat is a basic and generalized human ability that can reach sophisticated levels in professional percussionist (Honing et al., 2015; Merchant et al., 2015a). Indeed, humans are extremely sensitive to auditory regularities, can entrain to auditory beats across a wide range of tempos, and can use timed movements of different body parts (such as finger or foot taps, or body sway) to keep the beat (Mendoza & Merchant, 2014; Repp & Su, 2013). A classical task used to study rhythmic entrainment is the synchronization-continuation task (SCT), where subjects first entrain their tapping to an isochronous metronome and then continue tapping without the periodic stimulus using an internal clock (Merchant et al., 2008a,b;

---

\*Corresponding authors. Email addresses: [hugomerchant@unam.mx](mailto:hugomerchant@unam.mx), [lconcha@unam.mx](mailto:lconcha@unam.mx)

Wing, 2002; Wing & Kristofferson, 1973). In this type of tasks humans show negative asynchronies, namely they tap a few milliseconds before the metronome, supporting the notion that entrainment depends on a predictive internal pulse (Lenc et al., 2021; Repp & Su, 2013; Zarco et al., 2009). Neurophysiological and functional imaging studies have shown that the internal beat representation during SCT rests in the motor system including the basal ganglia and the supplementary motor regions (Bartolo et al., 2014; Merchant et al., 2013a, 2015a; Merchant & Averbeck, 2017; Rao et al., 2001; Sánchez-Moncada et al., 2020). These areas produce a regenerating rhythmic signal (Crowe et al., 2014; Gámez et al., 2019; Merchant et al., 2015b) that dynamically interacts with the auditory areas, creating audiomotor loops where the motor prediction of the pulse is flexibly compared and adjusted depending on the changes in the input stream of rhythmic stimuli (Comstock et al., 2018; Merchant & Honing, 2014; Patel & Iversen, 2014). In addition, the supplementary motor areas possess neurons that are tuned to the tempo duration of tapping (Merchant et al., 2014, 2013b) giving rise to interval specific circuits that define a chronotopic map, with short preferred intervals in the anterior portion and long preferred intervals in the posterior portion of the medial premotor areas (Protopapa et al., 2019).

The existence of neural circuits with preferred intervals conforming chronotopic maps goes in line with the notion of a spontaneous rhythmic tempo, an interval that is produced naturally when asked to tap in the absence of external cues (McAuley et al., 2006; Zamm et al., 2018). This spontaneous or preferred tempo is around 600 to 750 ms in human adults (Fraisse, 1963, 1978) being faster in early childhood and slower in elderly individuals (McAuley et al., 2006). Thus, the biological intrinsic periodicity may depend on a biased distribution of preferred tempos towards 1.5 Hz in the interval-tuned neurons of the motor system (Bartolo et al., 2014; Bartolo & Merchant, 2009; Merchant et al., 2013b).

Although rhythmic entrainment is prevalent across all human cultures and is a natural behaviour for social interaction (Nettle, 2000), there are wide individual differences in the precision, accuracy and predictability (measured with the asynchronies that correspond to the time between the stimulus and the response) of movement synchronization. There are subjects who, despite lacking musical training, spontaneously synchronize to rhythmic stimuli ranging from strictly periodic metronomes to complex musical pieces, with performance comparable to trained musicians (Scheurich et al., 2018; Tranchant et al., 2016). In contrast, there are also poor synchronizers (around ten percent of the population) that show low accuracy and large asynchronies to isochronous metronomes and musical excerpts (Phillips-Silver et al., 2011; Sowiński & Dalla Bella, 2013; Tranchant et al., 2016). Furthermore, non-musicians synchronize less flexibly and less precisely across tempi than musicians (Scheurich et al., 2018). Hence, both genetic and learning factors generate human populations with a wide variety of beat entrainment abilities. Notably, functional imaging has revealed that individual differences in beat perception depend on activation differences in SMA and the posterior auditory cortex (Grahn & McAuley, 2009). In addition, the putamen, SMA and auditory cortex show larger functional connectivity during rhythm perception, with larger modulation for musicians than non-musicians (Grahn & Rowe, 2009, 2013). Consequently, the magnitude of the anatomofunctional association between the auditory and motor control areas of the cerebral cortex seem to covary with the individual difference to perceive and entrain to simple regular beats. Such brain networks rely on the structural connectivity provided by white matter. Yet, little is known about whether the macro and microstructural properties in the superficial and deep white matter can explain the individual variations in rhythmic entrainment performance (Assaneo et al., 2019; Steele et al., 2013). We hypothesized that the properties in the microstructure of fibers running superficially to the cortical surface and of the deep tracts of the audiomotor system account for the

individual differences of the predictive abilities during tapping synchronization.

To test this hypothesis, we acquired diffusion-weighted images (DWI) from 32 subjects that previously performed a SCT using flashing visual or auditory tones as metronomes in the ranges of hundreds of milliseconds. SCT rhythmic performance across durations (550, 650, 750, 850 or 950 ms) and modalities (auditory and visual) was characterized using the absolute asynchronies, the autocorrelation of the inter-tap interval time series during the synchronization epoch, the constant error, and the temporal variability during both synchronization and continuation epochs. These parameters measure the prediction, error correction, accuracy, and precision of the rhythmic tapping of the subjects, respectively (Figure 2A). On the other hand, we evaluated deep white matter fascicles using a fixel-based approach (Dhollander et al., 2021), and developed two metrics of superficial white matter, namely fiber density corresponding to fibers entering/exiting the cortex and U-fibers running tangentially to the cortex. Crucially, we used a wide range of tempos in the SCT in order to test the existence of both a structural foundation for the preferred interval and an interval selective topographical arrangement in the behavioural and structural associations. Widespread correlations in the right audiomotor circuit were found between the tangential U-fibers density and the predictive abilities of subjects during synchronization epoch of the auditory condition for the intervals of 650 and 750ms. The interval specificity in these associations suggest that the preferred tempo for rhythmic entrainment has its origins in the structural properties of the U-fibers running superficially across the audiomotor circuit. In addition, there was a significant association between predictive rhythmic entrainment in the auditory condition and the density and bundle diameter of the corpus callosum, forming a chronotopic map with an anterior-posterior gradient, similar to the topography of interval tuned clusters observed with functional imaging. Crucially, the anatomobehavioural associations were negative, indicating that subjects with good predictive abilities and small asynchronies showed large superficial and deep apparent fiber densities, while subjects with large asynchronies showed low fiber densities.

## 2. Results

### 2.1. Rhythmic behaviour

Thirty-two subjects performed a modified version of the classical synchronization continuation task (SCT) that included the following three epochs: beat perception, synchronization, and continuation (Figure 1). This task starts with the active perception of the isochronous beat defined by visual stimuli alternating between a right and left position, followed by tapping synchronization to the alternating stimuli and the internally driven tapping continuation to the right or left targets without the metronome. The subjects also performed an auditory version of the SCT (Figure 1, see Methods).

Absolute asynchronies correspond to the time difference between each stimulus and response pair and are a measure of rhythmic prediction. Hence, this parameter can only be measured during the synchronization epoch of SCT. A repeated-measures ANOVA on absolute asynchronies with metronome modality (auditory and visual) and instructed interval duration as within-subject factor, showed significant main effects for duration ( $F(4,124) = 36.88$   $P < 0.0001$ ) and modality ( $F(1,31) = 20$   $P < 0.0001$ ), as well as a significant duration x modality interaction ( $F(4,124) = 32.3$   $P < 0.0001$ ). Tukey's honest significant difference (HSD) post hoc test showed significantly larger asynchronies for the visual than auditory modality across all durations ( $P < 0.0001$  for 550, 650, 750 and 850 ms;  $P = 0.006$  for 950 ms). These results confirm the preponderance of the auditory over

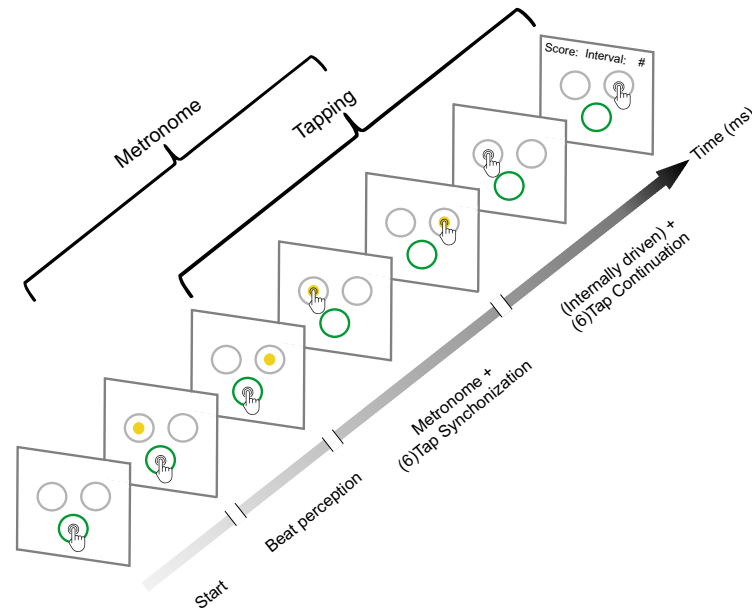


Figure 1: Synchronization-Continuation Task (SCT). Initially, the subjects placed their finger at the central-bottom target of a touch screen to start the trial and maintained the finger in this position while attending to a sequence of two to six right/left alternating visual stimuli with a constant interstimulus interval (target duration, perception epoch). Then, once they considered to have extracted the beat, subjects tapped the touch screen on the left or right target in synchrony with the alternating visual metronome during six intervals (synchronization epoch). Finally, they continued tapping on the right or left targets for another six intervals without the metronome (continuation epoch). The mean produced and the target interval were displayed at the end of each trial as feedback. Subjects also performed an auditory metronome version of the SCT, where the metronome consisted of 500 Hz tones that were presented on the right or left side of a headphone. Five target durations (550,650,750,850 or 950 ms) were presented pseudorandomly, with the visual and auditory conditions interleaved between subjects.



the visual modality to guide rhythmic temporal prediction (Comstock et al., 2018; Gámez et al., 2018; Merchant et al., 2015b). In addition, we computed the intersubject correlation matrix on the absolute asynchronies across instructed intervals and we found a significant correlation between 650 and 750 ms for both the auditory and visual conditions ( $r=0.67$ ,  $p=.000026$ ;  $r=0.69$ ,  $p=.000011$ , respectively) (Extended Data 1). This finding suggests the existence of a shared mechanism for rhythmic temporal prediction in these intermediate intervals.

A negative lag 1 autocorrelation of the produced intervals during the synchronization epoch indicates the involvement of an error correction mechanism that maintains tap synchronization with the metronome, since a longer produced interval tends to be followed by a shorter interval, while a shorter interval tends to be followed by a longer produced duration (Iversen et al., 2015; Repp & Su, 2013). The corresponding repeated-measures ANOVA on autocorrelation of the inter-tap interval time series (Figure 2C) showed no significant differences between modalities ( $F(1,31) = 1.6$ ,  $P = 0.21$ ), intervals ( $F(4,124) = 2.05$ ,  $P = 0.09$ ) or their interaction ( $F(4,124) = 2.34$ ,  $P = 0.06$ ). Thus, lag 1 autocorrelation across trials was negative and similar across modalities and target durations supporting the notion of a robust and amodal error correction mechanism during the SCT.

Constant error is the difference between the produced and the target intervals and is a measure of timing accuracy during the synchronization and continuation epochs. A repeated-measures ANOVA on constant error with modality, target duration and task epoch as within-subject factors, revealed statistically significant main effects for modality ( $F(1,31) = 46.05$ ,  $P < 0.0001$ ), target duration ( $F(4,124) = 29.15$ ,  $P < 0.0001$ ) and epoch ( $F(1,31) = 6.44$ ,  $P = 0.01$ ), as well significant interactions between: modality x interval ( $F(4,124) = 3.5$ ,  $P = 0.008$ ), epoch x interval ( $F(4,124) = 48.33$ ,  $P < 0.0001$ ), and epoch x modality ( $F(1,31) = 23.35$ ,  $P < 0.0001$ ). The post hoc Tukey HSD showed significant no differences for target duration in constant error during synchronization for both modalities, with accurate timing close to zero. In contrast, during continuation, the same post hoc test revealed significant differences between distant intervals within the auditory and visual modalities (Figure 2D). For the continuation epoch, the constant error followed the bias effect, with overestimation for short and underestimation for long durations, especially for the auditory condition (Jazayeri & Shadlen, 2010; Pérez & Merchant, 2018). Indeed, the indifference interval, that corresponds to the interval associated with zero constant error, was 654ms for the auditory (Figure 2D, blue arrow) and 420 ms for the visual condition. This finding suggests that our subjects had a clear preferred interval in the auditory condition that is close to the reported 2Hz in the literature (Zamm et al., 2016). Finally, temporal variability was defined as the standard deviation of the produced intervals and is a metric of timing precision. The same repeated-measures ANOVA on temporal variability showed significant main effects for target duration ( $F(4,124) = 110$ ,  $P < 0.0001$ ), and modality ( $F(1,31) = 58.06$ ,  $P < 0.0001$ ), but no significant main effect for task epoch ( $F(1,31)=0.94$ ,  $P = 0.33$ ). In addition, significant effects were revealed for the following interactions: epoch-interval ( $F(4,124) = 3.42$ ,  $P = 0.01$ ), epoch-modality ( $F(1,31) = 70.68$ ,  $P < 0.0001$ ), interval-modality ( $F(4,124) = 11.34$ ,  $P = 0.0001$ ), and epoch-modality-interval ( $F(4,124) = 14.74$ ,  $P = 0.0001$ ). The HSD post hoc tests showed significantly larger temporal variability in the visual than the auditory metronome (for the intervals 650, 750, 850, and 950:  $P < 0.0001$ ) during the synchronization but not the continuation epoch, confirming the larger precision for auditory metronomes, especially during synchronization (Gámez et al., 2018; Repp & Penel, 2004).

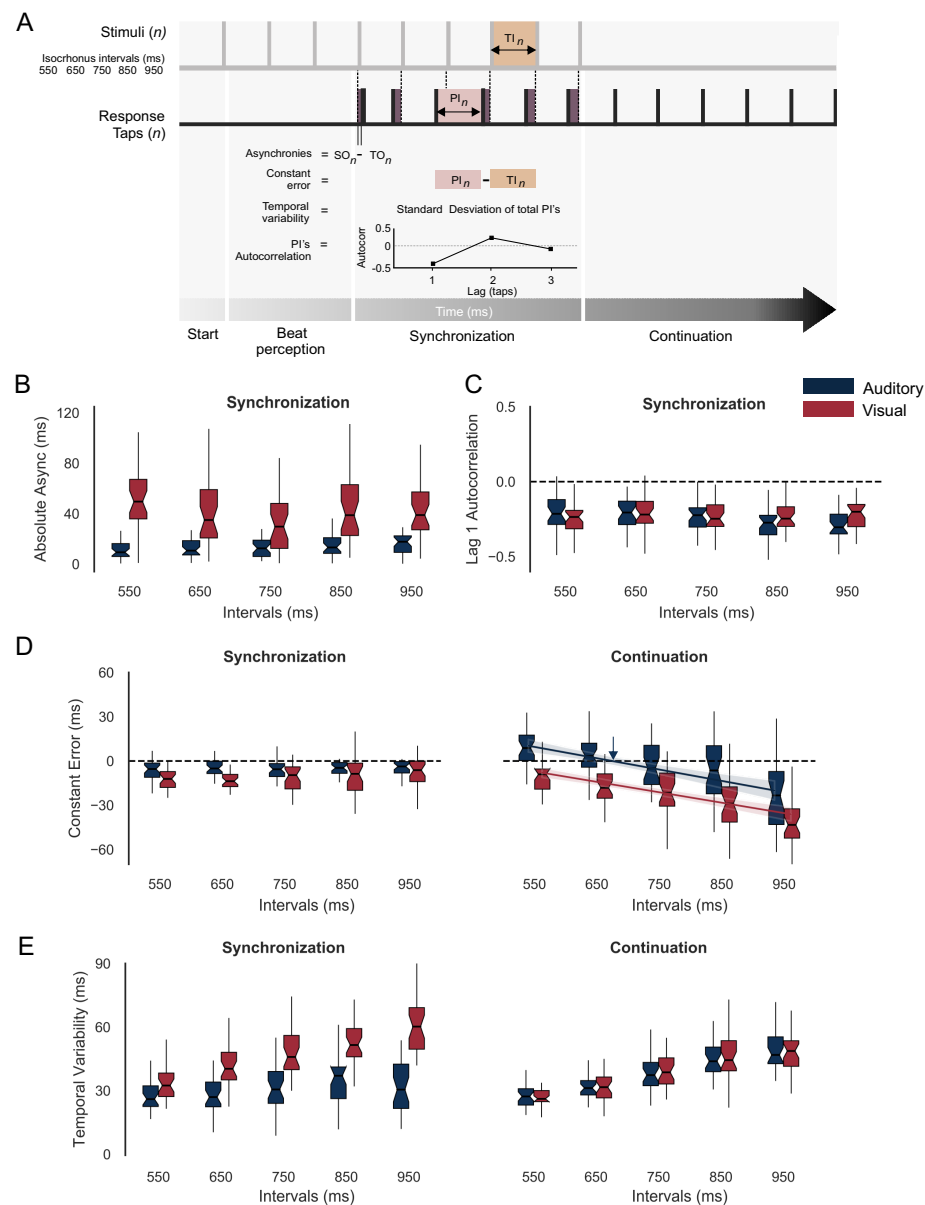


Figure 2: A. Parameters of rhythmic performance during SCT. Asynchronies correspond to the time difference between stimulus (SO) and tap onset (TO) across the  $n$  intervals of the synchronization epoch. The constant error is the difference between produced (PI) and target intervals (TI), and the temporal variability is the standard deviation of the PI. Finally, the autocorrelation of the PI during synchronization is computed and the lag 1 autocorrelation: a negative value indicates that the subject is using an error correction mechanism (see the text). B and C. Absolute asynchronies and lag 1 autocorrelation for each instructed interval and metronome modality (Auditory: blue, Visual: red) during the synchronization epoch. D and E. Constant error and temporal variability as a function of target interval for both metronome modalities and the two tapping epochs of the SCT. The coloured lines in D for the continuation correspond to the linear fit between the constant error and the target interval; the indifference interval corresponds to 680ms (blue vertical arrow) for the auditory condition.

## 2.2. Analysis of white matter

Before the task performance session, participants were scanned in a 3 T Philips Achieva TX MR scanner with a 32-channel head coil. T1-weighted volumes and diffusion-weighted images (DWI) were obtained (see Methods). For each subject, the gray/white matter interface was identified using a surface mesh (Garcia-Saldivar et al., 2021). The apparent fiber density (AFD) (Raffelt et al., 2012), derived from the DWI using constrained spherical deconvolution (CSD) (Tournier et al., 2004), was sampled at each vertex of this mesh. To evaluate white matter properties at different depths with respect to the cortical mantle, we created synthetic trajectories that organically extend from each vertex at the grey/white matter interface towards the ventricles and sampled diffusion metrics along these trajectories every 0.5 mm (Figure 3A). Leveraging the ability of CSD to disentangle crossing fiber populations, AFD (Figure 3B) was evaluated separately for those fibers that enter/exit the cortex and are, therefore, parallel to the virtual trajectories ( $_{par}AFD$ ), and those that extend tangentially to the cortex, perpendicular to the virtual trajectories ( $_{tan}AFD$ ). The latter corresponds to short-range cortico-cortical connections through U-fibers (Schüz & Braitenberg, 2002).

## 2.3. Correlations between behaviour and Superficial White Matter (SWM)

Next, we examined the association between behavioural performance and microstructural properties of superficial white matter, inferred from diffusion-weighted magnetic resonance images. The surface-based analysis of SWM (SBA, see Methods) was performed to determine the possible association between the different metrics of rhythmic timing performance and the metrics of the SWM sampled at five depths with respect to the grey/white matter interface (0, 0.5, 1, 1.5 and 2 mm). This analysis showed the existence of negative correlations between the auditory absolute asynchronies and values of the AFD maps in the five depths sampled (Figure 3 EFG; also see Extended Data 2). Thus, subjects with auditory asynchronies closer to zero, and hence with larger predictive abilities, had significantly higher values AFD and  $_{tan}AFD$  than subjects with less predictive performance. Notably, these significant associations were observed mainly for the auditory asynchronies of the 650 and 750 ms intervals and the values of  $_{total}AFD$  and  $_{tan}AFD$  (see Extended Data 3, 4 and 5). No significant correlations were found between the three AFD maps and the constant error, temporal variability, and lag 1 autocorrelation for the auditory condition. Furthermore, no correlations were observed between the three metrics of the SWM and all the behavioural parameters for the visual condition (see Extended Data 2). We did not find significant correlations between the parameters of SCT rhythmic performance and the  $_{par}AFD$ , except for few clusters with low correlation coefficients for the auditory asynchronies of the intervals of 550, 650 and 750 ms (Extended Data 2 and Figure 3H). Indeed, the level of association between predictive behaviour and the AFD maps was larger for  $_{tan}AFD$  than  $_{par}AFD$  (Figure 3H; compare Figure 5 with Extended Data 8). The average thickness of the U-fibre system is around 1.5 mm (Schüz & Braitenberg, 2002), and following analyses focus on this depth. Figure 4 shows the correlation coefficient values between the auditory asynchronies and  $_{tan}AFD$  at 1.5mm below the GM/WM interface across all the tested tempos. After correction for multiple comparisons ( $p_{FWE} < 0.001$  and  $p_{cluster} < 0.001$ ), only intervals of 650, 750 and 850ms showed a significant association between the behavioural and structural parameter. Indeed, nine (with 2232 vertices), twelve (with 2827 vertices), and two (with 374 vertices) clusters showed significant correlations between auditory asynchronies and  $_{tan}AFD$  for the 650, 750 and 850ms intervals, respectively.

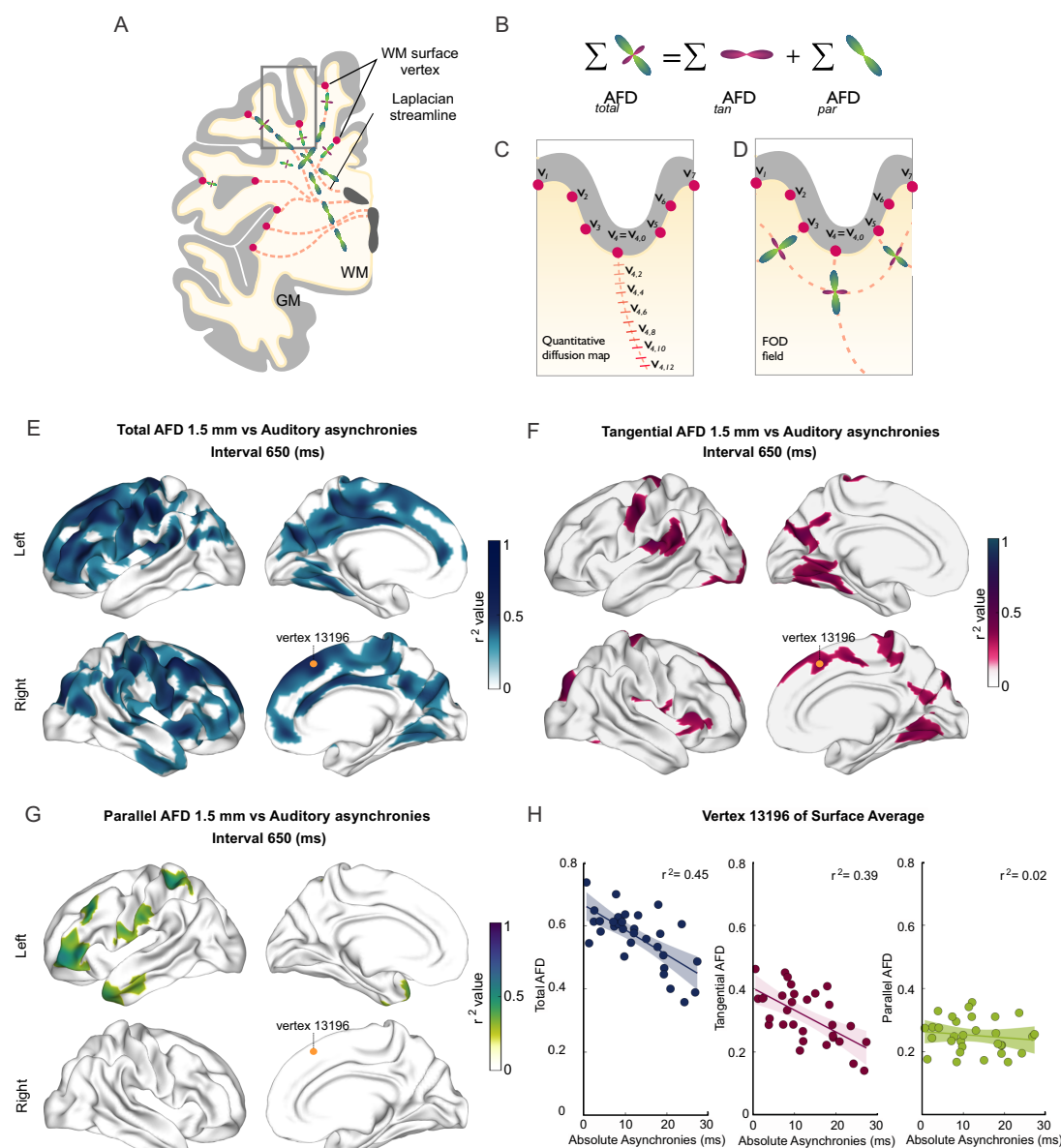


Figure 3: SWM analysis. Apparent fiber density (AFD) of superficial white matter was systematically sampled along synthetic streamlines that extend from each vertex of the gray/white matter interface surface towards the ventricles following a Laplacian field (A, C). The integral of all fiber orientation distribution functions (FOD) corresponds to the total apparent fiber density ( $totalAFD$ ), further separated (B) into fiber densities corresponding to fibers entering/exiting the cortex parallel to the Laplacian streamlines ( $parAFD$ ), or U-fibers running tangentially to the cortex ( $tanAFD$ ) (D). (E) There are widespread significant correlations between the subjects' asynchronies during the SCT with an auditory metronome and  $totalAFD$ , shown here for the 650 ms interval (E). (F) Large areas within the frontal, parietal and occipital lobes showed significant correlations between behaviour and  $tanAFD$ , while only restricted frontal and temporal regions showed correlation between  $parAFD$  and the asynchronies (G). Panel H shows the correlation across the 32 subjects between the three AFD metrics and asynchronies for one exemplary vertex (yellow dot in E, F, and G).

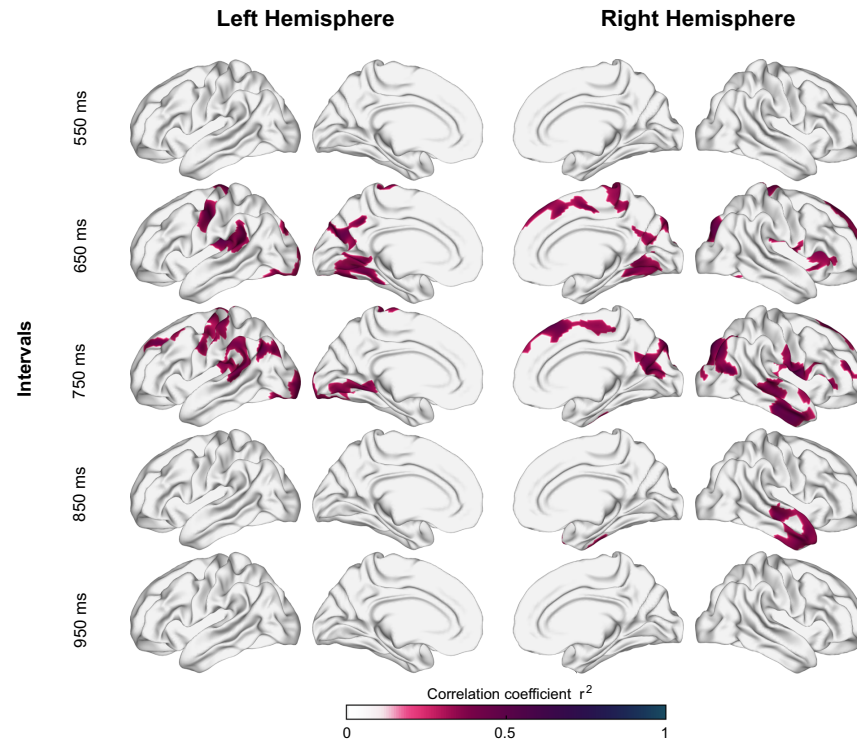


Figure 4: Correlation ( $r^2$ ) between auditory asynchronies and tangential superficial white matter apparent fiber density for each vertex (sampled at 1.5 mm below the gray/white matter interface) across all tested intervals in the SCT for both hemispheres. Significant correlations were localized in large clusters within motor, auditory, and visual areas, particularly for 650 and 750 ms intervals.

To identify the anatomical regions with significant clusters of vertices, we parcellated the SWM based on the Brainnetome atlas (Fan et al., 2016). Areas with significant vertices (after correction for multiple comparisons:  $p_{FWE} < 0.001$  and  $p_{cluster} < 0.001$ ), were grouped into 14 regions: primary motor, dorsolateral secondary motor, medial motor (supplementary and pre-supplementary motor area), dorsolateral prefrontal cortex (DLPFC), language motor, primary somatosensorial, parietal association, precuneus, primary and secondary auditory areas (auditory), inferior temporal cortex, object and face recognition areas, primary and secondary visual areas (visual), and limbic association areas (Figure 5C).

(Figure 5AB shows a systematic decrease in the percent of significant vertices as a function of depth across areas of both hemispheres. The most prominent effect is the interval selectivity of the correlation coefficients and percent of significant vertices, with a preferred interval between 650 and 750 ms, across many areas and depths (Figure 5ABDE). Specifically, at our target depth of 1.5 mm linked with the U-fiber system, the areas with the higher percent of significant vertices on the right hemisphere were pre-SMA (52.4%,  $r^2=.41$ ,  $p < 0.001$ ), auditory (49.2%,  $r^2=.38$ ,  $p < 0.001$ ) and language motor (40.2%,  $r^2=.45$ ,  $p < 0.001$ ) for the correlation between the 750 ms auditory asynchronies and  $_{tan}AFD$ . Similarly, at 650 ms the higher percent areas with significant vertices on the right hemisphere were the medial motor areas (pre-SMA, 41.8%,  $r^2=.36$ ,  $p < 0.001$  and SMA 28.1%,  $r^2=.39$ ,  $p < 0.001$ ) and language motor (28.1%,  $r^2=.40$ ,  $p < 0.001$ ). Figures 5D and E show the interval selectivity curves for the correlation coefficients and percent of significant vertices across  $_{tan}AFD$  depths (colour coded) for auditory cortex and preSMA, confirming the notion of interval preference for the behavioural and structural associations. In contrast, in the left hemisphere the percent areas with significant vertices were smaller, with significant effects for visual areas (31.8%,  $r^2=.45$ ,  $p < 0.001$ ), object and face recognition (29.9%,  $r^2=.44$ ,  $p < 0.001$ ) and primary motor areas (28.4%,  $r^2=.44$ ,  $p < 0.001$ ) for the 650 ms interval; and parietal association (35.1%,  $r^2=.52$ ,  $p < 0.001$ ) primary motor areas (28%,  $r^2=.40$ ,  $p < 0.001$ ) and visual areas (25%,  $r^2=.40$ ,  $p < 0.001$ ) for the 750 ms interval.

These results support the notion that the audiomotor system of the right hemisphere shows individual differences in the density of U-fibers with a gradient that correlates with the degree of prediction across subjects in the synchronization phase of the SCT. Crucially, this correlation is modality specific, suggesting that structural changes in the audiomotor system occur only for the modality that produces rhythmic timing with more precision, accuracy, and degree of predictability of the metronome (see Extended Data 2). Furthermore, the interval selectivity on the behavioural and structural correlation indicates that the timing circuits tuned at intervals of 650 and 750ms are the ones that are susceptible to present macro and microstructural differences in U-fibers of WM within the tested population of subjects. Finally, the visual areas and the motor cortex, contralateral (left hemisphere) to the tapping movements, also showed significant associations between rhythmic timing prediction and the U-fiber density.

## 2.4. Deep WM and structural chronotopic representation

We also evaluated the association between the precision, accuracy, prediction and error correction of rhythmic timing with the properties of the deep WM (dWM). Fixel- based analysis (FBA) (Dhollander et al., 2021; Raffelt et al., 2017) was used to estimate micro and macrostructural differences within-voxel of the dWM (Genc et al., 2018; Kelly et al., 2020; Mito et al., 2018; Rau et al., 2019). As with our analysis of superficial WM, this method is based on the constrained spherical deconvolution (CSD) of DWI data. FBA provides three fiber-specific indices (fiber density, fiber



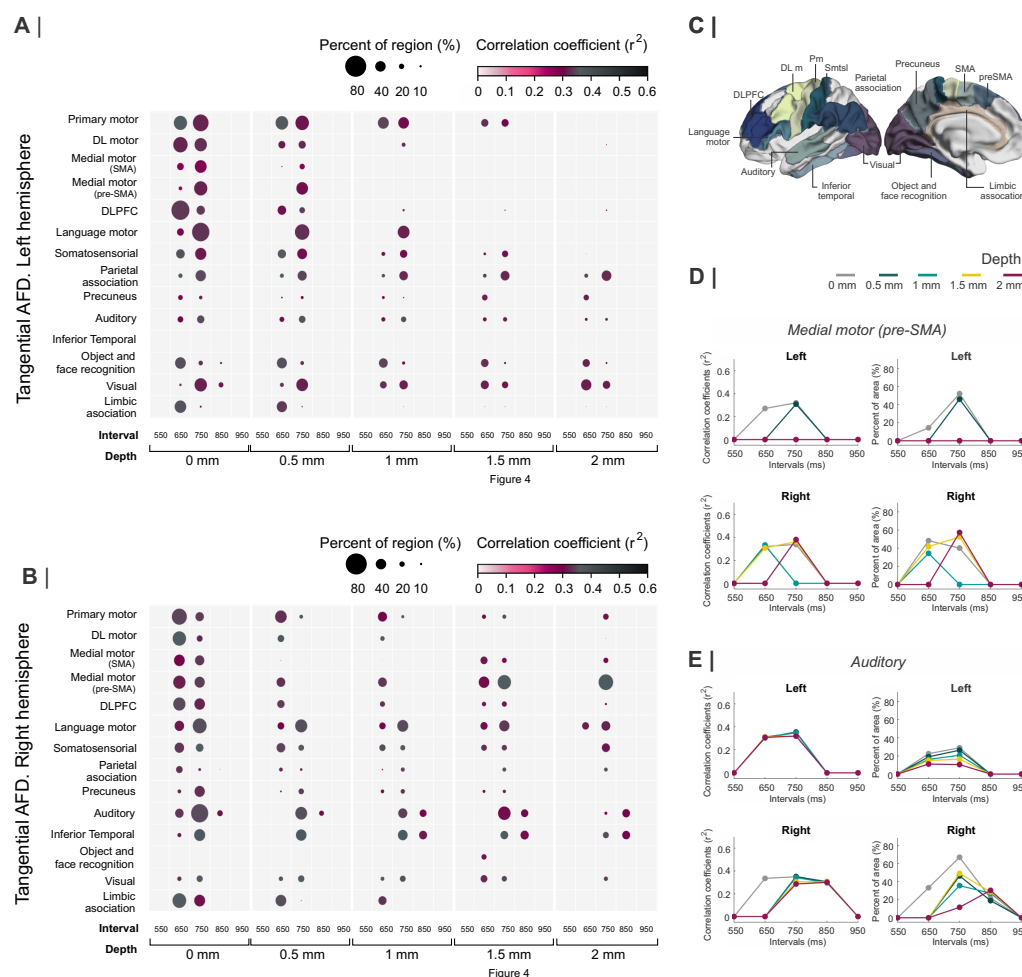


Figure 5: A and B. Interregional correlation plot of the auditory asynchronies and *tanAFD* as a function of the instructed interval and the depth of the *tanAFD* calculation. Significant clusters were aggregated into fourteen regions (y-axis) based on the Brainnetome atlas (C) (Fan et al., 2016). The colour and size of the circle for each cluster correspond to the correlation coefficient (critical value  $r^2 > 0.355$  at  $p < 0.02$ ,  $df=29$ ) and the percent of significant vertices in each area, respectively. The left hemisphere showed more areas with significant vertices than the right ( $17541 > 16641$ ). A systematic decrease in circle size as a function of depth was observed across areas of both hemispheres. At a depth of 1.5 mm the regions with a larger percent of significant vertices for the right hemisphere at 750ms were the medial premotor (SMA and preSMA), auditory, and language motor areas (B), while small number of significant vertices were observed across the regions of the left hemisphere (A). In A and B it is possible to see the cluster analysis per area of the data from (Figure 4). D. Interval selectivity curves for the correlation coefficients and percent of significant vertices across *tanAFD* depths (colour coded) for preSMA. E: same and D but for auditory cortex. Note that the preferred interval in the two areas is between 650 and 750 ms.



cross-section, and fiber density cross-section; FD, FC, and FDC, respectively) (Raffelt et al., 2017). FD is derived from the integral of the FOD lobes and is proportional to the total intra-axonal volume, thus reflecting the density of a population of fibers within a voxel (Rojas-Vite et al., 2019). If more than one fiber population coexist in a given voxel, the FOD is segmented, and a FD is assigned to each population referred to as fixel (fiber element). FC is a macroscopic metric of the fiber bundle diameter and, finally, FDC is a combination of FD and FC (see Methods). Briefly, the FBA analysis pipeline consists of five steps. First, the images of each subject are processed to obtain a WM Fiber Orientation Distribution (FOD) map in native space using CSD. Second, a template of the FODs and a fixel mask are built. Third, the AFD map and the corresponding fixels for each voxel are computed for each subject in native space. The FD metric is obtained for each fixel computed from the total DWI signal per voxel. Fourth the fixels and AFD map are reoriented to the template. Finally, fixel-wise statistics are performed at each spatial location in template space.

The FBA revealed significant negative correlations between the FDC in the corpus callosum (CC) and the predictive ability to synchronize to an auditory metronome: asynchronies for 650 ms, 750 ms, 850 ms and 950 ms intervals (Figure 6). Therefore, this analysis showed a tight relation between the density and bundle diameter of CC fibers and beat entrainment, with the subjects showing large predictive abilities and asynchronies close to zero also showing large FDC values, and the subjects with bad prediction and large asynchronies showing low FDC values (Extended Data 8). As in the case of the U-fiber metrics, the FBA values were not correlated with the other SCT behavioural parameters of the auditory and visual conditions during the synchronization and continuation epochs.

The association between rhythmic entrainment and white matter properties defined a chronotopic map in the CC, with FDC at different levels of CC showing significant correlations with the absolute asynchronies at specific intervals (Figure 6). This chronotopic map showed an anterior-posterior gradient, with behavioural and structural associations for short and long intervals in the anterior and posterior portions of the CC, respectively. Thus, the FDC fixel values of the posterior midbody of CC (interconnecting motor, pre-motor cortices and M1) showed a significant negative correlation with absolute auditory asynchronies for the 650 and 750 ms intervals (Figure 6B,C; FWE-corrected P-value < 0.05). For the asynchronies at the intermediate interval of 850 ms, the negative correlation was observed with FDC fixel values located in the isthmus and the splenium (Figure 6C; interconnecting primary motor, temporal, and visual cortices). Finally, the asynchronies of the interval of 950 ms showed a significant negative correlation with fixels located in forceps minor (FMI) and major (FMA) (Figure 6A,D; interconnecting prefrontal and visual cortices, respectively).

It is evident in Figure 6A-E that the streamlines segments of the fixels with entrainment correlations were located mainly at the joint of the two hemispheres across the CC. Nevertheless, a lateralization effect was found for the left hemisphere, with fixels associated with auditory asynchronies of 750 and 950 ms in the isthmus and splenium, respectively. Besides the CC, right fornix showed a significant association with the asynchronies of the 750 ms interval for the FD metric (see Extended Data 8).

Lastly, we carried out a correlation analysis between the mean absolute asynchronies across the five intervals and the FC (Figure 7) and FDC (Extended Data 8). Notably, for the auditory condition the tracts with significant FC fixels were: left Arcuate Fasciculus (AF; Figure 7A,E), CC M1 (Figure 7C,E), Forceps Mayor (FMA; Figure 7B,D), Superior Longitudinal Fasciculus 2 (SFL2; Figure 7F) and right Fornix (FX), (Extended Data 8).

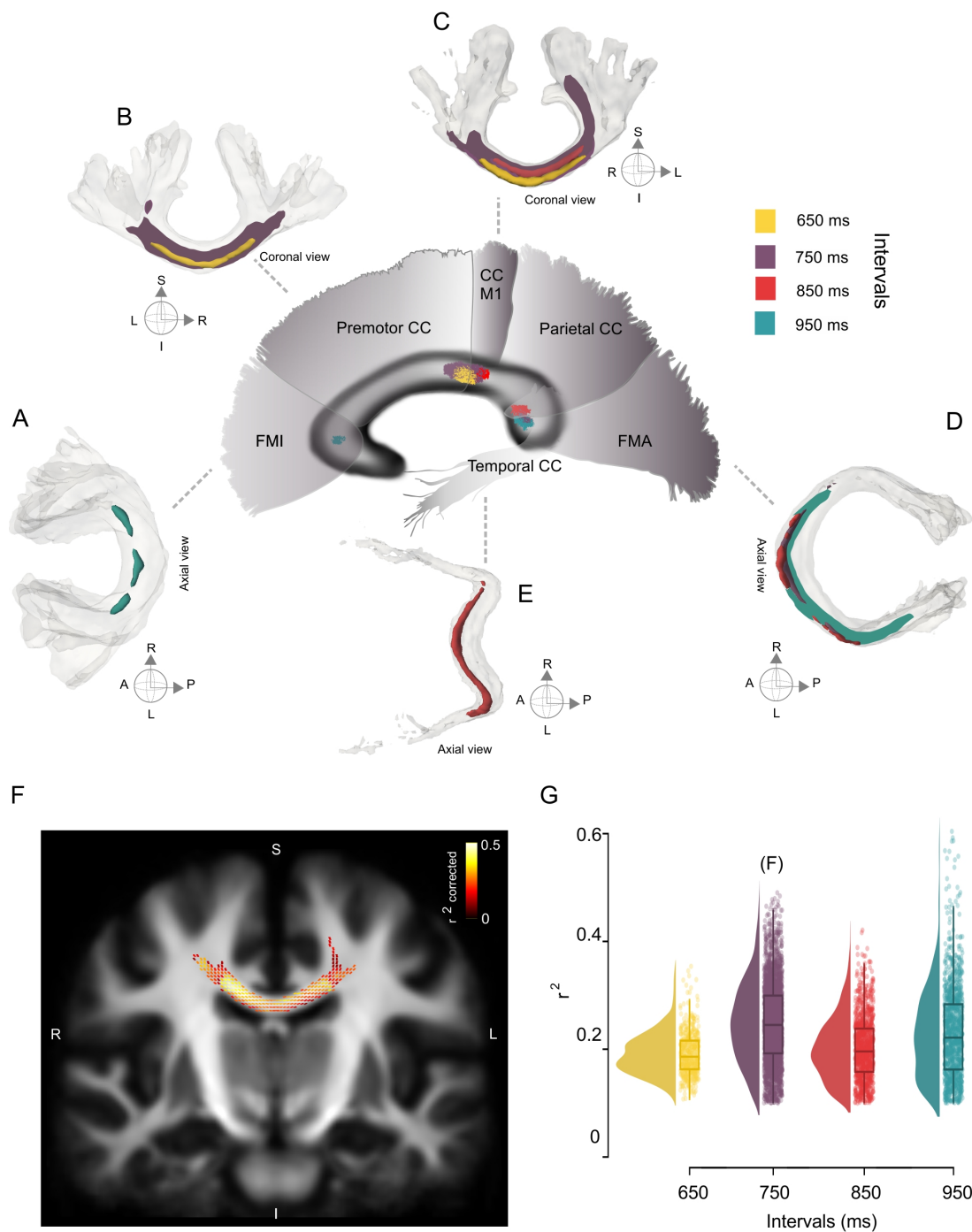


Figure 6: Chronotopic map of the correlation between the Deep White Matter (dWM). A-E. Fiber bundles that showed significant correlations between the asynchronies in the SCT auditory condition and the fiber density and cross-section (FDC) of the corpus callosum (CC). A chronotopic map with an anterior posterior gradient is evident. F. Coronal section of the CC showing the fixels with a significant correlation coefficient (colour coded  $r^2$  values; only fixels with  $p_{corr} < 0.05$  are shown) between asynchronies at the interval of 750 ms in the auditory condition and FDC. G. Distribution of correlation coefficient values ( $r^2$ ) of the FDC vs absolute asynchronies for the four intervals listed in the x-axis. On the right the interquartile box plots are depicted.

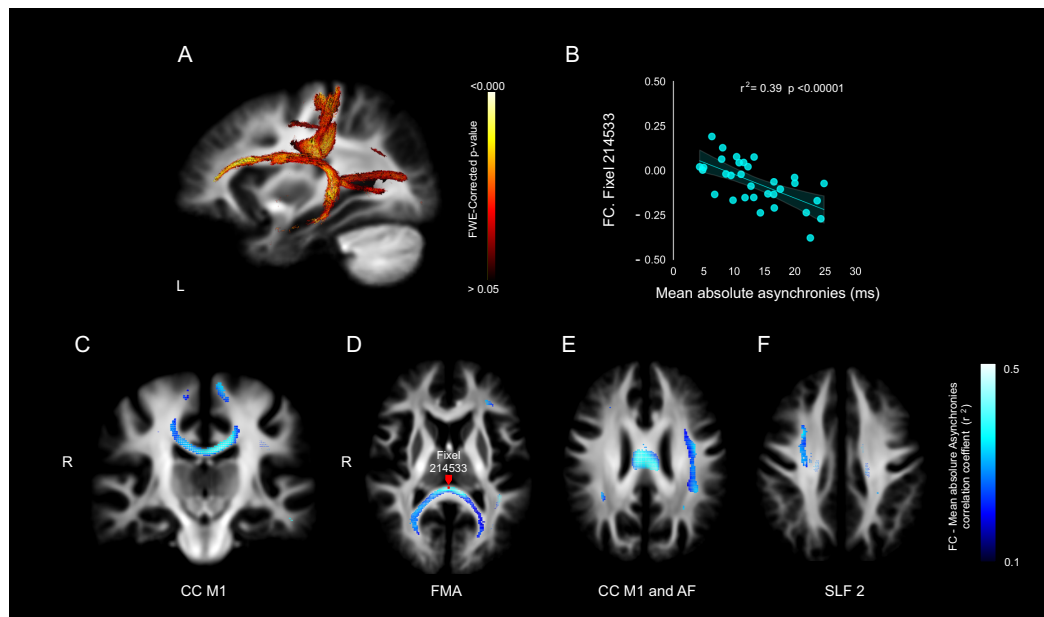


Figure 7: Association between Deep White Matter (dWM) and mean absolute asynchronies of the five tested intervals during SCT (550,650,750,850 and 950 ms). A. p-value of significant fixels after Family Wise Error (FEW) correction. B. Example of the significant correlation between the fiber cross-section (FC) metric and the mean absolute asynchronies in fixel 214533 (red dot in panel D). C, D, E and F. Negative correlation ( $r^2$ ) between FC and mean absolute asynchronies for significant fixels. CC= Corpus Callosum (C), FMA= Forceps Major (D), AF= Arcuate Fasciculus (E), SLF 2= Superior Longitudinal Fasciculus 2 (F).

### 3. Discussion

The present research determined the intersubject association between structural properties of the superficial and deep white matter and different measures of rhythmic timing during a synchronization-continuation tapping task. Our study supports five conclusions: First, healthy human subjects showed precision, accuracy, predictive entrainment, and error correction for rhythmic timing that was biased towards auditory instead of visual metronomes, confirming previous observations. Second, the right audiomotor system showed individual differences in the density of superficial white matter including the U-fibers that were correlated with the degree of predictive entrainment across subjects. Notably, these correlations were selective for the synchronization epoch of the auditory condition and were specific for the intervals of 650 and 750ms. Third, there was a significant association between predictive rhythmic entrainment and the density and bundle diameter of the corpus callosum, forming a chronotopic map with an anterior-posterior gradient. This implies that the behavioural and structural associations for short and long intervals were found in the anterior and posterior portions of the CC, respectively. Fourth, the fiber bundle diameter of the arcuate fasciculus, the Corpus Callosum, the Forceps Major and the Superior Longitudinal Fasciculus showed a significant correlation with the mean asynchronies across all tested tempos. Finally, no significant associations between the superficial and deep white matter properties were found with temporal variability, constant error, or lag one autocorrelation for the visual and auditory conditions during the synchronization and continuation epochs of the SCT. These last findings accentuate the fact that the micro and macro structural properties of the white matter are specifically linked with the predictive ability of the subjects to tap in synchrony with an auditory metronome.

Many studies have shown that performance tapping synchronized to an auditory metronome is more precise and accurate than synchronization to a flashing visual metronome with the same timing characteristics (Chen et al., 2002; Hove et al., 2013; Merchant et al., 2008c; Patel et al., 2005; Repp & Penel, 2004; Zarco et al., 2009). Since the first processing relays, the auditory system has higher temporal resolution compared to the visual system (Duysens J. & A., 1996; He et al., 1997; Sayegh et al., 2011), and plays a critical role on time perception and reproduction across many tasks, not only on tapping SCT (Grondin et al., 2005; Merchant et al., 2008b; Merchant & de Lafuente, 2014). For example, when audiovisual stimuli are used in an oddball paradigm, the perceived duration is dominated by the auditory modality (Chen & Yeh, 2009). In addition, TMS disruption of the auditory cortex impaired not only time estimation of auditory but also of visual stimuli, while disruption in the visual cortex only produced timing impairments for visual stimuli (Kanai et al., 2011). Consequently, a current hypothesis is that the auditory cortex is engaged in multimodal temporal processing and the interaction between the auditory and the motor system in the frontal lobe allows not only for time encoding but also time prediction (Merchant & Honing, 2014; Merchant & Yarrow, 2016; Patel & Iversen, 2014; Schwartz et al., 2012). Our present findings support this notion in different ways. First, the rhythmic timing performance of our subjects was biased towards the auditory condition across different behavioural measures. Second, the right audiomotor system showed widespread and significant correlations between the density of superficial U-fibers and the degree of auditory predictive entrainment across subjects. These anatomo-behavioural associations are selective to the intervals in the 650 to 750 range. Third, the fiber bundle diameter of the left arcuate fasciculus, a key tract connecting the parieto temporal auditory system with the frontal lobe, showed a significant correlation with the mean asynchronies across all tested tempos. The later results are in line with a recent study that showed that the left arcuate fasciculus show altered connectivity and fractional anisotropy reduction in stuttering subjects compared with fluent speakers (Jossinger et al., 2022). Furthermore, the clear clustering of

subjects as either good or bad synchronizers in a syllable isochronous entrainment task correlates with both the difference in the activation magnitude in frontal areas and changes in the white matter pathways (i.e. left arcuate fasciculus) that connect the auditory with the premotor cortical system (Assaneo et al., 2019). Hence, our results accentuate the audiomotor structural foundation for rhythmic entrainment (Honing & Merchant, 2014; Miyata et al., 2022). The lack of structural associations in the visual condition is probably due to the larger variability in rhythmic tapping for this modality, making difficult to infer statistical correlations between our metrics.

The observed associations between the measured properties of superficial and deep white matter and the rhythmic entrainment prediction were negative, indicating that the subjects with asynchronies close to zero and large predictive abilities showed large values in the white matter parameter, while the subjects with bad prediction and large asynchronies showed low values on the same metrics. These results support the theory that intersubject differences in rhythmic entrainment prediction depend on micro and macrostructural white matter properties, which could have a genetic and/or learned substrate. On the genetic side, we could speculate that the existence of poor and superior synchronizers (Blecher et al., 2016) may depend on the fiber density of superficial U-fibers in the right audiomotor system as well as on the density of deep tracts such as the CC and the arcuate fasciculus. On the training side, these superficial and deep white matter bundles may develop larger density and myelination during intense musical practice, distinguishing the audiomotor tracts between musicians and non-musicians (Palomar-García et al., 2020; Vaquero et al., 2018; Zatorre et al., 2007).

The correlations between the density of tangential U-fibers in the right audiomotor circuit and the asynchronies for the auditory condition were interval selective for the intermediate tested tempos. This interval specificity supports the idea that the spontaneous rhythmic tempo observed in many studies of rhythmic entrainment, and with values between 600 and 750 ms, (Drake et al., 2000a,b; McAuley et al., 2006) (see (Figure 2D) has its biological origins in the structural properties of the U-fibers running superficially across the audiomotor circuit. In this scenario, the cortico-cortical connectivity within the audiomotor system is specially designed to support the internal representation of pulse at the preferred tempo, providing larger predictive abilities across subjects for the spontaneous motor tempi (Balasubramaniam et al., 2021). Further anatomofunctional studies are required to test this hypothesis.

Ultra-high field (7 T) functional imaging revealed that the medial premotor areas (SMA-preSMA) of the human brain possess neural circuits that are tuned to different durations forming a topographical arrangement. This chronomaps show units with enhanced responses for the preferred interval and suppressed activity for the nonpreferred duration, and define a topographical gradient with short preferred intervals in the anterior portion and long preferred intervals in the posterior portion of the medial premotor areas (Protopapa et al., 2019; Schwartz et al., 2012). Our measurements on the FDA also revealed a topographical arrangement in the correlations between the density of deep fibers and in CC and the rhythmic predictive abilities of subjects, so that significant anatomobehavioural associations were found in the anterior part of the CC for short intervals and in the posterior CC for long tapping tempos (Schwartz et al., 2012). Nevertheless, our chronotopic map is defined by the correlation between the asynchronies and the FDA, with no functional basis. Moreover, it extends from the CC anterior portions that connect the two hemispheres of the premotor areas of the frontal cortex all the way to the occipital lobe. Hence, the present topography goes well beyond SMA. In line with our present results, a recent imaging study described the existence of large chronomaps covering the cortical mantle from the dorsal and ventral premotor areas to the occipital pole (Harvey et al., 2020; Hendrikx et al., 2022). Consequently, timing maps

define a cortical processing framework for efficient timing integration that has both functional and anatomical basis.

Previous studies have shown correlations between the performance of SMS tasks and microstructural characteristics of deep white matter (DWM) derived from diffusion-weighted magnetic resonance images (DWI). Blecher (Blecher et al., 2016) found significant correlations between asynchronies of participants that performed a finger-tapping task with an auditory metronome in the millisecond scale and fractional anisotropy (FA) of the left arcuate fasciculus and in the temporal segment of the corpus callosum. Better performance in the finger-tapping task was associated with higher values in FA. Nevertheless, the superficial white matter (SWM) immediately below the cortex has so far been neglected, despite its volume (60 % of the total WM volume) and its proximity to neocortical regions having a pivotal role in maintaining cortico-cortical connectivity (Schüz & Braitenberg, 2002; Schmahmann et al., 2006). The SWM is integrated for short-range association fibers that connect adjacent gyri (U-fibers) as well as for the initial or final part of the projection and association (large-range) tracts that form the deep white matter (Schüz & Braitenberg, 2002; Guevara et al., 2020; Kirilina et al., 2020; Yoshino et al., 2020). In recent years abnormalities of SWM have been described in epilepsy (Liu et al., 2016), autism spectrum disorder (Hong et al., 2019), Alzheimer's disease (Phillips et al., 2016), schizophrenia (Nazeri et al., 2013), and stroke (Stockert et al., 2021). Two limitations of studying the SWM in vivo have been constantly described: the fiber-crossing and lower values of anisotropy in comparison with the dWM (Guevara et al., 2020). In order to solve that, we joined two methods for the study of the SWM using synthetic streamlines to query the properties of SWM at different depths and the FD values associated with different orientations of the fiber. This approximation allowed us to perform sub-voxel analyses at the fiber-specific level, describing local macrostructure and microstructural properties. FBA, on the other hand, has been used to evaluate deep white matter in the context of healthy aging, childhood development, psychiatric disorders, fetal and neonatal development, and other disorders (Dhollander et al., 2021). Our results showed that the subjects' accuracy in the performance of SCT was associated with higher FC, FDC or FD values. This is in accordance with the literature that shows that better performance in different tasks is associated with higher values of FBA metrics. For example, better performance in a task of bimanual coordination was associated with higher values of the FBA in the corpus callosum (Adab et al., 2020). Furthermore, a global decrease in FBA has been observed in studies on the aging effect (Choy et al., 2020), memory decline in older and younger adults (Radhakrishnan et al., 2020), neurodegenerative and demyelinating disorders (Al-Amin et al., 2020; Carandini et al., 2021).

## Funding

This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT) Grant CONACYT: A1-S-8430, UNAM-DGAPA-PAPIIT IN201721, and SECITI 2342 to H. Merchant. L. Concha is partially funded by CONACYT (C1782) and UNAM-DGAPA-PAPIIT (IB201712, AG200117, IN204720). Pamela Garcia-Saldivar is a doctoral student from Programa de Maestría y Doctorado en Psicología, Universidad Nacional Autónoma de México (UNAM) and received a fellowship number 414029 from CONACYT through the scholarship holder number 280464. The National Laboratory for MRI is supported by CONACYT and UNAM.



## Acknowledgements

We thank Sonja Kotz and Florencia Assaneo for their fruitful comments on the manuscript. We also thank Luis Prado, Erick Pasaye, Juan Ortiz, Leopoldo González, for their technical assistance.

## Methods

*Participants.* Thirty-two healthy human subjects (age =  $25.37 \pm 3.21$  years; 19 females) without musical training volunteered to participate and gave informed consent, which complied with the Declaration of Helsinki and was approved by our Institutional Review Board. All participants were right-handed and native Spanish speakers. They did not have MRI contraindications nor neurological, psychiatric, or cognitive conditions.

*Apparatus.* Participants were seated comfortably in a quiet experimental room, facing a high-definition 23" touch screen (refresh rate: 60 Hz; ELO Touch solutions) located 50 cm away, which they were instructed to tap using the right index finger. Auditory stimuli were presented through noise-cancelling headphones (Sony, MDR-NC50).

*Experimental Task.* The SCT of the present study is similar to the standard SCT described elsewhere (Merchant et al., 2008c,b) however, instead of tapping a button, the subject tapped on the right or left halves of the touch screen. The task started when three empty white circles (radius 1 degree of visual angle) on a black background were presented simultaneously, forming an inverted isosceles triangle (2 degrees of visual angle on each side). Subjects were trained to place their finger at central-bottom target to start the trial, and to attend to a sequence of six right/left alternating stimuli with a constant interstimulus interval (perception epoch), tapping the touch screen on the corresponding target in synchrony with metronome during six intervals (synchronization epoch), and to continue tapping on the screen for other six intervals without the metronome (continuation epoch) (Figure 1). The mean asynchronies were presented at the end of the trial as feedback. A trial was considered correct when the asynchronies were below 30% of the target interval during the synchronization and below 40% during the continuation epoch. The task was programmed using Psychtoolbox for Matlab (2018, The Mathworks).

*Stimuli.* The visual metronome (33 ms, yellow circle, 0.5 degrees of visual angle) was alternatively displayed inside of the right or left empty circles. The auditory metronome (33 ms, 550 Hz, 80–85 dB SPL), was also presented alternatively on the right or left headphone. The isochronous target intervals were 550, 650, 750, 850, and 950 ms, which were pseudorandomly presented within a block. Each subject performed 10 correct trials for each target interval.

*Procedure.* The MRIs were collected in the first session. In a second visit, outside of the MRI scanner, the auditory and visual tasks were performed randomly across subjects.

*Analysis of behavioral data.* Four metrics were calculated to assess the subjects' performance during SCT (Figure 2A) (Gámez et al., 2018; Merchant et al., 2008c). During the synchronization epoch, we estimated the absolute asynchronies and autocorrelation of the inter-tap interval time series (Iversen et al., 2015; Wing, 2002). The constant error and temporal variability were calculated from the produced intervals during both the synchronization and continuation epochs. Absolute asynchronies were defined as the unsigned difference between stimulus onset and the tap onset. Constant error was calculated as the difference between the produced and the target interval and is a measure of timing accuracy. Temporal variability was defined as the standard deviation of the total produced intervals and is a metric of timing precision. The autocorrelation of the six inter-tap-intervals during a trial was calculated and averaged across trials. Thus, lag 1 autocorrelation is



normally negative for isochrone metronomes, meaning that a longer produced interval tends to be followed by a shorter interval and vice versa, reflecting an error correction mechanism used to not lose the beat of the metronome during synchronization (Iversen et al., 2015). A repeated-measures ANOVA, with two and three factors, was carried out for the analysis of asynchronies, constant error and temporal variability.

*Imaging protocol.* Images were acquired using a 3 T Philips Achieva TX scanner with a 32-channel head coil. T1-weighted volumes were obtained using a three-dimensionally encoded spoiled gradient echo sequence (TR/TE=8.2/3.7 ms, flip angle =  $8^\circ$ , field of view =  $256 \times 256 \times 176$  mm<sup>3</sup>, matrix size =  $256 \times 240 \times 176$  mm<sup>3</sup> yielding voxel resolution =  $1 \times 1 \times 1$  mm<sup>3</sup>). Diffusion-weighted images (DWI). Diffusion-weighted images (DWI) were acquired with echo-planar imaging (EPI),  $2 \times 2 \times 2$  mm<sup>3</sup> voxel resolution, FOV =  $256 \times 256$  mm<sup>3</sup>, 62 axial slices, TR/TE=16500/72 ms. Images were sensitized to diffusion with  $b=1000$  s/mm<sup>2</sup> (64 unique directions) and  $b=3000$  s/mm<sup>2</sup> (96 directions). Five volumes without diffusion weighting ( $b = 0$  s/mm<sup>2</sup>) were also acquired, along with an additional  $b = 0$  s/mm<sup>2</sup> volume obtained with reversed-phase encoding polarity for correction of geometric distortions.

#### *Image processing*

a) Cortical surface. T1-weighted images were used to estimate cortical surfaces. Images were first denoised (Coupe et al., 2008), and corrected for intensity inhomogeneities (Tustison et al., 2011). Binary masks of the brain were obtained with volBRAIN v.2.0 (Manjón & Coupé, 2016). Cortical surfaces were obtained through the FreeSurfer pipeline v.6.0 (Fischl, 2012). Individual surfaces were registered to the surface template with 20,484 vertex (fsaverage5).

b) Diffusion-weighted volumes (DWI) were first denoised (Manjon et al., 2013) and bias field corrected (Tustison et al., 2011), then corrected for geometric distortions and motion using FSL's topup-eddy algorithm (Andersson & Sotiropoulos, 2016).

*Fixel-based analysis (FBA)* We analyzed individual fiber-specific properties in the presence of crossing fiber populations ('fixels') (Raffelt et al., 2015), following the steps described in (Raffelt et al., 2017), using the tools available in MRtrix3 (Tournier et al., 2019). A WM mask was computed for each subject, followed by global signal intensity normalization of the DWI, which was performed across subjects by dividing all volumes by the median  $b = 0$  s/mm<sup>2</sup> intensity. Images were upsampled to 1 mm<sup>3</sup> isotropic resolution. WM Fiber Orientation Distributions (FODs) were estimated using the Multi-Shell Multi-Tissue Constrained Spherical Deconvolution (MSMT-CSD) algorithm (Jeurissen et al., 2014). Tissue-specific response functions were calculated for each subject, from which we derived group-averaged response functions that were used to estimate FODs ( $l_{\max}=8$ ) for each subject. A template of FODs was constructed through iterative nonlinear registration, using the FODs of all 32 subjects followed by the calculation of the intersection of masks of all subjects. Fixels were derived at each voxel by segmentation of the FODs and reoriented to the study template. Finally, FBA metrics (fiber density, fiber cross-section, and fiber density cross-section; FD, FC, and FDC, respectively) were calculated for each fixel. FBA Statistical analysis. A whole-brain probabilistic tractogram was calculated based on the FODs population template, seeded from a whole-brain white matter mask to produce a tractogram of 20 million streamlines. Next, the SIFT algorithm (Smith et al., 2013) was used to select a subset of streamlines ( $n=2$  million) that best fit the diffusion signal and therefore reduce tractography biases. Based on the probabilistic tractography, the structural connectivity metric between fixels was obtained by performing the Connectivity-based fixel enhancement (CFE) tool (Raffelt et al., 2015).

The measures of FD, FC, and FDC were correlated with the performance in the SCT for both conditions (visual and auditory) using a General Linear Model. Family-wise error (FWE)

permutation test was carried out for multiple comparisons correction (Nichols & Holmes, 2001).

Following statistical analysis, the identification of the tracts with significant fixels in the group space was done using the tract-selection regions from XTRACT tool (Warrington et al., 2020) included in FSL software (FMRIB’s Software Library – FSL, Oxford, U.K) (Smith et al., 2004), which were warped into our population template. Tracts with significant fixels were reconstructed using MRTrix (Tournier et al., 2019).

*Superficial White Matter surfaces.* To assess the properties of SWM we used a synthetic representation of axons based on cortical topology. We computed a Laplacian potential field between the ventricles and the GM-WM boundary (Jones et al., 2000; Lerch et al., 2008; Liu et al., 2016) using minclaplace (Lerch et al., 2008). Next, Laplacian-streamlines were seeded at each vertex of the WM surface and propagated towards the ventricles using Matlab, 2020A. Thus, one Laplacian streamline was obtained for every WM surface vertex (Figure 3A). The distance between each step of the Laplacian-streamline was 100  $\mu\text{m}$ , and streamlines were truncated at 5 mm. This resulted in smooth and non-overlapping pathways that respect topology and traverse the superficial white matter. Furthermore, the first segment of these streamlines is perpendicular to the GM-WM surface, with subsequent segments gradually bending as they extend away from it. These aspects make the Laplacian streamlines behave similarly to what is expected from the anatomy of the SWM. Finally, all the WM surfaces and synthetic streamlines were warped to their corresponding subject-specific DWI space via between-modality non-linear registration using ANTS (Avants et al., 2011).

We used the fixel-based information to independently evaluate the two fiber systems that coexist in the SWM: U-fibers subserving short-range cortico-cortical connectivity, and long-range projection, association and commissural fibers (Kirilina et al., 2020; Schüz & Braitenberg, 2002; Yoshino et al., 2020). With the assumption that U-fibers run tangentially to the GM-WM surface, while long-range fibers impinge on the surface in a perpendicular fashion, we attributed the fixels oriented parallel to the segments of the Laplacian streamlines to long-range fibers, and the remaining fixels to the U-fiber system. DWI-derived metrics were sampled along each Laplacian streamline at 0, 0.5, 1, 1.5 and 2 mm under the GM-WM interface. The metrics sampled were  $totalAFD$  (the integral of the FODs within a given voxel), and  $AFD$  attributed to either long-range fibers ( $parAFD$ ; the integral of the FOD of the fixel parallel to the Laplacian streamline), or U-fibers ( $tanAFD$ ; defined as  $totalAFD - parAFD$ ).

All metrics along the Laplacian streamlines were projected onto the GM-WM surface of the fsaverage5 template for visualization and statistical analyses and smoothed using a two-dimensional kernel of 15 mm of full width at half maximum.

*Surface-based analysis of SWM.* Analyses were performed by fitting a general linear model at each vertex. This analysis assessed the relation between the value of diffusion metrics in each vertex ( $i$ ) and the behavioral metrics from the SCT (absolute asynchronies, constant error, temporal variability, and lag 1 of the autocorrelation of the inter-tap-interval time series), as:

$$SWM_i = \beta_0 + \beta_1 * STC_{metric} \quad (1)$$

Surface vertex-wise analysis was corrected for multiple comparisons using random-field theory with a threshold of family-wise error to ( $p_{FWE} < 0.001$ ). Clusters with  $p_{cluster} < 0.001$  were deemed significant.

Significant clusters were anatomically identified using on the Brain Atlas Based on Connectional Architecture (Brainnetome) (Fan et al., 2016). All the analyses were carried out in Surfstat (Worsley et al., 2009) for Matlab (2018; The Mathworks).

## References

- Adab, H., Chalavi, S., Monteiro, T. S., Gooijers, J., Dhollander, T., Mantini, D., & Swinnen, S. P. (2020). Fiber-specific variations in anterior transcallosal white matter structure contribute to age-related differences in motor performance. *NeuroImage*, 209, 116530. URL: <https://doi.org/10.1016/j.neuroimage.2020.116530>. doi:10.1016/j.neuroimage.2020.116530.
- Al-Amin, M. M., Grandjean, J., Klohs, J., & Kim, J. (2020). Disrupted structural connectivity in arca $\beta$  mouse model of a $\beta$  amyloidosis, . (p. 1–26). URL: <https://doi.org/10.1101/2020.04.27.064634>. doi:10.1101/2020.04.27.064634.
- Andersson, J. L., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging. *NeuroImage*, 125, 1063–1078. URL: <http://dx.doi.org/10.1016/j.neuroimage.2015.10.019>. doi:10.1016/j.neuroimage.2015.10.019.
- Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience*, 22, 627–632. URL: <http://dx.doi.org/10.1038/s41593-019-0353-z>. doi:10.1038/s41593-019-0353-z.
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ants similarity metric performance in brain image registration. *NeuroImage*, 54, 2033–2044. URL: <http://dx.doi.org/10.1016/j.neuroimage.2010.09.025>. doi:10.1016/j.neuroimage.2010.09.025.
- Balasubramaniam, R., Haegens, S., Jazayeri, M., Merchant, H., Sternad, D., & Song, J. H. (2021). Neural encoding and representation of time for sensorimotor control and learning. *Journal of Neuroscience*, 41, 866–872. doi:10.1523/JNEUROSCI.1652-20.2020.
- Bartolo, R., & Merchant, H. (2009). Learning and generalization of time production in humans: Rules of transfer across modalities and interval durations. *Experimental Brain Research*, 197, 91–100. doi:10.1007/s00221-009-1895-1.
- Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience*, 34, 3910–3923. doi:10.1523/JNEUROSCI.2679-13.2014.
- Blecher, T., Tal, I., & Ben-Shachar, M. (2016). White matter microstructural properties correlate with sensorimotor synchronization abilities. *NeuroImage*, 138, 1–12. URL: <http://dx.doi.org/10.1016/j.neuroimage.2016.05.022>. doi:10.1016/j.neuroimage.2016.05.022.
- Carandini, T., Mancini, M., Bogdan, I., Rae, C. L., Barritt, A. W., Sethi, A., Harrison, N., Rashid, W., Scarpini, E., Galimberti, D., Bozzali, M., & Cercignani, M. (2021). Disruption of brainstem monoaminergic fibre tracts in multiple sclerosis as a putative mechanism for cognitive fatigue: a fixel-based analysis. *NeuroImage: Clinical*, 30, 102587. URL: <https://doi.org/10.1016/j.nicl.2021.102587>. doi:10.1016/j.nicl.2021.102587.
- Chen, K. M., & Yeh, S. L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica*, 130, 225–234. URL: <http://dx.doi.org/10.1016/j.actpsy.2008.12.008>. doi:10.1016/j.actpsy.2008.12.008.

- Chen, Y., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, 21, 515–532. doi:[10.1016/S0167-9457\(02\)00138-0](https://doi.org/10.1016/S0167-9457(02)00138-0).
- Choy, S. W., Bagarinao, E., Watanabe, H., Ho, E. T. W., Maesawa, S., Mori, D., Hara, K., Kawabata, K., Yoneyama, N., Ohdake, R., Imai, K., Masuda, M., Yokoi, T., Ogura, A., Taoka, T., Koyama, S., Tanabe, H. C., Katsuno, M., Wakabayashi, T., Kuzuya, M., Hoshiyama, M., Isoda, H., Naganawa, S., Ozaki, N., & Sobue, G. (2020). Changes in white matter fiber density and morphology across the adult lifespan: A cross-sectional fixel-based analysis. *Human Brain Mapping*, 41, 3198–3211. doi:[10.1002/hbm.25008](https://doi.org/10.1002/hbm.25008).
- 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, 1–8. doi:[10.3389/fncom.2018.00053](https://doi.org/10.3389/fncom.2018.00053).
- Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., & Barillot, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. *IEEE Transactions on Medical Imaging*, 27, 425–441. doi:[10.1109/TMI.2007.906087](https://doi.org/10.1109/TMI.2007.906087).
- Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *Journal of Neuroscience*, 34, 11972–11983. doi:[10.1523/JNEUROSCI.2177-14.2014](https://doi.org/10.1523/JNEUROSCI.2177-14.2014).
- Dhollander, T., Clemente, A., Singh, M., Boonstra, F., Civier, O., Duque, J. D., Egorova, N., Enticott, P., Fuelscher, I., Gajamange, S., Genc, S., Gottlieb, E., Hyde, C., Imms, P., Kelly, C., Kirkovski, M., Kolbe, S., Liang, X., Malhotra, A., Mito, R., Poudel, G., Silk, T. J., Vaughan, D. N., Zanin, J., Raffelt, D., & Caeyenberghs, K. (2021). Fixel-based analysis of diffusion mri: Methods, applications, challenges and opportunities. *NeuroImage*, 241, 118417. URL: <https://doi.org/10.1016/j.neuroimage.2021.118417>. doi:[10.1016/j.neuroimage.2021.118417](https://doi.org/10.1016/j.neuroimage.2021.118417).
- Drake, C., Jones, M. R., & Baruch, C. (2000a). *The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending* volume 77. doi:[10.1016/S0010-0277\(00\)00106-2](https://doi.org/10.1016/S0010-0277(00)00106-2).
- Drake, C., Penel, A., & Bigand, E. (2000b). Tapping in time with mechanically and expressively performed music. *Music Perception*, 18, 1–23. doi:[10.2307/40285899](https://doi.org/10.2307/40285899).
- Duysens J., S. G., & A., O. (1996). Cortical off response tuning fo stimulus duration. *Vision Research*, 36, 3243–3251.
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The human brainnetome atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex*, 26, 3508–3526. doi:[10.1093/cercor/bhw157](https://doi.org/10.1093/cercor/bhw157).
- Fischl, B. (2012). Freesurfer. *NeuroImage*, 62, 774–781. doi:[10.1016/j.neuroimage.2012.01.021](https://doi.org/10.1016/j.neuroimage.2012.01.021).
- Fraisse, P. (1963). *The psychology of time*. New York: Harper.
- Fraisse, P. (1978). *Time and rhythm perception*. In *Handbook of Perception*. New York: Academic.

- Garcia-Saldivar, P., Garimella, A., Garza-Villarreal, E. A., Mendez, F. A., Concha, L., & Merchant, H. (2021). Preemac: Pipeline for preprocessing and extraction of the macaque brain surface. *NeuroImage*, 227, 117671. URL: <https://doi.org/10.1016/j.neuroimage.2020.117671>. doi:10.1016/j.neuroimage.2020.117671.
- Genc, S., Smith, R. E., Malpas, C. B., Anderson, V., Nicholson, J. M., Efron, D., Sciberras, E., Seal, M. L., & Silk, T. J. (2018). Development of white matter fibre density and morphology over childhood: A longitudinal fixel-based analysis. *NeuroImage*, 183, 666–676. URL: <https://doi.org/10.1016/j.neuroimage.2018.08.043>. doi:10.1016/j.neuroimage.2018.08.043.
- Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, 47, 1894–1903. URL: <http://dx.doi.org/10.1016/j.neuroimage.2009.04.039>. doi:10.1016/j.neuroimage.2009.04.039.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, 29, 7540–7548. doi:10.1523/JNEUROSCI.2018-08.2009.
- Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, 23, 913–921. doi:10.1093/cercor/bhs083.
- Grondin, S., Roussel, M. E., Gamache, P. L., Roy, M., & Quillet, B. (2005). The structure of sensory events and the accuracy of time judgments. *Perception*, 34, 45–58. doi:10.1068/p5369.
- Guevara, M., Guevara, P., Román, C., & Mangin, J. F. (2020). Superficial white matter: A review on the dmri analysis methods and applications. *NeuroImage*, 212. doi:10.1016/j.neuroimage.2020.116673.
- Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology*, 17. doi:10.1371/journal.pbio.3000054.
- Gámez, J., Yc, K., Ayala, Y. A., Dotov, D., Prado, L., & Merchant, H. (2018). Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Annals of the New York Academy of Sciences*, 1423, 396–414. doi:10.1111/nyas.13671.
- Harvey, B. M., Dumoulin, S. O., Fracasso, A., & Paul, J. M. (2020). A network of topographic maps in human association cortex hierarchically transforms visual timing-selective responses. *Current Biology*, 30, 1424–1434.e6. URL: <https://doi.org/10.1016/j.cub.2020.01.090>. doi:10.1016/j.cub.2020.01.090.
- He, J., Hashikawa, T., Ojima, H., & Kinouchi, Y. (1997). Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *Journal of Neuroscience*, 17, 2615–2625. doi:10.1523/jneurosci.17-07-02615.1997.
- Hendrikx, E., Jacob M, P., van Ackooij, M., van der Stoep, N., & Ben M, H. (2022). Visual timing-tuned responses in human association cortices and response dynamics in early visual cortex. *Nature Communications*, 13. doi:10.1038/s41467-022-31675-9.



- Hong, S. J., Hyung, B., Paquola, C., & Bernhardt, B. C. (2019). The superficial white matter in autism and its role in connectivity anomalies and symptom severity. *Cerebral Cortex*, 29, 4415–4425. doi:[10.1093/cercor/bhy321](https://doi.org/10.1093/cercor/bhy321).
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370. doi:[10.1098/rstb.2014.0088](https://doi.org/10.1098/rstb.2014.0088).
- Honing, H., & Merchant, H. (2014). Differences in auditory timing between human and nonhuman primates. *Behavioral and Brain Sciences*, 37, 557–558. doi:[10.1017/S0140525X13004056](https://doi.org/10.1017/S0140525X13004056).
- 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. URL: <http://dx.doi.org/10.1016/j.neuroimage.2012.11.032>. doi:[10.1016/j.neuroimage.2012.11.032](https://doi.org/10.1016/j.neuroimage.2012.11.032).
- Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, 134, 232–244. URL: <http://dx.doi.org/10.1016/j.cognition.2014.10.018>. doi:[10.1016/j.cognition.2014.10.018](https://doi.org/10.1016/j.cognition.2014.10.018).
- Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, 13, 1020–1026. doi:[10.1038/nn.2590](https://doi.org/10.1038/nn.2590).
- Jeurissen, B., Tournier, J. D., Dhollander, T., Connelly, A., & Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. *NeuroImage*, 103, 411–426. URL: <http://dx.doi.org/10.1016/j.neuroimage.2014.07.061>. doi:[10.1016/j.neuroimage.2014.07.061](https://doi.org/10.1016/j.neuroimage.2014.07.061).
- Jones, S. E., Buchbinder, B. R., & Aharon, I. (2000). Three-dimensional mapping of cortical thickness using laplace’s equation. *Human Brain Mapping*, 11, 12–32. doi:[10.1002/1097-0193\(200009\)11:1<12::AID-HBM20>3.0.CO;2-K](https://doi.org/10.1002/1097-0193(200009)11:1<12::AID-HBM20>3.0.CO;2-K).
- Jossinger, S., Sares, A., Zislis, A., Sury, D., Gracco, V., & Ben-Shachar, M. (2022). White matter correlates of sensorimotor synchronization in persistent developmental stuttering. *Journal of Communication Disorders*, 95, 106169. URL: <https://doi.org/10.1016/j.jcomdis.2021.106169>. doi:[10.1016/j.jcomdis.2021.106169](https://doi.org/10.1016/j.jcomdis.2021.106169).
- Kanai, R., Lloyd, H., Buetti, D., & Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. *Experimental Brain Research*, 209, 465–471. doi:[10.1007/s00221-011-2577-3](https://doi.org/10.1007/s00221-011-2577-3).
- Kelly, C. E., Thompson, D. K., Genc, S., Chen, J., Yang, J. Y., Adamson, C., Beare, R., Seal, M. L., Doyle, L. W., Cheong, J. L., & Anderson, P. J. (2020). Long-term development of white matter fibre density and morphology up to 13 years after preterm birth: A fixel-based analysis. *NeuroImage*, 220, 117068. URL: <https://doi.org/10.1016/j.neuroimage.2020.117068>. doi:[10.1016/j.neuroimage.2020.117068](https://doi.org/10.1016/j.neuroimage.2020.117068).
- Kirilina, E., Helbling, S., Morawski, M., Pine, K., Reimann, K., Jankuhn, S., Dinse, J., Deistung, A., Reichenbach, J. R., Trampel, R., Geyer, S., Müller, L., Jakubowski, N., Arendt, T., Bazin, P. L., & Weiskopf, N. (2020). Superficial white matter imaging: Contrast mechanisms and whole-brain in vivo mapping. *Science Advances*, 6. doi:[10.1126/sciadv.aaz9281](https://doi.org/10.1126/sciadv.aaz9281).

- 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: Biological Sciences*, 376. doi:[10.1098/rstb.2020.0325](https://doi.org/10.1098/rstb.2020.0325).
- Lerch, J. P., Carroll, J. B., Dorr, A., Spring, S., Evans, A. C., Hayden, M. R., Sled, J. G., & Henkelman, R. M. (2008). Cortical thickness measured from mri in the yac128 mouse model of huntington's disease. *NeuroImage*, 41, 243–251. doi:[10.1016/j.neuroimage.2008.02.019](https://doi.org/10.1016/j.neuroimage.2008.02.019).
- Liu, M., Bernhardt, B. C., Hong, S. J., Caldaïrou, B., Bernasconi, A., & Bernasconi, N. (2016). The superficial white matter in temporal lobe epilepsy: A key link between structural and functional network disruptions. *Brain*, 139, 2431–2440. doi:[10.1093/brain/aww167](https://doi.org/10.1093/brain/aww167).
- Manjon, V., Concha, L., Buades, A., & Collins, D. L. (2013). Diffusion weighted image denoising using overcomplete local pca, . 8, 1–12. doi:[10.1371/journal.pone.0073021](https://doi.org/10.1371/journal.pone.0073021).
- Manjón, J. V., & Coupé, P. (2016). Volbrain: An online mri brain volumetry system. *Frontiers in Neuroinformatics*, 10, 1–14. doi:[10.3389/fninf.2016.00030](https://doi.org/10.3389/fninf.2016.00030).
- 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, 348–367. doi:[10.1037/0096-3445.135.3.348](https://doi.org/10.1037/0096-3445.135.3.348).
- Mendoza, G., & Merchant, H. (2014). Motor system evolution and the emergence of high cognitive functions. *Progress in Neurobiology*, 122, 73–93. doi:[10.1016/j.pneurobio.2014.09.001](https://doi.org/10.1016/j.pneurobio.2014.09.001).
- Merchant, H., & Averbek, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *Journal of Neuroscience*, 37, 4552–4564. doi:[10.1523/JNEUROSCI.0367-17.2017](https://doi.org/10.1523/JNEUROSCI.0367-17.2017).
- Merchant, H., Bartolo, R., Pérez, O., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2014). Neurophysiology of timing in the hundreds of milliseconds: Multiple layers of neuronal clocks in the medial premotor areas. *Advances in Experimental Medicine and Biology*, 829, 143–154. doi:[10.1007/978-1-4939-1782-2\\_8](https://doi.org/10.1007/978-1-4939-1782-2_8).
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015a). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370. doi:[10.1098/rstb.2014.0093](https://doi.org/10.1098/rstb.2014.0093).
- Merchant, H., Harrington, D. L., & Meck, W. H. (2013a). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience*, 36, 313–336. doi:[10.1146/annurev-neuro-062012-170349](https://doi.org/10.1146/annurev-neuro-062012-170349).
- Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7, 1–8. doi:[10.3389/fnins.2013.00274](https://doi.org/10.3389/fnins.2013.00274).
- Merchant, H., & de Lafuente, V. (2014). *Neurobiology of Interval Timing* volume 829. URL: <http://www.ncbi.nlm.nih.gov/pubmed/25358718>.



- Merchant, H., Luciana, M., Hooper, C., Majestic, S., & Tuite, P. (2008a). Interval timing and parkinson's disease: Heterogeneity in temporal performance. *Experimental Brain Research*, 184, 233–248. doi:[10.1007/s00221-007-1097-7](https://doi.org/10.1007/s00221-007-1097-7).
- Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2015b). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *European Journal of Neuroscience*, 41, 586–602. doi:[10.1111/ejn.12811](https://doi.org/10.1111/ejn.12811).
- Merchant, H., Pérez, O., Zarco, W., & Gámez, J. (2013b). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience*, 33, 9082–9096. doi:[10.1523/JNEUROSCI.5513-12.2013](https://doi.org/10.1523/JNEUROSCI.5513-12.2013).
- Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*, 8, 22–27. URL: <http://dx.doi.org/10.1016/j.cobeha.2016.01.006>. doi:[10.1016/j.cobeha.2016.01.006](https://doi.org/10.1016/j.cobeha.2016.01.006).
- Merchant, H., Zarco, W., Bartolo, R., & Prado, L. (2008b). The context of temporal processing is represented in the multidimensional relationships between timing tasks. *PLoS ONE*, 3. doi:[10.1371/journal.pone.0003169](https://doi.org/10.1371/journal.pone.0003169).
- Merchant, H., Zarco, W., & Prado, L. (2008c). Do we have a common mechanism for measuring time in the hundreds of millisecond range? evidence from multiple-interval timing tasks. *Journal of Neurophysiology*, 99, 939–949. doi:[10.1152/jn.01225.2007](https://doi.org/10.1152/jn.01225.2007).
- Mito, R., Raffelt, D., Dhollander, T., Vaughan, D. N., Tournier, J. D., Salvado, O., Brodtmann, A., Rowe, C. C., Villemagne, V. L., & Connelly, A. (2018). Fibre-specific white matter reductions in alzheimer's disease and mild cognitive impairment. *Brain*, 141, 888–902. doi:[10.1093/brain/awx355](https://doi.org/10.1093/brain/awx355).
- Miyata, K., Yamamoto, T., Fukunaga, M., Sugawara, S., & Sadato, N. (2022). Neural correlates with individual differences in temporal prediction during auditory-motor synchronization. *Cerebral Cortex Communications*, 3. doi:[10.1093/texcom/tgac014](https://doi.org/10.1093/texcom/tgac014).
- Nazeri, A., Chakravarty, M. M., Felsky, D., Lobaugh, N. J., Rajji, T. K., Mulsant, B. H., & Voineskos, A. N. (2013). Alterations of superficial white matter in schizophrenia and relationship to cognitive performance. *Neuropsychopharmacology*, 38, 1954–1962. doi:[10.1038/npp.2013.93](https://doi.org/10.1038/npp.2013.93).
- Nettle, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. *The Origins of Music*, . doi:[10.7551/mitpress/5190.003.0032](https://doi.org/10.7551/mitpress/5190.003.0032).
- Nichols, T. E., & Holmes, A. P. (2001). Nichols-2001.pdf, . 25, 1–25.
- Palomar-García, M. , Hernández, M., Olcina, G., Adrián-Ventura, J., Costumero, V., Miró-Padilla, A., Villar-Rodríguez, E., & Ávila, C. (2020). Auditory and frontal anatomic correlates of pitch discrimination in musicians, non-musicians, and children without musical training. *Brain Structure and Function*, 225, 2735–2744. URL: <https://doi.org/10.1007/s00429-020-02151-1>. doi:[10.1007/s00429-020-02151-1](https://doi.org/10.1007/s00429-020-02151-1).
- 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, 1–14. doi:[10.3389/fnsys.2014.00057](https://doi.org/10.3389/fnsys.2014.00057).

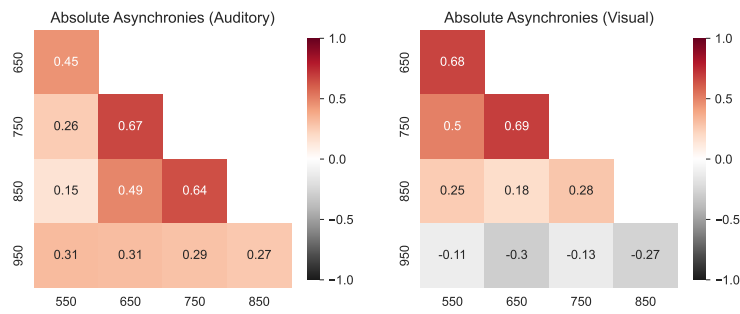
- Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238. doi:[10.1007/s00221-004-2159-8](https://doi.org/10.1007/s00221-004-2159-8).
- Phillips, O. R., Joshi, S. H., Piras, F., Orfei, M. D., Iorio, M., Narr, K. L., Shattuck, D. W., Caltagirone, C., Spalletta, G., & Di Paola, M. (2016). The superficial white matter in alzheimer’s disease. *Human Brain Mapping*, *37*, 1321–1334. doi:[10.1002/hbm.23105](https://doi.org/10.1002/hbm.23105).
- Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, *49*, 961–969. URL: <http://dx.doi.org/10.1016/j.neuropsychologia.2011.02.002>. doi:[10.1016/j.neuropsychologia.2011.02.002](https://doi.org/10.1016/j.neuropsychologia.2011.02.002).
- Protopapa, F., Hayashi, M. J., Kulashekhar, S., Van Der Zwaag, W., Battistella, G., Murray, M. M., Kanai, R., & Buetti, D. (2019). *Research article chronotopic maps in human supplementary motor area* volume 17. doi:[10.1371/journal.pbio.3000026](https://doi.org/10.1371/journal.pbio.3000026).
- Pérez, O., & Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *Journal of Neuroscience*, *38*, 4186–4199. doi:[10.1523/JNEUROSCI.2651-17.2018](https://doi.org/10.1523/JNEUROSCI.2651-17.2018).
- Radhakrishnan, H., Stark, S. M., & Stark, C. E. (2020). Microstructural alterations in hippocampal subfields mediate age-related memory decline in humans. *Frontiers in Aging Neuroscience*, *12*, 1–15. doi:[10.3389/fnagi.2020.00094](https://doi.org/10.3389/fnagi.2020.00094).
- Raffelt, A., D, Smith, R. E., Ridgway, G. R., Tournier, J. D., Vaughan, D. N., Rose, S., Henderson, R., & Connelly, A. (2015). Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion mri measures in the presence of crossing fibres. *NeuroImage*, *117*, 40–55. URL: <http://dx.doi.org/10.1016/j.neuroimage.2015.05.039>. doi:[10.1016/j.neuroimage.2015.05.039](https://doi.org/10.1016/j.neuroimage.2015.05.039).
- Raffelt, D., Tournier, J. D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O., & Connelly, A. (2012). Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, *59*, 3976–3994. URL: <http://dx.doi.org/10.1016/j.neuroimage.2011.10.045>. doi:[10.1016/j.neuroimage.2011.10.045](https://doi.org/10.1016/j.neuroimage.2011.10.045).
- Raffelt, D. A., Tournier, J. D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*, *144*, 58–73. URL: <http://dx.doi.org/10.1016/j.neuroimage.2016.09.029>. doi:[10.1016/j.neuroimage.2016.09.029](https://doi.org/10.1016/j.neuroimage.2016.09.029).
- Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, *4*, 317–323. doi:[10.1038/85191](https://doi.org/10.1038/85191).
- Rau, Y. A., Wang, S. M., Tournier, J. D., Lin, S. H., Lu, C. S., Weng, Y. H., Chen, Y. L., Ng, S. H., Yu, S. W., Wu, Y. M., Tsai, C. C., & Wang, J. J. (2019). A longitudinal fixel-based analysis of white matter alterations in patients with parkinson’s disease. *NeuroImage: Clinical*, *24*, 102098. URL: <https://doi.org/10.1016/j.nicl.2019.102098>. doi:[10.1016/j.nicl.2019.102098](https://doi.org/10.1016/j.nicl.2019.102098).
- Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, *68*, 252–270. doi:[10.1007/s00426-003-0143-8](https://doi.org/10.1007/s00426-003-0143-8).

- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review*, 20, 403–452. doi:[10.3758/s13423-012-0371-2](https://doi.org/10.3758/s13423-012-0371-2).
- Rojas-Vite, G., Coronado-Leija, R., Narvaez-Delgado, O., Ramírez-Manzanares, A., Marroquín, J. L., Noguez-Imm, R., Aranda, M. L., Scherrer, B., Larriva-Sahd, J., & Concha, L. (2019). Histological validation of per-bundle water diffusion metrics within a region of fiber crossing following axonal degeneration. *NeuroImage*, 201. doi:[10.1016/j.neuroimage.2019.116013](https://doi.org/10.1016/j.neuroimage.2019.116013).
- Sayegh, R., Aubie, B., & Faure, P. A. (2011). Duration tuning in the auditory midbrain of echolocating and non-echolocating vertebrates. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 197, 571–583. doi:[10.1007/s00359-011-0627-8](https://doi.org/10.1007/s00359-011-0627-8).
- Scheurich, R., Zamm, A., & Palmer, C. (2018). Tapping into rate flexibility: Musical training facilitates synchronization around spontaneous production rates. *Frontiers in Psychology*, 9, 1–13. doi:[10.3389/fpsyg.2018.00458](https://doi.org/10.3389/fpsyg.2018.00458).
- Schmahmann, J. D., Schmahmann, J., & Pandya, D. (2006). *Fiber pathways of the brain*. Oxford University Press, New York.
- Schüz, A., & Braitenberg, V. (2002). The human cortical white matter: quantitative aspects of cortico-cortical long-range connectivity. In *Cortical Areas* (pp. 389–398). CRC Press.
- Schwartz, M., Rothermich, K., & Kotz, S. A. (2012). Functional dissociation of pre-sma and sma-proper in temporal processing. *NeuroImage*, 60, 290–298. doi:[10.1016/j.neuroimage.2011.11.089](https://doi.org/10.1016/j.neuroimage.2011.11.089).
- Smith, R. E., Tournier, J. D., Calamante, F., & Connelly, A. (2013). Sift: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*, 67, 298–312. URL: <http://dx.doi.org/10.1016/j.neuroimage.2012.11.049>. doi:[10.1016/j.neuroimage.2012.11.049](https://doi.org/10.1016/j.neuroimage.2012.11.049).
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23, 208–219. doi:[10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051).
- Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, 51, 1952–1963. doi:[10.1016/j.neuropsychologia.2013.06.027](https://doi.org/10.1016/j.neuropsychologia.2013.06.027).
- Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: Evidence for a sensitive period. *Journal of Neuroscience*, 33, 1282–1290. doi:[10.1523/JNEUROSCI.3578-12.2013](https://doi.org/10.1523/JNEUROSCI.3578-12.2013).
- Stockert, A., Schwartz, M., Poeppel, D., Anwander, A., & Kotz, S. A. (2021). Temporo-cerebellar connectivity underlies timing constraints in audition. *eLife*, 10, 1–23. doi:[10.7554/eLife.67303](https://doi.org/10.7554/eLife.67303).
- Sánchez-Moncada, I., Concha, L., & Merchant, H. (2020). The timing network is engaged in the practice of internally driven tapping independently of the learning transfer from perceptual to motor timing. *bioRxiv*, (p. 1–31). doi:[10.1101/2020.12.17.423301](https://doi.org/10.1101/2020.12.17.423301).

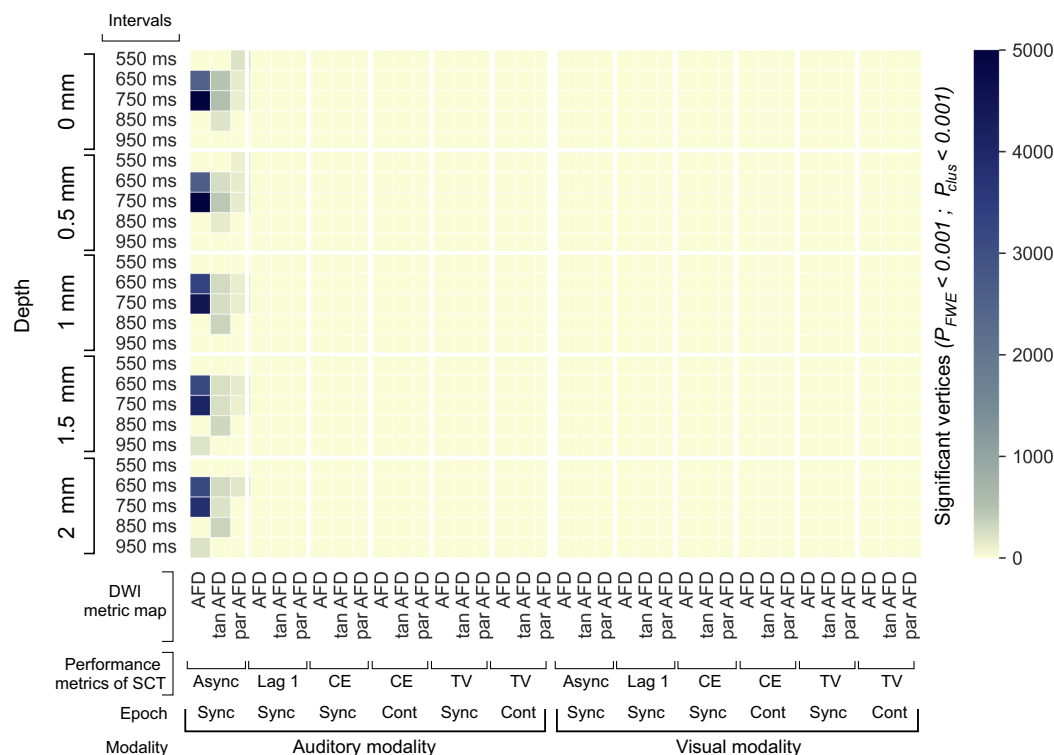
- Tournier, J. D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage*, 23, 1176–1185. doi:[10.1016/j.neuroimage.2004.07.037](https://doi.org/10.1016/j.neuroimage.2004.07.037).
- Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C. H., & Connelly, A. (2019). Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 202, 116137. URL: <https://doi.org/10.1016/j.neuroimage.2019.116137>. doi:[10.1016/j.neuroimage.2019.116137](https://doi.org/10.1016/j.neuroimage.2019.116137).
- Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the beat: A large sample study of bouncing and clapping to music. *PLoS ONE*, 11, 1–19. doi:[10.1371/journal.pone.0160178](https://doi.org/10.1371/journal.pone.0160178).
- Tustison, N. J., Cook, P. A., & Gee, J. C. (2011). N4itk, . 29, 1310–1320. doi:[10.1109/TMI.2010.2046908](https://doi.org/10.1109/TMI.2010.2046908).N4ITK.
- Vaquero, L., Ramos-Escobar, N., François, C., Penhune, V., & Rodríguez-Fornells, A. (2018). White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians. *NeuroImage*, 181, 252–262. doi:[10.1016/j.neuroimage.2018.06.054](https://doi.org/10.1016/j.neuroimage.2018.06.054).
- Warrington, S., Bryant, K. L., Khrapitchev, A. A., Sallet, J., Charquero-Ballester, M., Douaud, G., Jbabdi, S., Mars, R. B., & Sotiropoulos, S. N. (2020). Xtract - standardised protocols for automated tractography in the human and macaque brain. *NeuroImage*, 217, 1–15. doi:[10.1016/j.neuroimage.2020.116923](https://doi.org/10.1016/j.neuroimage.2020.116923).
- Wing (2002). Voluntary timing and brain function: An information processing approach. *Brain and Cognition*, 48, 7–30. doi:[10.1006/brcg.2001.1301](https://doi.org/10.1006/brcg.2001.1301).
- Wing, A. M., & Kristofferson, A. (1973). The timing of interresponse intervals. *Perception & Psychophysics*, 13, 455–460.
- Worsley, K., Taylor, J., Carbonell, F., Chung, M., Duerden, E., Bernhardt, B., Lyttelton, O., Boucher, M., & Evans, A. (2009). Surfstat: A matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *NeuroImage*, 47, S102. URL: [http://dx.doi.org/10.1016/S1053-8119\(09\)70882-1](http://dx.doi.org/10.1016/S1053-8119(09)70882-1). doi:[10.1016/S1053-8119\(09\)70882-1](https://doi.org/10.1016/S1053-8119(09)70882-1).
- Yoshino, M., Saito, K., Kawasaki, K., Horiike, T., Shinmyo, Y., & Kawasaki, H. (2020). The origin and development of subcortical u-fibers in gyrencephalic ferrets. *Molecular Brain*, 13, 1–9. doi:[10.1186/s13041-020-00575-8](https://doi.org/10.1186/s13041-020-00575-8).
- Zamm, A., Wang, Y., & Palmer, C. (2018). Musicians’ natural frequencies of performance display optimal temporal stability. *Journal of Biological Rhythms*, 33, 432–440. doi:[10.1177/0748730418783651](https://doi.org/10.1177/0748730418783651).
- Zamm, A., Wellman, C., & Palmer, C. (2016). Endogenous rhythms influence interpersonal synchrony. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 611–616. doi:[10.1037/xhp0000201](https://doi.org/10.1037/xhp0000201).

- Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, 102, 3191–3202. doi:[10.1152/jn.00066.2009](https://doi.org/10.1152/jn.00066.2009).
- 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, 547–558. doi:[10.1038/nrn2152](https://doi.org/10.1038/nrn2152).

#### 4. Extended data



Extended data 1: Correlation matrix for the Intersubject absolute asynchronies across instructed intervals for the auditory (left) and visual (right) conditions. Note that only the correlation between the intervals of 650 and 750 ms was significant for both the auditory and visual conditions.



Extended data 2: Characterization of the SWM properties associated with the performance of SCT. The Surface-based analysis (SBA) was used to assess the association between the performance metrics of the SCT and the metrics to characterize the SWM sampled in five depths with respect to the grey/white matter interface. The number of significant vertices (both hemispheres) after SBA are shown for the three DWI metric maps, through five depths sampled and the four behavioral metrics, in five intervals, two epochs, and two modalities. The number of vertices for the three AFD metrics showed association only with the absolute asynchronies of the different intervals of the auditory modality. The largest significant clusters were observed between the absolute asynchronies for the 650 and 750 ms intervals through the five depths sampled in the three AFD metrics. The SBA results were corrected to multiple comparisons, using as thresholds  $P_{FWE} < 0.001$  and  $P_{cluster} < 0.001$ . AFD (Apparent Fiber Density), *tanAFD* (Tangential Apparent Fiber Density), *parAFD* (Parallel Apparent Fiber Density). Async (absolute asynchronies), Lag 1 (Lag 1 of the autocorrelation for each instructed interval), CE (Constant Error), TV (Temporal Variability), Sync (Synchronization), Cont (Continuation).



### Extended data 3. Video

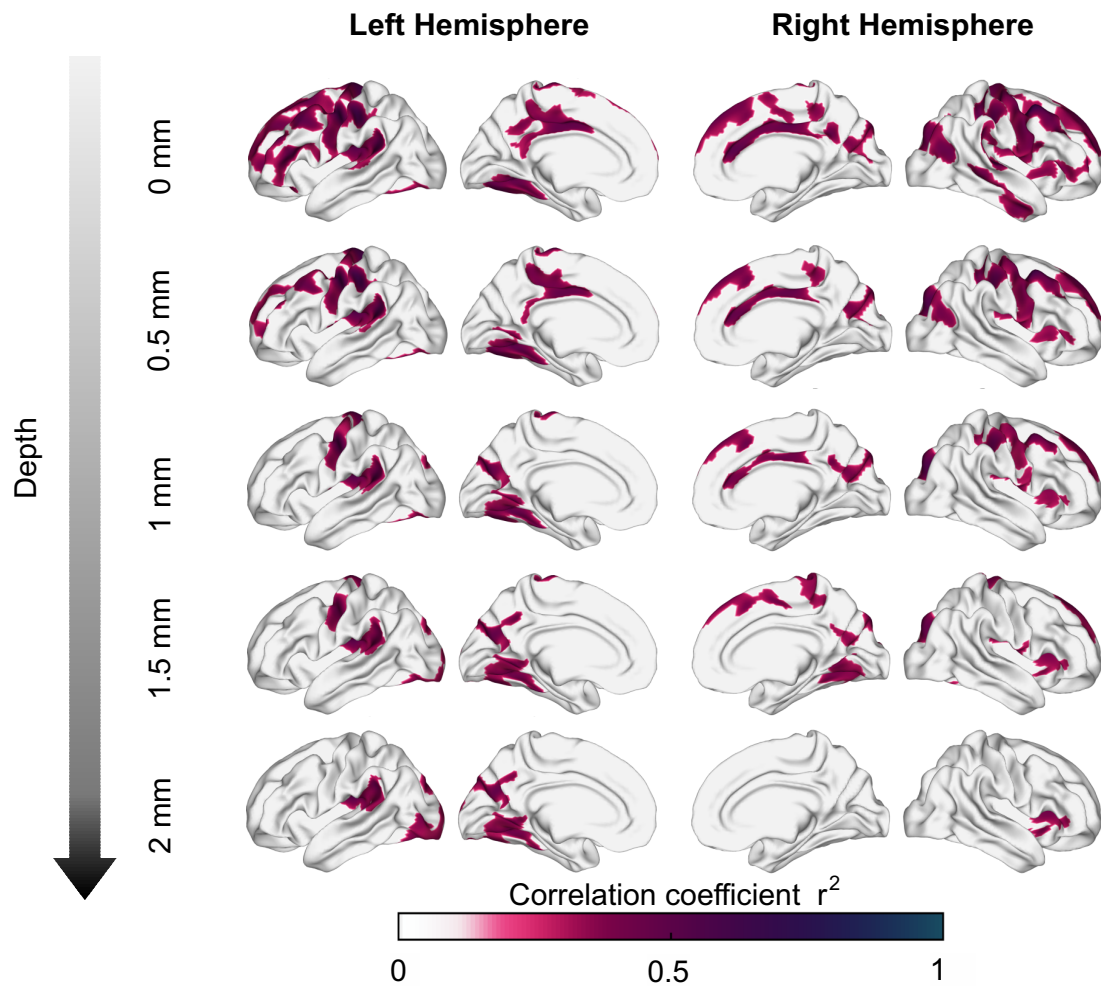
Extended data 3. Spatial correlations between the SWM Apparent Fiber Density and interval-specific absolute asynchronies of the auditory modality. Correlation coefficients ( $r^2$ ) of the significant cluster vertices after surface-based analysis through the depths sampled ( $p_{FWE} < 0.001$  and  $p_{cluster} < 0.001$ ).

### Extended data 4. Video

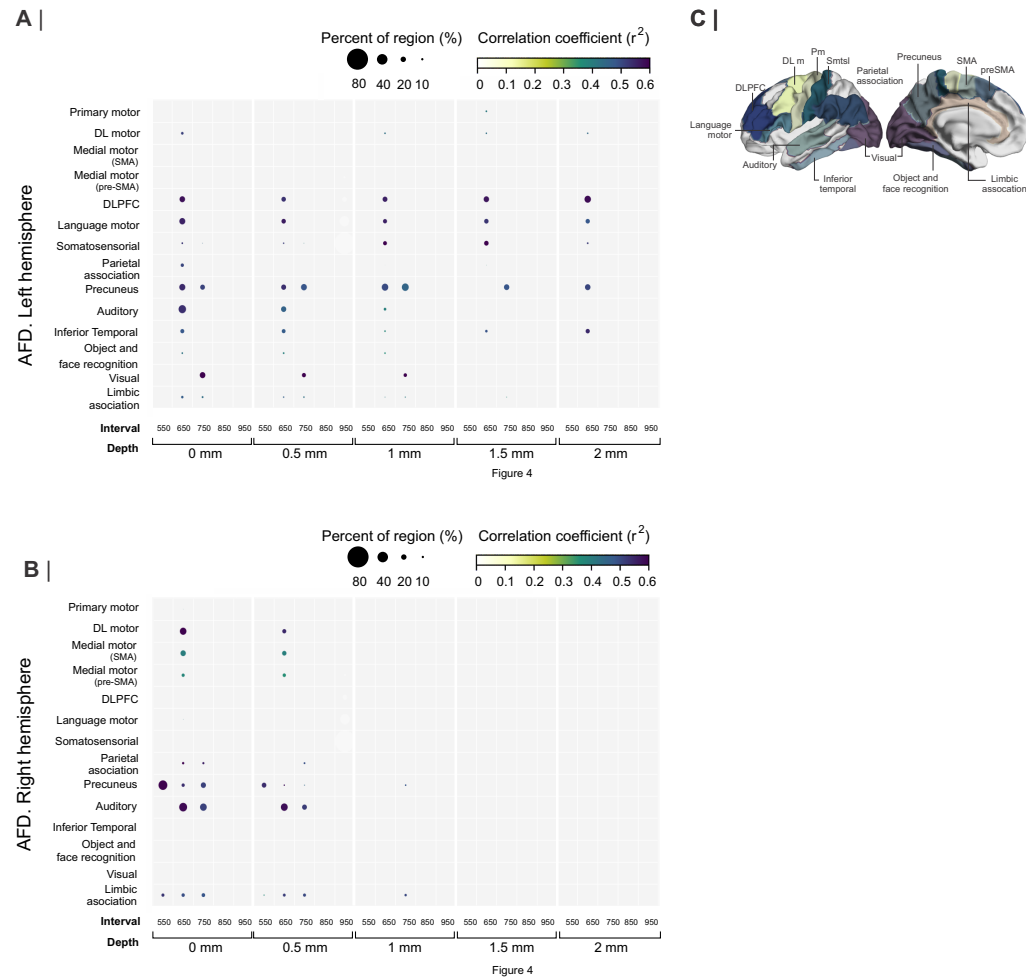
Extended data 4. Spatial correlations between the Tangential SWM Apparent Fiber Density and interval-specific absolute asynchronies of the auditory modality. Correlation coefficients ( $r^2$ ) of the significant cluster vertices after surface-based analysis through the depths sampled ( $p_{FWE} < 0.001$  and  $p_{cluster} < 0.001$ ).

### Extended data 5. Video

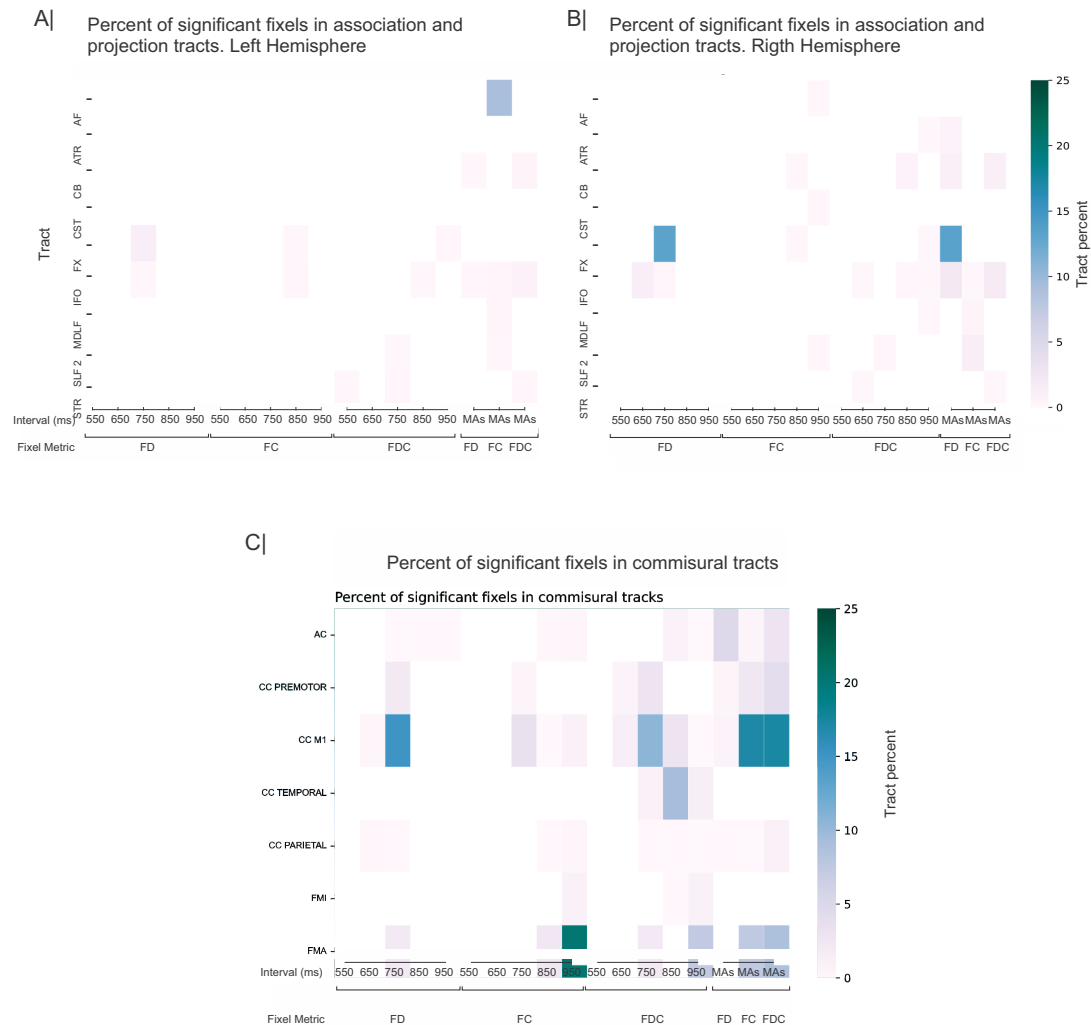
Extended data 5. Spatial correlations between the Parallel SWM Apparent Fiber Density and interval-specific absolute asynchronies of the auditory modality. Correlation coefficients ( $r^2$ ) of the significant cluster vertices after surface-based analysis through the depths sampled ( $p_{FWE} < 0.001$  and  $p_{cluster} < 0.001$ ).



Extended data 6: Association between asynchronies (650 ms intervals) and tangential superficial white matter apparent fiber density at various depths from the GM/WM interface. *tanAFD* sampled at 0 mm shows significant correlations distributed throughout most of the brain, but specific regions (e.g., motor, auditory and visual areas) consistently show significant correlations as depth increases.



Extended data 7: A and B. Interregional correlation plot of the auditory asynchronies and  $parAFD$  as a function of the instructed interval and the depth of the  $parAFD$  calculation.



Extended data 8: Percent of significant fixels in the Deep White Matter in projection tracts (A) Left, (B) Right and in commissural tracts (C). AF= Arcuate fasciculus; ATR = Anterior Thalamic Radiation ; CB = cingulum bundle; CST= Corticospinal Tract ; FX = fornix; IFO= Inferior Longitudinal; MDLF= Middle Longitudinal Fascicle ; SLF 2= Superior Longitudinal Fasciculus 2; STR ; AC= Anterior commissure; CC = Corpus Callosum; FMI= Forceps Minor; FMA= Forceps Major; FD= Fiber Density; FC = Fiber Cross; FDC= Fiber Density and Cross; MAs = Mean Asynchronies.