Revealing the spacetime hierarchical whole-brain dynamics of auditory predictive coding

Bonetti, L.[1,2,3,4*], *Fernández Rubio, G.*[1], *Carlomagno, F.*[1,5], *Pantazis, D.*[6], *Vuust, P.*[1], *Kringelbach, M. L.*[1,2,3]

[1] *Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music, Aarhus, Aalborg, Denmark*

[2] *Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, Oxford, United Kingdom*

[3] *Department of Psychiatry, University of Oxford, Oxford, United Kingdom*

[4] *Department of Psychology, University of Bologna*

[5] *Department of Education, Psychology, Communication, University of Bari Aldo Moro, Italy*

[6] *McGovern Institute for Brain Research, Massachusetts Institute of Technology (MIT), MA 02139, Cambridge, USA*

*Corresponding author: leonardo.bonetti@psych.ox.ac.uk

Abstract

To survive the brain must extract and predict information from key spacetime features of the physical world. While neural processing of visuospatial patterns has been extensively studied, much remains to be discovered about the hierarchical brain mechanisms underlying recognition of auditory sequences with associated prediction errors. We used magnetoencephalography (MEG) to study the temporal unfolding over milliseconds of brain activity in 83 participants recognising melodies and variations thereof. The results showed a hierarchy of processing in networks from the auditory to the ventromedial prefrontal and inferior temporal cortices, hippocampus and medial cingulate gyrus. Both original melodies and variations engaged the pathway from auditory cortex at the bottom of the hierarchy to upstream processing in hippocampus and ventromedial prefrontal cortex, but differed in terms of temporal dynamics, where the recognition of originals elicited stronger gamma power. Our results provide detailed spacetime insights into the hierarchical brain mechanisms underlying auditory sequence recognition.

Keywords

Recognition memory, Temporal sequences, Predictive coding (PC), Magnetoencephalography (MEG), Global neuronal workspace (GNW)

Introduction

In order to fully understand the neural substrate of perception and cognition in the human brain, we must reveal the hierarchical brain processing in both space and time [1–6], as suggested by several frameworks such as the global neuronal workspace hypothesis (GNW) [7–9] and the predictive coding theory [10,11]. To elucidate such brain mechanisms, much research has focused on vision, which is often said to be the key sensory modality for humans [12,13]. However, given that we do not have eyes in the back, auditory information is equally important for survival. In addition, while the visual system primarily relies on the recognition of patterns arranged in space, the auditory system extracts information from patterns and sequences over time [14], providing unique opportunities to understand the temporal hierarchies of the brain.

Decades of studies have clarified that auditory perception is hierarchically organised (originating in the periphery in the cochlea and proceeding, progressively, to the brainstem, pons, trapezoid body, superior olivary complex, lateral lemniscus, inferior, medial geniculate nucleus of the thalamus and finally primary auditory cortex [15,16]). However, little is known about the dynamics of higher-level integration of auditory information. Moreover, much remains to be discovered about the fast-scale, hierarchical brain mechanisms responsible for encoding and recognizing sequences of sounds extended over time.

Here, we took advantage of the unique opportunities offered by music. In fact, music is a highly prized artform providing pleasure and acquiring meaning through the combination of its constituent elements extended over time [17], and exactly for these reasons it provides an excellent tool for investigating the brain's temporal dynamics [18]. However, much remains to be learned about the fine-grained neural dynamics of the auditory system at the milliseconds level since most of the previous studies on music neuroscience have used functional magnetic resonance imaging (fMRI) with relatively poor temporal resolution on the scale of seconds. Still, much progress has been made through clever experimental designs. For instance, Gaab and colleagues observed the brain activity of participants who were requested to compare different simple melodies [19]. Successfully performing the task showed significant changes in activity mainly in the superior temporal, superior parietal, posterior dorsolateral frontal and dorsolateral cerebellar regions, supramarginal and left inferior frontal gyri. In another classic study, Zatorre and colleagues [20] investigated the brain activity related to the perception of melodies and the pitch comparison of particular tones. The results revealed a dissociation where melody perception is related to activity in the right superior temporal cortex, while

pitch comparison mainly involves the right prefrontal cortex. Similarly, a more recent study by Kumar and colleagues [21] showed the key role of the activity and connectivity between primary auditory cortex, inferior frontal gyrus and hippocampus for performing an auditory working memory (WM) task consisting of maintaining a series of single sounds.

These studies revealed the underlying brain networks for music processing but could not provide the precise dynamical unfolding of neural activity. To overcome this issue, here we used magnetoencephalography (MEG), which has excellent temporal resolution capable of tracking rapid brain responses happening at the milliseconds level [22]. For this reason, previous research has utilized MEG to reveal the lower levels of hierarchical processing in the auditory system by investigating the well-known components of the event-related potentials/fields (ERP/F), which occur in response to sounds and violation of expectations, such as the N100, mismatch negativity, and P3a [23–29]. Equally, and even more importantly, MEG allows for the study of higher cognitive processes, providing information on the rapid brain mechanisms associated with perception and manipulation of sounds. As an example, Albouy and colleagues [30] explored the brain activity underlying memory retention, showing that theta oscillations in the dorsal stream of the participants' brain anticipated their abilities to perform an auditory WM task which consisted of manipulating and maintaining sound information. Similarly, Bonetti and colleagues [31–33] revealed that encoding and recognition of sound information recruited a large network of brain areas, spanning from auditory cortex to medial cingulate, inferior temporal cortex, insula, frontal operculum, and hippocampus. Moreover, they showed that music complexity [34] and individual cognitive differences [35] modulated the activity recorded in the brain network. While such studies have provided a first glimpse of the neural basis of music processing and perception of auditory information, the rapid brain hierarchies underlying conscious recognition of temporal sequences and their associated prediction error have not yet been identified.

Here, moving beyond the state-of-the-art, we used source-localised MEG of a group of 83 participants as they recognised original melodies and variations thereof. This revealed the precise spatiotemporal unfolding of brain activity over milliseconds allowing us to map the hierarchical brain dynamics underlying the recognition of previously learned and varied temporal sequences. As such this provided novel insights into the fine-grained hierarchical dynamics of brain processing of spacetime information.

Results

Experimental design and behavioral results

Eighty-three participants performed an old/new auditory recognition task during magnetoencephalography (MEG) recordings. After learning a short musical piece, participants were presented with 135 five-tone musical excerpts lasting 1750 ms each and were requested to state whether each excerpt belonged to the original music ('memorised' sequence (M), old) or it was a varied musical sequence ('novel' sequence (N), new) (**Figure 1a**).

Twenty-seven excerpts were taken from the original musical piece and 108 were variations of the original melodies. We organized these variations in four categories depending on whether changes involved every musical tone of the sequence after the first (NT1), second (NT2), third (NT3) or fourth (NT4) tone (**Figure 1b**). Thus, all the original sequences and variations had the first same tone. **Figure S1** shows a depiction in musical notation of all the sequences used in the study.

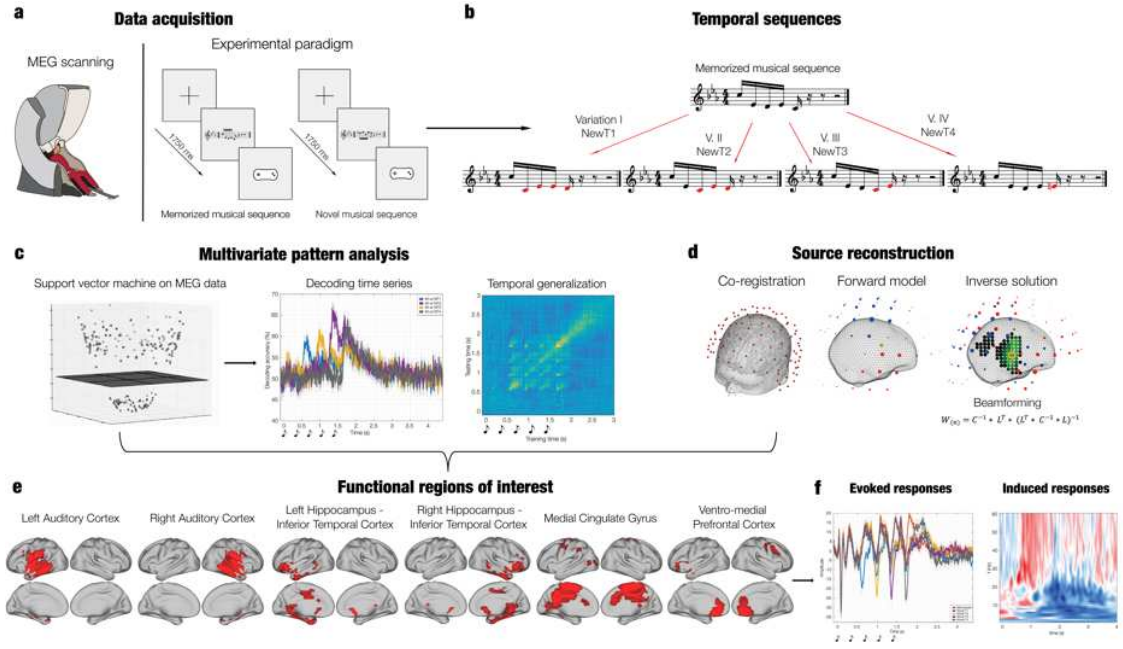


Figure 1. Experimental design, stimuli, and analysis pipeline.

***a** – The brain activity in 83 participants was collected using magnetoencephalography (MEG) while they performed an old/new auditory recognition task. One at a time, five-tone temporal sequences (i.e., melodies) were presented in randomized order and participants were instructed to respond with button presses whether*

*they were 'old' (memorized musical sequences, 'M') or 'new' (novel musical sequences, 'N'). **b** – Five types of temporal sequences (M, NT1, NT2, NT3, NT4) were used in the study. There were three M sequences that comprised the first five tones of the first three measures of the memorised musical piece. The figure shows one of them, as an example (top row). The N sequences were created through systematic variations of each of the three M sequences. This procedure consisted of changing every musical tone of the sequence after the first (NT1), second (NT2), third (NT3) or fourth (NT4) tone, as illustrated in the example reported in the bottom row (red musical tones). **c** – After MEG data pre-processing, multivariate pattern analysis (decoding) was used to assess whether it was possible to discriminate the experimental conditions based on the neural activity recorded with the MEG. **d** – The MEG data was co-registered with the individual anatomical MRI data, and source reconstructed using a beamforming algorithm. This procedure returned one time series for each of the 3559 reconstructed brain sources. **e** – Based on the outcome of the multivariate pattern analysis and of the activity of the source reconstructed data, six main functional brain regions (ROIs) were derived. **f** - We studied the evoked (left plot) and induced (right plot) responses for each ROI and experimental condition.*

Before focusing on the recorded brain data, we performed statistical analyses on the MEG task behavioral data (see **Table 1** for descriptive statistics). We computed two independent Kruskal-Wallis H tests (non-parametric one-way analysis of variance) to assess whether the five categories of temporal sequences (M, NT1, NT2, NT3, and NT4) differed in terms of response accuracy and reaction times (**Figure 2a**).

The Kruskal-Wallis H test for response accuracy was significant ($H(4) = 36.38$, $p < .001$), indicating a difference between categories in the number of correct responses. The Tukey-Kramer correction for multiple comparisons highlighted that NT4 trials were correctly identified with a lower frequency than M ($p = .001$), NT1 ($p = .001$), NT2 ($p = .0003$), NT3 ($p < .0001$).

The Kruskal-Wallis H test for the reaction times was also significant ($H(4) = 22.53$, $p = .0002$).The Tukey-Kramer correction for multiple comparisons highlighted that NT4 trials were correctly identified with a greater reaction time than M ($p = .0016$), NT1 ($p = .0013$), NT2 ($p = .0054$), NT3 ($p = .0008$).

Behavioral variables	M	NT1	NT2	NT3	NT4
Correct recognition	22.33 ± 5.30	22.36 ± 4.27	21.58 ± 5.31	21.66 ± 5.34	17.04 ± 7.12
Reaction times (ms)	2426 ± 226	2407 ± 284	2431 ± 282	2415 ± 272	2578 ± 259

Table 1. MEG task behavioral results show differences between NT4 and all the other conditions.

Mean and standard deviations across participants of number of correctly recognised trials and reaction times (ms) for the five experimental conditions (previously memorised ['M'], novel ['NT1'], novel ['NT2'], novel ['NT3'], novel ['NT4']).

Multivariate pattern analysis

Using a support vector machine (SVM) classifier (see details in Methods, and **Figure 1c**), we performed multivariate pattern analyses to decode different neural activity associated with the recognition of M versus N. Specifically, we computed four independent analyses, decoding M versus each of the four categories of novel sequences (i.e., M versus NT1, M versus NT2, M versus NT3, and M versus NT4).

As shown in **Figure 2b** and **Figure S3**, each of these analyses returned a decoding time series showing how the neural activity differentiated the pair of experimental conditions. Overall, the results showed that the SVM was able to detect significant differences between memorised and novel sequences. Specifically, decoding time series were significantly different from chance level in several time windows. As illustrated in **Figure 2b**, decoding M versus NT1 returned the following main significant time windows: 0.53-0.73 sec; 0.91 – 0.95 sec; 1.27 – 1.30 sec; 1.62 – 1.88 sec (q < .012, false-discovery rate [FDR]-corrected). Decoding M versus NT2 gave rise to the following main significant time windows: 0.89 – 1.18 sec; 1.26 – 1.42 sec; 1.54 – 1.89 sec (q < .012, FDR-corrected). Decoding M versus NT3 returned one main significant time window: 1.25-2.07 sec (q < .012, FDR-corrected). Finally, decoding M versus NT4 showed the following main significant time window: 1.64-2.07 (q < .012, FDR-corrected). Detailed statistical results are reported in **Table S1** and illustrated in **Figure S3**.

To evaluate the persistence of discriminable neural information over time, we used a temporal generalization approach by training the SVM classifier at a given timepoint t and testing it across all timepoints. This was calculated for the four pairs of experimental conditions described above. The signed-rank test against chance level and cluster-based Monte-Carlo simulations [42-44, 46, 47] (MCS; α = .01, MCS *p-value* = .001) showed that the performance of the classifier was significantly above chance even a few hundreds of milliseconds beyond the diagonal, for all pairs of conditions. Notably, the neural difference between M and N was comparable across diverse musical tones, as shown by the recurrent patterns depicted in **Figure 2c** and highlighted by the graphs and stars. Detailed statistical results are reported in **Table S2**.

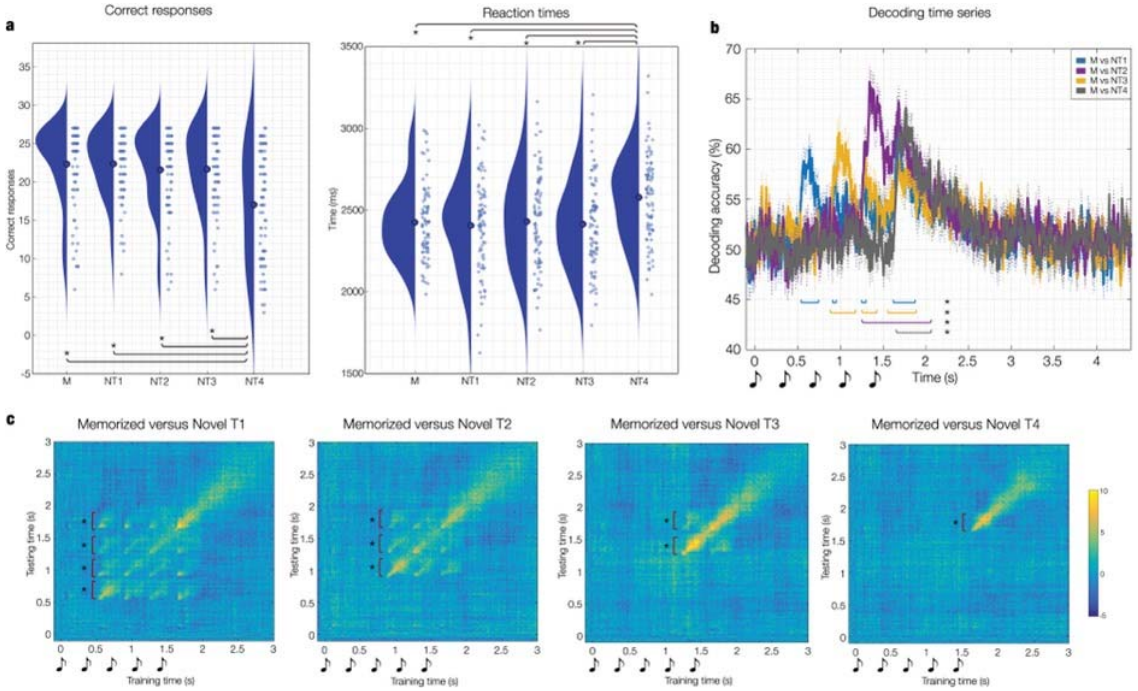


Figure 2. Machine learning allows for accurate decoding of experimental conditions.

a *– Scatter and violin plots for the correct responses (left) and reaction times (right) for the experimental task performed by the participants in the MEG. The plots separately illustrate each experimental condition. Each dot represents a participant. The graphs and stars indicate that both accuracy and reaction times for NT4 were significantly different ($p < .01$) from M, NT1, NT2, and NT3.* ***b*** *- Multivariate pattern analysis decoding the different neural activity associated with memorised versus novel musical sequences. The plot shows the decoding time series for four rounds of pairwise decoding (M vs NT1, M vs NT2, M vs NT3, M vs NT4). The sketch of the musical tones represents the onset of the sounds forming the temporal sequences. The graphs and stars indicate the significant time-points when the algorithm successfully decoded the experimental conditions ($q < .012$, false-discovery rate [FDR]-corrected)* ***c*** *- Temporal generalization of pairwise decoding of the same conditions. The graphs and stars indicate the significant time-points when the experimental conditions were successfully decoded (MCS; $\alpha = .01$, MCS p-value = .001). The sketch of the musical tones represents the onset of the sounds forming the temporal sequences.*

Neural sources of the differential brain activity between M and N

We employed a local-spheres forward model and a beamforming approach as inverse solution in an 8-mm grid (**Figure 1d**). To detect the brain sources underlying the differential signal

observed for M and N, we considered both the results from the decoding analyses and the inspection of the MEG data after preprocessing. This was necessary since decoding can capture differences between conditions, but it does not discriminate which condition is associated with a larger neural signal (see Methods for details).

This procedure returned the following significant time windows: 0.50 – 0.60 sec; 0.70 – 0.80 sec; 0.98 – 1.02 sec; 1.05 – 1.15 sec; 1.33 – 1.39 sec; 1.45 – 1.55 sec; 1.70 – 1.75 sec; 1.75 – 1.85 sec. For each time window and condition, we averaged the time series of all brain sources over time and computed t-tests contrasting each combination of M versus Ns. Finally, we corrected for multiple comparisons using a 3D cluster-based MCS (MCS, $\alpha = .003$, MCS *p-value* = .001).

As reported in **Table 2** and illustrated in **Figure S4**, this analysis returned several significant clusters of differential brain activity between M and N, primarily located in the bilateral medial cingulate gyrus (MC), left (HITL) and right hippocampal area and inferior temporal cortex (HITR), left (ACL) and right auditory cortex (ACR), and bilateral ventromedial prefrontal cortex (VMPFC).

Detailed statistical results are extensively reported in **Table S3**.

Time-windows	Contrast	Cluster size	Main ROI(s)	Peak t-value	MCS p-value
0.50 - 0.60	M vs NT1	10	MC	-4.18	< .001
0.70 - 0.80	M vs NT1	132	MC	5.59	< .001
1.05 - 1.15	M vs NT1	162	MC	5.38	< .001
1.45 – 1.55	M vs NT1	270	ACR MC	4.91	< .001
1.75 – 1.85	M vs NT1	175	HITR	5.94	< .001
0.98 – 1.02	M vs NT2	94	VMPFC	-4.33	< .001
1.05 – 1.15	M vs NT2	57	MC	4.47	< .001
1.45 – 1.55	M vs NT2	9	MC	3.7	< .001
1.75 – 1.85	M vs NT2	28	HITR	4.16	< .001
1.33 – 1.39	M vs NT3	300	HITR VMPFC	-5.34	< .001
1.45 – 1.55	M vs NT3	58	ACL	-4.57	< .001
1.75 – 1.85	M vs NT3	79	HITR	4.42	< .001
1.70 – 1.75	M vs NT4	584	VMPFC	-5.57	< .001
1.75 – 1.85	M vs NT4	197	ACL	-5.39	< .001

Table 2. Brain sources of decoding significant time windows.

Significant clusters of brain sources of the significant time windows returned by the decoding analysis. The table shows the time windows, contrasts, cluster size (number of voxels forming the cluster), main ROI(s) involved in the cluster, peak t-value, and MCS p-value.

Functional regions of interests (ROIs)

Expanding on the previous analyses, we wished to better define the spatial extent of the significant brain areas based on their functional profile and computed their associated time series (see Methods for details).

This procedure returned six broad ROIs which were particularly active during the recognition task: left *(i)* and right auditory cortex *(ii)*, left *(iii)* and right hippocampus-inferior temporal cortex *(iv)*, ventromedial prefrontal cortex *(v)*, and medial cingulate gyrus *(vi)*. Then, we computed one t-test for each time-point and each combination of M versus Ns. Finally, we corrected for multiple comparisons using one-dimensional (1D) cluster-based MCS (MCS, $\alpha = .05$, MCS *p-value* = .001).

This analysis returned several significant clusters of differential brain activity over time between M and Ns. As shown in **Figure 3**, M versus N was characterized by stronger activity in VMPFC, ACL, and HITR after 350 – 450 ms from the onset of each tone. Similarly, M presented stronger negative activity ($p < .001$) than N in the MC after 400 – 500 ms from the onset of each tone.

Contrast	ROI	Temporal extent of the largest clusters from the 1st tone of the sequence	Peak t-value	P-value
Positive activity				
M vs NT_1	VMPFC	900 – 1190	6.13	< .001
	ACL	900 – 1040	6.26	< .001
	HITR	1660 – 1930	5.45	< .001
M vs NT_2	VMPFC	1680 – 1930	5.94	< .001
	ACL	1260 – 1390	3.56	< .001
	HITR	1640 – 1930	6.12	< .001
M vs NT_3	VMPFC	1310 – 1540	7.04	< .001
	ACL	1250 – 1390	4.80	< .001
	HITR	1310 – 1530	6.69	< .001
M vs NT_4	VMPFC	1680 – 1910	8.92	< .001
	ACL	1640 – 1790	6.93	< .001
	HITR	1680 – 1900	7.45	< .001

Negative activity				
M vs NT1	MC	680 – 860	-5.18	< .001
M vs NT2	MC	980 – 1180	-5.77	< .001
M vs NT3	MC	1360 – 1550	-5.41	< .001
M vs NT4	MC	1730 – 1920	-5.40	< .001

Table 3. Largest clusters of stronger activity of M versus Ns.

Largest clusters of significantly stronger activity of M versus Ns computed for the six ROIs considered in the study. The table shows the contrast, the correspondent ROI, the temporal extent (in ms) of the largest cluster, the peak t-value of the cluster and the associated MCS p-value. The MC shows stronger negativity since the polarity of the MC signal was negative. All clusters are reported in detail in ***Table S4****.*

Conversely, late N100 responses localized in AC were stronger for N versus M. For instance, the temporal extent of the larger cluster of ACL for M versus NT1 was 520 – 700 ms from the onset of the first tone of the sequence (peak *t-value* = 6.75, $p < .001$). Moreover, HIT and VMPFC showed a stronger response for N versus M occurring about 250 – 300 ms after altering the original sequences. For instance, the temporal extent of the larger cluster of HITR for M versus NT4 was1680 – 1900 ms from the onset of the first tone of the sequence (peak *t-value* = 7.45, $p < .001$); the temporal extent of the larger cluster of VMPFC for M versus NT4 was 1680 – 1910 ms from the onset of the first tone of the sequence (peak *t-value* = 8.92, $p < .001$). **Table 3** reports the main significant clusters, while complete statistical results are reported in **Table S4**.

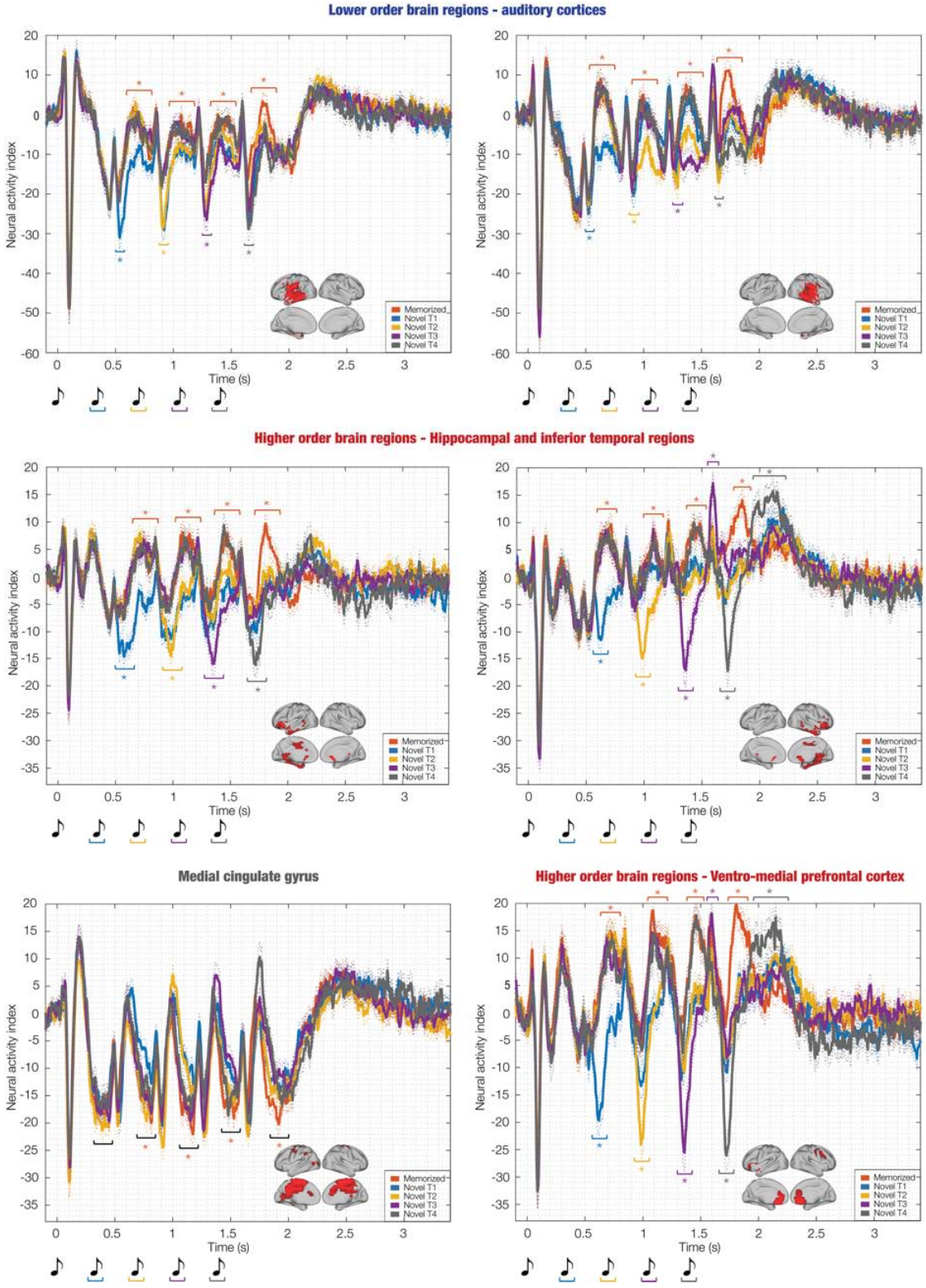


Figure 3. Revealing the hierarchy of temporal sequence recognition in the human brain.

Source localized brain activity illustrated for each experimental condition (M, NT1, NT2, NT3, NT4) and ROI.

The sketch of the musical tones represents the onset of the sounds forming the temporal sequences. The brain templates illustrate the spatial extent of the ROIs. Graphs and star indicate significant differences. The red ones refer to the contrast 'M' versus 'NT1', the blue to 'NT1' versus 'M', the yellow to 'NT2' versus 'M', the purple to 'NT3' versus 'M' and finally the grey to 'NT4' versus 'M'.

The first row shows the lower order brain regions (auditory cortices) involved in this study. The bottom graphs highlight the fast N100 responses to each tone of the sequences, which were stronger for the novel conditions. The top graphs indicate the stronger and slower responses for the memorised sequences. The second row and the right plot of the third row illustrate higher-order brain regions, corresponding to the hippocampal area and inferior-temporal cortices, and ventromedial prefrontal cortex. Also in this case, the bottom graphs indicate the stronger responses for the novel conditions, while the top graphs the stronger activity recorded for the memorised sequences. To be noted, the time series systematically change depending on the variation introduced in the melodic sequences. Moreover, the responses of these higher-order brain regions are slightly delayed (about 80 ms) compared to the lower-order regions. This suggests a hierarchical processing happening in the brain to extract meaning form the musical sequences and recognise them. Finally, the left plot of the third row indicates the medial cingulate gyrus, which presents an overall similar activity in response to each experimental condition. This suggests that the medial cingulate may be implicated in the general auditory processing and not specifically in the memory recognition and evaluation of the temporal sequences.

Of particular interest is the brain response occurring when the original musical sequence was varied. Here, AC presented a rapid, sharp signal (150 ms after the altered tone), while HIT and VMPFC responded slightly later (around 200-250 ms from the altered tone). As an example, **Figure 4b, 4c** and **S5** depicts the cross-correlation between the time series of left AC and VMPFC averaged over participants, highlighting the 80 ms lag between the two signals.

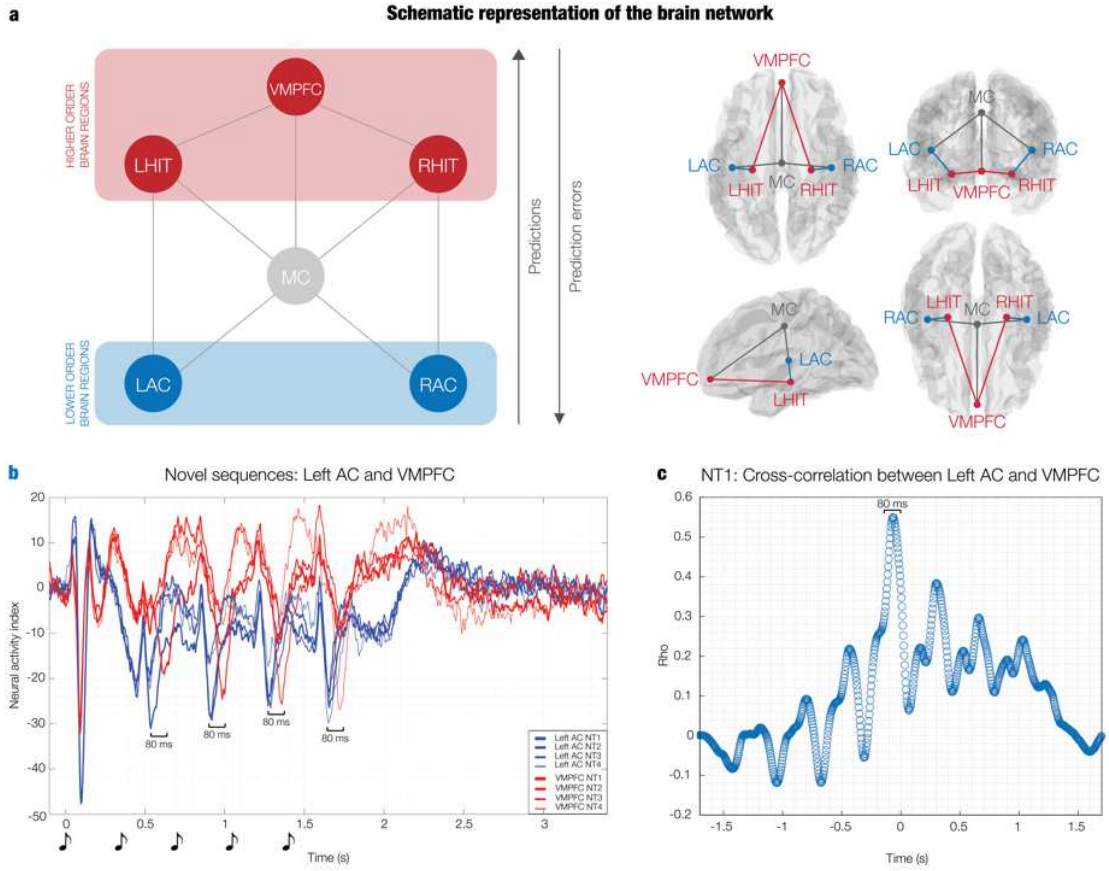


Figure 4. Brain network underlying temporal sequence recognition.

a *– Schematic representation (left) of the brain network underlying the recognition of the temporal sequences. The schema illustrates the lower-order brain regions at the bottom of the hierarchy (left and right auditory cortex) and the higher-order brain regions (left and right hippocampus and inferior-temporal cortices, and ventromedial prefrontal cortex) at the top of the hierarchical brain processing. In the middle, the medial cingulate gyrus is thought to play an associative role, possibly mediating between lower- and higher-order brain regions. The right plots show the same representation within brain templates.* ***b*** *- The left plot shows the source localized brain activity illustrated for each category of N (i.e., NT1, NT2, NT3, NT4) and both left auditory cortex (AC) and ventromedial prefrontal cortex (VMPFC). Of particular interest it is the sharp peak occurring after the onset of each tone where left AC precedes VMPFC of about 80 ms. Moreover, while the strength of the signal increases over time for VMPFC, this does not happen for left AC. This evidence suggests a hierarchical brain processing underlying the prediction error occurring for the novel conditions. The sketch of the musical tones below the first two plots represents the onset of the sounds forming the temporal sequences.* ***c*** *– The plot shows that the strongest cross-correlation computed between the time series averaged over participants of left AC and VMPFC occurred with a time-lag of approximately 80 ms.*

Time-frequency analysis for induced responses

We computed time-frequency analysis using complex Morlet wavelet transform (from 1 to 60 Hz with 1-Hz intervals). This analysis was conducted for induced responses. First, we estimated independent time-frequency decompositions for each voxel of the six ROIs and for each trial. Then, the computed power spectra were averaged over voxels within each ROI and over trials. Finally, in line with the previous analyses, we calculated four contrasts (M versus NT1, M versus NT2, M versus NT3, and M versus NT4). Specifically, we computed a t-test for each frequency and timepoint and corrected for multiple comparisons using 2D cluster-based MCS. As shown in **Figure 5**, results were similar across ROIs and displayed a generalized increased gamma power for M versus N ($p < .001$). Conversely, N versus M presented a stronger power between 2 and 20 Hz (corresponding to, approximately, theta, alpha, and beta bands), in the time window 1.0 – 3.0 seconds ($p < .001$). Detailed statistical results about this procedure are extensively reported in **Table S5** and depicted in **Figure S6**.

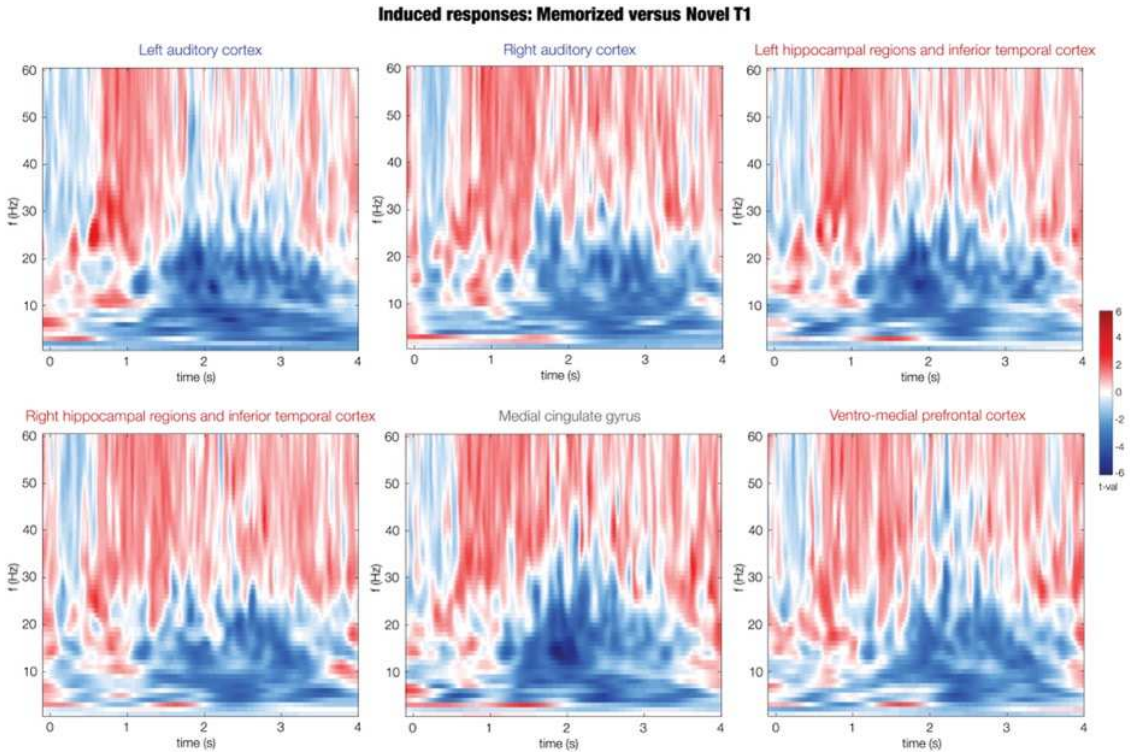


Figure 5. Significant differences in gamma power when recognising melodies compared to variations.

Contrasts between the source localized induced responses of M versus NT1. The plots indicate a stronger power for gamma in M. Moreover, theta was overall stronger for M versus NT1 during the presentation of the sounds,

while alpha, beta and theta were stronger for NT1 versus M after the end of the temporal sequences. The colorbar indicates the t-values obtained by contrasting M versus NT1.

Discussion

Using a music recognition paradigm in MEG, we were able to reveal the fine-grained hierarchical spacetime brain dynamics underlying the recognition of previously memorised musical sequences, as well as the prediction error arising when varying these auditory sequences.

The results show that the recognition of the auditory sequence is built over time through a rapid hierarchical pathway progressing from the auditory cortex to the ventromedial prefrontal cortex, hippocampus, and inferior temporal cortex. This provides important evidence strengthening the case made by other studies suggesting a hierarchical organization of the brain in terms of predictive coding [1–4]. Importantly, the results reported here expands the previous literature by providing novel information on the spatiotemporal dynamics of the hierarchical organization of the auditory system linked to crucial memory processes [11,14,30,36].

More specifically, we observed a hierarchical pathway which characterized the recognition of the original sequences. Initially, each tone of the sequence showed the expected well-known N100 negative peak after 100 ms in the auditory cortex. Notably, we also found a novel positive component peaking after 300 ms. These peaks were followed by a similar component originated in ventromedial prefrontal cortex, hippocampus and inferior temporal cortex which achieved a maximum 400 ms after each tone. In addition, the medial cingulate presented a negative peak 400 – 450 ms after each tone. The variations could then be characterised in terms of how these components were either absent or strongly reduced. Interestingly, while the response of the medial cingulate was the same to all notes, the activity of ventromedial prefrontal cortex, hippocampus and inferior temporal cortex increased over time only for the original sequences.

In addition, our results showed a different hierarchical pathway for the prediction error arising in the responses to the systematic variations of the sequences. Here, compared to the original sequences, the variations were associated with a stronger N100 response in the auditory cortex for each tone. In contrast, the later responses in the ventromedial prefrontal cortex and hippocampus occurred only after the tone that disrupted the original sequences (see for example the response in Figure 3 for [NT1], i.e. after tone two of the melody). Here, a first sharp negative peak appeared 250 – 300 ms after the tone altering the sequence, while a second positive component peaked 500 – 550 ms after the tone. The strength of the responses was progressively greater depending on the tone used to introduce the variation

(i.e., a variation in tone five [NT4] elicited stronger responses than a variation in tone two [NT1]).

The induced-response analysis showed that activity in the gamma band was stronger for the original sequences compared to variations, especially during the presentation of the sequences. Conversely, after the end of the sequence, activity in alpha and beta bands was stronger for variations compared to originals. These results were very similar to those found across the brain network highlighted by the evoked-response analyses.

Finally, in terms of behavioural responses, the accuracy and reaction times were nearly equal for the original sequences and the systematic variations (NT1, NT2, NT3). Notably, when varying only the last tone of the sequence (NT4), the reaction times were suddenly significantly larger and the accuracy lower.

These results revealed the hierarchical brain processing occurring when listening to auditory sequences and the spacetime differences in hierarchy when listening to systematic variations thereof. Broadly, they are consistent with previous findings [31,32,34,35], showing a large network of brain areas involved in the recognition of previously memorised auditory sequences.

Moreover, our findings support previous studies which showed the auditory cortex as the primary neural core for processing auditory information [36,37]. Interestingly, although the stimuli used in this study were musical, the recognition of the temporal sequences required the involvement of the inferior temporal cortex and hippocampus, which were previously associated mainly to linguistic elaborations [37,38] and abstract memory [39–41].

Similarly, we found that the recognition of the auditory sequences required the involvement of the ventromedial prefrontal cortex, whose role in auditory processing has not been clearly established yet. Although it is not possible to make definite claims at this stage, we argue that the ventromedial prefrontal cortex, together with the hippocampus, may track the progress of the sequence to make the evaluation about the sounds. In this view, it may be the primary responsible of extracting the meaning of the sequence.

Another relevant brain region playing a prominent role at the top of the hierarchy is the medial cingulate gyrus, which presented a negative peak 400 – 450 ms after each tone. The activity was overall stronger for original sequences versus variations, but the underlying activity remained there after the sequences were altered. Moreover, the strength did not increase over time. This constant response strongly suggests that the medial cingulate is unlikely to be involved in the recognition of the sequences. Instead, in accordance with

previous research [42], this brain region may be required for the orchestration of allocating attentional resources. It could also be argued that the medial cingulate might be specifically involved in auditory processing, joining and unifying the perception of musical tones, and providing the perception of the musical sequences as a whole pattern rather than a (disjointed) sequential group of single tones. This interpretation would be coherent with the specific involvement of medial cingulate in a variety of musical tasks, as highlighted by two recent meta-analyses [43,44], but it calls for future studies specifically designed to test this hypothesis.

Interestingly, our findings provide new insights into the results previously reported [31,32,34,35] on the frequency of evoked responses to melodies. In fact, those studies showed that the recognition of the temporal sequences is associated with activity in a network including the cingulate, hippocampus, insula, and inferior temporal cortex in a very slow frequency (0.1 – 1 Hz) accompanying the whole duration of the sequences. Here, however, we observed a distinct, faster response to each tone of the sequence. The frequency of these responses was of approximately 1 – 1.5 Hz. This new finding clearly shows that the recognition of the sequence is associated with responses to each of the musical tones. As such, it could suggest that the slow signal previously observed [31,32,34,35] could simply be the summation of the responses to each sound of the sequences. Lending support to this interpretation is the fact that the duration of each tone was 250 ms in these previous studies rather than the 350 ms used here, and so the individual responses were closer in time and therefore added up to build a slower and stronger signal (0.1 – 1 Hz). To further understand this effect, future studies could modulate the speed of the musical sequences to detect whether 1 – 1.5 Hz is a specific rhythm of the brain associated to recognition or whether it is driven by the speed of the stimuli.

Overall, our results on temporal sequences recognition are consistent and provide insights into the GNW hypothesis proposed by Changeux and Dehaene [7,9]. The authors hypothesized that processing privileged categories of stimuli such as meaningful temporal sequences activate the brain areas comprised in the GNW. The GNW was defined as a network of brain areas responsible for consciously processing information in terms of attention, memory, and valence, and subsequently for making it available to the whole brain [7,9]. As hypothesised, in our study the recognition of memorised versus novel musical sequences led to stronger activity in putative regions of the GNW such as the cingulate gyrus, hippocampus, and ventromedial prefrontal cortex [5]. These areas could be necessary to extract a meaningful

representation of the sequences and match them with previously acquired memory traces. Importantly, our study also potentially expands the theory of the GNW by revealing the fine-grained hierarchical dynamics between the underlying brain regions while participants recognised memorised and novel temporal sequences.

Our findings are also providing support for the well-known predictive coding theory [10,11], which states that the brain is constantly updating internal models to predict external information and stimuli. Recently, this theory has been connected to complex cognitive processes, finding a remarkable example in the neuroscience of music [11,45]. Both Vuust and colleagues [45] and Koelsch and colleagues [11] suggested that, while processing music, the brain constantly generates hypothesis about the upcoming development of musical sentences. Our results support and expand predictive coding theory for the recognition of both previously memorised and novel sequences in terms of identifying the underlying hierarchical processing. On the one hand, when the upcoming sound was matched with the predicted sound based on the previously stored memory trace, first the auditory cortex and then hippocampus, inferior temporal cortex and ventromedial prefrontal cortex respond. These responses increase over time, showing stronger neural activity after each successful prediction of the upcoming sounds. On the other hand, the present study revealed the changes in hierarchical processing associated with prediction errors when the melodies were systematically altered. This indicates that when the upcoming sound was incoherent with the prediction made by the brain, a network of hierarchical areas was recruited, with the information flowing from auditory cortex to ventromedial prefrontal cortex and hippocampal regions. Notably, this brain network was similar to the one employed for the recognition of previously memorised sequences, but their temporal dynamics sharply differed.

This latter finding is also coherent with the plethora of studies investigating automatic prediction error in audition indexed by N100 and mismatch negativity (MMN) [24,26,29,46]. Previous research has revealed the primary involvement of auditory cortex in the generation of the prediction error signal [46,47], reporting a complementary yet much reduced recruitment of the medial cingulate, frontal and hippocampal areas [28,48]. Conversely, in our study we investigated the prediction error and revealed the hierarchical organization of the brain which recruited first auditory cortex (100 – 150 ms) and then, with a stronger activity, the ventromedial prefrontal cortex and hippocampus (250 – 500 ms). Moreover, our results showed that auditory cortex discriminated melodies versus systematic variations but did not distinguish the strength of the errors (i.e., errors happening later in the sequence). For

instance, the response to the variation inserted at tone number two was the same to the variation at tone number five. In addition, the auditory cortex responded to novel tones for the entire duration of the sequence in a similar manner (i.e., for NT1, the auditory cortex responded in the same way to tones number two, three, four, and five). Remarkably, the prediction error observed in the hippocampal regions and ventromedial prefrontal cortex showed a different strength depending on when the sequence was altered (e.g., the response to the variation inserted at tone number five was much stronger than to the variation at tone number two). Moreover, these areas responded primarily to the first tone where we introduced the variation and very little to the subsequent sounds. These findings suggest that the brain signature underlying the awareness of the variation may be represented by the responses recorded in the ventromedial prefrontal cortex and hippocampus and their specific temporal dynamics.

Along this line, our findings showed a potential relationship between reaction times and accuracy in the recognition task and the second response of the prediction error occurring in the right hippocampus and inferior temporal cortex and in the ventromedial prefrontal cortex. Indeed, both reaction times and accuracy were approximately the same for original sequences and NT1, NT2, and NT3. However, accuracy was significantly reduced, and reaction times increased for NT4. Similarly, while the second component of the prediction errors (occurring after approximately 500 ms from the onset of the varied tone) was rather sharp for NT1, NT2, and NT3, its frequency was much slower for NT4. There are at least two possible explanations for this phenomenon: 1) The variation of the last tone of the sequence elicited a slower prediction error, both at a neural and behavioral level or, alternatively, 2) a bolder and more intriguing hypothesis relates to musical chunking and the beat used to present the stimuli. We would argue that the lower accuracy and higher reaction times for NT4 was due to the chunking occurring when listening to the musical stimuli presented with a beat every four tones. In this view, after listening to four tones of the original sequence (corresponding to a full beat), the perception that the sequence belonged to the group of previously learned sequences was very strong and, especially much stronger than after only three tones. For this reason, we did not observe a linear increase in reaction times and accuracy but only a strong difference between all categories and NT4 and a much slower prediction error only for NT4 compared to the other categories of N. Currently, we do not have enough data to make definitive claims and future studies are needed where the length of the sequences is systematically varied. This could, for example, be achieved by having sequences with a beat

every four tones and a length of seven, eight or nine tones. Similarly, future studies should use musical stimuli with a beat every three tones and thus reveal which interpretation is correct.

Finally, the induced-response analysis showed that the gamma band was stronger for original compared to varied melodies, especially during the presentation of the sequences. Conversely, after the end of the sequence, alpha and beta bands were stronger for the variations compared to the originals. This result is coherent with previous studies which reported increased gamma power during recognition of target stimuli [49,50] and, more generally, a modulation of the brain oscillations associated with memory load and complex cognitive functions [51–52]. In addition, our findings expand on previous literature by providing evidence that bursts of gamma activity are associated with recognition of temporal sequences built upon musical sounds. The induced-response analysis also showed stronger power for alpha and beta in varied compared to originals after the end of the sequence. Arguably, this result may represent the higher processing required by the brain after listening to novel temporal sequences, possibly to store the new information carried by the unfamiliar sounds. Future studies are necessary to further clarify this interpretation. Moreover, further research employing MEG and additional tools such as intracranial EEG (iEEG) should conduct cross-frequency coupling analysis, testing whether gamma-theta coupling is connected to recognition of temporal sequences.

Overall, the results presented here reveal the hierarchical dynamics of the brain underlying processing of auditory sequences extended over time. The results provide pertinent evidence on the neural basis of memory recognition and prediction error, and provides new insights into the brain mechanisms responsible for making temporal information available to humans.

Materials and methods

Participants

The participant sample consisted of 83 volunteers (33 males and 50 females) aged 19 to 63 years old (mean age: 28.76 ± 8.06 years). The sample was recruited in Denmark and participants came from Western countries. All participants were healthy and reported normal hearing. Their educational background was homogeneous.

The project was approved by the Institutional Review Board (IRB) of Aarhus University (case number: DNC-IRB-2020-006). The experimental procedures were carried out in compliance with the Declaration of Helsinki – Ethical Principles for Medical Research. All participants gave the informed consent before starting the experimental procedure.

Experimental stimuli and design

In this study, we used an old/new paradigm auditory recognition task [31,32,34,35] during magnetoencephalography (MEG) recordings. First, participants listened to a short (approximately 25 seconds long) musical piece twice and were asked to memorise it as much as possible. The musical piece consisted of the first four measures of the right-hand part of Johann Sebastian Bach's Prelude No. 2 in C Minor, BWV 847. A MIDI version of the piece was created using Finale (MakeMusic, Boulder, CO). Each tone of the piece had the same duration (approximately 350 ms). Second, participants were presented with 135 five-tone musical excerpts that lasted 1750 ms each. Participants were requested to state whether each excerpt belonged to the original music ('memorised' sequence [M], old) or were a varied musical sequence ('novel' sequence [N], new) (**Figure 1a**). Twenty-seven excerpts were drawn from the original musical piece and 108 were variations of the original melodies (**Figure S1** shows all the sequences used in the study). The two categories of stimuli (M and N) were created as follows. The M sequences were comprised by the first five tones of the first three measures of the musical piece. These sequences were presented nine times each, for a total of 27 trials. The N sequences were created through systematic variations of the three M sequences (**Figure 1b**). This procedure consisted of changing every musical tone of the sequence after the first (NT1), second (NT2), third (NT3) or fourth (NT4) tone. We created nine variations for each of the original M sequences and each of the four categories of N. This resulted in 27 N sequences for each category, and 108 N in total. To be noted, the variations were created according to the following rules:

- Inverted melodic contours (used twice): the melodic contour of the variation was inverted with respect to the original M sequence (i.e., if the M sequence had the following melodic contour: down-down-up-down, the N sequence would be: up-up-down-up)
- Same tone scrambled (used three times): the remaining tones of the M sequence were scrambled (e.g., M sequence: C-E-D-E-C, was converted into NT1 sequence: C-C-E-E-D). When this was not possible (e.g., in the case of NT4, where only the last tone is different from the M sequence), we substituted the last tone of the M sequence with a random tone.
- Same tone (used three times): the same tone was repeatedly used, in some cases varying only the octave (e.g., M sequence: C-E-D-E-C, was transformed into NT1 sequence: C-E^{8-} E^{8-} E_8^- E_8).
- Scrambling intervals (used once): the intervals between the tones were scrambled (e.g., M sequence: 6^{th}m - 2^{nd}m – 2^{nd}m – 3^{rd}m, was adapted to NT1 sequence: 2^{nd}m, 6^{th}m, 3^{rd}m, 2^{nd}m).

This procedure allowed us to investigate *(i)* the brain dynamics underlying the recognition of previously memorised auditory sequences and (*ii*) the conscious detection of the sequence variation.

Data acquisition

The MEG recordings were acquired in a magnetically shielded room at Aarhus University Hospital (AUH), Aarhus, Denmark, using an Elekta Neuromag TRIUX MEG scanner with 306 channels (Elekta Neuromag, Helsinki, Finland). The data was recorded at a sampling rate of 1000 Hz with an analogue filtering of 0.1 – 330 Hz. Before the recordings, the head shape of the participants and the position of four Head Position Indicator (HPI) coils were registered with respect to three anatomical landmarks using a 3D digitizer (Polhemus Fastrak, Colchester, VT, USA). This recording was later used to co-register the MEG data with the MRI anatomical scans. For the entire duration of the MEG recordings, the HPI coils registered the continuous localization of the head, which was subsequently employed for movement correction. In addition, two sets of bipolar electrodes were used to record cardiac rhythm and eye movements. This allowed us to remove the electrocardiography (ECG) and electrooculography (EOG) artifacts in a later stage of the analysis pipeline.

The MRI scans were recorded on a CE-approved 3T Siemens MRI-scanner at AUH. The recorded data consisted of structural T1 (mprage with fat saturation) with a spatial resolution of 1.0 x 1.0 x 1.0 mm and the following sequence parameters: echo time (TE) = 2.61 ms, repetition time (TR) = 2300 ms, reconstructed matrix size = 256 x 256, echo spacing = 7.6 ms, bandwidth = 290 Hz/Px.

The MEG and MRI recordings were acquired in two separate days.

Behavioral data

We obtained behavioral data (number of correctly recognised trials and correspondent reaction times) from the experimental task carried out during the MEG recording.

Since the data was not normally distributed, we computed two independent Kruskal-Wallis H tests [6] (non-parametric one-way analysis of variance) to assess whether the five categories of temporal sequences (M, NT1, NT2, NT3, NT4, NT5) differed in terms of correct responses and reaction times. Multiple comparisons were corrected using the Tukey-Kramer correction [53].

MEG data preprocessing

The raw MEG sensor data (204 planar gradiometers and 102 magnetometers) was first pre-processed by MaxFilter [54] to attenuate external interferences. We applied signal space separation (MaxFilter parameters: spatiotemporal signal space separation [SSS], down-sample from 1000Hz to 250Hz, movement compensation using cHPI coils [default step size: 10 ms], correlation limit between inner and outer subspaces used to reject overlapping intersecting inner/outer signals during spatiotemporal SSS: 0.98).

The data was then converted into Statistical Parametric Mapping (SPM) format and further preprocessed and analyzed in MATLAB (MathWorks, Natick, MA, USA) using a combination of in-house-built codes (LBPD, https://github.com/leonardob92/LBPD-1.0.git) and the Oxford Centre for Human Brain Activity (OHBA) Software Library (OSL) [55] (https://ohba-analysis.github.io/osl-docs/), a freely available software that builds upon Fieldtrip [56], FSL [57], and SPM [58] toolboxes.

The continuous MEG data was visually inspected to identify and remove large artifacts using the OSLview tool. The data that was removed was less than 0.1% of the amount of collected data. Independent component analyses (ICA) were used to discard the interference of eyeblinks and heart-beat artefacts from the brain data [59]. First, we decomposed the original

signal into independent components. Second, we isolated and discarded the components that picked up eyeblink and heart-beat activities. Third, the signal was rebuilt using the remaining components. Finally, the signal was epoched in 135 trials (27 M, 27 NT1, 27 NT2, 27 NT3, 27 NT4) and baseline-corrected by removing the mean signal recorded in the baseline from the post-stimulus brain signal. Each trial lasted 4500 ms (4400 ms plus 100 ms of baseline time).

Multivariate pattern analysis (decoding)

We performed multivariate pattern analyses to decode different neural activity associated with the recognition of M versus N. Here, we computed four independent analyses, decoding M from each of the four categories of the novel sequences (i.e., M versus NT1, M versus NT2, M versus NT3, M versus NT4).

We used support vector machines (SVMs) [60] and calculated independent analyses for each participant. The MEG data was rearranged in a 3D matrix (channels x timepoints x trials) and submitted to the SVM algorithm. To avoid overfitting, a leave-one-out cross-validation approach was adopted to train the SVM classifier to decode the two experimental conditions. This procedure divided the trials into N different groups (here N = 8). Then, for each timepoint, it assigned N − 1 groups to the training set and the remaining N_{th} group to the testing set. After that, the classifier ability to separate the two conditions was evaluated. This process was performed 100 times with random reassignment of the data to training and testing sets. To summarize, the decoding accuracy time series were averaged to obtain a final time series showing the performance of the classifier for each participant.

To test the significance of the decoding results (chance level set at 50%), we employed a sign permutation test against the chance level for each timepoint and then corrected for multiple comparisons using false-discovery rate (FDR) correction ($\alpha = .05$; FDR-adjusted $q < .012$).

To assess whether each pair of conditions were differentiated by neural patterns which were stable over time, we computed four temporal generalization multivariate analyses. The algorithm was the same as the one previously described. However, in this case we utilized each timepoint of the training set to predict not only the same timepoint in the testing set, but all timepoints [61,62]. Here, the significance was tested using a signed permutation test against the chance level (50%) for each timepoint, as the previous analyses. Then, we corrected for multiple comparisons using two-dimensional (2D) cluster-based Monte-Carlo simulations (MCS, $\alpha = .01$, MCS p-value = .001) 1-5, 17. First, we computed the clusters size of the

continuous, binarized, significant values in time. Second, we made 1000 permutations of these binarized values. For each permutation, we computed the size of the maximum emerging cluster and built a reference distribution using those values. Finally, we considered significant the original clusters that were bigger than the 99.99% of the permuted data maximum cluster sizes.

Source reconstruction

MEG is a powerful tool to detect the whole-brain activity with excellent temporal resolution. However, to obtain a complete picture of the whole-brain activity underlying complex cognitive tasks the spatial component of the brain activity must be also identified. Here, we employed the established beamforming method [63–65], built upon a combination of in-house-built codes and codes available in OSL, SPM, and FieldTrip.

To reconstruct the brain sources that generated the MEG signal, an inverse problem must be solved. The MEG recording shows the activity of the neural signals outside the head but cannot provide information on the specific brain sources which generated it. Thus, we used beamforming algorithms to solve this problem, implementing the two following steps: (i) designing a forward model and (ii) computing the inverse solution.

The forward model is a theoretical model which considers each brain source as an active dipole (brain voxel). It describes how the unitary strength of each dipole would be reflected over all MEG sensors. Here, we employed magnetometer channels and an 8-mm grid, which returned 3559 dipole locations (voxels) within the whole brain. After co-registering the individual structural T1 data with the fiducial points (i.e., information about head landmarks), we computed the forward model by adopting the widely used method called “Single Shell”, which is presented in detail in Nolte [66]. The output of this computation, referred to as “leadfield model”, was stored in the matrix L (sources x MEG channels). In the three cases where the structural T1 was not available we performed the leadfield computation using a template (MNI152-T1 with 8-mm spatial resolution).

Then, we computed the inverse solution. As mentioned above, we chose the beamforming, which is one of the most popular and effective algorithms available in the field. This procedure employs a different set of weights which are sequentially applied to the source locations for isolating the contribution of each source to the activity recorded by the MEG channels. This is done for each timepoint of the recorded brain data. The beamforming inverse solution can be summarized by the following main steps.

694 The data recorded by MEG sensors (*B*) at time *t*, can be described by the following
695 equation (1):
696

$$B_{(t)} = L * Q_{(n_i,t)} + \quad (1)$$

697
698 where *L* is the leadfield model, *Q* is the dipole matrix carrying the activity of each active
699 dipole (*q*) over time and is noise (see Huang and colleagues [64] for details). To solve the
700 inverse problem, *Q* must be computed. In the beamforming algorithm, weights are computed
701 and then applied to the MEG sensors at each timepoint, as shown for the single dipole *q* in
702 equation (2):
703

$$q_{(t)} = W^T * B_{(t)} \quad (2)$$

704
705 To obtain *q*, the weights *W* should be computed (the subscript *T* refers to transpose matrix).
706 To this goal, the beamforming relies on the matrix multiplication between *L* and the
707 covariance matrix between MEG sensors (*C*), which is calculated on the concatenated
708 experimental trials. Specifically, for each brain source *n*, the weights W_n are computed as
709 shown in equation (3):
710

$$W_{(n)} = (L_{(n)}{}^T * C^{-1} * L_{(n)})^{-1} * L_{(n)}{}^T * C^{-1} \quad (3)$$

711
712 To be noted, the computation of the leadfield model was performed for the three main
713 orientations of each brain source (dipole), according to Nolte [66]. Before computing the
714 weights, the orientations were reduced to one using the singular value decomposition
715 algorithm on the matrix multiplication reported in equation (4). This procedure is widely
716 adopted to simplify the beamforming output [67,68].
717

$$L = svd(l^T * C^{-1} * l)^{-1} \quad (4)$$

718
719 Here, *l* represents the leadfield model with the three orientations, while *L* is the resolved one-
720 orientation model that was utilized in (3). Finally, the weights were applied to each brain
721 source and timepoint. To be noted, the covariance matrix C was computed on the continuous
722 signal, which was estimated by concatenating the trials of all experimental conditions. The

weights were applied to the brain data associated with each condition and normalized according to Luckhoo et al. [68] for counterbalancing the reconstruction bias towards the centre of the head. The weights were applied to the neural activity averaged over trials for the evoked responses and to the neural activity of each independent trial for the induced responses. This procedure returned a time series for each of the 3559 brain sources (and each trial in the case of induced responses). The sign ambiguity of the evoked responses time series was adjusted for each brain source using its sign in correspondence with the N100 response to the first tone of the auditory sequences [31,34,35].

Neural sources of the differential brain activity between M and N

To detect the brain sources underlying the differential signal observed for M and N, we considered both the results from the decoding analyses and the inspection of the MEG data after preprocessing. We calculated this to identify which condition was association with a larger neural signal. If condition one is significantly stronger than condition two at time t = x and then condition two is significantly stronger than condition one at time t = x + 1, the decoding will return an overall significant difference between conditions from t = x until t = x + 1, even if such difference is qualitatively not the same at t = x and at t = x +1. Thus, to define the time windows for inspecting the neural sources, it is good practice to look both at significant results from decoding and the correspondent brain activity in the MEG data after preprocessing.

This procedure, applied to all our four contrasts, returned the following time windows: 0.50 – 0.60 sec; 0.70 – 0.80 sec; 0.98 – 1.02 sec; 1.05 – 1.15 sec; 1.33 – 1.39 sec; 1.45 – 1.55 sec; 1.70 – 1.75 sec; 1.75 – 1.85 sec. For each time window and condition, we averaged the time series of all brain sources over time and computed t-tests contrasting M versus N (t-tests were computed independently for M versus each of the four categories of N). Finally, we corrected for multiple comparisons using a 3D cluster-based MCS (MCS, α = .003, MCS *p-value* = .001). Here, we calculated the sizes of the clusters of neighbouring brain voxels which were significant. Then, we computed 1000 permutations of the original data. For each permutation, we estimated the sizes of the clusters of neighbouring permuted brain voxels which were significant. This returned a reference distribution of the biggest cluster sizes observed in the permutated data. Finally, we considered significant the original clusters that were bigger than the 99.99% of clusters forming the reference distribution. Additional details on the MCS algorithm can be found in [31,34,35,69].

Functional regions of interests (ROIs)

Our previous analyses highlighted a network mainly comprising four broad brain areas which were involved in the task. These areas roughly corresponded to the bilateral medial cingulate gyrus (MC), right hippocampal area and inferior temporal cortex (HITR), left auditory cortex (ACL), and bilateral ventromedial prefrontal cortex (VMPFC).

Then, we wished to refine the spatial extent of those areas based on their functional profile and obtain their associated time series. Thus, first we computed *t*-values for each brain voxel and each timepoint contrasting M versus N. Second, we isolated the strongest *t*-value in absolute terms for each of the four broad regions identified in our previous analysis. This allowed us to identify the peaks of differential activity occurring between M and N for each ROI. Third, we used those peaks (averaged in a time window of ± 20 ms) and strict *t*-value thresholds (abs(*t*) > 3) to isolate the brain voxels that were mainly contributing to discriminate M versus N. This procedure refined the spatial extent of the four broad ROIs that we previously identified. Finally, to cover potential hemispheric differences, we created two more ROIs which mirrored HITR and ACL in the opposite hemisphere (HITL and ACR, respectively). Once we defined these six broad ROIs, we computed the time series showing their activity over time by averaging the time series of each of the brain voxels forming every ROI. To be noted, the spatial accuracy of the reconstructed MEG signal cannot be completely accurate, thus it is good practice employing such broad ROIs [70,71].

Statistical analysis on ROIs time series

We employed the time series of the previously identified ROIs to compute additional statistics between M and N conditions. Here, we computed one t-test for each timepoint and each combination of M versus Ns (i.e., M versus NT1, M versus NT2, M versus NT3, M versus NT4). Then, we corrected for multiple comparisons using a one-dimensional (1D) cluster-based MCS (MCS, $\alpha = .05$, MCS *p-value* = .001). First, we identified the clusters of significant continuous values in time. Second, we computed 1000 permutations, randomizing the significant values obtained from the t-tests. For each permutation, we extracted the maximum cluster size and built their reference distribution. Finally, we considered significant the original clusters that were larger than 99.99% of the permuted ones.

Time-frequency analysis for induced responses

We computed a time-frequency analysis using complex Morlet wavelet transform (from 1 to 60 Hz with 1-Hz intervals) [72]. This analysis was conducted for induced responses, independently for the six ROIs previously described and for the four contrasts considered in this study (i.e., M versus NT1, M versus NT2, M versus NT3, M versus NT4). Specifically, the time-frequency decomposition was done independently for each trial, brain voxel, and participant. Then, the power spectrum of each trial and each brain voxel was averaged within each of the six ROIs.

Finally, we computed a t-test for each frequency and timepoint, making four contrasts: M versus NT1, M versus NT2, M versus NT3, M versus NT4. The emerging *p*-values were binarized (α = .05) and then submitted to a 2D MCS (MCS *p*-value = .001). Here, we calculated the clusters size of continuous significant values in time and frequency. Then, we made 1000 permutations of the binarized *p*-values. For each permutation, we measured the size of the maximum emerging cluster and built a reference distribution with one value for each permutation. Finally, the original clusters were considered significant when they were bigger than the 99.99% of the permuted data maximum cluster sizes.

Data availability

The codes are available at the following link: https://github.com/leonardob92/LBPD-1.0.git. The multimodal neuroimaging data related to the experiment is available upon reasonable request.

Acknowledgements

The Center for Music in the Brain (MIB) is funded by the Danish National Research Foundation (project number DNRF117). Additionally, we thank the Fundación Mutua Madrileña for the economic support provided to author GFR.

LB is supported by Carlsberg Foundation (CF20-0239), Lundbeck Foundation (Talent Prize 2022), Center for Music in the Brain, Linacre College of the University of Oxford (Lucy Halsall fund), and Society for Education and Music Psychology (SEMPRE's 50th Anniversary Awards Scheme).

MLK is supported by Center for Music in the Brain and Centre for Eudaimonia and Human Flourishing, which is funded by the Pettit and Carlsberg Foundations.

Author contributions

LB, GFR and MLK conceived the hypotheses. LB designed the study. LB, MLK and PV recruited the resources for the experiment. LB, GFR and FC collected the data. LB, GFR and performed pre-processing and statistical analysis. DP provided MATLAB codes for decoding analysis. DP, MLK and PV provided essential help to interpret and frame the results within the neuroscientific literature. GFR and LB wrote the first draft of the manuscript. LB, FC, GFR and MLK prepared the figures. All the authors contributed to and approved the final version of the manuscript.

Competing interests' statement

The authors declare no competing interests.

References

1. Scalabrini, A., Mucci, C. & Northoff, G. The nested hierarchy of self and its trauma: In search for a synchronic dynamic and topographical re-organization. *Front. Hum. Neurosci.* **2**, 980353 (2022).
2. Mesulam, M. M. From sensation to cognition. *Brain* (1998). doi:10.1093/brain/121.6.1013
3. Preti, M. G., Bolton, T. A., & Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage*, *160*, 41-54
4. Deco, G., Tononi, G., Boly, M. & Kringelbach, M. L. Rethinking segregation and integration: Contributions of whole-brain modelling. *Nature Reviews Neuroscience* (2015). doi:10.1038/nrn3963
5. Deco, G., Vidaurre, D. & Kringelbach, M. L. Revisiting the global workspace orchestrating the hierarchical organization of the human brain. *Nat. Hum. Behav.* (2021). doi:10.1038/s41562-020-01003-6
6. Kringelbach, M. L. *et al.* Dynamic coupling of whole-brain neuronal and neurotransmitter systems. *Proc. Natl. Acad. Sci. U. S. A.* (2020). doi:10.1073/pnas.1921475117
7. Dehaene, S., Changeux, J. P. & Naccache, L. The global neuronal workspace model of conscious access: From neuronal architectures to clinical applications. *Res. Perspect. Neurosci.* (2011). doi:10.1007/978-3-642-18015-6_4
8. Dehaene, S., Kerszberg, M. & Changeux, J. P. A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U. S. A.* (1998). doi:10.1073/pnas.95.24.14529
9. Dehaene, S. & Changeux, J. P. Experimental and Theoretical Approaches to Conscious Processing. *Neuron* (2011). doi:10.1016/j.neuron.2011.03.018
10. Friston, K. Predictive coding, precision and synchrony. *Cognitive Neuroscience* (2012). doi:10.1080/17588928.2012.691277
11. Koelsch, S., Vuust, P. & Friston, K. Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences* (2019). doi:10.1016/j.tics.2018.10.006
12. Ip, I. B. & Bridge, H. Investigating the neurochemistry of the human visual system using magnetic resonance spectroscopy. *Brain Structure and Function* (2022). doi:10.1007/s00429-021-02273-0
13. King, J. R., Wyart, V. & King, J. R. The human brain encodes a chronicle of visual events at each instant of time through the multiplexing of traveling waves. *J. Neurosci.* (2021). doi:10.1523/JNEUROSCI.2098-20.2021
14. Zatorre, R. J. Sound analysis in auditory cortex. *Trends in Neurosciences* (2003). doi:10.1016/S0166-2236(03)00074-2
15. Moore, J. K. Organization of the human superior olivary complex. *Microsc. Res. Tech.* (2000). doi:10.1002/1097-0029(20001115)51:4<403::AID-JEMT8>3.0.CO;2-Q
16. Pandya, D. N. Anatomy of the auditory cortex. in *Revue Neurologique* (1995). doi:10.4324/9780203933107-7
17. Peretz, I. & Zatorre, R. J. *The Cognitive Neuroscience of Music*. *The Cognitive Neuroscience of Music* (2012). doi:10.1093/acprof:oso/9780198525202.001.0001
18. Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*(1), 37-48.
19. Gaab, N., Gaser, C., Zaehle, T., Jancke, L. & Schlaug, G. Functional anatomy of pitch memory - An fMRI study with sparse temporal sampling. *Neuroimage* (2003). doi:10.1016/S1053-8119(03)00224-6
20. Zatorre, R. J., Evans, A. C. & Meyer, E. Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* (1994). doi:10.1523/jneurosci.14-04-01908.1994

21. Kumar, S. *et al.* A brain system for auditory working memory. *J. Neurosci.* (2016). doi:10.1523/JNEUROSCI.4341-14.2016
22. Hansen, P., Kringelbach, M. & Salmelin, R. *MEG: An introduction to methods*. *MEG: An Introduction to Methods* (2010). doi:10.1093/acprof:oso/9780195307238.001.0001
23. Conley, E. M., Michalewski, H. J. & Starr, A. The N100 auditory cortical evoked potential indexes scanning of auditory short-term memory. *Clin. Neurophysiol.* **110**, 2086–2093 (1999).
24. Näätänen, R. & Picton, T. The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology* (1987). doi:10.1111/j.1469-8986.1987.tb00311.x
25. Bonetti, L. *et al.* Brain predictive coding processes are associated to COMT gene Val158Met polymorphism. *Neuroimage* (2021). doi:10.1016/j.neuroimage.2021.117954
26. Bonetti, L., Haumann, N. T., Vuust, P., Kliuchko, M. & Brattico, E. Risk of depression enhances auditory Pitch discrimination in the brain as indexed by the mismatch negativity. *Clin. Neurophysiol.* **128**, 1923–1936 (2017).
27. Bonetti, L. *et al.* Auditory sensory memory and working memory skills: Association between frontal MMN and performance scores. *Brain Res.* **1700**, 86–98 (2018).
28. Bonetti, L. *et al.* Whole-brain computation of cognitive versus acoustic errors in music: A mismatch negativity study. *Neuroimage: Reports* **2**, 100145 (2022).
29. Brattico, E., Winkler, I., Näätänen, R., Paavilainen, P. & Tervaniemi, M. Simultaneous storage of two complex temporal sound patterns in auditory sensory memory. *Neuroreport* (2002). doi:10.1097/00001756-200210070-00011
30. Albouy, P., Weiss, A., Baillet, S. & Zatorre, R. J. Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. *Neuron* (2017). doi:10.1016/j.neuron.2017.03.015
31. Bonetti, L. *et al.* Brain recognition of previously learned versus novel temporal sequences: a differential simultaneous processing. *Cereb. Cortex* **bhac439**, (2022).
32. Bonetti, L. *et al.* Spatiotemporal brain dynamics during recognition of the music of Johann Sebastian Bach. *bioRxiv* (2020). doi:10.1101/2020.06.23.165191
33. Bonetti, L. *et al.* Rapid encoding of musical tones discovered in whole-brain connectivity. *Neuroimage* (2021). doi:10.1016/j.neuroimage.2021.118735
34. Fernández Rubio, G. *et al.* Magnetoencephalography recordings reveal the spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences. *Commun. Biol.* **5**, (2022).
35. Fernández Rubio, G., Carlomagno, F., Vuust, P., Kringelbach, M. L. & Bonetti, L. Associations between abstract working memory abilities and brain activity underlying long-term recognition of auditory sequences. *PNAS Nexus* **1**, pgac216 (2022).
36. Brattico, E. & Pearce, M. The neuroaesthetics of music. *Psychol. Aesthetics, Creat. Arts* (2013). doi:10.1037/a0031624
37. Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences* (2002). doi:10.1016/S1364-6613(00)01816-7
38. Limongi, R., Sutherland, S. C., Zhu, J., Young, M. E. & Habib, R. Temporal prediction errors modulate cingulate-insular coupling. *Neuroimage* (2013). doi:10.1016/j.neuroimage.2012.12.078
39. Hasselmo, M. E. A Handbook for Modeling Hippocampal Circuits. *Front. Comput. Neurosci.* (2011). doi:10.3389/fncom.2011.00002
40. Squire, L. R. *Encyclopedia of Neuroscience*. *Encyclopedia of Neuroscience* (2010). doi:10.1016/C2009-1-03742-3

41. Stern, C. E., Sherman, S. J., Kirchhoff, B. A. & Hasselmo, M. E. Medial temporal and prefrontal contributions to working memory tasks with novel and familiar stimuli. *Hippocampus* (2001). doi:10.1002/hipo.1048
42. Merkley, T. L., Larson, M. J., Bigler, E. D., Good, D. A. & Perlstein, W. M. Structural and functional changes of the cingulate gyrus following traumatic brain injury: Relation to attention and executive skills. *J. Int. Neuropsychol. Soc.* (2013). doi:10.1017/S135561771300074X
43. Pando-Naude, V., Patyczek, A., Bonetti, L. & Vuust, P. An ALE meta-analytic review of top-down and bottom-up processing of music in the brain. *Sci. Rep.* (2021). doi:10.1038/s41598-021-00139-3
44. Criscuolo, A., Pando-Naude, V., Bonetti, L., Vuust, P. & Brattico, E. An ALE meta-analytic review of musical expertise. *Sci. Rep.* **12**, (2022).
45. Vuust, P. Music in the Brain. *Heggli, O.A. Friston, K. Kringelbach, M.L.* **23**, 287–305 (2022).
46. Brattico, E. *et al.* Neural discrimination of nonprototypical chords in music experts and laymen: an MEG study. *J. Cogn. Neurosci.* **21**, 2230–2244 (2009).
47. Näätänen, R., Paavilainen, P., Rinne, T. & Alho, K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology* (2007). doi:10.1016/j.clinph.2007.04.026
48. Korzyukov, O. A., Winkler, I., Gumenyuk, V. I. & Alho, K. Processing abstract auditory features in the human auditory cortex. *Neuroimage* (2003). doi:10.1016/j.neuroimage.2003.08.014
49. Slobounov, S., Tutwiler, R., Slobounova, E., Rearick, M. & Ray, W. Human oscillatory brain activity within gamma band (30-50 Hz) induced by visual recognition of non-stable postures. *Cogn. Brain Res.* (2000). doi:10.1016/S0926-6410(99)00055-5
50. Lenz, D., Schadow, J., Thaerig, S., Busch, N. A. & Herrmann, C. S. What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. *Int. J. Psychophysiol.* (2007). doi:10.1016/j.ijpsycho.2006.07.008
51. Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews*, *34*(7), 1015-1022
52. Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, *53*(1), 63-88
53. Tukey, J. W. Comparing Individual Means in the Analysis of Variance. *Biometrics* (1949). doi:10.2307/3001913
54. Taulu, S. & Simola, J. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Phys. Med. Biol.* (2006). doi:10.1088/0031-9155/51/7/008
55. Woolrich, M., Hunt, L., Groves, A. & Barnes, G. MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization. *Neuroimage* (2011). doi:10.1016/j.neuroimage.2011.04.041
56. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci. CIN, 2011* 1–9 (2011). doi::10.1155/2011/156869
57. Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* (2009). doi:10.1016/j.neuroimage.2008.10.055
58. Penny, W., Friston, K., Ashburner, J., Kiebel, S. & Nichols, T. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (2007). doi:10.1016/B978-0-12-372560-8.X5000-1

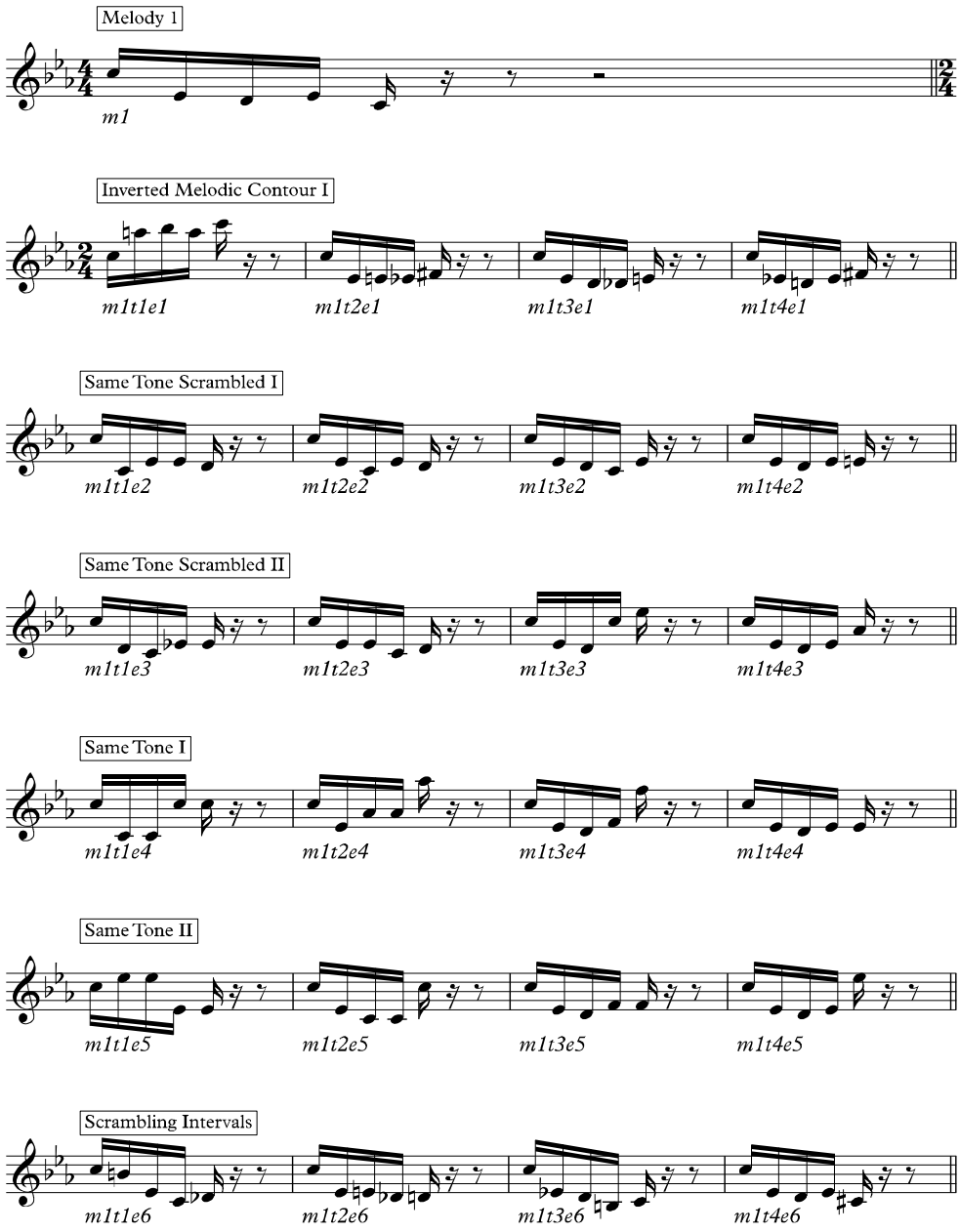
59. Mantini, D. *et al.* A signal-processing pipeline for magnetoencephalography resting-state networks. *Brain Connect.* (2011). doi:10.1089/brain.2011.0001
60. Wilson, M. D. Support Vector Machines. in *Encyclopedia of Ecology, Five-Volume Set* (2008). doi:10.1016/B978-008045405-4.00168-3
61. Cichy, R. M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and time. *Nat. Neurosci.* (2014). doi:10.1038/nn.3635
62. King, J. R. & Dehaene, S. Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences* (2014). doi:10.1016/j.tics.2014.01.002
63. Hillebrand, A. & Barnes, G. R. Beamformer Analysis of MEG Data. *International Review of Neurobiology* (2005). doi:10.1016/S0074-7742(05)68006-3
64. Huang, M. X., Mosher, J. C. & Leahy, R. M. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys. Med. Biol.* (1999). doi:10.1088/0031-9155/44/2/010
65. Brookes, M. J. *et al.* Beamformer reconstruction of correlated sources using a modified source model. *Neuroimage* (2007). doi:10.1016/j.neuroimage.2006.11.012
66. Nolte, G. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoenchephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* (2003). doi:10.1088/0031-9155/48/22/002
67. Huang, M. X. *et al.* Commonalities and Differences among Vectorized Beamformers in Electromagnetic Source Imaging. *Brain Topogr.* (2004). doi:10.1023/B:BRAT.0000019183.92439.51
68. Luckhoo, H. T., Brookes, M. J. & Woolrich, M. W. Multi-session statistics on beamformed MEG data. *Neuroimage* (2014). doi:10.1016/j.neuroimage.2013.12.026
69. Hoegholt, N. F. *et al.* A magnetoencephalography study of first-time mothers listening to infant cries. *Cereb. cortex* **bhac469**, (2022).
70. Gross, J. *et al.* Good practice for conducting and reporting MEG research. *NeuroImage* **65**, 349–363 (2013).
71. Pernet, C. *et al.* Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat. Neurosci.* (2020). doi:10.1038/s41593-020-00709-0
72. Daubechies, I. *Ten Lectures on Wavelets*. *Ten Lectures on Wavelets* (1992). doi:10.1137/1.9781611970104

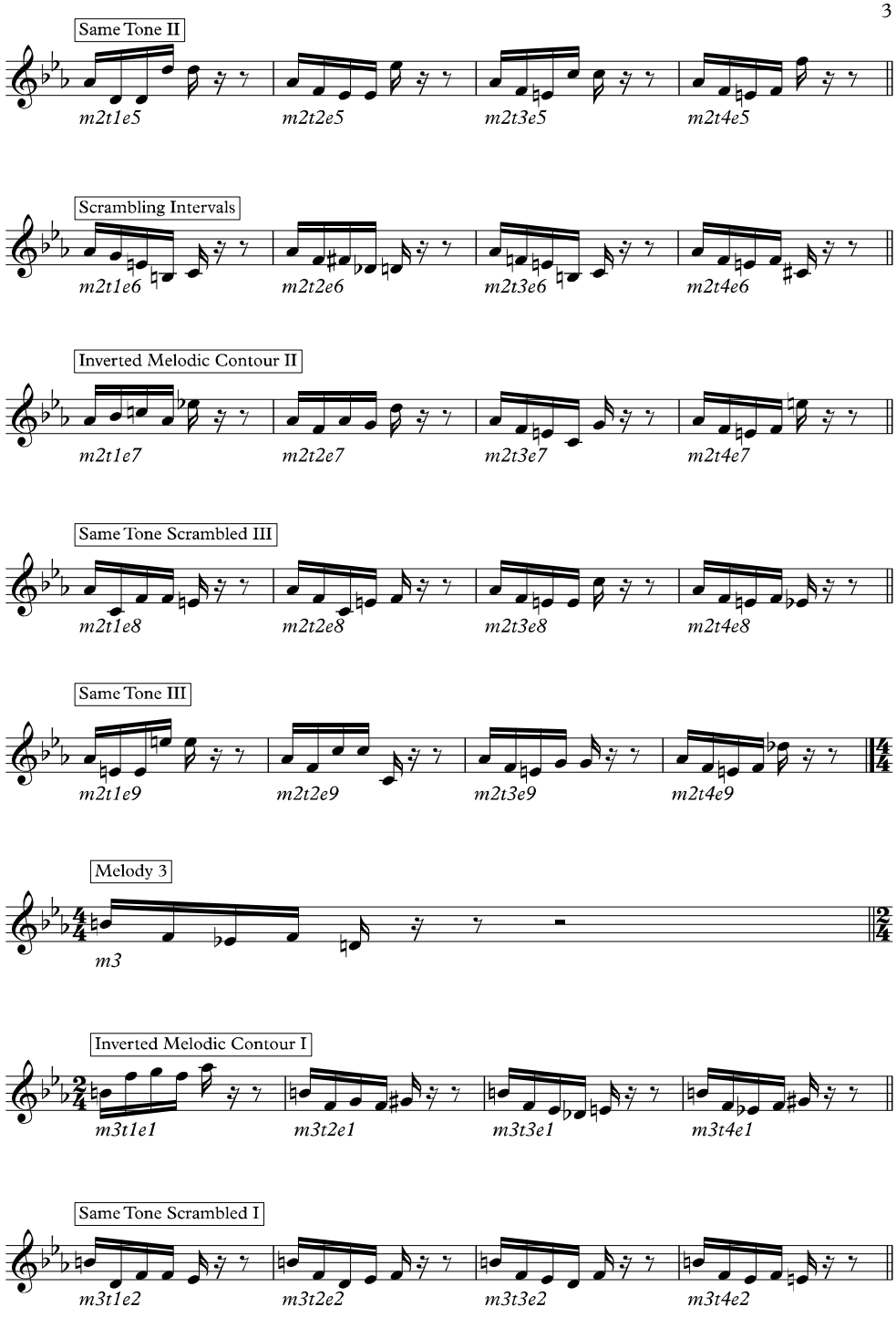
SUPPLEMENTARY MATERIAL

Supplementary materials related to this study and organised as supplementary figures (*i*) and tables (*ii*). In the cases when the supplementary tables were too large to be reported in the current document, they have been exported to Excel files that can be found at the following link:

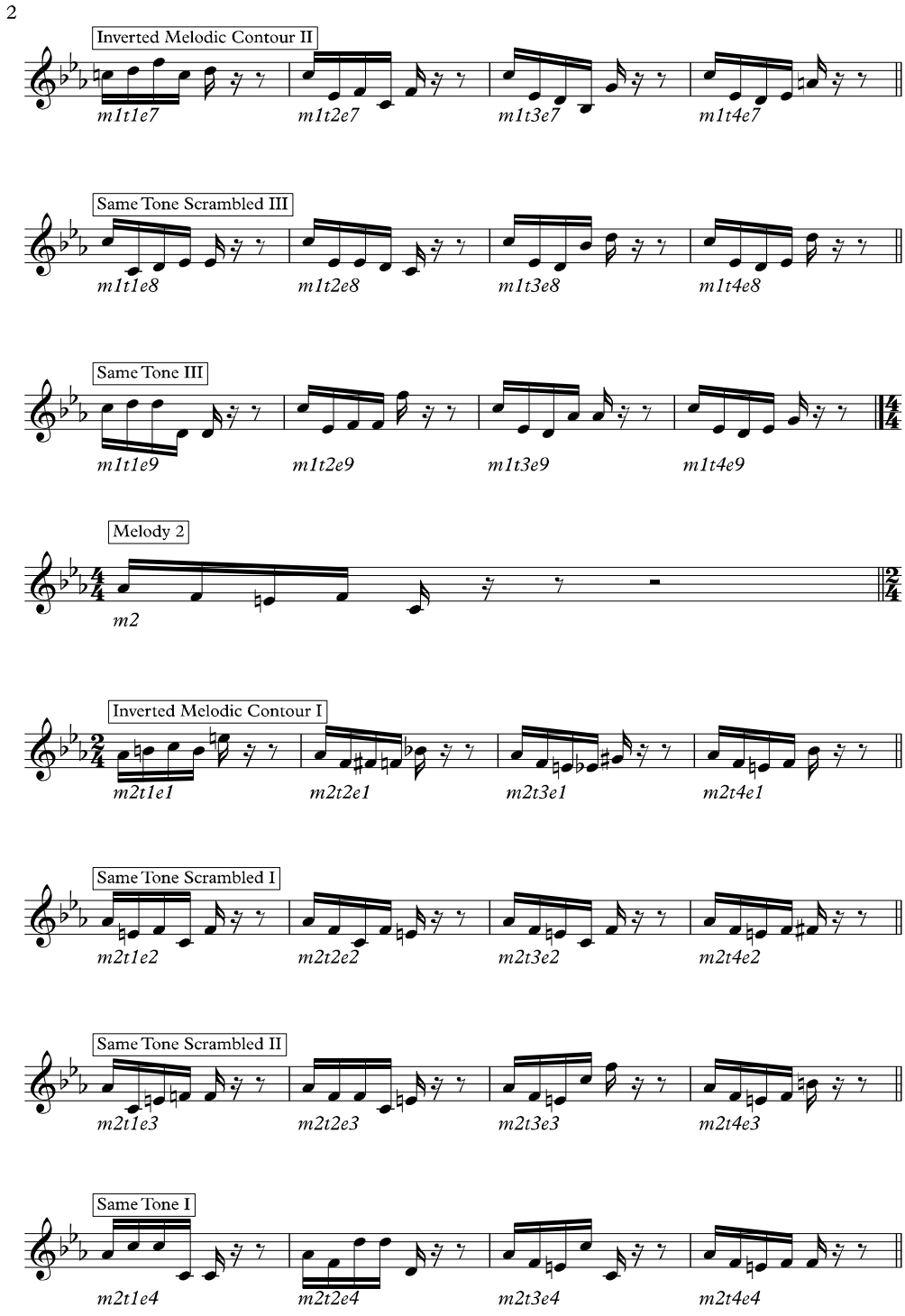
https://drive.google.com/drive/folders/1W1w8UpPKnyp0RMjksKmxi3XqC6UwseBY?usp=sharing

SUPPLEMENTARY FIGURES



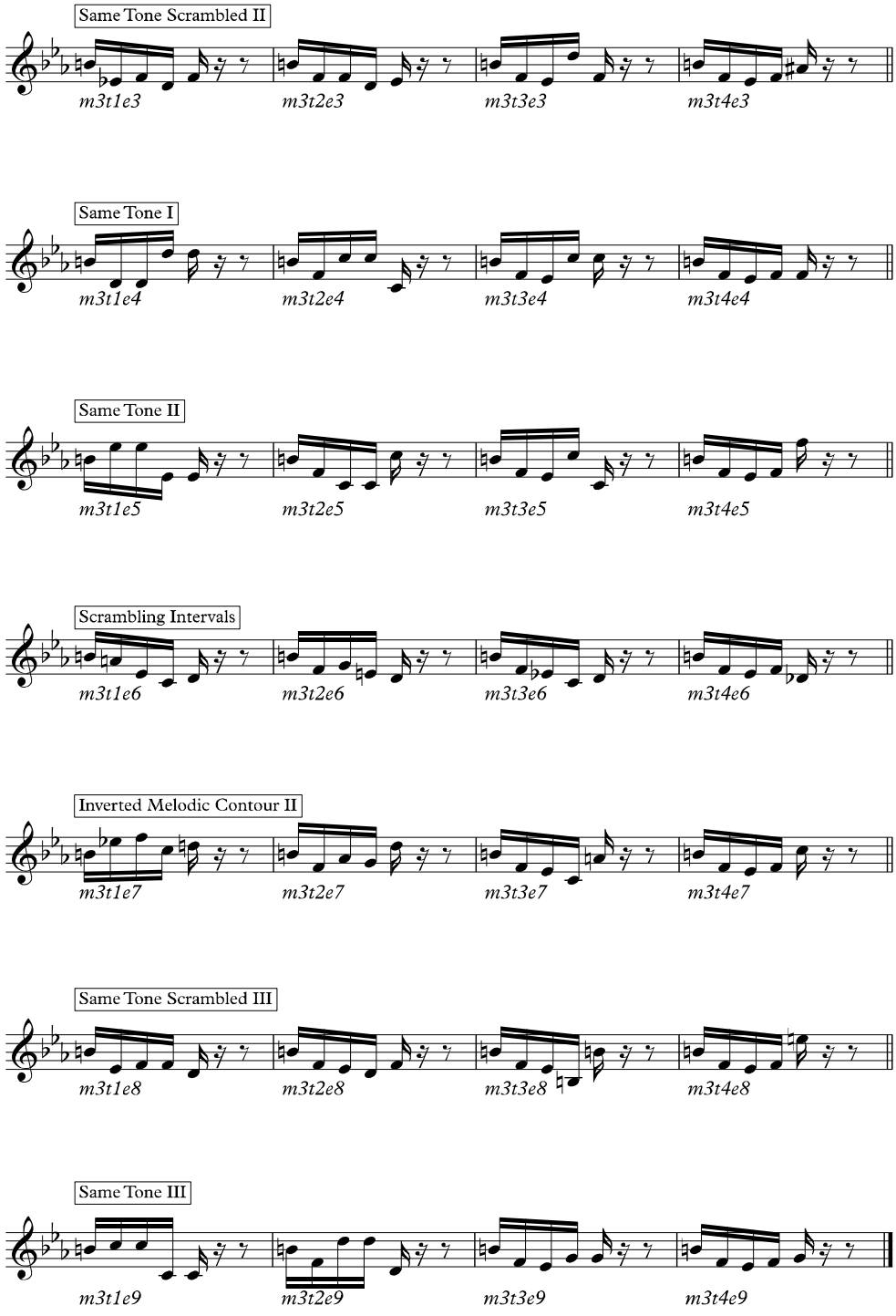


1030



1031

4



1032

1033

1034

Figure S1. Temporal sequences used in the experiment.

The figure shows all temporal sequences used in the experiment, providing detailed information on how they were created. The M sequences were three and comprised the first five tones of the first three measures of the musical piece. These three sequences were presented nine times each, for a total of 27 trials. The N sequences were created through systematic variations of the three M sequences. This procedure consisted of changing every musical tone of the sequence after the first (NT1), second (NT2), third (NT3) or fourth (NT4) tone. We created nine variations for each of the original M sequences and each of the four categories of N. This resulted in 27 N sequences for each category, and 108 N in total. To be noted, as shown in this figure, the variations were created according to the following rules: (i) Inverted melodic contours (used twice): the melodic contour of the variation was inverted with respect to the original M sequence (i.e., if the M sequence had the following melodic contour: down-down-up-down, the N sequence would be: up-up-down-up); (ii) Same tone scrambled (used three times): the remaining tones of the M sequence were scrambled (e.g., M sequence: C-E-D-E-C, was converted into NT1 sequence: C-C-E-E-D). When this was not possible (e.g., in the case of NT4, where only the last tone is different from the M sequence), we substituted the last tone of the M sequence with a random tone; (iii) Same tone (used three times): the same tone was repeatedly used, in some cases varying only the octave (e.g., M sequence: C-E-D-E-C, was transformed into NT1 sequence: C-E^{8-} E^{8-} E_{8}^{-} E_{8}); (iv) Scrambling intervals (used once): the intervals between the tones were scrambled (e.g., M sequence: 6^{th}m - 2^{nd}m – 2^{nd}m – 3^{rd}m, was adapted to NT1 sequence: 2^{nd}m, 6^{th}m, 3^{rd}m, 2^{nd}m).

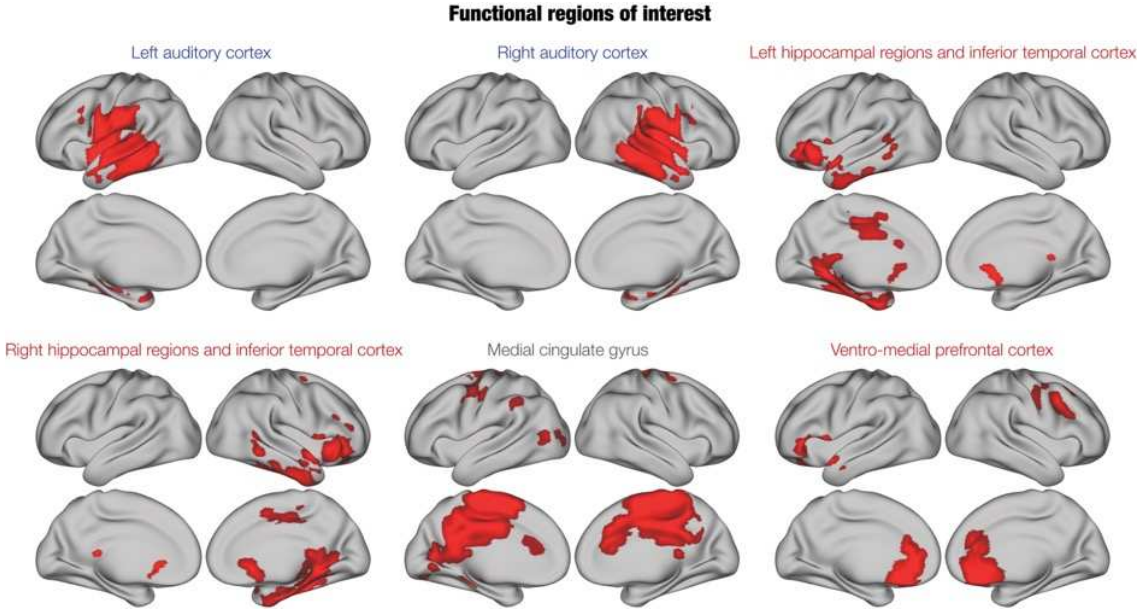


Figure S2. Functional parcels (ROIs) derived from the brain activity underlying the task.

The main activity during recognition of the previously memorised and novel auditory sequences gave rise to the following six functional ROIs: left (i) and right auditory cortex (ii); left (iii) and right hippocampal regions and inferior temporal cortex (iv); medial cingulate gyrus (v), and ventromedial prefrontal cortex (vi).

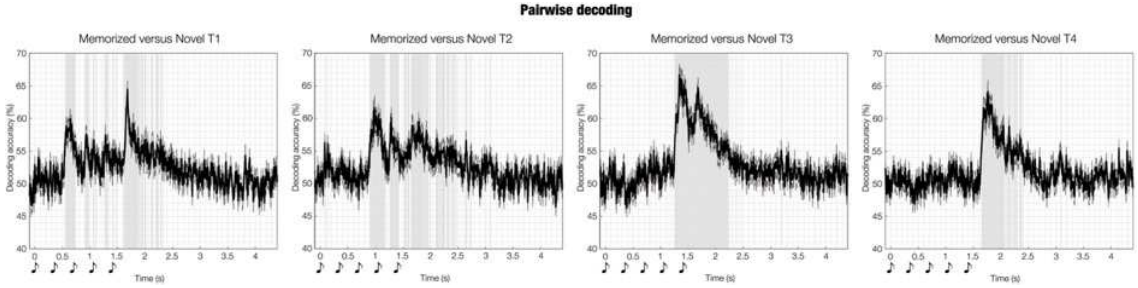


Figure S3. Pairwise decoding time series.

Multivariate pattern analysis decoding the different neural activity associated with memorised versus novel musical sequences. Each plot shows the decoding time series for one of the four rounds of pairwise decoding that we computed (M vs NT1, M vs NT2, M vs NT3, M vs NT4). The sketch of the musical tones represents the onset of the sounds forming the temporal sequences.

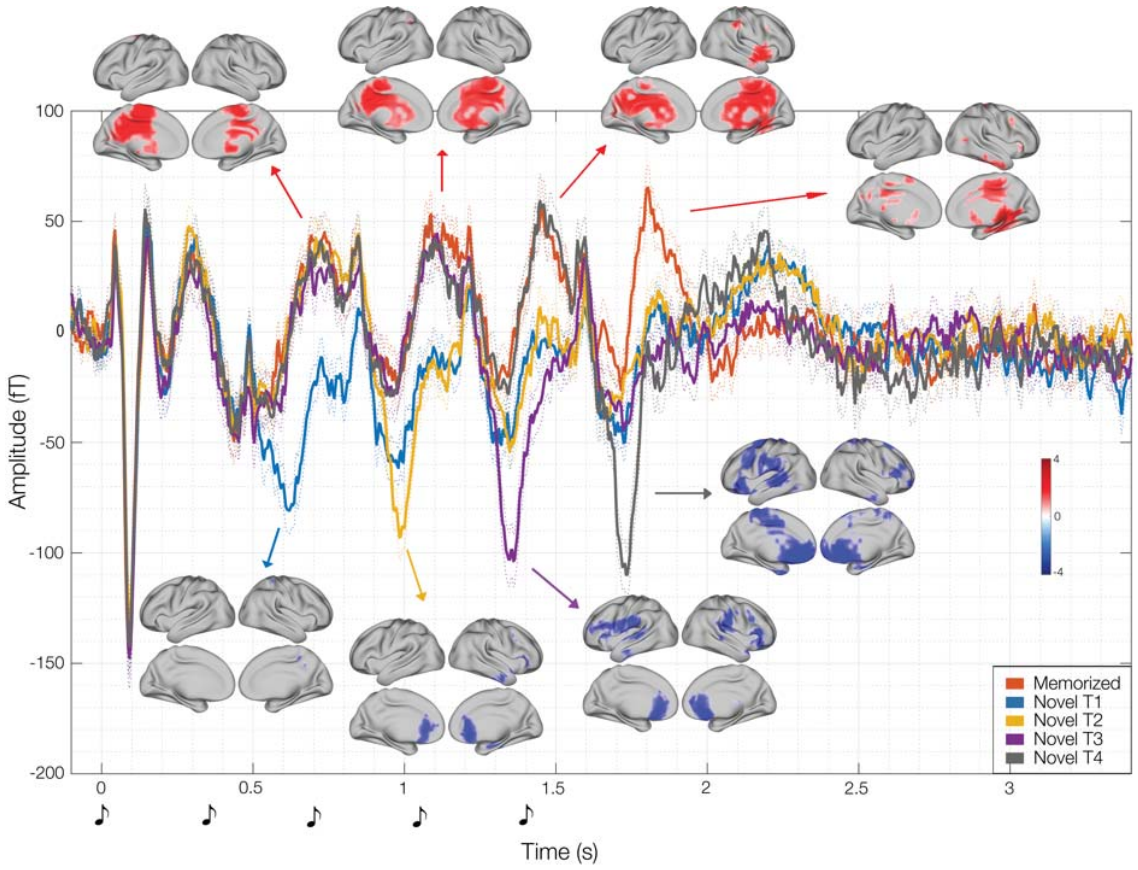


Figure S4. Brain activity recorded at a prototypical magnetometer channel (MEG 0211) and source reconstruction of the main components.

Brain activity recorded over time by the fronto-temporal left MEG channel 0211 showing the five experimental conditions. The sketch of the musical tones represents the onset of the sounds forming the musical sequences. For each of the main positive components, contrasts between the source reconstruction of M versus NT1 have been computed and corrected for multiple comparisons using cluster-based MCS. Results are reported in the brain template above the waveforms. With regards to the negative component indexing the prediction error associated to the disruption of the original sequences, we computed contrasts between the source reconstruction of M versus each category of N (i.e., M vs NT1, M vs NT2, M vs NT3, M vs NT4, respectively) and corrected for multiple comparisons using cluster-based MCS. Results are reported in the brain template below the waveforms. The colour of the arrows illustrates what contrast was performed (e.g., the blue arrow indicates that we contrasted M versus NT1, while the yellow arrow refers to M versus NT2, etc.). The colorbar shows the t-values obtained from the contrasts.

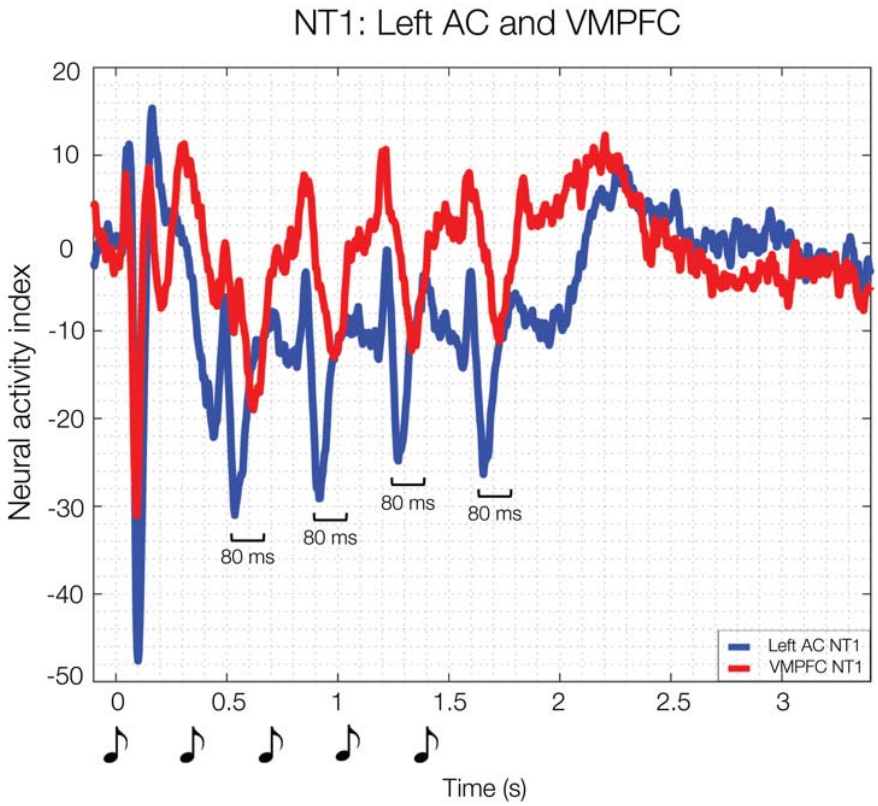


Figure S5. Focus on the NT1 for left auditory cortex (AC) and ventro-medial prefrontal cortex (VMPFC).

The left plot shows the source localized brain activity illustrated for NT1 for left auditory cortex (AC) and ventro-medial prefrontal cortex (VMPFC). Of particular interest it is the sharp peak occurring after the onset of each tone where left AC precedes VMPFC of approximately 80 ms, suggesting a hierarchical processing in the brain. The sketch of the musical tones below the first two plots represents the onset of the sounds forming the temporal sequences.

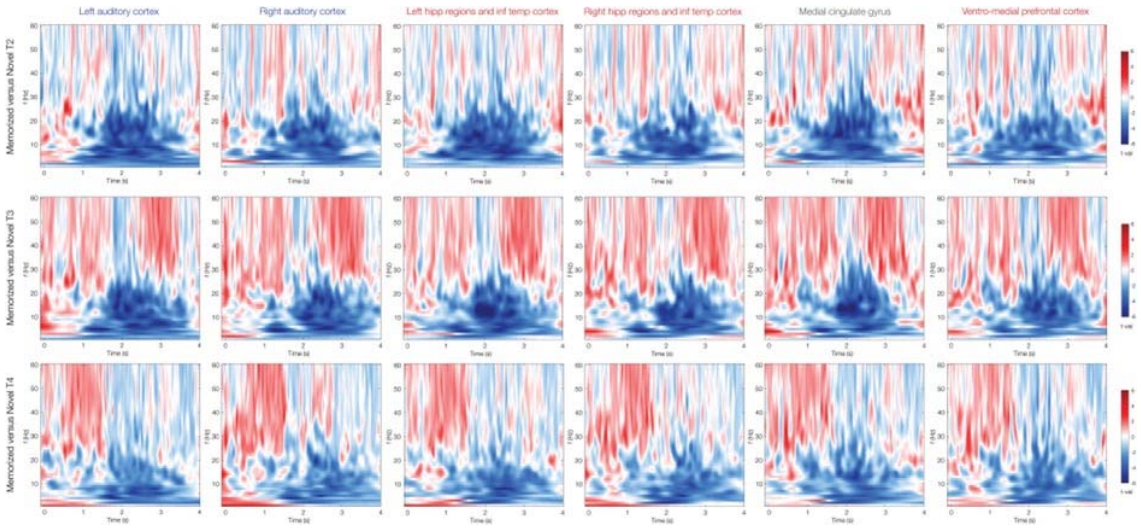


Figure S6. Source localized induced responses – M versus NT2, NT3, NT4.

Contrasts between the source localized induced responses of M versus NT2, NT3, NT4, respectively. The plots indicate a stronger power for gamma in M. Moreover, theta was overall stronger for M versus NT1 during the presentation of the sounds, while alpha, beta and theta were stronger for N versus M after the end of the temporal sequences. The colorbar indicates the t-values obtained by contrasting M versus N.

SUPPLEMENTARY TABLES

Table S1. Pairwise decoding.

Binary time series showing the FDR-corrected significant timepoints (1s) of the decoding time series (i.e., when the algorithm successfully classified M versus N). The first row shows time (in seconds), while the other rows refer to the four contrasts of this study (M versus NT1, M versus NT2, M versus NT3, M versus NT4).

Table S2. Temporal generalization.

Cluster-based MCS on temporal generalization decoding results computed independently for the four following contrasts: M versus NT1, M versus NT2, M versus NT3, M versus NT4. The table shows size, MCS p-value and temporal extent of the cluster (both training and testing sets).

Table S3. Brain source of decoding time windows.

Significant brain sources (after cluster based MCS correction for multiple comparisons) of the significant time windows emerged from the decoding analysis. Results are reported with the correspondent AAL label of each of the significant voxel, as well as their hemisphere, t-value and MNI coordinates. Results are provided for the following contrasts: M versus NT1, M versus NT2, M versus NT3, M versus NT4.

Table S4. ROIs time series.

Significant clusters of differential brain activity between M and N in the six broad functional ROIs isolated in the previous analyses. Results are reported independently for the six ROIs and for the four contrasts (M versus NT1, M versus NT2, M versus NT3, M versus NT4), and comprise cluster size, p-value, temporal extent of the clusters and peak t-value within the cluster.

Table S5. Time-frequency results of induced responses.

Significant clusters of differential power in different frequency bands (1 – 60Hz) computed using complex Morlet wavelet transform. Results are reported independently for the six ROIs and for the four contrasts (M versus NT1, M versus NT2, M versus NT3, M versus NT4), and comprise cluster size, p-value, temporal, and frequency extent of the clusters.