Current Research in Neurobiology 4 (2023) 100075

Contents lists available at ScienceDirect

Current Research in Neurobiology

journal homepage: www.sciencedirect.com/journal/current-research-in-neurobiology

# Layer-specific correlates of detected and undetected auditory targets during attention

Miriam Heynckes [a], Agustin Lage-Castellanos [a], Peter De Weerd [a], Elia Formisano [a,b], Federico De Martino [a,*]

[a] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6229 ER, Maastricht, the Netherlands
[b] Maastricht Centre for Systems Biology, Maastricht University, Universiteitssingel 60, 6229 ER, Maastricht, the Netherlands

ARTICLE INFO

*Keywords:*
7T fMRI
Depth-dependent
Layer-specific
Human primary auditory cortex

ABSTRACT

In everyday life, the processing of acoustic information allows us to react to subtle changes in the auditory scene. Yet even when closely attending to sounds in the context of a task, we occasionally miss task-relevant features. The neural computations that underlie our ability to detect behavioral relevant sound changes are thought to be grounded in both feedforward and feedback processes within the auditory hierarchy. Here, we assessed the role of feedforward and feedback contributions in primary and non-primary auditory areas during behavioral detection of target sounds using submillimeter spatial resolution functional magnetic resonance imaging (fMRI) at high-fields (7 T) in humans. We demonstrate that the successful detection of subtle temporal shifts in target sounds leads to a selective increase of activation in superficial layers of primary auditory cortex (PAC). These results indicate that feedback signals reaching as far back as PAC may be relevant to the detection of targets in the auditory scene.

## 1. Introduction

In a movie with a bank robbery scene, a criminal tries cracking a safe by carefully listening to any sounds while turning the wheels of the locking mechanism. By listening carefully, the robber detects when notches align, permitting the removal of a locking bar and opening of the safe. This highlights the auditory system's remarkable ability to process subtle acoustic information, upon which we base decisions and actions. Notably though, despite the robber keenly attending the sounds, an alternative scenario can be imagined, in which the same acoustic change is not detected, delaying or preventing the robbery. Here, we investigated why changes in the soundscape close to detection threshold, with the same behavioral relevance are sometimes detected and sometimes missed (despite being physically identical). We hypothesized that these changes in behavior are due to momentary fluctuations in attention and reflected in modulations of feedback signals.

In line with the known segregation of feedforward and feedback processes within the laminar organization of the cortex (Douglas et al., 1989; Douglas and Martin, 2004), previous studies in primary visual and auditory cortices demonstrated enhanced activity in superficial layers with attention (Lawrence et al., 2019; Liu et al., 2021; De Martino et al., 2015; Gau et al., 2020). In particular, electrophysiological research in animals has investigated the neural correlates of attention and highlighted changes in cortical oscillations in superficial layers (Lakatos et al., 2013; O'Connell et al., 2014). In humans, the modulation of cortical layers by attention in both vision and audition has been probed non-invasively using high-field functional magnetic resonance imaging (fMRI) (Lawrence et al., 2019; Liu et al., 2021; De Martino et al., 2015; Gau et al., 2020; Klein et al., 2018). In these studies, attentional modulation was probed by either drawing attention towards or away from the relevant stimulus (or stimulus feature). In particular, in the auditory domain, attending to an auditory stimulus (compared to a concurrently presented visual stimulus) has highlighted changes in frequency tuning (i.e. tuning width) (De Martino et al., 2015) and an increase in activation in superficial layers of the (primary) auditory cortex (Gau et al., 2020). In the visual domain, within-modality attentional manipulations (spatial or feature based attention) have been used in layer-specific studies, which demonstrated activity modulations in superficial layers (Lawrence et al., 2019; Liu et al., 2021) as well as changes in population receptive fields in deep layers of primary visual cortex (V1) (Klein et al., 2018). Altogether, these data indicate that the presence or absence of attention to stimuli modulates activity in superficial (and in some cases

* Corresponding author. Federico De Martino Department Cognitive Neuroscience Oxfordlaan 55, 6229EV, Maastricht, the Netherlands.
*E-mail address:* f.demartino@maastrichtuniversity.nl (F. De Martino).

https://doi.org/10.1016/j.crneur.2023.100075
Received 15 July 2022; Received in revised form 24 November 2022; Accepted 12 January 2023
Available online 25 January 2023

deep) layers. Here we asked where in the auditory cortex and in which cortical layers, neural activity variations would be present that could explain why identical auditory stimuli presented under identical attentional instructions would be detected in some trials, and not in others. In line with literature ascribing a role of superficial layers in receiving attentional feedback, we hypothesized increased activity in superficial layers of auditory cortex may be related to fluctuations in attention thus variations in the perception (detection) of physically identical stimuli.

Apart from the segregation of feedforward and feedback signals across cortical depths, the auditory cortex has a tonotopic organization (Merzenich and Brugge, 1973; Merzenich et al., 1973; Formisano et al., 2003). Attention to frequency specific targets gain-modulates frequency-specific (tonotopic) regions in a layer dependent manner (O'Connell et al., 2014). Task-dependent changes in the receptive fields have been shown using invasive electrophysiology in animals in superficial layers of the auditory cortex (Francis et al., 2018) and have been suggested as neural correlates of selective attention. Similarly, frequency-specific effects of attention have been shown non-invasively in humans at a macroscopic level (De Martino et al., 2015; Riecke et al., 2018; Da Costa et al., 2013). We presented narrowband stimuli at two distinct frequencies (high and low), to understand whether the topographic organization of human auditory cortical areas interacts with laminar processing when detecting a relevant sound.

In particular, we asked human listeners to perform an auditory temporal detection task while concurrently acquiring layer-specific fMRI data. We hypothesized that, the change in soundscape (presence of a target) may be reflected in a modulation of middle cortical layers (for both detected and undetected targets). By comparing responses to (acoustically identical) perceptually detected and undetected targets, we localized responses related to the detection of sounds under constant, demanding attentional conditions. We hypothesized that, the behavioral relevance of attention is reflected in the change of population level activity in superficial cortical layers of the primary auditory cortex (Fig. 1C) and further hypothesized that this effect may be tonotopic.

## 2. Materials and methods

### 2.1. Participants

Ten healthy participants (4 females, 6 males; median age 28,5; range = (Huber et al., 2017; Moerel et al., 2021; Snyder et al., 2012; van Mourik et al., 2021; Henry and Herrmann, 2014; Giani et al., 2015; Gutschalk et al., 2008; Wiegand and Gutschalk, 2012; Cusack, 2005; Micheyl et al., 2005; Kilian-Hütten et al., 2011; Hill et al., 2011; Brainard, 1997; Moore, 2003)) were recruited. All participants were students or employees of Maastricht University. The study was approved by the research ethics committee of the Faculty of Psychology and Neuroscience at Maastricht University. For every participant we acquired 1 run of the tonotopic localizer, between 3 and 9 runs of the target detection experiment (median number of runs = 6; range (Lawrence et al., 2019; Liu et al., 2021; De Martino et al., 2015; Gau et al., 2020; Lakatos et al., 2013; O'Connell et al., 2014; Klein et al., 2018)) and a high-resolution anatomical scan. Each volunteer participated in either one or two sessions depending on the number of functional runs collected in the main experiment (for a total of 14 sessions across all participants). Most participants had previous experience of high-resolution fMRI studies.

### 2.2. Experimental design and stimuli

All stimulus presentation scripts were written in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA), using the Psychophysics toolbox (Brainard, 1997) and custom-code. Participants underwent a training session (~20 min) followed by a scanning session (~2 ½ hours). Participants 01, 02, 03 and 08 underwent two scan sessions, to acquire additional functional runs and reach the target of at least six functional runs of the main experiment. Prior to each scan session the sound intensity of stimuli was adjusted individually to (perceptually) equalize the loudness of the experimental stimuli in the two stimulus conditions, and between outside the scanner (training session) and inside (main experiment).

Target detection experiment. Participants were asked to detect a target constructed by temporally shifting (TS) one of the narrow-band sounds forming the quintets (Fig. 1A). Narrowband sounds were centered around carrier frequencies of 200 Hz or 1100 Hz. The passbands around the carriers were constructed using equivalent rectangular bandwidths (ERBS = 4 (Moore, 2003);). Each passband consisted of a summation of 21 sinusoids with amplitude normalized to 1 and a random onset phase. A quintet consisted of five 10 ms narrowband sounds, each separated by 10 ms (see Fig. 1A, inset 1). Targets were constructed by shifting in time the third sound in a quintet (see Fig. 1A, inset 2). During a training session participants' 70% detection threshold was determined by means of a 2 down 1 up staircase, in which the size of a temporal shift (TS; ranging between 1 and 9 ms) determined the difficulty of a detection task. A 70% detection threshold outside the scanner was used as a starting threshold during the scanning session. The more challenging (i.e. louder scan environment) led to a (desirable) detection accuracy ~ 50% during scanning. Maintaining task difficulty to achieve a detection of 50% required adjusting the TS individually after every run by the experimenters, to ensure an approximately equal number of detected and undetected trials per participant, to be contrasted later on in the analysis. Supplementary Fig. S1 displays the behavioral detection rates per participant for high and low sounds separately. All sounds were presented in silent intervals between

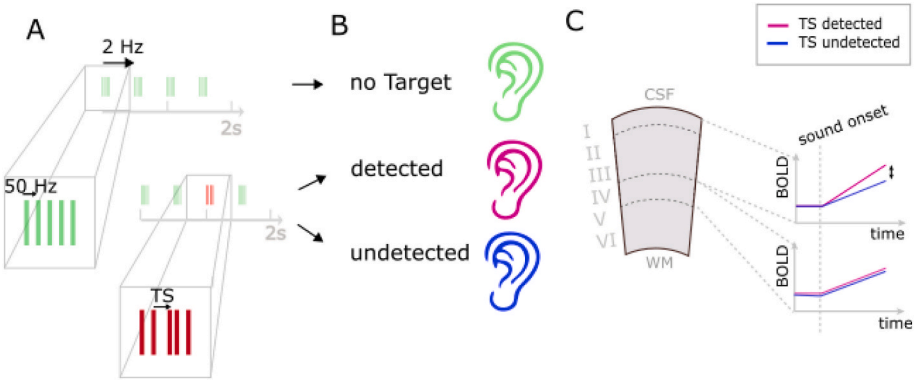


**Fig. 1. Task and Hypotheses. (A)** Stimuli were periodic sequences of narrowband quintets repeating at 2 Hz. Two narrowband frequency ranges around 200 Hz and 1100 Hz were used to create low and high pitch sounds. Five sounds repeating at 50 Hz (10 ms ISI) formed a quintet (inset 1). 75% of the stimuli contained a target. Target sounds (TS; inset 2) had a different temporal structure: the third sound in a quintet was temporally shifted between 1.5 and 7 ms, depending on participants' perceptual (detection) threshold. Figure S1.1 shows behavioral detection rates per participant. **(B)** Target trials were sorted based on the behavioral response. **(C)** Expected laminar BOLD response profile. Both detected and undetected TS would entail a feedforward BOLD increase in middle layers, but a detected TS (magenta line) additionally increases the BOLD response in superficial layers compared to undetected TS (blue line). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

acquisitions. After the sound finished, participants were cued by a green fixation cross to respond whether they had detected a target or not and instructed to press 1 or 2 on the button box. The cue for a button press was randomly jittered on each trial in the interval [0–200 ms] after the sound offset. Each run consisted of a total of 30 trials, 15 per carrier frequency, of which 3 trials per carrier were without a target and 12 containing a target. In addition, 3 silent trials per run were randomly interspersed functioning as baseline for sound vs silence contrasting.

Tonotopic localizer. To map tonotopic organization in the AC, a frequency localizer was performed (Formisano et al., 2003). We presented 7 center frequencies (130 Hz, 200 Hz, 306 Hz, 721 Hz, 1100 Hz, 1700 Hz and 4000 Hz) in blocks. Each block consisted of three amplitude modulated tones centered on one of the center frequencies (center frequencies ± 0.1 octaves). Five center frequencies were log-spaced between 130 Hz and 4000 Hz, and two additional center frequencies were inserted (200 and 1100 Hz, the carrier frequencies employed in the target detection experiment). Tones were amplitude modulated (8 Hz, modulation depth of 1) and presented for 800 ms. During the localizer, participants were asked to fixate and passively listen to the sounds. The duration of the localizer was 7 ½ min.

### 2.3. MRI acquisition

Data acquisition was performed on a whole-body Magnetom scanner (nominal field strength 7 T (T) (Siemens Medical Systems, Erlangen, German) at the Maastricht Brain Imaging Center, The Netherlands. All images were acquired using a 32-channel head coil (Nova Medical Inc. Wilmington, MA, USA).

Target detection experiment. For the sub-millimeter measurements, we used an event-related (sparse) design, with a 2D-GE-EPI sequence (TE/TR = 25/3500 ms, TA = 1400, silent gap = 2100; in-plane FoV 1120 × 1120 mm; matrix size 200 X 200; slices = 42; GRAPPA factor = 3; partial Fourier = 6/8; phase-encoding direction anterior - posterior, with multiband factor = 2, and ascending slice order, yielding a nominal resolution of 0.8 mm isotropic - see Fig S2.1 for design and coverage). Before acquisition of the first functional run, we acquired 10 vol for distortion correction (5 vol with opposite phase-encoding directions AP and PA).

Tonotopic localizer. We acquired the tonotopic localizer using a block design with a 2D-GE-EPI sequence (TE/TR = 21.2/2600 ms; TA = 1200 ms; silent gap = 1400 ms; in-plane FoV 1140 × 1140 mm; matrix size 136 X 136; slices = 46, GRAPPA factor = 2, multiband factor = 2, partial Fourier = 6/8, phase-encoding direction anterior-posterior - yielding a voxel resolution of 1.2 mm isotropic). Preceding the localizer, 10 vol in opposite phase-encoding direction (5 vol AP and PA each) were acquired for distortion correction.

*Anatomical scans.* For visualization of the functional results and to obtain high-quality segmentations of the gray and white matter we obtained anatomical scans at a nominal voxel resolution of 0.65 mm isotropic. For this we used a MP2RAGE (Marques et al., 2010) sequence (TR = 5000 ms; TE = 2.5 ms; TI1 = 900 ms; TI2 = 2700 ms; FoV 207 × 207 mm; matrix size 320 X 320; FA1 = 5°, FA2 = 3°; GRAPPA factor = 3 with an overall TA = 10:55 min). For four participants a second scan session was performed, in which a lower resolution, hence faster, MPRAGE sequence at 1 mm isotropic, was used to acquire T1-weighted images for in-session alignment of functional data (TR = 2370 ms; TE = 2.3 ms; TI = 1500 ms; FoV 256 × 256 mm; matrix size 256x256; FA = 5°; GRAPPA factor = 3 with an overall TA = 05:03 min).

### 2.4. Behavioral data analysis

Behavioral data were analyzed in Matlab (The MATHWORKS Inc., Natick, MA, 234 USA). For every participant we determined the number of detected and undetected trials for the low and high carrier sounds separately (see Fig. S1.1 for behavioral performance per participant). Reaction times have not been analyzed as participants were cued to respond.

### 2.5. Anatomical data processing

*Preprocessing.* Anatomical images were processed using the advanced segmentation tools in BrainVoyager 21.4 (Brain Innovation, Maastricht, Netherlands), SPM's bias correction (Ashburner and Friston, 2005), ITK SNAP (Yushkevich, 2006) and FSL BET (Jenkinson et al., 2012). If not otherwise indicated, default parameters were used.

The second inversion image of the MP2RAGE was subjected to the automated segmentation in SPM to obtain tissue-probability maps. The non-brain tissue-probability maps (C3, C4, C5) were manually thresholded and combined with a brain mask, obtained from the second inversion image using FSL BET. By combining these, we obtain a brain-mask that allows removing non-brain tissue and large veins (for a stepwise procedure see (Kashyap et al., 2019)). This anatomical pre-processing workflow was developed particularly to work well for MP2RAGE data (https://github.com/srikash/presurfer). The resulting mask was inspected, had the cerebellum manually removed and was further manually polished using ITK SNAP in combination with a graphics tablet (Intuos Art; Wacom Co.). The resulting mask was applied to the T1w image (UNI) of the MP2RAGE. We then used BrainVoyager's intensity inhomogeneity correction and up-sampled the image to a resolution of 0.4 mm isotropic, using the spatial transformation option in BrainVoyager's 3D Volume tools. Lastly, the image was transformed (only translation and rotations, no scaling) from native space into a space in which the anterior and posterior commissure were in the same plane (ACPC space). We refer to this space as the voxel space.

*Segmentation.* The resulting image was input to BrainVoyager's advanced segmentation routine to obtain a white matter (WM) mask. This initial WM mask was inspected and manually polished in ITK SNAP, where emphasis was placed on corrections in the region of interest (bilateral auditory cortex [AC]). The polished WM mask was input to the subsequent step of the advanced segmentation routine in BrainVoyager to obtain a GM mask. This GM mask tended to be too inclusive, containing blood vessels, posing a challenge especially around the strongly vascularized AC. Therefore, we manually polished the GM definition and GM/CSF boundary in ITK SNAP. As a last step the obtained GM/WM segmentation was manually split into two hemispheres.

*Cortical depth sampling.* Using the GM/WM segmentation at 0.4 mm isotropic resolution, we measure the cortical thickness of individual segmented cortical hemispheres in volume space. Based on the cortical thickness we can perform whole-mesh cortical depth sampling, where we create surface meshes at equivolume cortical depth levels between the WM/GM boundary and the GM/CSF boundary (Waehnert et al., 2014). The created set of meshes at different cortical depth were then used to sample the functional data using trilinear interpolation. Surface visualizations are always based on the mid GM surface reconstruction.

*Anatomical ROI selection.* Based on macro-anatomical landmarks (sulci and gyri) and following the definition reported in (Kim et al., 2000), the temporal lobe of each participant was divided into three anatomical ROIs in each hemisphere: Heschl's gyrus (HG), planum temporale (PT), planum polare (PP), drawn onto the inflated hemispheric surfaces, see Fig. 2B.

### 2.6. Functional data processing - tonotopic localizer

*Preprocessing. P*reprocessing of the localizer data was performed in BrainVoyager 21.4, the NeuroElf toolbox in Matlab, as well as custom code in Matlab R2017a (The MathWorks Inc). Where not specified otherwise, default settings were used. Slice-scan-time correction, 3D motion correction (with sinc interpolation), linear trend removal and high-pass filtering (7 cycles) was performed. BrainVoyager's COPE plugin was used to correct EPI geometric distortions using a pair of opposite-phase encoded data.

*Statistical analysis.* All statistical computations were performed at the

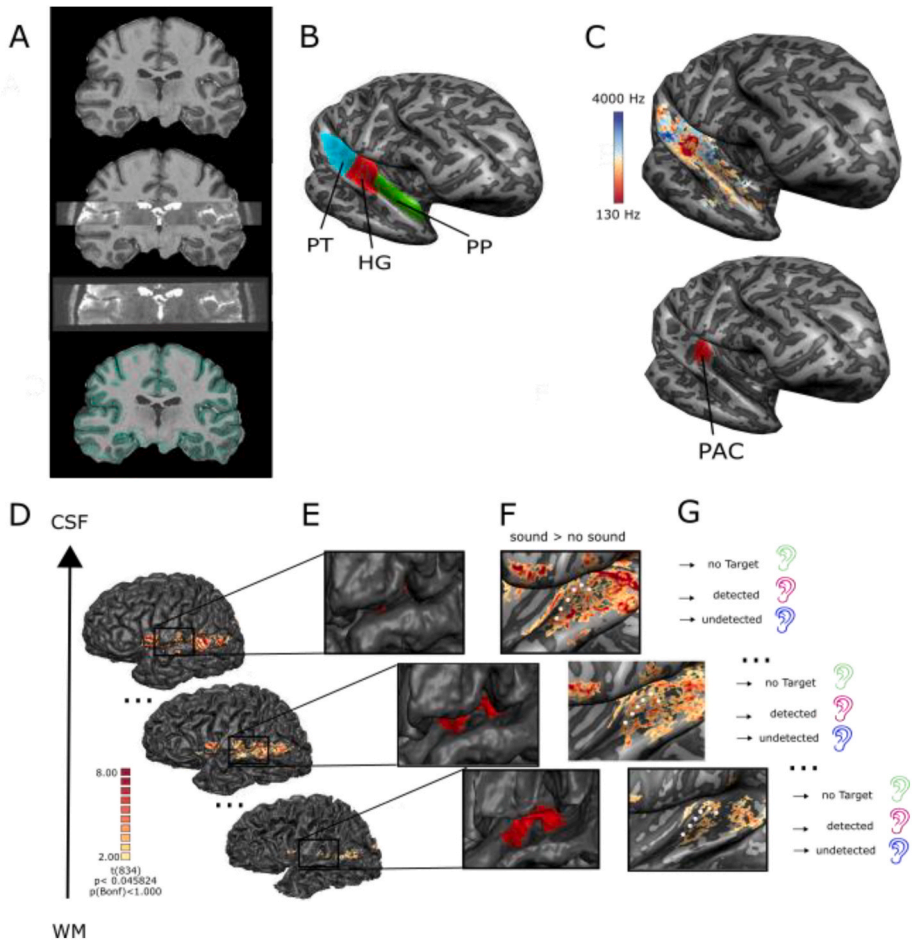


**Fig. 2. Analysis approach (previous page). (A)** Interleaved anatomical image and functional volume, highlighting correspondence between datasets, anatomical images are segmented (and manually corrected around the ROI) to identify white and gray matter. See Figure S2.5 for enlarged view. **(B)** The segmentation is used to reconstruct cortical surfaces (inflated view, with cortical curvature; light gray, gyrus; dark gray, sulcus). Anatomical ROIs (planum temporale (PT), Heschl's gyrus (HG) and planum polare (PP) are defined based on major anatomical landmarks (Kim et al., 2000) in every participant. **(C)** Right hemisphere showing tonotopic map and primary auditory cortex (PAC) in a single participant. PAC is functionally defined using the tonotopic localizer (Moerel et al., 2014). **(D)** We sampled eleven equivolume surfaces along the depth of cortical gray matter from white matter (WM) to cerebrospinal fluid (CSF). The most superficial, middle and deepest surface are shown. Overlaid activation map depicting the overall response to sounds (Ftest) in the main experiment. **(E)** Zoomed view onto temporal lobe with HG shown in red. Note the varying curvature across depth. **(F)** Inflated surfaces at each depth, functional map same as in D. The typical GE-BOLD increase of activation towards the cortical surface is visible. White dotted line demarks HG. **G)** The analysis focuses on the response per perceptual condition at each cortical depth. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

level of single participants, by fitting a general linear model with a predictor for each center frequency to the data of the tonotopic localizer, obtaining a beta (response-strength) for every predictor and computing a statistical activation maps (FMap) of all predictors combined (contrast: sound > no sound). Fig. S3 shows the overall response to sounds in the localizer at statistical significance threshold qFDR >0.05 for every participant.

*Tonotopic maps.* Tonotopic maps were derived following the standard procedure of z-scoring the response of voxels on the temporal lobe per frequency predictor, thereby removing a response bias towards low frequencies, and then color coding each voxel according to the frequency to which it best responded (i.e. its preferred frequency, indicated by the beta value – (Formisano et al., 2003)).

*Functional ROI definition.* In addition to dividing the human auditory cortex in terms of its major anatomical landmarks, we define primary auditory cortex functionally using the main tonotopic gradient obtain in the localizer (Moerel et al., 2014), as the auditory cortex in humans displays large macro-anatomical variability (Heynckes, 2022; Heschl, 1878).

The statistical activation map (FMap) in response to sounds, and the tonotopic map derived from the localizer were up-sampled from their native resolution at 1.2 mm isotropic by linearly interpolating to 0.8 mm isotropic to match the high-resolution functional data of the main experiment. The obtained up-sampled tonotopic map was then projected on the individual's reconstruction of the inflated mid-GM surface for each hemisphere, which allowed locating the main tonotopic gradient. The most likely position of the primary auditory cortex was localized using the tonotopic gradient of high frequency (posteromedial HG) to low frequency (medial portion HG) and back to high frequency. Supplementary figure S2.3 displays the tonotopic maps of all ten participants.

### 2.7. Functional data processing - target detection experiment

Functional data were processed using BrainVoyager 21.4, the NeuroElf toolbox in Matlab, as well as custom code in Matlab R2017a (The MathWorks Inc). Where not specified otherwise, default settings were used.

*Preprocessing.* Preprocessing for all high-resolution functional data was performed in the default order in BrainVoyager (slice-scan time correction, 3D motion correction [with sinc interpolation and across runs] and linear trend removal and high-pass filtering (7 cycles). We corrected all functional images for EPI geometric distortions using BrainVoyager's COPE plugin based on the AP/PA images.

*Co-registration of functional to anatomical images.* The functional data of the first run were registered to the pre-processed anatomical data in native space using BrainVoyager's FMR-VMR co-registration. The positional information provided in the header is used for an initial alignment followed by fine-tuning co-registration using boundary-based registration. The result for the first run was visually inspected by overlaying the functional and anatomical images acquired in the same session and manually improved where necessary. The obtained initial alignment and fine-tuning alignment transformation files were used for the remaining runs within a session in combination with an ACPC transformation file to create a volume timecourse per run in the voxel space, using sinc interpolation. When a second session was acquired, co-registration of functional images was performed to in-session MPRAGE anatomical data. In a second step, between session anatomical data were then aligned using BrainVoyager's vmr-vmr co-registration and the resulting transformation matrix applied when creating volume time courses.

### 2.8. Functional data – statistical analysis

*Statistical analysis* per *ROI.* We computed a GLM with a separate predictor for every trial, classified as either being low detected, low undetected, low no Target or high detected, high undetected or high no Target, where high and low refers to the carrier frequency of the sound. Fig. 2A shows the overall response to sounds compared to baseline silence, corrected at qFDR <0.05 for an exemplary participant. (See Fig. S2.2 for all participants).

In a second step we sampled these single trial beta maps on 11 reconstructed depth dependent surfaces and averaged across trials of the same perceptual condition (Fig. 2D–G). To obtain laminar profiles multiple inclusion criteria guided the selection of vertices for sampling the mean beta surface maps (see S3.1). Vertices had to be within a particular ROI (PAC, HG, PP, PT). Their statistical F-value in response to sounds in the localizer needed to exceed F > 2 and statistical F-value in response to sounds in the main experiment exceeded F > 0.1, thereby ensuring that voxels with an (average) positive BOLD response to sounds in the main experiment were included, independent of depth. For each participant we extracted the mean (beta) across these vertices, per perceptual condition per depth. The perceptual conditions depended on the behavior of the participant and could lead to unequal condition size (see Fig.S1.1). Therefore, we bootstrapped a 95% confidence interval of the mean of trials per perceptual condition per depth (by bootstrapping 100 times the mean percent signal per perceptual condition).

Second-level group statistics for each ROI (n participants = 10) were carried out on the mean differences of the bootstrapped betas between the perceptual conditions extracted from each participant (detected minus no Target and undetected minus no Target. Fig S3.4 and 3.5). By subtracting the response to the no target condition from both the responses to detected and undetected sounds we aimed to control for the layer dependent signal increase towards the superficial gray matter elicited by the draining vasculature (Turner, 2002; Polimeni et al., 2010). We binned the data across 11 depth levels as follows: Deep – depth 1:3, middle – depth 5:7, superficial – depth 9:11, thereby ensuring equal sized depth bins. We used three predictors (depth [deep; middle; superficial], condition [detected minus no target; undetected minus no target] and their interaction) in a separate generalized linear mixed effects (GLME) model for each ROI. Model fits were compared between a fixed effect and a random effect model using likelihood ratio tests.

*Assessing the frequency selectivity of the effect of detection in PAC.* We expected voxels to retain their frequency preference (high vs low frequency) across the localizer and main experiment. To test this, we selected voxels whose time courses were modulated in response to sounds, exceeding a statistical threshold of F > 2 in the tonotopic localizer. In the localizer we performed best frequency (BF) mapping (i.e. tonotopic mapping). Using a GLM with only two predictors (200 Hz and 1100 Hz), we defined voxels as preferring low or high frequency. This allowed us to directly compare the preference from the tonotopy to the main experiment. For these groups of voxels (i.e., labeled as preferring low or high frequency in the localizer), we then plotted the response (after z-scoring as customary in tonotopic mapping) to the high and low preferring sounds (separately) in the main experiment (see Fig. S4.1).

*Tonotopic analysis of main experiment.* For the tonotopic analysis of the data we selected vertices in PAC as outlined in the previous section. In particular, we extracted the mean (beta) across vertices per perceptual condition (detected, undetected, no target), per depth (11 levels), in low- and high-preferring groups of voxels within PAC, for low and high presented sounds in the main experiment.

Second-level group statistics (n = 10) were carried out on the differences between perceptual conditions extracted from each participant (detected minus no Target and undetected minus no Target, Fig S4.2). We binned the data from 11 depth levels into three depth bins. Deep – depth 1:3, middle – depth 5:7, superficial – depth 9:11, thereby ensuring equal sized depth bins. For the tonotopic analysis of PAC we used the predictors (*depth* [deep; middle; superficial], *condition* [detected minus no target; undetected minus no target], *BFandSound* [highSoundHighBF, highSoundLowBF, lowSoundHighBF, lowSoundlowBF] and their *interactions;* where BF stands for best frequency and was defined on the basis of the localizer [see above]) in a generalized linear mixed effects model (GLME).

## 3. Results

We examined the laminar response profile of human auditory cortex (AC), using 2-D gradient echo (GE) blood oxygen level dependent (BOLD) fMRI at 7 T, with sub-millimeter resolution, during perceptual detection of temporally shifted target sounds (TS) embedded in rhythmic sound sequences (Fig. 1A). Specifically, we contrasted different percepts of acoustically identical sound sequences containing a target (Fig. 1B).

Based on participants' responses, we labeled trials as detected (i.e., target present and detected), undetected (target present and not detected) and no target (target not present). In individual participants' data we estimated responses for every condition, focusing on primary and non-primary areas of human auditory cortex (Fig. 2B–C) and sampled them onto 11 equivolume depth surfaces (Waehnert et al., 2014) (Fig. 2D). Figure S3.2 shows the laminar profile per participant ROI and condition.

Second-level group statistics (n = 10) were carried out on the differences of the mean (across participants) betas between the perceptual conditions across cortical depths (Fig. 3D–F). Specifically, we statistically assessed whether detected and undetected sounds differentially modulated the response across depths. By subtracting the response to the no target condition from both the responses to detected and undetected sounds we aimed to control for the layer dependent signal increase towards the superficial gray matter elicited by the draining vasculature (Turner, 2002; Polimeni et al., 2010).

### 3.1. No significant frequency-specific modulation of the layer-dependent detection effect

The best frequency definition within PAC was stable between the localizer and the main experiment (Fig. S4.1). To assess the frequency-specificity of detection effect we fitted a generalized linear mixed effects (GLME) model, splitting each ROI according to best frequency (BF) as defined in the localizer, with four predictors (*depth* [categorical], *condition* [detected minus no target; undetected minus no target], *BFandSound* [highSoundHighBF, highSoundLowBF, lowSoundHighBF, lowSoundlowBF] and their *interactions*). We did not detect a significant 3-way interaction between condition depth and frequency (Depth:Condition:BFandSound ($F(6,696) = 0.17$, $p > 0.05$), indicating that the detection effect was not significantly different for high and low preferring targets in high and low preferring sub-regions of PAC. In follow-up analyses we collapsed data across frequency-preferring voxel populations and the frequency carrier of the sounds.

### 3.2. Detection of a target selectively increases activation in superficial layers of PAC

In PAC all three perceptual conditions show an increase in response from deep to superficial layers (Fig. 3 A). When subtracting the no Target condition, the additional modulation induced by detection of a target is apparent as an increase from deep to superficial layers (Fig. 3D - red line), while non detected targets do not result in a significant change in response compared to no target trials (Fig. 3D - black line). We fitted a generalized linear mixed effects (GLME) model for each ROI with three predictors (depth [categorical, 3 levels], condition [detected minus no target; undetected minus no target] and their interaction) to test for the non-frequency-specific effect of target detection and its interaction with cortical depth. We detected a significant interaction between depth and

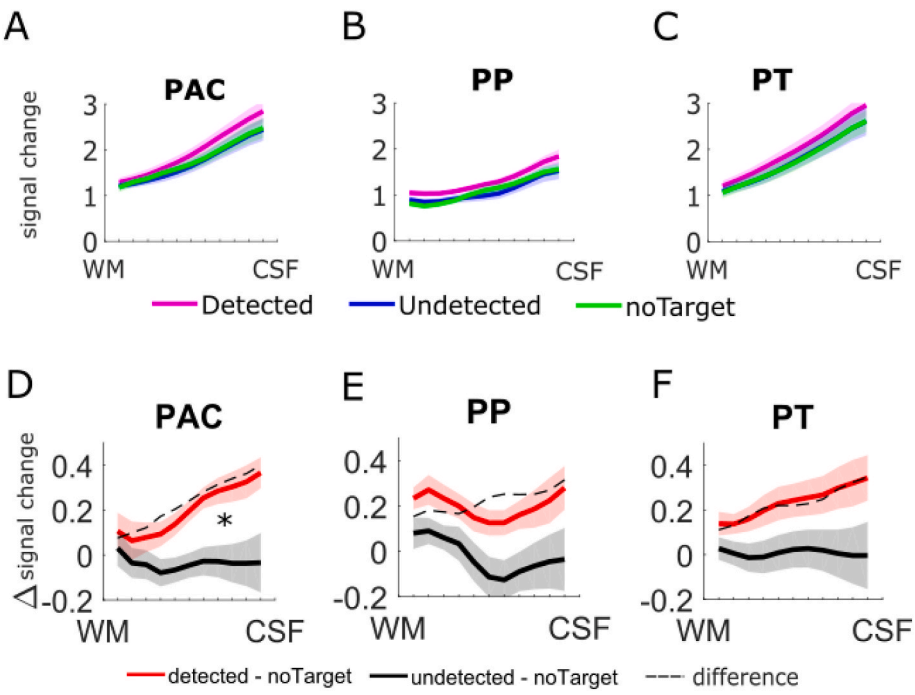


**Fig. 3. Layer-specific BOLD response for the different perceptual conditions per ROI. A.** BOLD response to detected (magenta), undetected (blue) and no Target (green) sounds in the different layers of PAC, averaged over trials and participants. **B.** Same as A but in planum polare (PP). **C.** Same as A but in planum temporale (PT). Laminar profiles in all ROIs show an increase towards the cortical surface (closer to CSF). **D.** Difference in BOLD response between detected and no Target sounds (red; detected - no Target) and, undetected and no Target sounds (black; undetected - no Target) show a modulation of the BOLD response towards superficial layers of PAC driven by detection. Dashed line depicts difference between red and black line. **E.** Same as in D but for PP. No significant differences between BOLD responses across depth to detected and undetected targets are observed in area PP. **F.** Same as in D but for PT. No significant differences between BOLD responses across depth to detected and undetected targets are observed in area PT. Shading indicates the standard error of the mean across participants. Figures S3.1-S3.6 show single participant plots and the results for the HG ROI (not depicted here). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

condition in PAC ($F(2,174) = 6.63$, $p < 0.01$). Follow-up ANOVAs per detection condition showed a significant effect of depth for the condition represented by the subtraction of detected and no Target trials ($F(1, 87) = 17.72$, $p < 0.001$) and no significant effect of depth for the condition represented by the subtraction of undetected and no target trials ($F(1,87) = 0.22$, $p = 0.8$). The significant main effect of depth for the condition represented by the subtraction of detected and no target trials warranted the comparison in activation across the three depth levels. After correcting for the multiple comparisons (Bonferroni, across cortical depths) we observed significant differences between the deep and superficial ($t(1,174) = 22.71$, $p < 0.01$), as well as middle and superficial ($t(1,174) = 6.8$, $p < 0.05$) depths. The difference between deep and middle depth was not significant ($t(1,174) = 4.65$, $p > 0.05$). This result is indicative of feedback related signals affecting the processing of superficial layers of PAC in relation to detected targets. When considering the whole of Heschl's gyrus as region of interest ($F(1,216) = 6.06$, $p < 0.05$; Fig. S3.6), only the comparison between the deep and superficial depth bin ($t(1,174) = 9.53$, $p < 0.01$) was significant, but not between middle and superficial depth ($t(1,174) = 3.19$, $p > 0.05$) or between deep and middle depth ($t(1,174) = 1.69$, $p > 0.05$).

### 3.3. Detection effects in non-primary areas

In planum polare (PP) responses to detected and undetected sounds did not differ across depth (Fig. 3 C & F – non-significant interaction ($F(2,174) = 1.00$, $p > 0.05$), while a main effect of detection was present ($F(2,174) = 9.55$, $p < 0.01$). Similarly, in planum temporale (PT) no significant interaction of condition and depth was observed ($F(2,174) = 2.75$, $p > 0.05$). A main effect of detection ($F(1,174) = 6.29$, $p < 0.05$) as well as a main effect of depth was present ($F(2,174) = 5.07$, $p < 0.01$).

## 4. Discussion

### 4.1. Layer-specific effects of detection in PAC

Previous (human) fMRI research has highlighted the modulation of superficial cortical layers of (primary) auditory cortex when attending and responding to auditory stimuli (and ignoring visual ones) (De Martino et al., 2015; Gau et al., 2020). Here, we aimed to understand whether feedback mechanisms can explain why physically identical stimuli presented under identical attentional instructions are detected in some trials and not in others. To do so, we measured laminar fMRI responses from human auditory cortex, while participants performed a temporal target detection task at perceptual threshold. This allowed us to contrast the response to detected and undetected targets, while the bottom-up acoustic information remained identical. We showed that detected targets elicited a comparatively stronger response in superficial layers of the primary auditory cortex, indicating the relevance of feedback processing. In non-primary regions (PP & PT) detecting a target resulted in a stronger response (compared to non-detected targets), yet this differential response did not vary across layers.

We have reported our results after subtracting the response to no target trials from the responses to detected and undetected sounds. By doing so, we were able to control for offset effects induced by local vascular contributions to the BOLD signal, which should be consistent across the experimental conditions. This permitted highlighting the modulation induced by the detection of a target, despite the overall increase of the GE-EPI signal towards the pial surface (Uludağ and Blinder, 2018). Acquisition techniques such as 3D-GRASE (Oshio and Feinberg, 1991) and VASO (Huber et al., 2017) which are not (or less) affected by vascular draining exist, nevertheless, GE-EPI offers increased sensitivity (compared to both 3D-GRASE and VASO), coverage (compared to 3D GRASE) and temporal efficiency (compared to VASO) all of which were essential to our study (Moerel et al., 2021).

It is conceivable that the increase in response we observed in superficial layers of PAC could have been the result of a fluctuation of attentional sampling, which is known to modulate long-latency sensory responses (Snyder et al., 2012). Multiple recent laminar fMRI studies located top-down effects of attention in superficial layers of human visual and auditory cortex either by attending to different modalities (auditory and visual) or by studying feature-based attention within a modality (Lawrence et al., 2019; Liu et al., 2021; De Martino et al., 2015; Gau et al., 2020), but see (van Mourik et al., 2021). Invasive electrophysiological studies have also related changes in sensory gain of superficial layers to fluctuations of attention (Lakatos et al., 2013; O'Connell et al., 2014; Henry and Herrmann, 2014). Our results are thus consistent with the idea that attention modulates superficial layers of

(auditory) cortical areas and provide first evidence that these small fluctuations can make the difference between detecting or not detecting an otherwise identical stimulus.

Our results identify PAC as a target of such feedback signals. Previous MEG research has also suggested that feedback to PAC (a unique long-latency response ranging between 50 ms and 300 ms) may be relevant to the detection of target sounds (Giani et al., 2015; Gutschalk et al., 2008). At the macroscopic level, increased fMRI BOLD responses in PAC in response to detected targets have been suggested to be the result of feedback signals (Wiegand and Gutschalk, 2012), potentially originating in parietal areas (Giani et al., 2015; Cusack, 2005). The plausible involvement of feedback to primary (auditory) cortical areas in determining the detectability of a stimulus is also corroborated by studies on bistable perception, or auditory streaming. These studies reported variations in (primary) auditory cortex responses to changes in percept evoked by identical physical stimuli (Micheyl et al., 2005). Using fMRI, for example, responses in regions adjoining PAC have been associated with the perceptual interpretation of acoustically identical sounds (Kilian-Hütten et al., 2011) as well as to perceptual streaming (Hill et al., 2011).

### *4.2. Responses to detected targets are not modulated by frequency of the sounds*

Contrary to previous invasive electrophysiology studies and non-invasive human studies (Lakatos et al., 2013; O'Connell et al., 2014; Riecke et al., 2018), we did not find a significant effect when considering separately cortical regions whose preference was maximal for the carrier frequency of the sounds (e.g. high vs. low frequency). While the absence of evidence is not evidence of the absence, a possible explanation for such inconsistency may stem from the nature of the task we employed. In previous research reporting frequency specific effects in auditory cortical regions, the task entailed focusing attention to the spectral content of the sounds (Lakatos et al., 2013; O'Connell et al., 2014; Riecke et al., 2018). In our task, the carrier frequency of the sounds was not the target of attention as participants were instructed to detect temporal shifts embedded in the stream of sounds. This line of reasoning, and our results are in line with previous investigations showing an attentional enhancement in layer 2/3, independent of the preferred frequency of the recording site when sound frequency was not task-relevant (Francis et al., 2018).

In conclusion, the current study shows that when detecting a temporally shifted target, the response of neural populations in superficial layers of primary auditory cortex increases (in a non-frequency specific manner). This modulation is compatible with feedback signals targeting the primary auditory cortex Future studies may be directed at identifying the source of the feedback signal we identified in auditory cortex by assessing laminar resolved functional connectivity after data acquisition with larger brain coverage. Invasive investigation in animals (with micro stimulation or optogenetics used to modulate activity in layer-specific sources of feedback) could also be directed to the investigation of the causal relationships among feedback sources, superficial layers responses, and behavior.

## Credit author statement

Miriam Heynckes: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization; Agustin Lage-Castellanos: Methodology, Software; Peter De Weerd: Conceptualization,Writing – review & editing, Funding acquisition, Supervision; Elia Formisano: Conceptualization, Writing – review & editing, Funding acquisition, Supervision; Federico De Martino: Conceptualization, Validation, Writing – review & editing, Funding acquisition, Supervision

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Miriam Heynckes, Elia Formisano, Peter De Weerd reports financial support was provided by Dutch Research Council. Federico De Martino reports financial support was provided by European Research Council.

## Data availability

Data will be made available on request.

## Acknowledgements

We thank Sriranga Kashyap for help with the anatomical preprocessing, and our participants for their endurance during long scan sessions. This work was funded by The Netherlands Organization for Scientific Research (NWO) Research Talent Grant 406.17.200 awarded to MH, EF and PDW. FDM has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. ERC - CoG 2020–101001270). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crneur.2023.100075.

## References

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26 (3), 839–851.
Brainard, D.H., 1997. The Psychophysics toolbox. Spatial Vis. 10 (4), 433–436. https://doi.org/10.1163/156856897X00357.
Cusack, R., 2005. The intraparietal sulcus and perceptual organization. J. Cognit. Neurosci. 17 (4), 641–651. https://doi.org/10.1162/0898929053467541.
Da Costa, S., van der Zwaag, W., Miller, L.M., Clarke, S., Saenz, M., 2013. Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. J. Neurosci. 33 (5), 1858–1863. https://doi.org/10.1523/JNEUROSCI.4405-12.2013.
De Martino, F., Moerel, M., Ugurbil, K., Goebel, R., Yacoub, E., Formisano, E., 2015. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proc. Natl. Acad. Sci. USA 112 (52), 16036–16041. https://doi.org/10.1073/pnas.1507552112.
Douglas, R.J., Martin, K.A., 2004. Neuronal circuits of the neocortex. Annu. Rev. Neurosci. 27, 419–451. https://doi.org/10.1146/annurev.neuro.27.070203.144152.
Douglas, R.J., Martin, K.A., Whitteridge, D., 1989. A canonical microcircuit for neocortex. Neural Comput. 1 (4), 480–488. https://doi.org/10.1162/neco.1989.1.4.480.
Formisano, E., Kim, D.S., Di Salle, F., Van de Moortele, P.F., Ugurbil, K., Goebel, R., 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40 (4), 859–869. https://doi.org/10.1016/S0896-6273(03)00669-X.
Francis, N.A., Elgueda, D., Englitz, B., Fritz, J.B., Shamma, S.A., 2018. Laminar profile of task-related plasticity in ferret primary auditory cortex. Sci. Rep. 8 (1), 1–10.
Gau, R., Bazin, P.L., Trampel, R., Turner, R., Noppeney, U., 2020. Resolving multisensory and attentional influences across cortical depth in sensory cortices. Elife 9, e46856.
Giani, A.S., Belardinelli, P., Ortiz, E., Kleiner, M., Noppeney, U., 2015. Detecting tones in complex auditory scenes. Neuroimage 122, 203–213. https://doi.org/10.1016/j.neuroimage.2015.07.001.
Gutschalk, A., Micheyl, C., Oxenham, A.J., 2008. Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol. 6 (6), e138.
Henry, M.J., Herrmann, B., 2014. Low-frequency neural oscillations support dynamic attending in temporal context. Timing & Time Percept. 2 (1), 62–86.
Heschl, R.L., 1878. - Über die vordere quere Schläfenwindung des menschlichen Großhirns. URL: https://nbn-resolving.org/urn:nbn:de:gbv:9-g-4879281.
Heynckes, M., 2022. English translation of Heschl's 'On the anterior transverse temporal gyrus of the human cerebrum. February 3). https://doi.org/10.31234/osf.io/ahny9.
Hill, K.T., Bishop, C.W., Yadav, D., Miller, L.M., 2011. Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. BMC Neurosci. 12 (1), 1–8.
Huber, L., Handwerker, D.A., Jangraw, D.C., Chen, G., Hall, A., Stüber, C., et al., 2017. High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. Neuron 96 (6), 1253–1263.
Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage 62 (2), 782–790.

Kashyap, S., Ivanov, D., Havlicek, M., Poser, B., Uludag, K., 2019. Laminar CBF and BOLD fMRI in the human visual cortex using arterial spin labelling at 7T. In: Proceedings of the 27th Scientific Meeting of ISMRM, p. 609.

Kilian-Hütten, N., Valente, G., Vroomen, J., Formisano, E., 2011. Auditory cortex encodes the perceptual interpretation of ambiguous sound. J. Neurosci. 31 (5), 1715–1720. https://doi.org/10.1523/JNEUROSCI.4572-10.2011.

Kim, J.J., Crespo-Facorro, B., Andreasen, N.C., O'Leary, D.S., Zhang, B., Harris, G., Magnotta, V.A., 2000. An MRI-based parcellation method for the temporal lobe. Neuroimage 11 (4), 271–288.

Klein, B.P., Fracasso, A., van Dijk, J.A., Paffen, C.L., Te Pas, S.F., Dumoulin, S.O., 2018. Cortical depth dependent population receptive field attraction by spatial attention in human V1. Neuroimage 176, 301–312.

Lakatos, P., Musacchia, G., O'Connell, M.N., Falchier, A.Y., Javitt, D.C., Schroeder, C.E., 2013. The spectrotemporal filter mechanism of auditory selective attention. Neuron 77 (4), 750–761. https://doi.org/10.1016/j.neuron.2012.11.034.

Lawrence, S.J., Norris, D.G., De Lange, F.P., 2019. Dissociable laminar profiles of concurrent bottom-up and top-down modulation in the human visual cortex. Elife 8, e44422. https://doi.org/10.7554/eLife.44422.

Liu, C., Guo, F., Qian, C., Zhang, Z., Sun, K., Wang, D.J., et al., 2021. Layer-dependent multiplicative effects of spatial attention on contrast responses in human early visual cortex. Prog. Neurobiol. 207, 101897.

Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage 49 (2), 1271–1281.

Merzenich, M.M., Brugge, J.F., 1973. Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Res. 50 (2), 275–296. https://doi.org/10.1016/0006-8993(73)90731-2.

Merzenich, M.M., Knight, P.L., Roth, G.L., 1973. Cochleotopic organization of primary auditory cortex in the cat. Brain Res. 63, 343–346. https://doi.org/10.1016/0006-8993(73)90101-7.

Micheyl, C., Tian, B., Carlyon, R.P., Rauschecker, J.P., 2005. Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 48 (1), 139–148. https://doi.org/10.1016/j.neuron.2005.08.039.

Moerel, M., De Martino, F., Formisano, E., 2014. An anatomical and functional topography of human auditory cortical areas. Front. Neurosci. 8 (8 JUL), 1–14. https://doi.org/10.3389/fnins.2014.00225.

Moerel, M., Yacoub, E., Gulban, O.F., Lage-Castellanos, A., De Martino, F., 2021. Using high spatial resolution fMRI to understand representation in the auditory network. Prog. Neurobiol. 207 (January 2020), 101887 https://doi.org/10.1016/j.pneurobio.2020.101887.

Moore, B.C.J., 2003. An Introduction to the Psychology of Hearing, vol. 3. Boston Academic Press, p. 413. https://doi.org/10.1016/j.tins.2007.05.005.

O'Connell, M.N., Barczak, A., Schroeder, C.E., Lakatos, P., 2014. Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. J. Neurosci. 34 (49), 16496–16508. https://doi.org/10.1523/JNEUROSCI.2055-14.2014.

Oshio, K., Feinberg, D.A., 1991. GRASE (gradient-and spin-echo) imaging: a novel fast MRI technique. Magn. Reson. Med. 20 (2), 344–349.

Polimeni, J.R., Fischl, B., Greve, D.N., Wald, L.L., 2010. Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. Neuroimage 52 (4), 1334–1346.

Riecke, L., Peters, J.C., Valente, G., Poser, B.A., Kemper, V.G., Formisano, E., Sorger, B., 2018. Frequency-specific attentional modulation in human primary auditory cortex and midbrain. Neuroimage 174, 274–287. https://doi.org/10.1016/j.neuroimage.2018.03.038.

Snyder, J.S., Gregg, M.K., Weintraub, D.M., Alain, C., 2012. Attention, awareness, and the perception of auditory scenes. Front. Psychol. 3 (FEB), 1–17.

Turner, R., 2002. How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. Neuroimage 16 (4), 1062–1067.

Uludağ, K., Blinder, P., 2018. Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. Neuroimage 168, 279–295.

van Mourik, T., Koopmans, P.J., Bains, L.J., Norris, D.G., Jehee, J.F., 2021. Investigation of Layer Specific BOLD in the Human Visual Cortex during Visual Attention. bioRxiv.

Waehnert, M.D., Dinse, J., Weiss, M., Streicher, M.N., Waehnert, P., Geyer, S., et al., 2014. Anatomically motivated modeling of cortical laminae. Neuroimage 93, 210–220.

Wiegand, K., Gutschalk, A., 2012. Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. Neuroimage 61 (1), 62–69. https://doi.org/10.1016/j.neuroimage.2012.02.067.

Yushkevich, P., 2006. ITK-SNaP integration, NLM insight.