

Cross-spatial scale processing of hierarchical auditory sequences in human brains revealed using 7 T magnetic resonance imaging

Received: 23 June 2025

Accepted: 28 November 2025

Published online: 10 December 2025

 Check for updatesYangjiayi Mu^{1,2}, Qingyang Liu^{1,2}, Yinghua Chu³, Zhao Xu^{1,2}, Weiyang Ding^{1,2}, He Wang^{1,2} & Yuwei Jiang^{1,2}✉

Human auditory cognition involves multiregional hierarchical processing, yet the neural mechanisms integrating cortical layers and large-scale networks remain unclear. Using high-resolution fMRI, we investigate cross-spatial scale dynamics during hierarchical auditory processing. Distinct activations across layers of superior temporal gyrus (STG) and inferior frontal gyrus (IFG) are detected during standard stimuli and violation. We identify the layer-specific effective connectivity among key nodes in the auditory hierarchy using dynamic causal modelling. The results are consistent with hierarchical predictive coding schemes: (i) tone and sequence processing provide input to the superficial and middle layers of STG, respectively. (ii) Forward connections came from the superficial layers of STG, while (iii) backward connections implicated the deep layers of IFG. The integration of mesoscale and macro-scale activities confirms the interaction of superficial/middle layers of STG with higher-level regions. These findings elucidate a multiscale mechanism coordinating cortical layers and distributed networks in hierarchical auditory processing.

A key component of human cognition is the hierarchical processing of auditory information, which enables humans to perceive, comprehend, and react to complex soundscapes¹. In the auditory domain, diverse sounds could be considered as hierarchical auditory sequences with multiple timescales, such as language and music^{2,3}. As an example, human language can be divided into phoneme sequences with a shorter timescale and word sequences with a longer timescale according to its temporal and spatial features^{4,5}. Evidences demonstrate that the brain is hierarchically organized in structure and function⁶, while the auditory cortex participates in cortical hierarchies responsible for sensory integration on multiple timescales⁷. As suggested by predictive coding theory^{8–10}, the auditory information on shorter- and longer-timescales successively induces bidirectional processes of hierarchical neural circuits, involving the primary auditory cortex, downstream auditory areas, and higher-level cognitive

regions. In particular, the canonical predictive coding framework, which is built upon microcircuits^{11–15}, suggests that sensory input arrives in middle layers and is compared with predictions in deep layers to form prediction errors in superficial layers. These prediction errors are then passed forward to the middle layer of the next hierarchical level, while backward predictions are conveyed from deep layers. The sources of forward and backward projections are typically considered to be superficial and deep pyramidal cells, respectively, while spiny stellate cells in the middle layer receive prediction errors from the superficial layers of the level below. Numerous studies focusing on mismatch negativity or oddball paradigms across species have shown that the responses to violations, deviants, or oddballs as eliciting prediction errors that are then broadcast through forward connections to the auditory hierarchy, which supports the hierarchical processing of auditory information in the auditory hierarchy^{16–22}.

¹Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China. ²Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University, Ministry of Education, Shanghai, China. ³MR Research Collaboration Team, Siemens Healthineers Ltd., Shanghai, China. ✉e-mail: yuwjiang@fudan.edu.cn

Therefore, we proposed to use dynamic causal modelling (DCM) to estimate intrinsic (within-level, between layers) and extrinsic (between levels) effective connectivity under different architectures, to assess the evidence for a canonical predictive coding architecture.

Research using intracranial recordings has shown that the hierarchical patterns of activation in the auditory cortex reflect distinct representations of sound properties^{7,23}. The cerebral cortex, also known as gray matter, can be subdivided into six distinct layers of nerve cells. From superficial to deep cortical layers, layer 1 is adjacent to the cerebrospinal fluid (CSF), while layer 6 is adjacent to the white matter (WM)²⁴. Sound encoding across different depths of superior temporal cortex (STG) in humans showed that pitch encoding was stronger in superficial layers, whereas encoding of acoustic-phonetic features was significant in middle-deep layers²⁵. Extensive studies spanning species have reported bidirectional microcircuits across different layers of the auditory cortex. In mice, the corticocollicular neurons from layer 5 and 6 receive distinct cortical and thalamic inputs²⁶, while layer 5 and 6 provide unimodal and heteromodal modulations to the inferior colliculus, respectively, revealing different auditory microcircuits for separate roles²⁷. Macaque studies showed

that the sensory responses activated the middle layer first, but the unpredicted auditory input activated superficial layers most, suggesting feedforward processing in the microcircuit, and the prediction signals feedback from deep layers^{22,28}. However, limited by invasive recording technology, the neural circuits across the cortical layers of the human auditory cortex are less well studied.

To this end, we investigated the cross-spatial scale cerebral processing of hierarchical auditory sequences in humans by combining whole-brain functional magnetic resonance imaging (fMRI) and layer-fMRI (Fig. 1A, B). The layer-fMRI is also known as laminar-fMRI, in that it resolves cortical layers or lamina at each level in a cortical hierarchy. The emergence of non-invasive layer-fMRI with submillimeter-resolution has allowed researchers to explore functional circuits within the human cortex at a mesoscopic scale, even though it has not reached the microscopic scale of neurons²⁹. For example, the layer-fMRI can decompose the activity of each finger from the primary motor cortex³⁰, and reveal somatosensory and premotor input in superficial layers and cortico-spinal motor output in deep layers³¹. Visual studies with layer-fMRI have shown that the superficial layers of the primary visual cortex are more sensitive to the recognition of facial

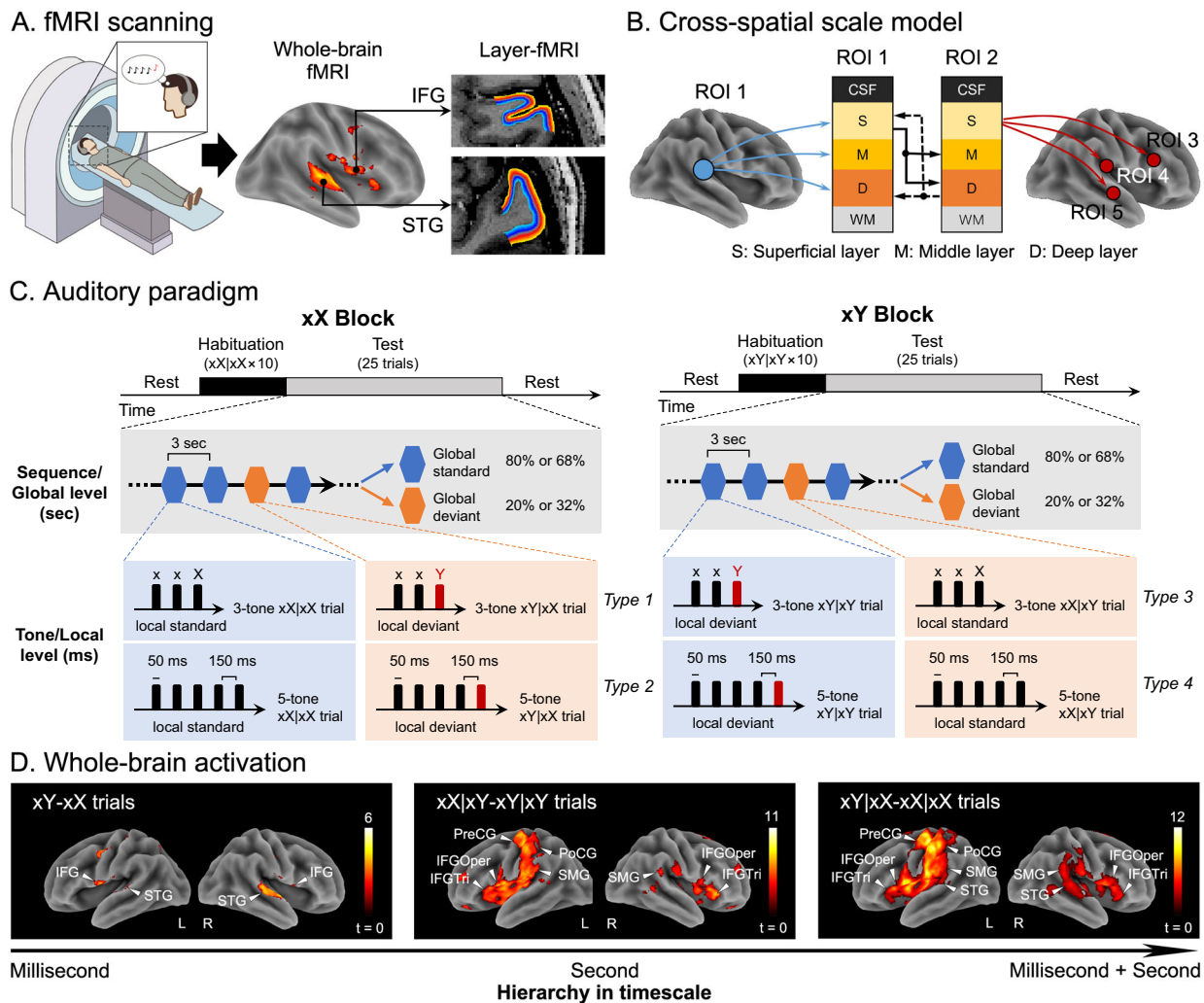


Fig. 1 | Experimental design and fMRI activations across the whole brain. **A** Whole-brain and layer-fMRI scanning, when participants were performing auditory tasks. IFG, inferior frontal gyrus; STG, superior temporal gyrus. **B** Schematic of cross-spatial scale model with whole-brain and layer-fMRI data. ROI, region of interest; CSF, cerebrospinal fluid; WM, white matter. **C** Experimental paradigm of auditory stimuli (for details, see Methods). **D** Whole brain activation of auditory processing in hierarchical ways ($n = 30$ participants). Left, fMRI responses to local

violation of the last tone on the millisecond timescale. Middle, fMRI responses to global violation of sequence on the second timescale. Right, fMRI responses to the violation of both tone and sequence. The threshold for significance was set at $p < 0.05$, cluster-wise FDR corrected. Colorbars indicate t values. Horizontal arrow indicates the hierarchy along the timescale. SMG, supramarginal gyrus, PreCG, precentral gyrus, PoCG, postcentral gyrus, IFGTri, triangular part of IFG, IFGOper, opercular part of IFG.

expression³², and suggested feedforward signals in middle layers and feedback signals in superficial and deep layers³³. From these, we can see that layer-fMRI has a potentially important role in refining functional localization and probing cortical microcircuits. Auditory processing necessitates the involvement of multiple brain hierarchies at different spatial scales, ranging from sub-millimeter cortical layers to cortical regions³⁴. However, little research has investigated hierarchical processing across cortical layers between the auditory cortex and higher-level regions.

In this study, we adopted a modified local-global paradigm³⁵ to explore the neural circuits and microcircuits of hierarchical auditory processing. Using 7 T MRI, we detected the layer-specific activities on two hierarchical areas engaged in human auditory processing, including STG and inferior frontal gyrus (IFG)^{36,37} (Fig. 1A). We first assessed the hierarchical activities of auditory sequences based on whole-brain fMRI data, which confirmed distinct brain networks for auditory processing on two timescales. Next, with layer-fMRI data collected on STG and IFG from the same participant, we detected layer-specific responses to deviants at tone and sequence levels to identify the effective connectivity within and between levels in the auditory hierarchy. Finally, by using searchlight approaches, we searched for regions with the same pattern of layer-specific activities in the whole brain, in order to build a correlation between layer and whole-brain activation, which enabled us to combine mesoscopic and macroscopic findings to investigate the cross-spatial scale representation of hierarchical auditory processing (Fig. 1B).

Results

Experimental design and whole-brain activation

Whole-brain and layer-specific fMRI data were collected in humans during a task of hierarchical auditory sequences (Fig. 1C). An auditory sequence comprised either 3 or 5 pure tones whose carrier frequency was randomly selected from six frequencies (low frequency: 500, 650, 800 Hz; high frequency: 1500, 1800, 2000 Hz). At the local level corresponding to the millisecond timescale, a sequence containing 3 or 5 identical tones was regarded as a local standard, including xxX and xxxxX (both were defined as xX). When the last tone of a sequence was a violation using a carrier frequency selected from the opposite frequency set, the sequence was regarded as local deviant (xxY and xxxxY, both were defined as xY). In the habituation stage, the participants passively listened to 10 sequences that had the same regularity. During the following test stage that corresponds to the second timescale, the participants heard 25 trials, which were either frequent sequences respecting the habituated sequence (defined as global standard) or rare sequences with the same tone number but distinct regularity (defined as global deviant). The probability of sequence violation could be 32% or 20%, and the participants were requested to press a button when they detected the violated sequence. In both 3-tone and 5-tone xX blocks, the xY trial (defined as xY|xX) violated both tone- and sequence-level rules of the xX trial (defined as xX|xX). In both 3-tone and 5-tone xY blocks, the xX trial (defined as xX|xY) only violated the sequence-level rule compared to the xY trial (defined as xY|xY). By randomly selecting the carrier frequencies of the sequences, we can explore brain networks that respond to abstract auditory sequences.

We first determined the activated brain networks of tone and sequence violations using a paired-sample t-test analysis ($p < 0.05$, cluster-wise false-discovery rate [FDR] corrected for multiple comparisons across the brain, see Methods). To detect the response to local deviation, corresponding to the tone violation on a millisecond timescale, we compared xY with xX trials, regardless of tone number. Significant increases in activity induced by the violation of the last tone were found in bilateral STG and IFG (Fig. 1D left). By comparing xY|xX with xX|xX trials, we probed responses to the violations of both tone and sequence, which reflected the auditory processing on both

millisecond and second timescales. The results showed that the higher activations were identified along bilateral temporal-frontal networks, including STG, supramarginal gyrus (SMG), opercular part of IFG (IFGOper), and triangular part of IFG (IFGTri) (Fig. 1D right). In contrast, the effects of sequence violation only (xX|xY > xY|xY) were focused on higher-level regions, exhibiting in bilateral SMG, IFGOper, and IFGTri (Fig. 1D middle). The three types of contrast results localized the hierarchical brain networks for auditory processing of abstract sequences along a timescale from milliseconds to seconds. Besides, for the two types of global deviation, we also found activations in the left precentral gyrus (PreCG) and postcentral gyrus (PoCG), which were evoked by button responses of the right hand.

Hierarchical dynamics across whole brain

To explore the hierarchical dynamics of large-scale brain regions during auditory processing, we applied DCM in each hemisphere separately to assess the causal interactions between activated regions of the corresponding hierarchy and examine the effective connections. To compare the designed models, we used a procedure of random-effects Bayesian model selection, which provided the exceedance probability of each model to indicate the winning model. For auditory processing on a millisecond timescale, we conducted DCM analysis on local deviant trials by assessing four models comprised of STG and IFG (Fig. 2A top). These models could be classified in two ways: 1) STG or IFG input; 2) without, only, or with feedback connections. In both hemispheres, the best model that fitted processing of local/tone information was characterized by signal input from STG and having feedback connections from IFG (Fig. 2A, middle and bottom). Moving to auditory processing on the second timescale, we performed DCM analysis on xX|xY trials to evaluate sequence violation only. Five models that comprised SMG, IFGOper, and IFGTri areas were designed based on input and feedback types (Fig. 2B top). The winning model for bilateral hemispheres indicated that during processing of global/sequence information, the downstream sensory signals were input from SMG and feedforward to IFGOper and IFGTri, and then feedback in the opposite direction (Fig. 2B middle and bottom).

Combining the results of local deviation and global deviation only, we hypothesized that there were two pathways for encoding of local and global auditory information, respectively. One was the ventral pathway, in which the local information propagated from STG to IFGTri; and the other was the dorsal pathway, in which the global information propagated from STG to SMG, IFGOper, and IFGTri. Thus, we performed DCM analysis on xY|xX trials to examine the effective connections across STG, SMG, IFGOper, and IFGTri areas in processing of both tone and sequence violations. Figure 2C showed the specified models and the results of model comparison. The winning model was consistent in bilateral hemispheres, illustrating the bidirectional propagation of neural signals along both dorsal and ventral pathways, which confirmed two processes of tone and sequence information. Referring to the best-fit models obtained under three auditory stimulation conditions, our results decomposed hierarchical architectures of the brain in processing abstract auditory sequences on a tone (millisecond timescale) and a sequence (second timescale), respectively.

Layer-specific activity in temporal and frontal cortices

Consistent with our hypothesis, the STG and IFG served as two core brain regions in the processing of hierarchical auditory sequences. To investigate the layer-specific activation of hierarchical auditory information in these regions, we employed submillimeter ($0.7 \times 0.7 \text{ mm}^2$) layer-fMRI technique at 7 T during the task of auditory sequences. Figure 3A illustrated the scanning localization of layer-fMRI, which simultaneously covered both STG and IFG. Each area was segmented into 20 cortical depths, enabling us to probe neural activities from superficial (close to CSF) to deep (WM) layers in more

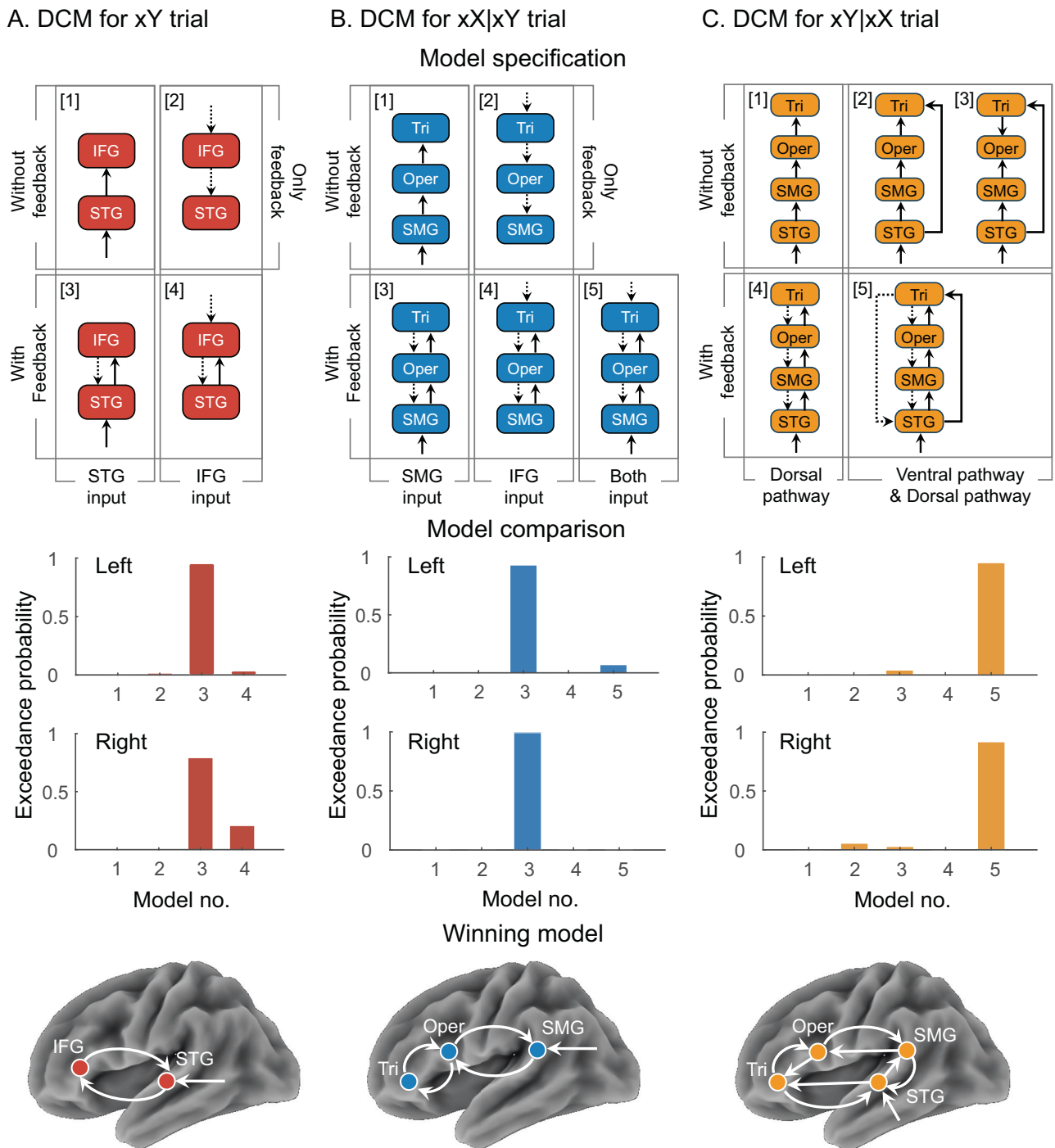


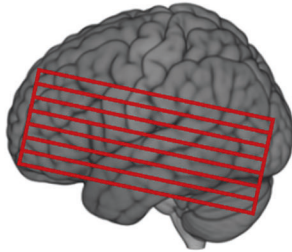
Fig. 2 | Dynamic causal modeling (DCM) of whole brain during hierarchical auditory processing. **A** DCM analysis was conducted on xY trial, assessing the neural connections during the tone violation (n = 28 participants). Top, graphic description of specified dynamic causal models employed for assessment. The models were illustrated by groups of input and connectivity types. The selected regions of interest were based on the fMRI activation of the corresponding comparison. Lines indicate feedforward connections; dashed lines indicate feedback connections. Middle, model comparison results of the left and right hemispheres.

The bars indicate the exceedance probability of each model. The model with the highest probability is the winning model that reflects the directions of effective connections in xY trials. Bottom, the illustration of the winning model. **B, C** DCM analyses were conducted on xX|xY (**B**) and xY|xX (**C**) trials, which evaluated the neural connections during the sequence violation only, and during both tone and sequence violations, respectively. **B, C** with the same format as (**A**). STG superior temporal gyrus, IFG inferior frontal gyrus, SMG supramarginal gyrus, Tri triangular part of IFG, Oper opercular part of IFG.

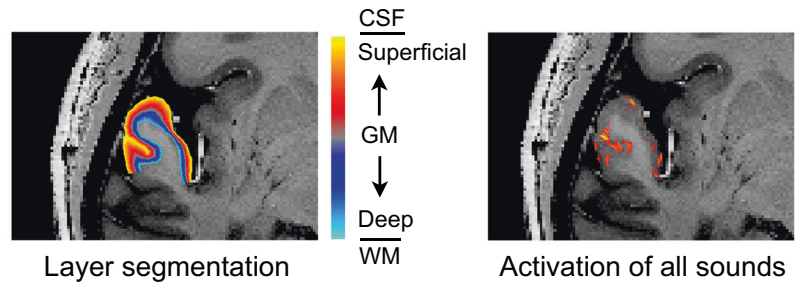
refined structures (Fig. 3B). To ensure that button-pressing responses did not influence neural activities related to auditory stimuli, our further analyses engaging laminar profiling were performed solely on the right hemisphere.

During the auditory processing of local information, the layer-specific activities on STG were shown in Fig. 3C. We observed that the responses to local standard xX trials were strongest in the superficial layers and declined in the middle and deep layers, indicating that

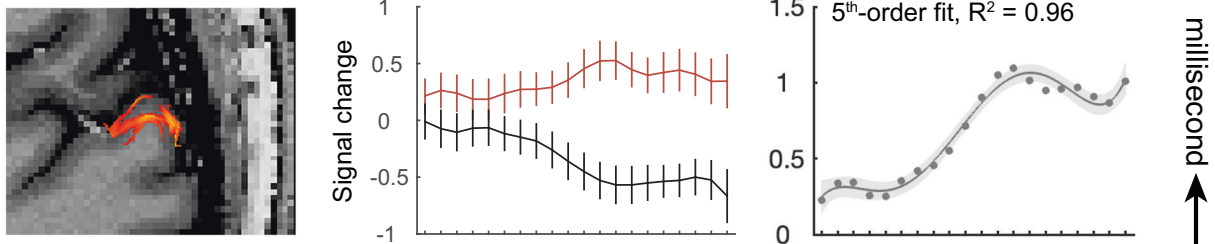
A. Layer-fMRI localization



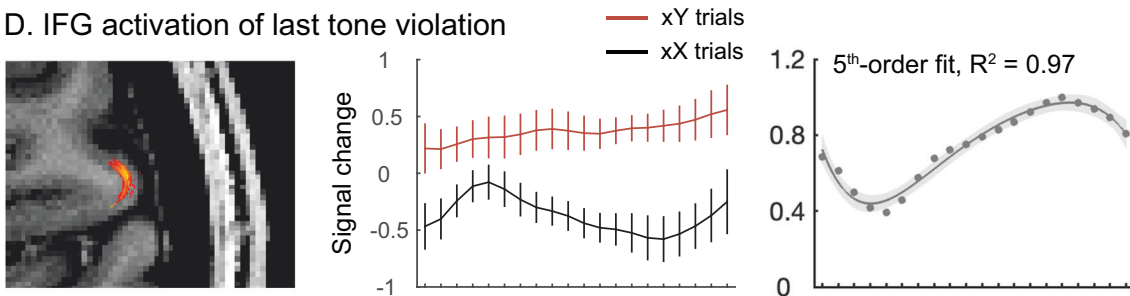
B. Layer segmentation and activation



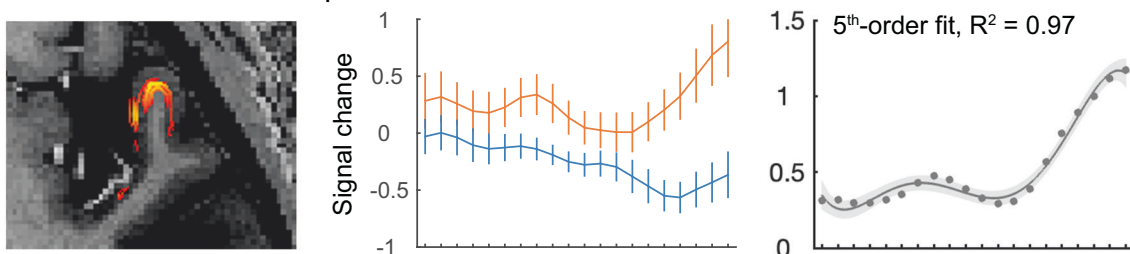
C. STG activation of last tone violation



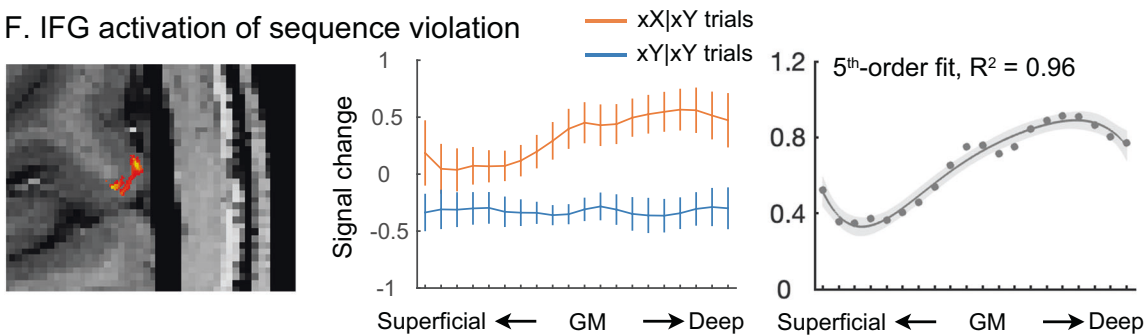
D. IFG activation of last tone violation



E. STG activation of sequence violation



F. IFG activation of sequence violation



auditory input or feedforward signals were activated more in the superficial layers. When comparing xY with xX trials, we performed a fifth-order polynomial model ($R^2 = 0.96$) to fit the activation differences between local deviant and local standard trials across cortical depths, whose predicted results were the best. The fitting curve exhibited that the response differences were primarily over the middle and deep layers. In the higher-level areas of the temporal-frontal

network, the activity profile of tone violation showed a clear increase from superficial to deep layers in IFG (Fig. 3D). The fifth-order polynomial fitting ($R^2 = 0.97$) was also used to predict the activation differences, and we found a stable increase from the middle to deep layers.

Turning to the auditory processing of global information, we compared the layer-specific activities between xX|xY and xY|xY trials,

Fig. 3 | Layer-specific activations of auditory processing in auditory and frontal cortices. **A** The scanning localization of layer-fMRI. The localization of layer-image acquisition was manually adjusted for each participant. **B** An example of layer segmentation and the activation of all sound stimuli across auditory cortical layers, regardless of sequence type. Left, the ROI was drawn on the structural image, and the cortical regions were segmented into 20 layers using the LayNii package. Right, the activation results of auditory cortical layers at a representative slice. CSF, cerebrospinal fluid; WM, white matter; GM, grey matter. **C, D** The layer activation of local contrasts on regions of superior temporal gyrus (STG, C) and inferior frontal gyrus (IFG, D), comparing the local deviant with the local standard trials ($n = 28$ participants). Left, the layer activation of a representative slice ($p < 0.001$,

uncorrected for multiple comparisons, one-sided). Middle, the activity profile of each trial type, using smoothed signal data with z-score normalization. Error bars indicate the SE. Data are presented as mean values \pm SE. Right, the activity profile of the difference between local deviant and standard trials, using unsmoothed signal data with z-score normalization for fitting. Dots show averaged values of individual differences. Shaded areas indicate the CI. Polynomial models explain the changes in difference as a function of layer. The coefficient of determination (R^2) shows the fitting performance. **E, F** The layer activation of global contrasts on regions of STG (**E**) and IFG (**F**). The results were calculated by comparing $xX|xY$ with $xY|xY$ trials. **E, F** with the same format as (**C, D**).

which corresponded to the sequence violation only. The responses on the STG (Fig. 3E) showed that the averaged activity profile for global standard sequences was strongest in the superficial layers, followed by the middle layers, confirming the input way of auditory signals across cortical depths. However, the activation differences between global deviant and global standard trials (fifth-order polynomial fitting, $R^2 = 0.97$) were greatest in the deep layers, but relatively weak in the superficial and middle layers. Similar to the responses at the local level, the difference between $xX|xY$ and $xY|xY$ trials (fifth-order polynomial fitting, $R^2 = 0.96$) activated more in the middle and deep layers of IFG (Fig. 3F). Combining the results of local and global violations, the middle and deep layers of IFG can be associated with the targets of forward and backward signals, respectively.

Bidirectional dynamics across cortical layers

Although we observed layer-specific responses to local standard, local deviant, global standard, and global deviant conditions on STG and IFG areas, respectively, we were still unable to determine the role of different cortical depths in bidirectional effective connectivity. We thus performed DCM analyses on layer signals to estimate the bidirectional dynamics across STG and IFG layers. We split both the STG and IFG regions into three parts: the 7 cortical depths closest to the WM were identified as the deep layers, the 7 cortical depths closest to the CSF were identified as the superficial layers, and the other 6 cortical depths were identified as the middle layers. Then we designed 9 models according to previous auditory research^{22,28,38}, which potentially reflected the dynamic connections across the layers from two hierarchical regions.

The common features of these designed models (Fig. 4A) could be summarized by: 1) bidirectional connections across the superficial, middle, and deep layers within the STG; 2) bidirectional connections across cortical depths within the IFG; 3) feedback connections from the deep layer of IFG to the superficial and deep layers of STG. Besides, two factors were proposed to determine the classification of these models, resulting in a 3×3 pattern. One factor was related to the input on the STG layers, including signals input from superficial, middle, or deep layers; the other one was related to the feedforward connection from STG to IFG, including feedforward signals from superficial layers of STG to middle layers of IFG, to deep layers of IFG, or to both middle and deep layers of IFG. This led to a comprehensive set of candidates to explore the layer-specific connections between STG and IFG during auditory processing at local or global levels. Therefore, we applied these models to both local deviant (xY) sequences and global deviant ($xX|xY$) sequences.

Our results showed that at different levels of hierarchical auditory processing, the difference was mainly in the manner of signal input, while the patterns of feedforward connections were consistent. Specifically, the winning model at the tone level showed that the signal input from the superficial layer of STG took dominance, while the deep layer may also contribute to the input (Fig. 4B). Then, the information of local violation propagated from the superficial layers of STG to the middle and deep layers of IFG. At the sequence level (Fig. 4C), the signals were changed to the input from the superficial and middle

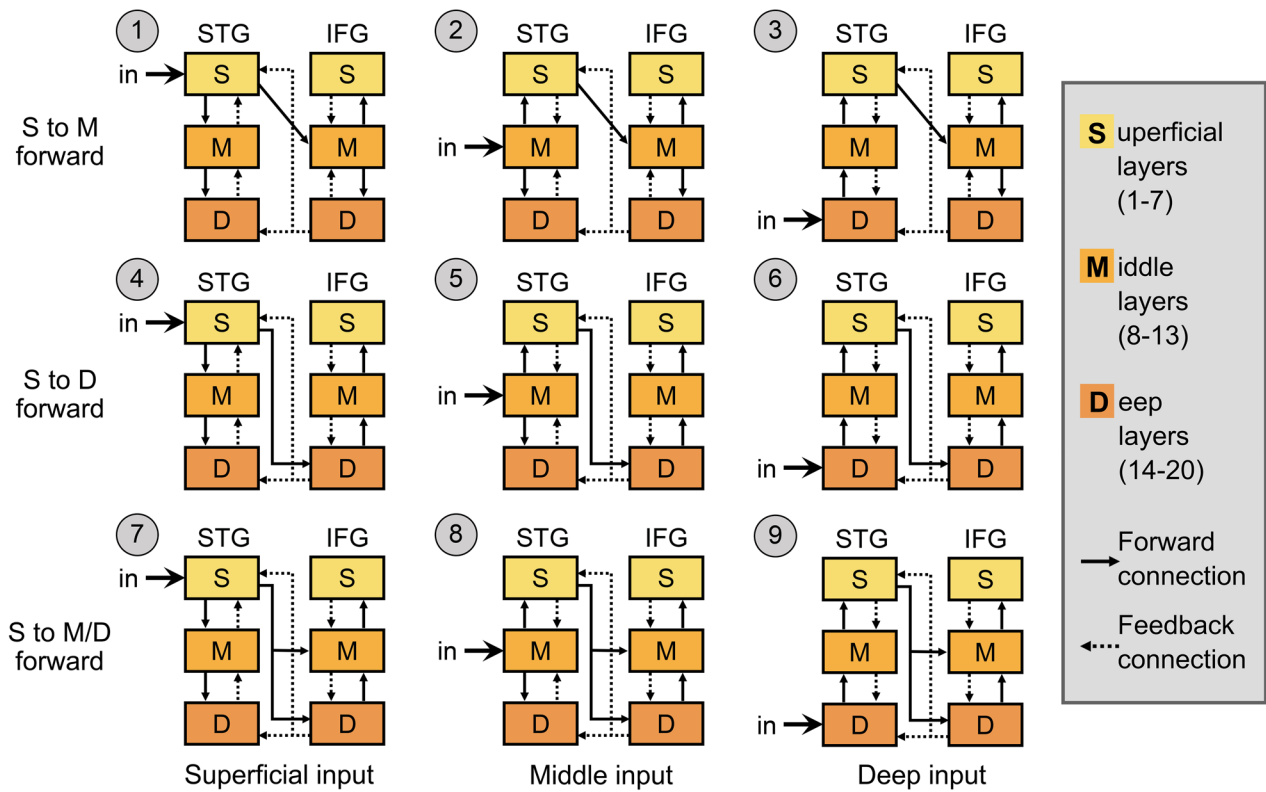
layers of STG, where the middle layer input was more dominant, suggesting that compared with the reception of various auditory signals in the superficial layer of STG, the middle layers were mainly responsible for receiving higher-order information of auditory sequences. However, the information of global violation still propagated through feedforward connections from the superficial layers of STG to the middle and deep layers of IFG, which confirmed the roles of the middle and deep layers of IFG in the processing of auditory violations. These findings suggest that, in the fine structure, the difference in hierarchical information propagation is reflected in the layer-specific processing of sensory input; however, the propagation pattern is similar across brain regions.

Cross-spatial scale correlation between cortical layer and whole brain

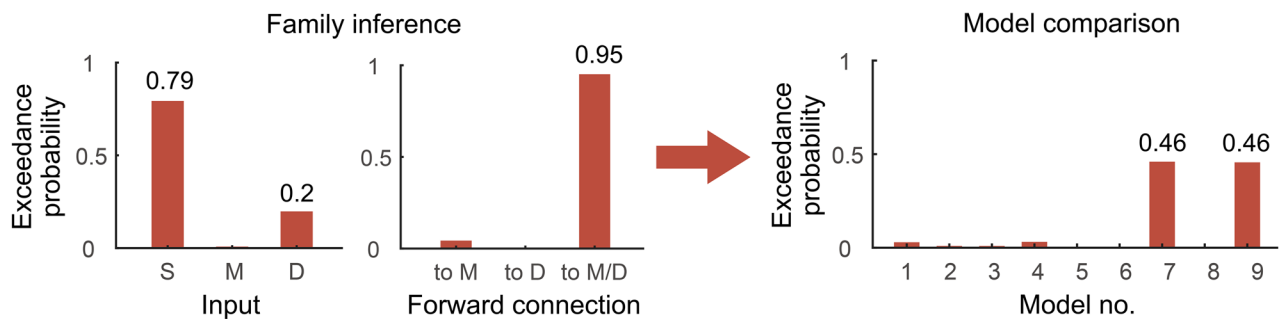
To bridge the neural activities of cortical layers on the mesoscale and those of whole-brain areas on the macroscale, we then applied the searchlight-based representational similarity analysis (RSA) between whole-brain and layer-fMRI data (Fig. 5A), which enabled us to figure out the relationship of functional anatomies or connection architectures at different spatial scales. For both neural activities at local and global levels, we computed the representational dissimilarity matrices (RDMs) across corresponding conditions at each layer depth (divided into superficial, middle, and deep layers), respectively. By using a searchlight method on whole-brain activation data, we estimated the similarity of neural responses on different spatial scales by computing Spearman's rho between the RDMs extracted from layer-fMRI data and those from whole-brain fMRI data for each participant. The brain regions that showed significant similarities to layer-specific activities were determined by a one-sample t-test ($p < 0.05$, cluster-wise FDR corrected for multiple comparisons across the brain). Due to weak responses to local standard and local deviant trials on the millisecond timescale, no significant region was found by the searchlight approach at local level.

We next explored the cross-spatial scale responses to global violation by testing the similarity between layer-specific and whole-brain activities. The searchlight-based RSA showed that the whole-brain activities detected by the superficial, middle, and deep layers of STG had similar patterns (Fig. 5B left). All layers had significant similarities of activations with STG, SMG, IFGOper, and IFGTri. In the investigation of IFG layers, the similarity patterns of the whole brain detected by the superficial and deep layers were more robust, which were observed in STG, SMG, and IFGOper (Fig. 5B right). Sequentially, we tested the difference in similarities of activation patterns among three depths of the layers. For STG layers (indicated the feedforward flow), the activation similarity between middle layers and IFGTri was significantly higher than that of superficial layers (Fig. 5C top), corresponding to the dominant role of the middle layers of STG in the input of sequence information. For IFG layers (indicated the feedback flow), the activation similarity between deep layers and SMG was significantly higher than that of middle layers (Fig. 5C bottom), highlighting the communication of deep layers with the primary auditory cortex.

A. Model specification



B. DCM results at tone level



C. DCM results at sequence level

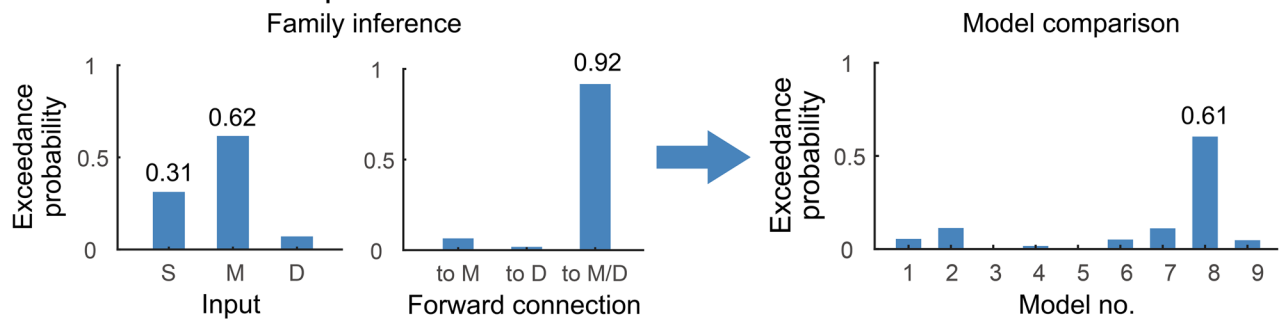


Fig. 4 | Layer-specific dynamic causal modeling (DCM) of hierarchical auditory processing across STG and IFG. A Graphic description of specified models employed for DCM analysis across cortical layers. STG, superior temporal gyrus; IFG, inferior frontal gyrus. **B** Layer-specific DCM results of tone-level violation (n = 26 participants). The analysis was conducted on xY trials. Left, family-wise exceedance probability for network architectures characterized by three types of signal input. ‘S’, superficial input; ‘M’, middle input; ‘D’, deep input. Middle, family-

wise exceedance probability for network architectures characterized by two types of feedforward connection. ‘to M’, superficial to middle layer feedforward connection; ‘to D’, superficial to deep layer feedforward connection; ‘to M/D’, superficial to middle and deep layers feedforward connections. Right, model exceedance probability of DCM analysis, which showed the best model with the highest probability. **C** Layer-specific DCM results of sequence-level violation, performing the analysis on xx|xY trials only. **C** with the same format as (B).

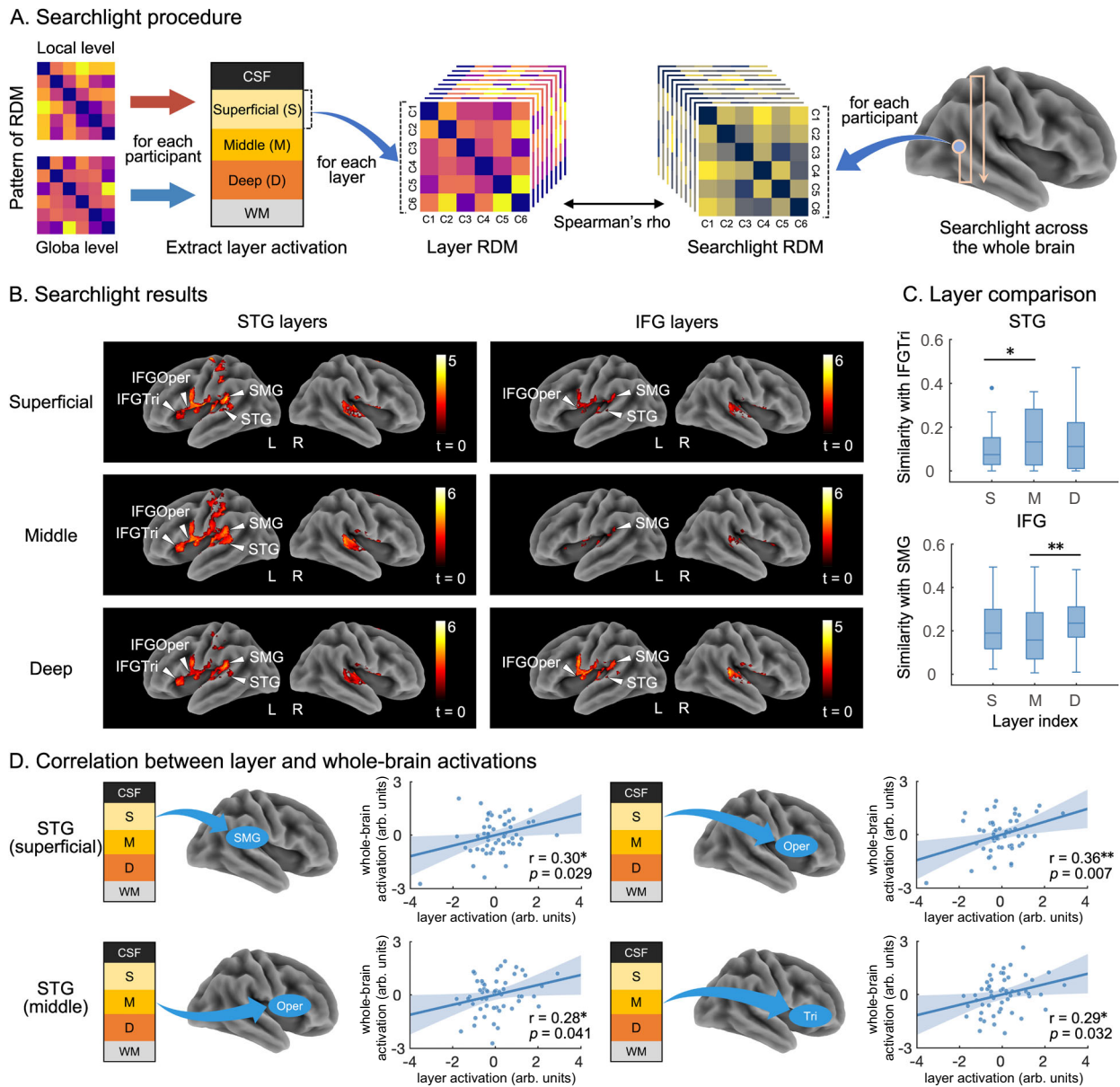


Fig. 5 | Cross-spatial scale correlation between layer-specific and whole-brain activities. **A** Searchlight procedure for detecting the ROIs across the whole brain with similar features to the layer-specific activity (for details, see Methods). The correlation coefficient used for statistics was computed between representational dissimilarity matrices (RDMs) extracted from layer-fMRI and whole-brain fMRI data for each participant. **B** Searchlight results of the whole brain using RDMs extracted from each type of layer ($n = 27$ participants). The threshold was set at $p < 0.05$ (cluster-wise FDR corrected). Colorbars indicate t values. STG, superior temporal gyrus; IFG, inferior frontal gyrus; IFGTri, triangular part of IFG; IFGOper, opercular part of IFG. **C** Comparison of correlation coefficients across the superficial, middle,

and deep layers using paired-sample t-test (two-side). The individual correlation coefficient was averaged from searchlight results. Boxes represent interquartile range (IQR), central lines indicate the median, whiskers indicate $1.5 \times$ IQR, and dot symbols indicate outliers. Between STG superficial and middle layers $p = 0.0442$. Between IFG middle and deep layers $p = 0.0057$. **D** Activation correlations between each type of layer and whole-brain ROIs. The ROIs were defined by layer-specific and whole-brain activations. Dot indicates the activation value extracted from each participant. Tri triangular part of IFG, Oper opercular part of IFG. Shaded areas indicate the CI. Data are presented as mean values \pm CI.

We can see that the whole-brain similarity patterns detected by the searchlight-based RSA using layer-specific RDMs were similar to the neural activities of global violation; therefore, we further examined the correlation between layer-specific and whole-brain activations. By extracting the activated difference between $xY|xY$ and $xY|xY$ trials, we found that the activation of the superficial layers of STG was positively correlated with SMG and IFGOper (Fig. 5D top), while the activation of the middle layers of STG was positively correlated with IFGOper and IFGTri (Fig. 5D bottom). Combining with the whole-brain and layer-specific DCM results, our correlation results implied that the

superficial layers of STG, which were more responsible for feedforward connections, modulated the activities of the brain areas at the next level. The middle layers of STG, which were more responsible for the input of higher-order auditory processing, had a greater impact on higher-level brain regions.

Discussion

Our study employed ultra-high field 7 T MRI to investigate the hierarchical processing of abstract auditory sequences across spatial scales in the human brain. We showed distinct layer-specific activities

for the processing of standard and novel stimuli in STG and IFG. By estimating the bidirectional dynamics across STG and IFG layers, we found that sensory input of local information to STG was from the superficial layers, while global information was predominantly received in the middle layers. At both local and global levels, the feedforward connections propagated from the superficial layers of STG to the middle and deep layers of IFG, whereas feedback connections from the deep layers of IFG reciprocally modulated superficial and deep layers of STG. Furthermore, the integration of whole-brain and layer-specific activation bridged mesoscopic laminar activity to macroscale networks, highlighting the roles of superficial and middle layers of STG in the interaction with higher-level frontal regions, as well as deep layers of IFG engaged with auditory processing in sensory areas. Together, these findings elucidate a multiscale hierarchical architecture of brain organization, in which laminar microcircuits and distributed networks dynamically interact to support bidirectional processes of auditory information.

Hierarchical auditory processing across cortical layers

We examined refined auditory processing at submillimeter scale across cortical layers of the STG and IFG, both involved in whole-brain networks identified on millisecond (local, tone-level) and second (global, sequence-level) timescales²¹. Our observation in the STG revealed distinct layer-specific activations for standard and violated stimuli. The standard stimuli evoked the strongest activity in the superficial layers, regardless of whether they were at local or global levels, which may indicate the signal input of auditory information based on the predictive coding hypothesis³⁹. Similar results were reported in previous laminar findings, with auditory responses (noise burst or auditory motion) increased from deep to superficial layers in Heschl's gyrus, Heschl's sulcus, planum temporale, and posterior part of STG^{40,41}, which confirmed the functional activity pattern of superficial layers⁴². Nevertheless, the activation difference between deviant and standard trials was larger in middle and deep layers of STG at the local level and the greatest in deep layers at the global level, suggesting that the deep layers of STG are associated with awareness of violations on both time scales. These findings are in line with a laminar study of the auditory perceptual awareness, which demonstrates that the detection of target sounds in noise leads to increased activity in deep layers of the auditory cortex⁴³.

To our knowledge, the layer-specific activity of the frontal cortex has seldom been investigated in sensory processing, which is an important part of hierarchical organization⁴⁴. We identified a stable increase from middle to deep layers of IFG in the activation difference between deviant and standard stimuli at both local and global levels. According to previous studies, the IFG has bidirectional connections with auditory cortex in hierarchical auditory processing^{45,46}, which has also been confirmed by our whole-brain results. Thus, the middle and deep layers of IFG should play important roles in receiving feedforward signals from the auditory cortex and outputting feedback signals simultaneously. Studies on working memory suggest that the increased activity in deep layers of IFG may also be related to behavioral responses^{47,48}. However, the raised activation difference of deep layers during local violation verified its function in auditory processing, in which no behavioral response was required.

Layer-specific connections across hierarchical brain areas

An important key of our study is to evaluate the dynamic connections across cortical layers, which provide refined neural architectures for hierarchical auditory processing. The fMRI signal itself is more sensitive to synaptic input than to output, meaning that the fMRI detects receiving activity more clearly. By performing DCM on fMRI signals, we can assess the effective connectivity between paired regions in both feedforward and feedback directions, thus describing the causality of interactions between neurons or populations of neurons. The DCM

also enabled us to clarify that the neural activity we observed was not induced by the draining vein effect. We revealed different arrangements of sensory input to the cortical layers of STG for auditory processing at hierarchical levels. Our results are consistent with other research that the local/tone signals are input from superficial layers of STG^{22,25}, and correspond to results that local input yielded activity in superficial layers. Furthermore, our findings extend the evidence of auditory input of higher-order information by demonstrating that the global/sequence signals are mainly input from middle layers, and may also be input from superficial layers, which is in keeping with results that global input evoked activities in superficial and middle layers. The distinct contributions of cortical layers in the processing of sensory information on different timescales could also be observed in visual and somatosensory cortices^{49,50}. Particularly, in the task of finger tapping, which has a shorter duration, the sensory input was detected in the superficial layers of primary motor cortex; while in the task of index finger stroking prediction, which has a longer duration, the sensory input was detected in the middle layers of primary somatosensory cortex²⁹.

Layer-dependent analysis enabled us to observe cortical circuit models of human cognition²⁹. However, to date, most research is limited to one unique sensory cortex, including auditory⁴², visual⁵¹, and somatosensory cortices⁵². In the present study, we showed the cortical circuit models across STG (lower-level sensory area) and IFG (higher-level cognitive area), in which the feedforward and feedback connections were the same at local and global levels. By comparing three classes of feedforward connections in layer-specific models, we confirmed that the feedforward signals originate from the superficial layers of STG and transmit to the middle and deep layers of IFG; meanwhile, the feedback connections propagate from the deep layers of IFG to the superficial and deep layers of STG. Moreover, we also found that the probability of a feedforward connection reaching the middle layers of IFG was higher than that reaching the deep layers, indicating that the feedforward connection reaching the middle layers was more dominant. The determined model of bidirectional connections in our study is in accordance with cortical circuit models proposed by other studies for hierarchical processing in sensory areas^{38,53}, implying conserved features of sensory processing across brain regions. Even though the feedback connections were predefined, according to previous studies of predictive coding, the superficial and deep layers of somatosensory⁵² and visual⁵³ cortices were also modulated by feedback signals from higher-level areas. Our findings provided potential cortical circuit models between lower-level sensory and higher-level cognitive cortices.

Integration of mesoscale and macroscale neural activities in auditory processing

Although the feedforward and feedback connections among hierarchical regions have been verified by predictive coding and related theories for decades^{45,54,55}, the integrative processes coordinating cortical layers with cerebral cortices are rarely studied⁵⁶. A crucial goal of this research was to link the mesoscopic operations within cortical layers to the macroscopic engagement of large-scale brain networks during hierarchical auditory processing. By using searchlight-based RSA, we identified whole-brain regions whose activity patterns were associated with those observed within specific cortical layers of the STG and IFG. Our observation at the global level uncovered that activity patterns across all layers of STG were significantly similar to the activity in the auditory network comprising the STG, SMG, IFGOper, and IFGTri. Similar results were also found in superficial and deep layers of IFG, which is a classical pathway for sequence processing^{57,58}, and consistent with our whole-brain activation at the global level. Further, the direct comparison of similarity of activity patterns across cortical layers provided evidence to support our layer-specific DCM findings that the middle layers of STG play a dominant role for the input of higher-order

sequence information, while the deep layers of IFG are associated with feedback projections to auditory processing areas⁵⁹.

Moreover, the direct correlational analyses between layer-specific and whole-brain activations further proved these cross-spatial scale links. Activity in the superficial layers of STG was positively correlated with that in the SMG and IFGOper, corresponding to its role in feed-forward processing, which modulated activity in the cortical areas at the next level of the dorsal pathway. On the other hand, activity in the middle layers of STG, which was considered critical for the input of higher-order auditory information, showed positive correlations with IFGOper and IFGTri activations, indicating its considerable influence on higher-level frontal regions during the processing of abstract auditory sequences. Together, our findings demonstrate that distinct cortical layers within auditory and frontal cortices are associated with neural dynamics across distributed brain networks, revealing how fine-grained laminar computations contribute to hierarchical sensory processing in large-scale brain networks.

In summary, our study integrated macroscopic and mesoscopic neural networks to reveal the cross-spatial scale processing of hierarchical auditory sequences in human brains, providing a new insight for research of multi-scale interactions in auditory cognition and non-invasive imaging technology.

Methods

Participants

Thirty participants (ages 18–36 years, 13 female) were recruited from local communities. All participants reported normal hearing and had no history of psychiatric or neurological illness. In whole-brain analyses, all participants were included in the activation analysis; 2 participants were dropped in the DCM analysis due to insufficient runs. In layer-fMRI analyses, 2 participants voluntarily exited layer-fMRI scanning, resulting in 28 effective participants included in laminar activation profiling; 2 additional participants were excluded from laminar DCM analysis due to voluntary early termination of the experiment. For cross-spatial scale MVPA analysis, there were a total of 27 participants who had participated in both whole-brain fMRI and layer-fMRI scanning. The study was conducted in accordance with a protocol approved by the ethics committee of Fudan University [no. FE23164]. All participants gave written informed consent.

Experimental design and stimuli

All participants were asked to complete a modified local-global paradigm (Fig. 1C) task during fMRI scanning. A single trial consisted of a sequence of either 3 or 5 tone stimuli, each lasting for 50 ms with an interval of 150 ms. A flexible duration of silence was set at the end of each sequence, making each trial last for 3 seconds. The carrier-frequency of tone stimuli was randomly selected from six frequencies (low frequency: 500, 650, 800 Hz; high frequency: 1500, 1800, 2000 Hz) and presented using Psychtoolbox (<http://psychtoolbox.org>). At the local on-millisecond timescale, a sequence containing tones with an identical carrier frequency was defined as a local standard trial, e.g., xxX and xxxxX (both referred to as xX). When the carrier-frequency of the last tone was selected from the opposite frequency set and thus became a violation, the sequence was defined as a local deviant trial, e.g., xxY and xxxxY (both referred to as xY). Each block was divided into the habituation stage and the test stage, followed by a 12-second rest. In the habituation stage, a participant passively heard 10 trials of abstract sequences that followed the same regularity, which determined the global standard trials (e.g., xX|xX trials in 3-tone xX block) in the test stage. In the following test stage that corresponds to global level on second timescale, the participant heard 25 sequences. A rare sequence with the same tone number but violating the abstract regularity of habituation trials was defined as a global deviant trial (e.g., xY|xX trials in 3-tone xX block). The global deviant trials were presented randomly, having a probability of either 20% or 32%. By using 4

types of regularities and 2 types of deviant probabilities, the paradigm had 8 conditions of blocks. Each run had 4 blocks, and 4 runs were collected for both whole-brain fMRI and layer-fMRI scanning, respectively. Thus, each session could collect 2 blocks of each condition, and the order of blocks was randomized and counterbalanced within and across subjects. To avoid sleeping, the participants were instructed to press a button when they detected a sequence violation.

Data acquisition

MRI data were acquired using a 7 T MRI system (Siemens, Terra, Erlangen, Germany) equipped with a 32-channel head coil (Nova Medical, Wilmington, MA, USA). For whole-brain scanning, anatomical images were acquired using an MP2RAGE sequence (TR = 3800 ms, TE = 2.31 ms, 0.70 mm isotropic voxels, 256 slices) at two inversion times (TI of 800 ms, 2700 ms with a flip angle of 7°, 5°, respectively); functional images were acquired using T2*-weighted gradient-echo EPI sequence (1.6 mm isotropic voxels, TR = 1000 ms, TE = 12.8 ms, flip angle 39°, 100 slices, 483 scans). In addition, a B0 field map was collected for unwarping distortions (4 mm isotropic voxels, 40 slices, TR = 620 ms, TE1 = 4.08 ms, TE2 = 5.1 ms, flip angle 39°). For layer-fMRI scanning, anatomical images were acquired using MP2RAGE sequence (TR = 5200 ms, TE = 3.35 ms, 0.3 × 0.3 × 1 mm voxels, 36 slices) at two inversion times (TI of 1090 ms, 3250 ms with a flip angle of 6°, 7°, respectively); functional images were acquired using 3D-VASO sequence⁶⁰ (0.7 × 0.7 × 2 mm voxels, TR = 2905 ms, TE = 32.10 ms, TI = 1210 ms/2210 ms with a flip angle of 60°, 20 slices, 167 scans), which acquired the VASO and blood oxygen level dependent (BOLD) signals in an interval way.

Data analysis

Whole-brain fMRI data were preprocessed with MATLAB R2022a (MathWorks, Natick, MA, USA), Statistical Parametric Mapping (SPM12, <http://www.fil.ion.ucl.ac.uk/spm>) and the CONN-fMRI toolbox (version 22.a, <https://www.nitrc.org/projects/conn>) using the standard pre-processing pipeline⁶¹. The initial 3 non-steady scans in each functional run were discarded first. The remaining scans were slice-time corrected, realigned to the first scan of the series, and integrated field-maps for susceptibility distortion correction. Potential outlier scans were identified using ART⁶² for scrubbing. Functional and anatomical data were coregistered, segmented into WM, GM, and CSF, and normalized to the Montreal Neurology Institute (MNI) space with isotropic voxels of 1.5 mm³. The normalized fMRI data were spatially smoothed using a 6-mm full-width at half maximum Gaussian kernel. Finally, the functional data were denoised by the regression of potential confounding effects, including WM timeseries, CSF timeseries, head motion parameters, outlier scans, and linear trends within each functional run, followed by high-pass frequency filtering of the functional timeseries above 1/128 Hz.

Layer-fMRI data were preprocessed with NiPype⁶³, SPM, and AFNI^{64,65} for a customized pipeline. Considering the impact of the short duration of stimuli on signal intensity, we only used BOLD signals for further analyses. The functional images were realigned to the first scan of the series for correction of head motion. The corrected functional data were denoised. The structural images were coregistered to the first scan of functional images. Then, the structural and functional images were upsampled to 0.2 mm³ voxels using AFNI. Afterwards, the regions of interest (ROIs) on GM were manually drawn individually based on the structural image using FSLeves⁶⁶ around IFG and STG regions. Lastly, the cortical cortices were segmented into 20 layers using LayNii⁶⁷ and used as the laminar mask for profile extraction of layer-specific activity.

Statistical analysis of fMRI responses

For both whole-brain fMRI and layer-fMRI data, a general linear model (GLM) for individual analyses was established using stimulus types,

including 3-tone and 5-tone habituation, $xX|xX$, $xY|xX$, $xY|xY$, $xX|xY$, together with head motion parameters as regressors. Subsequently, the beta-weighted images yielded by the GLM, which represented activation for each stimulus type relative to rest, were used to fit a paired-sample t-test second-level SPM model for group-level analyses of whole-brain fMRI data. The contrasts were defined as follows: the responses to local/tone violation on millisecond timescale, $xY - xX$ trials combining 3-tone and 5-tone sequences; the response to both local and global violations, $xY|xX - xX|xX$ trials combining 20% and 32% probabilities; the response to global/sequence violation on second timescale only, $xX|xY - xY|xY$ trials combining 20% and 32% probabilities. The statistical thresholds of all contrasts were set at $p < 0.05$ (cluster-wise FDR-corrected).

To describe the group-level changes of laminar-specific fMRI responses to each stimulus type, we extracted the laminar activation by applying the laminar mask to layer-fMRI data individually. According to the GLM analyses, there was an individual difference in the location of layer-specific activation; thus, the activation extraction was conducted on the target slice for each participant, where the maximum sum of difference between corresponding violation and standard trials was detected. Then, the laminar activation was averaged for group-level profiling. Finally, to characterize the changes of activation differences (including $xY-xX$ trials and $xX|xY-xY|xY$ trials) along with cortical depths (from superficial to deep layers), we performed polynomial models to fit the averaged differences of activation with 20 layers. To reduce the influence of outliers, we used a robust regression with a bisquare kernel function. The coefficient of determination (R^2) was used to indicate the fitting performance. The 5th-order polynomial model with the highest R^2 was determined as the final fitting model.

Dynamic causal modeling analysis

In our study, the DCM analyses were conducted using SPM and followed the standard procedure: ROI definition, extraction of time series, model specification, model estimation, and model comparison. For the whole-brain DCM analyses, the ROIs were defined by the corresponding results of hierarchical activations. For each participant, we searched for the nearest voxel with peak intensity, which was inside the corresponding activated area, and extracted a sphere of 5-mm radius using the nearest voxel as center. Then, a conjunction of this sphere with the anatomical area in MNI space was defined as the ROI for DCM analysis. For the local level, we designed four models using STG and IFG (Fig. 2A). The models were classified as: 1) STG or IFG input; 2) without, only, or with feedback connections. For global level only, we designed five models using SMG, IFGOper, and IFGTri (Fig. 2B). The models were classified as: 1) SMG input, IFG input, or both input; 2) without, only, or with feedback connections. For both local and global levels, we designed five models using STG, SMG, IFGOper, and IFGTri (Fig. 2C). The models were classified as: 1) dorsal pathway only or both dorsal and ventral pathways; 2) without or with feedback connections.

For the layer-specific DCM, the ROIs were defined by splitting the 20 layers of STG and IFG into 3 parts, respectively, including superficial layer: 1-7; middle layer: 8-13; deep layer: 14-20. Each part was merged into a single mask. Then, the time series was extracted using the defined mask on the preprocessed coregistered functional images. According to previous auditory research^{22,28,38}, we designed nine DCMs (Fig. 4A) to test cross-layer connections between STG and IFG for each hierarchical response. The common features of these models were described as follows: 1) bidirectional connections from superficial to deep layers within STG and IFG, respectively; 2) feedback connections from deep layers of IFG to the superficial and deep layers of STG. The distinct features used to classify these models were described as follows: 1) signal input from superficial, middle, or deep layers; 2) feed-forward connections from superficial layers of STG to middle layers of IFG, to deep layers of IFG, or to both middle and deep layers of IFG.

Finally, according to the standard procedure, for both whole-brain and layer-specific DCM analyses, the DCMs were estimated individually. Then, the random-effects Bayesian model selection was used to test the more reasonable class of models and identify an optimal model for each hierarchy.

Statistical analysis of cross-spatial scale correlation

In order to bridge the layer-specific and whole-brain activities, we performed searchlight-based RSA (Fig. 5A) by using RSAToolbox (<https://github.com/rsagroup/rsatoolbox>) to search for whole-brain voxels with similar activation patterns to each layer. By using one out of three types of layer masks (superficial, middle, deep layers), we extracted voxel-wise activities, and then computed the correlation coefficients across contrasts to establish RDM, resulting in a 6×6 matrix with 0 on the diagonal. Since the contrasts were generated by combining all runs of the same type of fMRI data, it could help to eliminate the noise hidden in each run. At local level, the contrasts used were 3-tone xX , 5-tone xX , 3-tone xY , 5-tone xY , all xX , and all xY trials. At global level, the contrasts used were $xX|xY$ (20%), $xX|xY$ (32%), $xY|xY$ (68%), $xY|xY$ (80%), all $xY|xY$, all $xX|xY$ trials. The last 2 contrasts of each matrix were appended. Since the contrasts were intrinsically coupled (e.g., 3-tone xX always presented with 3-tone xY), we added these run/block-irrelevant features to avoid overfitting on within-run/block noises other than the targeted contrasts and reduce the noise impact on the correlation coefficient. Whole-brain ROIs were extracted by searchlight method with a 5-mm radius sphere, and then the RDMs were computed with the same method. Subsequently, we computed Spearman's rho between layer-specific and whole-brain RDMs, and applied Fisher's Z transformation to convert the results to a normal distributed z-score. We used SPM to conduct one-sample t-test for group-level analysis. The threshold was set at $p < 0.05$ (cluster-wise FDR corrected). By using the whole-brain ROIs with significantly positive coefficients, we extracted the averaged values from searchlight results of each layer. A paired-sample t-test was conducted to test the difference in activation similarities among the three types of layers.

Finally, we performed Pearson's correlation to examine the relationship between layer-specific and whole-brain activations. The ROIs were defined by the fMRI responses of the corresponding contrast. We extracted the activation values of each contrast individually and computed Pearson's correlation between one type of layer and a whole-brain ROI.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw data of this study are available from the corresponding author upon request. The source data that support the findings in main figures are provided with this paper. Source data are provided with this paper.

Code availability

fMRI data analyses were performed in the freely available toolbox Matlab R2022a, CONN-fMRI v.22.v2407, SPM 12, NiPy 1.9.1, LayNii 2.7.0, FSLeves 1.11.0, AFNI 24.3.06, and RSAToolbox 0.2.0. The software code that supports the findings of this study are available from the corresponding author upon request.

References

- de Heer, W. A., Huth, A. G., Griffiths, T. L., Gallant, J. L. & Theunissen, F. E. The Hierarchical Cortical Organization of Human Speech Processing. *J. Neurosci.* **37**, 6539–6557 (2017).
- Dehaene, S., Meyniel, F., Wacongne, C., Wang, L. & Pallier, C. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* **88**, 2–19 (2015).

3. Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S. & Sablé-Meyer, M. Symbols and mental programs: a hypothesis about human singularity. *Trends Cogn. Sci.* **26**, 751–766 (2022).
4. Schmitt, L. M. et al. Predicting speech from a cortical hierarchy of event-based time scales. *Sci. Adv.* **7**, eabi6070 (2021).
5. Uddén, J., de Jesus Dias Martins, M., Zuidema, W. & Tecumseh Fitch, W. Hierarchical structure in sequence processing: how to measure it and determine its neural implementation. *Top. Cogn. Sci.* **12**, 910–924 (2020).
6. Sydnor, V. J. et al. Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* **109**, 2820–2846 (2021).
7. Norman-Haignere, S. V. et al. Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nat. Hum. Behav.* **6**, 455–469 (2022).
8. Denham, S. L. & Winkler, I. Predictive coding in auditory perception: challenges and unresolved questions. *Eur. J. Neurosci.* **51**, 1151–1160 (2020).
9. Donhauser, P. W. & Baillet, S. Two distinct neural timescales for predictive speech processing. *Neuron* **105**, 385–393.e389 (2020).
10. Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P. & de Lange, F. P. A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl. Acad. Sci. USA* **119** <https://doi.org/10.1073/pnas.2201968119> (2022).
11. Shipp, S. Neural Elements for Predictive Coding. *Front. Psych.* **7** <https://doi.org/10.3389/fpsyg.2016.01792> (2016).
12. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
13. Mumford, D. On the computational architecture of the neocortex. *Biol. Cybern.* **66**, 241–251 (1992).
14. Lee, T. S. & Mumford, D. Hierarchical Bayesian inference in the visual cortex. *J. Optical Society of America A* **20** <https://doi.org/10.1364/josaa.20.001434> (2003).
15. Bastos, A. ndreM. et al. Canonical Microcircuits for Predictive Coding. *Neuron* **76**, 695–711 (2012).
16. Uhrig, L., Dehaene, S. & Jarraya, B. A hierarchy of responses to auditory regularities in the macaque brain. *J. Neurosci.* **34**, 1127–1132 (2014).
17. Jiang, Y. et al. Constructing the hierarchy of predictive auditory sequences in the marmoset brain. *Elife* **11** <https://doi.org/10.7554/eLife.74653> (2022).
18. Gabhart, K. M., Xiong, Y. & Bastos, A. M. Predictive coding: a more cognitive process than we thought? *Trends in Cognitive Sciences* <https://doi.org/10.1016/j.tics.2025.01.012> (2025).
19. Wacongne, C., Changeux, J.-P. & Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **32**, 3665–3678 (2012).
20. Liukovich, K., Ukrainteva, Y. & Martynova, O. Implicit auditory perception of local and global irregularities in passive listening condition. *Neuropsychologia* **165**, 108129 (2022).
21. Grundei, M., Schmidt, T. T. & Blankenburg, F. A multimodal cortical network of sensory expectation violation revealed by fMRI. *Hum. Brain Mapp.* **44**, 5871–5891 (2023).
22. Lakatos, P. et al. The Thalamocortical Circuit of Auditory Mismatch Negativity. *Biol. Psychiatry* **87**, 770–780 (2020).
23. Keshishian, M. et al. Joint, distributed and hierarchically organized encoding of linguistic features in the human auditory cortex. *Nat. Hum. Behav.* **7**, 740–753 (2023).
24. Swenson, R. S. & Gullledge, A. T. in *Conn’s Translational Neuroscience* (ed P. Michael Conn) 263–288 (Academic Press, (2017).
25. Leonard, M. K. et al. Large-scale single-neuron speech sound encoding across the depth of human cortex. *Nature* **626**, 593–602 (2024).
26. Slater, B. J., Sons, S. K., Yudintsev, G., Lee, C. M. & Llano, D. A. Thalamocortical and intracortical inputs differentiate layer-specific mouse auditory corticocollicular neurons. *J. Neurosci.* **39**, 256–270 (2019).
27. Yudintsev, G. et al. Evidence for layer-specific connective heterogeneity in the mouse auditory corticocollicular system. *J. Neurosci.* **41**, 9906–9918 (2021).
28. Xiong, Y. S. et al. Propofol-mediated loss of consciousness disrupts predictive routing and local field phase modulation of neural activity. *Proc. Natl. Acad. Sci. USA* **121**, e2315160121 (2024).
29. Yang, J., Huber, L., Yu, Y. & Bandettini, P. A. Linking cortical circuit models to human cognition with laminar fMRI. *Neurosci. Biobehav. Rev.* **128**, 467–478 (2021).
30. Huber, L. et al. Sub-millimeter fMRI reveals multiple topographical digit representations that form action maps in human motor cortex. *Neuroimage* **208**, 116463 (2020).
31. Huber, L. et al. High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. *Neuron* **96**, 1253–1263.e1257 (2017).
32. Liu, T. T. et al. Layer-specific, retinotopically-diffuse modulation in human visual cortex in response to viewing emotionally expressive faces. *Nat. Commun.* **13**, 6302 (2022).
33. Haarsma, J., Deveci, N., Corbin, N., Callaghan, M. F. & Kok, P. Expectation cues and false percepts generate stimulus-specific activity in distinct layers of the early visual cortex. *J. Neurosci.* **43**, 7946–7957 (2023).
34. Ahveninen, J. et al. Intracortical depth analyses of frequency-sensitive regions of human auditory cortex using 7TfMRI. *NeuroImage* **143**, 116–127 (2016).
35. Bekinschtein, T. A. et al. Neural signature of the conscious processing of auditory regularities. *Proc. Natl. Acad. Sci.* **106**, 1672–1677 (2009).
36. Howard, M. A. et al. Auditory cortex on the human posterior superior temporal gyrus. *J. Comp. Neurol.* **416**, 79–92 (2000).
37. Doeller, C. F. et al. Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *Neuroimage* **20**, 1270–1282 (2003).
38. Heilbron, M. & Chait, M. Great expectations: is there evidence for predictive coding in auditory cortex?. *Neuroscience* **389**, 54–73 (2018).
39. Giraud, A. L. & Arnal, L. H. Hierarchical predictive information is channeled by asymmetric oscillatory activity. *Neuron* **100**, 1022–1024 (2018).
40. Chai, Y. et al. Topographical and laminar distribution of audiovisual processing within human planum temporale. *Prog. Neurobiol.* **205**, 102121 (2021).
41. Lankinen, K. et al. Cortical depth profiles of auditory and visual 7 T functional MRI responses in human superior temporal areas. *Hum. Brain Mapp.* **44**, 362–372 (2023).
42. De Martino, F. et al. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. *Proc. Natl. Acad. Sci. USA* **112**, 16036–16041 (2015).
43. Fernandez Pujol, C., Blundon, E. G. & Dykstra, A. R. Laminar specificity of the auditory perceptual awareness negativity: A biophysical modeling study. *PLoS Comput Biol.* **19**, e1011003 (2023).
44. Huntenburg, J. M., Bazin, P. L. & Margulies, D. S. Large-scale gradients in human cortical organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
45. Cope, T. E. et al. Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat. Commun.* **8**, 2154 (2017).
46. Chao, Z. C., Huang, Y. T. & Wu, C. T. A quantitative model reveals a frequency ordering of prediction and prediction-error signals in the human brain. *Commun. Biol.* **5**, 1076 (2022).

47. Finn, E. S., Huber, L., Jangraw, D. C., Molfese, P. J. & Bandettini, P. A. Layer-dependent activity in human prefrontal cortex during working memory. *Nat. Neurosci.* **22**, 1687–1695 (2019).
48. Degutis, J. K. et al. Dynamic layer-specific processing in the prefrontal cortex during working memory. *Commun. Biol.* **7**, 1140 (2024).
49. Bijanzadeh, M., Nurminen, L., Merlin, S., Clark, A. M. & Angelucci, A. Distinct laminar processing of local and global context in primate primary visual cortex. *Neuron* **100**, 259–274.e254 (2018).
50. Yu, Y. et al. Layer-specific activation in human primary somatosensory cortex during tactile temporal prediction error processing. *NeuroImage* **248**, 118867 (2022).
51. Muckli, L. et al. Contextual feedback to superficial layers of V1. *Curr. Biol.* **25**, 2690–2695 (2015).
52. Yu, Y. et al. Layer-specific activation of sensory input and predictive feedback in the human primary somatosensory cortex. *Sci. Adv.* **5**, eaav9053 (2019).
53. Lawrence, S. J. D. et al. Laminar organization of working memory signals in human visual cortex. *Curr. Biol.* **28**, 3435–3440.e3434 (2018).
54. Teufel, C. & Fletcher, P. C. Forms of prediction in the nervous system. *Nat. Rev. Neurosci.* **21**, 231–242 (2020).
55. Pezzulo, G., Zorzi, M. & Corbetta, M. The secret life of predictive brains: what's spontaneous activity for?. *Trends Cogn. Sci.* **25**, 730–743 (2021).
56. de Lange, F. P., Schmitt, L. M. & Heilbron, M. Reconstructing the predictive architecture of the mind and brain. *Trends Cogn. Sci.* **26**, 1018–1019 (2022).
57. Uhrig, L., Janssen, D., Dehaene, S. & Jarraya, B. Cerebral responses to local and global auditory novelty under general anesthesia. *NeuroImage* **141**, 326–340 (2016).
58. Huang, Y. T. et al. Crossmodal hierarchical predictive coding for audiovisual sequences in the human brain. *Commun. Biol.* **7** <https://doi.org/10.1038/s42003-024-06677-6> (2024).
59. Malmierca, M. S. et al. Attention to audiovisual speech shapes neural processing through feedback-feedforward loops between different nodes of the speech network. *PLOS Biology* **22** <https://doi.org/10.1371/journal.pbio.3002534> (2024).
60. Huber, L. et al. Slab-selective, BOLD-corrected VASO at 7 Tesla provides measures of cerebral blood volume reactivity with high signal-to-noise ratio. *Magn. Reson. Med.* **72**, 137–148 (2014).
61. Nieto-Castanon, A. in *Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN* 3–16 (Hilbert Press, (2020).
62. Whitfield-Gabrieli, S., Nieto-Castanon, A. & Ghosh, S. *Artifact detection tools (ART)*, (2009).
63. Ghosh, S. S. et al. Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinform.* **5** <https://doi.org/10.3389/fninf.2011.00013> (2011).
64. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed. Res* **29**, 162–173 (1996).
65. Cox, R. W. & Hyde, J. S. Software tools for analysis and visualization of fMRI data. *NMR Biomed.* **10**, 171–178 (1997).
66. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. Fsl. *NeuroImage* **62**, 782–790 (2012).
67. Huber, L. R. et al. LayNii: A software suite for layer-fMRI. *NeuroImage* **237**, 118091 (2021).

Acknowledgements

The authors thank the MRI platform of Zhangjiang Brain Imaging Center of Fudan University for data acquisition. This work was supported by a grant from the National Natural Science Foundation of China (No. 32471091 to Y.J).

Author contributions

Y.J. conceptualized the study. Y.M., Q.L., Y.C., and Z.X. collected the data. Y.M. analyzed the data. Y.J. and Y.M. wrote the manuscript. H.W. and W.D. reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-67380-6>.

Correspondence and requests for materials should be addressed to Yuwei Jiang.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025