Title: Auditory predictions shape the neural responses to stimulus repetition and sensory change

Authors: Raffaele Cacciaglia, PhD[1,2], Jordi Costa-Faidella, PhD[1,2], Katarzyna Zarnowiec, MSc[1,2], Sabine Grimm, PhD[1,2], Carles Escera, PhD*[1,2,3]

[1]Brainlab - Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, 08035, Barcelona, Catalonia, Spain.

[2]Institute of Neurosciences, University of Barcelona, 08035, Barcelona, Catalonia, Spain.

[3]Institut de Recerca Sant Joan de Déu, 08950, Esplugues de Llobregat, Catalonia, Spain.

***Corresponding author:**

Carles Escera, PhD. Department of Clinical Psychology and Psychobiology, University of Barcelona, 08035, Barcelona, Catalonia, Spain. Electronic address: cescera@ub.edu.

Declarations of interest: none

Number of words in the Abstract: 230

Number of words in the Introduction: **716**

Number of words in the Discussion: **2248**

Number of Tables: **4**

Number of Figures: **5**

Number of Supplementary Materials: 1

ABSTRACT

Perception is a highly active process relying on the continuous formulation of predictive inferences using short-term sensory memory templates, which are recursively adjusted based on new input. According to this idea, earlier studies have shown that novel stimuli preceded by a higher number of repetitions yield greater novelty responses, indexed by larger mismatch negativity (MMN). However, it is not clear whether this MMN memory trace effect is driven by more adapted responses to prior stimulation or rather by a heightened processing of the unexpected deviant, and only few studies have so far attempted to characterize the functional neuroanatomy of these effects. Here we implemented a modified version of the auditory frequency oddball paradigm that enables modeling the responses to both repeated standard and deviant stimuli. Fifteen subjects underwent functional magnetic resonance imaging (fMRI) while their attention was diverted from auditory stimulation. We found that deviants with longer stimulus history of standard repetitions yielded a more robust and widespread activation in the bilateral auditory cortex. Standard tones repetition yielded a pattern of response entangling both suppression and enhancement effects depending on the predictability of upcoming stimuli. We also observed that regularity encoding and deviance detection mapped onto spatially segregated cortical subfields. Our data provide a better understanding of the neural representations underlying auditory repetition and deviance detection effects, and further support that perception operates though the principles of Bayesian predictive coding.

Keywords: Auditory deviance detection; Repetition suppression; Repetition enhancement; Predictive coding, event-related fMRI

INTRODUCTION

Perception is a highly active process that involves the generation of sensory predictions in high-order cortical fields based on previous experience (Bar, 2009). Recent formulations suggest that predictions serve to generate hypotheses about upcoming information, which undergo recursive updates along increasingly complex levels of the hierarchical neural processing (Kanai et al., 2015). This reiterative comparison between sensory evidence and internal models generates prediction errors, which are sent to higher cortical levels in order to optimize representational templates (Friston, 2005). In the last decade, this view of the brain as a Bayesian predictive machine has gained increasing neurophysiological support and has been fostered by a number of empirical studies in visual (Hughes and Waszak, 2014; Dunovan et al., 2014) and auditory (Wacongne et al., 2011; SanMiguel et al., 2013) modalities. Predictions are formulated on the basis of extracted sensory features, the most important being statistical regularities inferred from stimulus repetition (Grill-Spector et al., 2006).

The auditory system represents an ideal machinery for probing repetition effects, as the encoding of statistical regularities is crucial for an efficient processing of serial information, since sounds are transient in nature (Bendixen, 2014). The mismatch negativity (MMN) represents the most well-studied index of deviance detection in the auditory domain (Näätänen et al., 1978, 2007), and it has commonly been interpreted as a brain signature of the prediction error (Bendixen et al., 2009; Friston, 2005; Garrido et al., 2009; Grimm and Escera, 2012; but see: May and Tiitinen, 2010; Fishman, 2014). MMN is typically elicited by presenting sequences of identical stimuli in a row and by occasionally replacing them with a deviant sound that does not match the previous chain of stimuli, therefore violating the regularity (Escera and Malmierca, 2014; Escera et al., 2014). MMN generators have been located in supratemporal (Recasens et al., 2014; Maess et al., 2007) and frontal (Deouell, 2007; Rinne et al., 2000) areas. Its amplitude increases with increased number of standard repetitions (*i.e.*, decreased local deviant probability), suggesting that a sensory memory trace is strengthened with a longer regular stimulus history that precedes the deviant (Cowan et al.,

1993; Imada et al., 1993; Javitt et al., 1998; Sabri and Campbell, 2001; Sams et al., 1983). Accordingly, several studies using electroencephalography (EEG) have shown the existence of a repetition positivity (RP), a combined modulation of the P50, N1 and P2 event-related potentials (ERPs) which is positively related to MMN amplitude and that represents a direct signature of auditory sensory memory (Baldeweg et al., 2004; Haenschel et al., 2005; Costa-Faidella et al., 2011a; 2011b).

In line with earlier EEG experiments, functional magnetic resonance imaging (fMRI) studies have shown that violations of acoustic regularity trigger significant activations in auditory as well as inferior frontal regions (Doeller et al., 2003; Opitz et al., 2002, 2005; Sabri et al., 2004; Schall et al., 2003). Additionally, recent fMRI investigations have provided evidence for the involvement of subcortical stations of the auditory pathway in processing deviant sounds (Cacciaglia et al., 2015; Gao et al., 2014).

However, no fMRI study has so far attempted to characterize the functional neuroanatomy of the responses preceding a sensory change, that is, examining how the statistical regularity is being represented in the auditory brain. More specifically, how such echoic memory is dynamically modulated as a function of recent stimulus history is yet to be addressed.

In the present study, we capitalized on the superior spatial resolution of the fMRI, to tap into the spatial encoding of stimulus repetition and to explore how previous stimulus history modulates the response to a deviant sound. Additionally, we aimed to study the cerebral topology associated with the effects of stimulus repetition on standard and deviant tones processing separately. We took advantage of an experimental paradigm that allows modeling the responses to both deviant as well as the repeated standard stimuli, the roving standard paradigm (Cowan et al., 1993), which has been already implemented in EEG experiments (Baldeweg et al., 2004; Costa-Faidella et al., 2011a; Haenschel et al., 2005; Spriggs et al., 2018). We hypothesize to find a progressive decrease of the hemodynamic response along with standard stimuli presentation, involving repetition suppression effects as typically observed in fMRI adaptation paradigms (Grill-Spector et al., 2006).

Additionally, we predict to find greater and more widespread brain responses for deviants preceded by a larger number of standard sounds.

METHODS AND MATERIALS

Study participants

15 healthy participants (9 female, mean age = 25.6, standard deviation [SD] = 4.3, 3 left handed) took part in our experiment upon monetary compensation. None of them was under current or chronic medication. All participants had normal hearing, with a mean hearing threshold below 25 dB sound pressure level (dB-SPL), as assessed with binaural audiometric test using pure tones at five frequencies (250, 500, 1000, 3000, and 8000 Hz). The experimental protocol was approved by the Ethical Committee of the University of Barcelona and was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written informed consent was obtained for all participants before the experiment.

Stimuli and experimental design

All stimuli were generated with Matlab R2014a (The MathWorks Inc., Natick, MA, USA) and were binaurally delivered using the software Presentation (Version 0.70, www.neurobs.com) through a MR-compatible headset which attenuates scanner noise by ~ 15 dB-SPL (VisuaStim digital, Resonance Technology Inc., Northridge, CA, USA). Stimuli were 9 pure tones (PTs) frequency sampled at 44.1 kHz, of 100-ms duration with 5 ms of rise and 5 ms of fall time. Frequency spacing between adjacent tones was calculated according to the formula $\Delta f = (f2-f1)/(f2 \times f1)^{½}$ (Ulanovsky et al., 2003). The resulting selected frequencies were 300, 399, 530, 704, 935, 1242, 1651, 2193, and 2914 Hz. Stimulus intensity (sound pressure level, SPL) was individually calibrated as being 20% above the discrimination level with respect to the scanner noise. PTs were arranged in a roving standard fashion (Cowan et al., 1993), similarly to previous studies conducted in our laboratory (Costa-Faidella, et al., 2011a; Recasens et al., 2015). More specifically, tones were

arranged in trains of the same frequency but different lengths (4, 12, 24 or 36 repetitions), and delivered at a constant SOA of 500-ms with no inter-train pauses (Fig.1). This way, the first tone of a given train always acted as low-probability event, or deviant (DEV), while the subsequent repeated tones acted as standard stimuli (STD). Such an experimental design allows to simultaneously assess the neural correlates of repetition effects as well as the response to infrequent tones (Baldeweg et al., 2004, 2006; Costa-Faidella et al., 2011a). Our event of interest - hereinafter referred to as "trial" - consisted of a sequence of 4 consecutive PTs, which covered the duration of one scan repetition time (TR: 2000 ms) throughout the fMRI acquisition (see below). Stimulus frequency was equally represented in each of the four train lengths. Hence, we included 6 trains for each of the 9 frequencies yielding 54 trains of four different lengths, resulting in a total of 216 trains. To control for the effect of frequency spacing on deviance detection (Yago et al., 2001; Novitski et al., 2004) and to assure that any observed difference between the four DEV categories was attributable to the different number of preceding STD stimuli, inter-train frequency spacing was balanced across the four train lengths.

Prior to the beginning of the session, subjects were instructed to passively listen to the stimuli while watching a silent subtitled movie, in order to divert their attention from the auditory stimulation. The second session consisted of the actual experimental paradigm.

fMRI data acquisition

Functional magnetic resonance images were collected with a 3T full body scanner (Magnetom Siemens Trio, Siemens Medical Solutions, Erlangen, Germany), equipped with a phased-array transmit/receive head coil. Blood oxygenation level-dependent (BOLD) contrast images were acquired using a T2*-weighted gradient-echo Echo Planar Imaging (EPI) sequence (echo time [TE] = 40 ms, repetition time [TR] = 2000 ms, flip angle = 90°, field of view [FoV] = 220 mm, matrix size = 128 x 128 mm, voxel size = 1.7 x 1.7 x 3.5 mm; interslice gap = 0.8 mm, N. of slices = 24). A total of 1029 functional volumes were acquired, each covering the whole brain excluding a small

portion of the most posterior dorsal aspect of the parietal lobes. Slice orientation of the axial plane was set by forming a 45 ^{o} angle with respect to the longitudinal axis of the brainstem. This minimizes the heartbeat-related motion along the dorso-ventral and rostro-caudal axes of the brainstem and allows for better image quality of the midbrain auditory nuclei, without affecting image quality in cortical areas (Slabu et al., 2010; Cacciaglia et al., 2015). Prior to scanning, 3 dummy functional volumes were acquired and discarded in order to allow for T_1 saturation effects. For anatomical reference, structural images were acquired using a T_1-weighted high-resolution 3D gradient echo pulse sequence (TE = 2.98 ms, TR = 2300 ms, flip angle = 9°, voxel size = 1 x 1 x 1mm). The experimental session lasted approximately 40 minutes.

fMRI data analysis

Preprocessing, first and second level analyses were conducted with Statistical Parametric Mapping (SPM 12, Wellcome Department of Imaging Neuroscience, London, UK). Time-series were slice-time corrected to reference slice 12 for difference in acquisition timing and realigned with a two-pass procedure in which functional volumes were registered to the first volume in the series, and to the mean image of all the realigned volumes. None of the study participants exceeded motion estimates of 2 mm and 2 degrees. After realignment, images were co-registered with the individual structural volumes and normalized to the EPI template provided by SPM. Finally, images were spatially smoothed with a 6-mm full-width at half maximum (FWHM) Gaussian kernel. To remove low-frequency noise, a high-pass filter (cutoff 1/128Hz) was applied and the time-series were corrected for serial autocorrelations using first-order autoregressive functions AR(1). At single-subject level, a fixed effects analysis was conducted by setting up a general linear model (GLM) including 12 orthogonal regressors each coding the conditions for STD trials (STD_4, STD_8, STD_{12}, STD_{16}, STD_{20}, STD_{24}, STD_{28}, STD_{32}) and 4 conditions for DEV trials (DEV_4, DEV_{12}, DEV_{24}, DEV_{36}), together with the six motion parameters generated during realignment. Regressors were constructed with event onsets, where each event corresponded to a mini-sequence of four

consecutive tones lasting 2 seconds. These inputs were convolved with a canonical hemodynamic response function (first order expansion) to form the design matrix. Parameters estimates were computed by extracting the mean value within each significant cluster with the MARSeille Boîte À Région d'Intérêt (MARSBAR) toolbox (http://marsbar.sourceforge.net/).

Statistical Analyses

In order to measure auditory deviance detection, we performed a one sample *t*-test assessing the BOLD response to all DEV compared to all STD trials (DEV_{ALL}>STD_{ALL}). Next, to assess auditory change detection as a function of different number of preceding STD, we performed a within-subject analysis of variance (ANOVA) in SPM, where individual beta images corresponding to the four DEV trial types derived from the 1st level design matrices were introduced as within-subject factor. Modulation of the BOLD response as a function of the number of preceding STDs was then assessed by performing an omnibus *F*-test interrogating for any differences among the experimental conditions. Upon significance of the *F*-test, pair-wise *t*-tests were subsequently performed to directly compare the responses to each DEV type. Next, to rule out any potential influence of residual BOLD response of preceding sounds in the activity captured by any given DEV regressor, we broke down the DEV_{ALL}>STD_{ALL} comparison in distinct contrasts capturing the difference between each DEV type against the corresponding STD trial, preceded by an identical number of stimuli (DEV_x>STD_x, with *x* denoting the number of preceding stimuli). The resulting contrasts were DEV_4>STD_4, DEV_{12}>STD_{12}, and DEV_{24}>STD_{24}. Note that the STD_{36} condition to be compared to DEV_{36} was missing in our design, because the longest train of 36 repetitions allowed modelling up the trial STD_{32}. Similarly to analysis describe above, we then compared the activity captured by any of these differential contrasts among each other, using pair-wise *t*-tests. For these contrasts, effects sizes were computed voxel-wise by computing the difference between the mean of the respective contrast images across subjects and dividing it by the pooled standard deviation.

In order to track the hemodynamic response along repeated STD trials, we performed a within-subjects ANOVA including beta images retrieved from single-subjects' design matrices, similar to the model described above for DEV trials.

Finally, to determine whether stimulus repetition differentially modulated the cortical encoding of STD and DEV trials, we compared the differential response between any given DEV category preceded by different number of stimuli against the correspondent differential contrast involving STD trials (*e.g.*, [DEV_{12} > DEV_4] *vs.* [STD_{12} > STD_4]). These comparisons were assessed using unbiased *F*-tests.

For all comparisons, we firstly performed a hypothesis-driven analysis that was restricted to all voxels within *a-priori* defined anatomical regions of interest (ROI). Masks for the ROI analysis included the superior temporal gyrus (STG), the Heschl's gyrus (HG) and the inferior frontal gyrus (IFG) bilaterally (Doeller et al., 2003; Opitz et al., 2002, 2005; Schall et al.,2003), and were generated using the Automated Anatomical Labeling (Tzourio-Mazoyer et al., 2002), implemented in the Wake Forest University Pick Atlas toolbox (Maldjian et al., 2003).

The IFG ROI comprised three separate masks according to its cytoarchitectonic subdivisions, that is, the *pars triangularis*, *pars orbitalis* and *pars opercularis* (Anwander et al., 2007; Petrides et al., 2012). Masks for subcortical areas were defined by generating a 5-mm radius sphere centered around standardized coordinates previously reported for the inferior colliculus (Mühlau et al., 2006) and medial geniculate body (von Kriegstein et al., 2008).

For the ROI analysis, we selected a primary voxel-wise threshold of uncorrected $P<0.001$ and considered significant results that survived a correction for multiple testing using a family-wise error rate approach ($P_{FWE}<0.05$) on the cluster level. Finally, to detect potential effects in additional brain regions, we performed an additional separate unconstrained analysis. For this whole-brain analysis, results were considered significant if surviving a voxel-wise statistical threshold of uncorrected $P<0.001$, with a cluster-extent threshold of 15 voxels (Woo et al., 2014).

RESULTS

Auditory deviance detection

When contrasting all DEV versus all STD trials (DEV_{ALL}>STD_{ALL}) using the *a priori* defined ROIs, we found significant responses in the bilateral STG, as well as bilateral HG but no significant activations were detected in the IFG, or subcortical stations (*i.e.*, MGB or IC) (Supplementary Fig. 1). The exploratory whole-brain analysis confirmed significant activations in two auditory cortical areas corresponding to the right (t_{14} = 8.68, cluster size [k] = 2121, P < 0.001, [x = 66, y = -28, z = 16]) and left (t_{14} = 7.99, k = 1183, P < 0.001, [x = -50, y = -28, z = 10]) STG extending to the bilateral HG. Additionally, this unbiased analysis revealed a significant activation in one cluster comprising the dorsal middle cingulate cortex extending to the superior frontal gyrus (t_{14} = 5.83, k = 23, P < 0.001, [x = 14, y = 14, z = 34]).

Effects of number of preceding STD stimuli on DEV detection response

Next, we examined whether the number of preceding STD stimuli modulated the response magnitude to DEV tones. First, the ANOVA comprising four regressors each encoding a specific class of DEV trial (DEV_4, DEV_{12}, DEV_{24}, DEV_{36}), yielded a significant main effect of DEV position in the bilateral STG (right: $F_{3,42}$ = 20.87, k = 496, P_{FWE} < 0.001, [x = 62, y = -14, z = 2]; left: $F_{3,42}$ = 15.32, k = 265, P_{FWE} < 0.001, [x = -50, y = -28, z = 10]), as well as the bilateral HG (right: $F_{3,42}$ = 21.79, k = 119, P_{FWE} < 0.001, [x = 50, y = -10, z = 4]); left: $F_{3,42}$ = 11.01, k = 17, P_{FWE} = 0.009, [x = -42, y = -26, z = 10]) (Fig. 2a-2b). Pairwise post-hoc t-tests revealed that DEV_4 yielded a significant reduced response in auditory areas compared to any other DEV type and that DEV_{24} yielded the highest response magnitude (Table 1). No further significant responses in any region emerged in the whole-brain analysis. In the subsequent analysis controlling for the number of preceding tones across stimulus type (*i.e.*, DEVx>STDx) we observed significant responses in the bilateral STG for all the three contrasts (Table 2; Fig3a-3d). The comparisons DEV_{12}>STD_{12}, and DEV_{24}>STD_{24} additionally revealed a significant response in the bilateral HG, but no

significant responses in IFG, MGB or IC were detected. The whole-brain analysis for these contrasts did not return any additional significant activation. When assessing statistical differences between the three differential contrasts, we found significant responses in the STG with the contrast [DEV_{24}>STD_{24}] > [DEV_4>STD_4] additionally yielding activation in the bilateral HG (Table 3, Fig 4a-4c, Fig. 4g). For each comparison, the effect size computed in the respective local maxima for the STG and HG was relatively high with values between 0.77 and 2.27, indicating robust differences among each condition (Table 3; Fig. 4d-5f). The unbiased whole-brain analysis did not retrieve additional responses in any areas. Overall these results indicate that the magnitude of auditory deviance detection was significantly modulated by the number of preceding STD stimuli, with increased number of repetition yielding stronger DEV-related responses.

Cortical encoding of statistical auditory regularities

The within-subjects ANOVA conducted on the estimated beta images corresponding to all STD trials revealed a significant main effect of tone repetition in the right ($F_{7,98}$ = 13.59, k = 1173, P_{FWE} < 0.001, [x = 66, y = -22, z = 6]) and left ($F_{7,98}$ = 10.26, k = 724, P_{FWE} < 0.001, [x = -48, y = -24, z = 4]) STG, as well as the right HG ($F_{7,98}$ = 9.70, k = 107, P_{FWE} < 0.001, [x = 48, y = -21, z = 6]). No further brain regions showed significant activity in the unbiased analysis.

Fig. 5a-5b shows the spatial topography as well as the temporal evolution of the hemodynamic response magnitude along the STD tones repetition. We observed a pattern of variability where the response intensity appeared to be driven by the predictability of a DEV trial occurrence. Specifically, we observed response decrease along those STD trials which were not followed by a DEV, entangled with an enhanced activity for those STDs which were potentially followed by DEV trials. A confirmatory analysis was conducted by performing a t-test comparing the brain activity to STDs immediately followed by a probable DEV trial (STD_8, STD_{20} and STD_{32}), against the response to STDs which were never followed by DEV trials (STD_4, STD_{16} and STD_{28}).

The ROI analysis revealed significant activations in the bilateral STG (right: t_{98} = 5.64, k = 403, P_{FWE} < 0.001, [x = 56, y = -20, z = 8]; left: t_{98} = 5.02, k = 261, P_{FWE} < 0.05, [x = -44, y = -24, z = 0]), as well as the right HG (t_{98} = 5.78, k = 118, P_{FWE} < 0.05, [x = 40, y = -22, z = 12]). The whole-brain analysis did not returned significant responses in any other brain region. To rule out the possibility that some residual activity encapsulated in DEV regressors may affect the response to STD trials, we set up a new model which included only trains of 36 repetitions, where no physical DEV was ever delivered, and repeated the same analysis. We found a pattern of responses which was highly consistent with the model including all trains (Fig. 5c-5d), Specifically, the omnibus F-test interrogating for any differences among STD trials revealed a significant effect of STD repetition in the bilateral STG (right: $F_{7,98}$ = 11.30, k = 1058, P_{FWE} < 0.001, [x = 58, y = -22, z = 8]; left: $F_{7,98}$ = 8.56, k = 652, P_{FWE} < 0.001, [x = -44, y = -30, z = 8]). Again, the post-hoc t-test yielded significant activations in the bilateral STG (right: t_{98} = 4.90, k = 295, P_{FWE} < 0.001, [x = 56, y = -20, z = 8]; left: t = 5.36, k = 200, P_{FWE} < 0.001, [x = -44, y = -30, z = 8]). These confirmatory results suggest that the dynamically modulated response along STD trial repetitions relies on the auditory prediction of the upcoming sensory change

Impact of stimulus repetition on STD and DEV processing

F-contrasts comparing the differential response to DEV and STD trials preceded by different number of stimuli (*e.g.*, [DEV_{12} > DEV_4] *vs.* [STD_{12} > STD_4]) revealed significant effects in auditory areas (Table 4), indicating that stimulus repetition yielded significantly different changes in the response to STD and DEV.

In order to pinpoint the cortical topology associated with changes in STD and DEV processing along stimulus repetition, we conducted a one-way ANOVA on the individual MNI coordinates retrieved from single-subject *F*-contrasts testing for any difference across all STD and all DEV separately. We found a significant effect of stimulus type in the right sagittal plane ($F_{1,28}$ = 7.94, P = 0.009), which revealed a significantly different topological pattern. Specifically, the differential

response to distinct DEV trials mapped onto a more medial aspect of the *planum temporale* in the close proximity of primary auditory cortex, with respect to the modulation of response along STD trials (Supplementary Fig. 2).

DISCUSSION

In the present study, we have implemented a frequency roving standard paradigm while recording brain hemodynamic responses in healthy young individuals in order to characterize the functional neuroanatomy of auditory stimulus repetition effects and to explore how previous stimulus history modulates the response to deviant sounds. Notably, since we controlled for inter-train frequency spacing, our paradigm allowed isolating the unique effects of recent stimuli history and the resulting change in predictability of upcoming deviant sounds.

First, we found that regardless of the number of preceding stimuli, the comparison between all DEV against all STD trials yielded significant activations in the STG and HG, bilaterally. This is in line with several previous reports using fMRI in classic frequency oddball paradigms and therefore corroborates the effectiveness of our experimental paradigm (Doeller et al., 2003; Opitz et al., 2002, 2005; Sabri et al., 2004; Schall et al., 2003). Surprisingly however, we did not find significant activations in the IFG, a result that was observed in previous studies examining auditory deviance detection (Auksztulewicz and Friston, 2015; Moran et al., 2014; Rinne et al., 2000; Doeller et al., 2003; Opitz et al., 2002). Such a lack of IFG activation is not highly exceptional in fMRI research on auditory deviance detection. Using frequency oddball paradigms, some earlier fMRI studies reported an involvement of the IFG (e.g. Molholm et al., 2005; Rinne et al., 2007; Yucel et al., 2005), while others did not (Szycik et al., 2014; Sabri et al., 2004; Sabri et al., 2006; Opitz et al., 2005). Moreover, in two fMRI studies (Doeller et al., 2003; Opitz et al., 2002) the involvement of inferior frontal areas was only found for those DEV>STD contrasts which were maximal in frequency separation, a factor that we treated as confounder in our analyses. Yet Spriggs et al. (2018) found a significant IFG involvement in auditory deviance detection using EEG recordings in

a protocol which was very similar to the one tested here (*i.e.*, frequency roving standard paradigm). Overall, this suggests that compared to superior temporal activations, inferior frontal activation is less consistently detected in fMRI studies, which has led to the suggestion that IFG response to auditory change might involve an increase in synchronization of neurons rather than an increase in the number or their firing rates (Deouell, 2007). Most importantly, we should underline that ours is the first study implementing a roving standard paradigm using event-related fMRI and therefore a straightforward comparison with earlier studies, which all used the classic oddball design, remains somehow difficult. In addition to auditory areas, the whole brain analysis performed for all DEV_{ALL}>STD_{ALL} retrieved a significant activation in the middle cingulate gyrus extending to the superior frontal gyrus. The cingulate cortex, although most typically its anterior subdivision, has been included in the network subserving auditory deviance detection and its function is related to automatic error detection as well as conflict monitoring (Kiehl et al., 2005). Its middle portion displays reciprocal connection with the insula and other areas within the salience network (Menon and Uddin, 2010) and its activity in our paradigm may reflect a general mechanism of novelty detection.

Next, we found progressively enhanced and more spatially extended responses for those DEV preceded by a higher number of STD trials. Such evidence was corroborated by two findings. First, the ANOVA conducted on DEV trials yielded greater activation along with the number of preceding STD, revealing a maximal response for DEV_{24} (Fig. 1). Post-hoc comparisons confirmed that DEV_{24} condition yielded a larger response compared to any other DEV condition (*i.e.*, DEV_4, DEV_{12}) and that DEV_4 retrieved significantly reduced responses compared to both DEV_{12} and DEV_{36} (Table 1). No further response increase was observed for DEV_{36} trials, that is, both the DEV_{36}>DEV_{24} and DEV_{36}>DEV_{12} comparisons did not return significant activations. This may be related to a saturation of the echoic sensory memory register, or may alternatively be explained by the fact that, after the 36th STD repetition, the DEV_{36} occurrence was highly expected and therefore it did not elicit a stronger response. Second, when comparing each DEV trial category against the

respective STD with identical stimulus history, we found progressively increased auditory responses as the number of preceding STD increased. We directly tested whether these differential responses were significant and found that the right STG was activate din both [DEV_{12}>STD_{12}] > [DEV_4>STD_4] and [DEV_{12}>STD_{12}] > [DEV_4>STD_4] comparisons. Additionally the comparison [DEV_{24}>STD_{24}] > [DEV_4>STD_4] yielded a significant response in the left STG and the bilateral HG. The effect sizes computed voxel-wise for each differential contrast were relatively high, further suggesting that the magnitude of deviance detection was moderated by the number of preceding STDs. Our data provide the first evidence using BOLD fMRI that the response magnitude of auditory deviance detection depends on the number of preceding stimuli. This is consistent with the view that increasing repetitions of preceding STD stimuli strengthen a sensory memory trace, leading to larger responses to upcoming deviants (Cowan et al., 1993; Imada et al., 1993; Javitt et al., 1998; Sabri and Campbell, 2001; Sams et al., 1983). However, it is still possible that the observed effects partially rely on increasingly adapted responses of feature-selective neurons along with STD repetition, whereas a different neural population tuned to features present in the DEV would yield stronger responses, particularly with larger time gaps between single DEV events (Fishman et al., 2014). Nevertheless, as we will review below, we did not observe a ubiquitous adapted hemodynamic response along sequential STD trials, and this seems to counter such a theoretical explanation.

We next examined the temporal development of the hemodynamic response along stimulus repetition. Our initial hypothesis was to find progressively more adapted auditory responses along with the unfolding of STD stimulus delivery, in line with previous fMRI adaptation models (Grill-Spector et al., 2006). However, the results showed a complex but coherent response pattern which embedded a local decrease of activity for those STD trials that never predicted a DEV occurrence (*i.e.*, STD_4, STD_{16} and STD_{28}), together with a response increase for those STD potentially followed by a DEV trial (*i.e.*, STD_8, STD_{20} and STD_{32}). Such a differential response was further corroborated by our finding on a greater auditory cortex response to DEV-predicting against DEV-

unpredicting STD trials (*i.e.*, never followed by a DEV). Importantly, when repeating the same analysis in trains of 36 repetitions only, where no DEV was physically delivered, we observed a highly consistent response pattern, suggesting that such modulatory effect was at least partially driven by auditory predictions.

These results suggest that repetition effects are not merely driven by purely bottom-up processing, such as neural adaptation, but also by a process of perceptual inference driven by stimulus predictability, which generates a probabilistic expectation of the stimuli up to come (Friston, 2005; Baldeweg, 2006; Grotheer and Kovacs, 2016). In support of our interpretation, earlier studies reported that the magnitude of repetition effects could be modulated by stimulus expectation and precisely that repetition suppression was reduced when a repetition was less likely to occur (Summerfield et al., 2008; Todorovic et al., 2011). Similarly, we have observed that increased predictability of a deviant sound (*i.e.*, decreased probability of STD repetition) shifted repetition suppression to enhancement (*i.e.*, produced a reduction in repetition suppression). Our interpretation is consistent with previous data showing that prior expectation of a specific stimulus evokes a feature-specific pattern of activity in sensory cortices similar to that evoked by the corresponding actual stimulus (Kok et al., 2017). Such an increase in the preparatory response of sensory neurons induced by stimulus expectation optimizes processing of the predicted stimuli (Kok et al., 2014; Kok et al., 2012; Hindy et al., 2016). Our data further suggest that sensory predictions may occur in absence of modulatory top-down afferences, as we did not find here significant responses in frontal areas.

Our findings on increased responses to DEV preceded by high number of STD seems countering earlier reports on smaller activity elicited by DEV tones which were arranged in a predictable fashion (Lecaignard et al., 2015). However, unlike these earlier reports our paradigm does not render the DEV fully predictable because trains of different frequencies and different number of repetitions appear in a randomized fashion. This implies that the response to DEV preceded by more STD tones does not have to be necessarily smaller, because subjects cannot be certain it will

appear. With this respect, it is worth noting that in our paradigm, DEV predictability may be subject to two distinct probabilistic inferences, which respond to either a local or global rule. On one side, DEV preceded by a higher number of STD trials (*i.e.*, DEV_{24} or DEV_{36}) are less probable because of the longer stimulus history, but the same time they are more predictable because of the increased cumulative probability due to putative sequential learning effects. Thus, one intriguing possibility is that the cortical computational mechanisms underlying auditory predictive coding operate through a dual yet parallel probabilistic inference processes, the former being nested in primary sensory regions and accounting for local stimulus probability, the latter recruiting more widespread associative areas and accounting for more complex rules. In the context of our experiment, the former mechanism would explain the higher response magnitude we found for DEV with a longer stimulus history, while the latter generates the modulated response for repeated STD tones entangling both suppression and enhancement effects, depending on stimulus cumulative probability. In support of this, we found that stimulus repetition had a differential impact on the modulated response along STD and DEV trials (Table 4; Supplementary Fig. 2). Additionally, when comparing the cortical topology associated to the effect of stimulus repetition for STD and DEV trials, we found that the two classes of stimuli mapped onto spatially segregated regions within the auditory cortex. Specifically, DEV sounds were represented more medially within the *planum temporale*, in the close proximity of the primary auditory cortex, while STD trials engaged a more posterior lateral region, encompassing the entire STG. Importantly, in recent a study we have reported a significant effect of stimulus repetition when processing DEV tones in primary auditory regions (Recasens et al., 2015), which was in the same direction as the results in the present study, namely, a greater response to DEV with longer stimulus history. Further, the spatial dissociation we found between STD and DEV trials, suggests that even for a slow dynamic response such as the BOLD signal, specific neuronal populations are devoted to encoding statistical regularities, which are distinct from those detecting a sensory change.

The effects we report seem to occur preattentively. Even though our participants' attention was diverted from the auditory stimulation, they may still have noticed the pattern of stimulation and the temporal occurrence of the frequency change along the paradigm. This scenario seems however unlikely given that tone change was subject to pseudo-randomization, making it hard to actively follow the sound patterns. Furthermore, the sequence of STD stimuli were relatively long, with durations being 12 and 18 seconds in trains of 24 and 36 repetitions, respectively. Unless our study participants were actively counting all the stimuli for the entire duration of the experiment, it is improbable that the observed modulation of the BOLD along STD trials would be the result of a conscious anticipation. This claims for future studies, which shall implement more stringent attentional control procedures and compare attended *vs.* unattended condition in paradigms similar to ours.

In the present study, we found no significant responses in subcortical stations of the auditory pathway. This result seems to contradict our previous data showing the involvement of the MGB and IC in auditory novelty processing using fMRI (Cacciaglia et al., 2015). However, the two studies employed different stimuli (bandpass-filtered noises vs. pure tones) delivered at different SOA (150 vs. 500 ms). These differences might have contributed to the differences in the results. Indeed, neurons of IC are more responsive to stimuli which are rich in their frequency spectrum and have different adaptation rates with respect to cortical neurons (Skoe and Kraus, 2010).

In considering the present work one should be aware of the following limitations.

In the current design, the TR (2000ms) corresponded to four times the SOA (500ms) and event onsets were not randomized across the experiment. This may have led to suboptimal parameter estimation in the GLM and likely to overlap of the hemodynamic response across consecutively modeled regressors. Nevertheless, the number of trials entered in each condition was relatively high (*i.e.*, up to 163 for STD_{24} and STD_{36}), which generally provides a more stable parameters estimation. Again, due to the modest temporal resolution of the fMRI, we were not able to capture auditory response variations that occur across shorter latencies than four tone repetition. This fosters

future studies to combine EEG with fMRI in study designs similar to the one implemented here. Finally, we shall underline a limitation common to fMRI studies. The neurophysiological bases of the fMRI-BOLD response have been clarified over the past decades and strong temporal associations have been documented with intracranially recorded local field potentials (LFP) (Goense and Logothetis, 2008). As the LFP reflects a neuromodulatory activity, larger BOLD amplitudes may reflect an actual increase of neuronal activity as well as other modulatory processes. Thus, variability of the BOLD response is only interpretable with respect to a baseline represented by the net average across the experiment.

Taken together, our data improve our understanding on the neural representations underlying auditory repetition and deviance detection effects. Future studies shall consider the role of additional parameters such as varying the temporal predictability or attentional resource manipulation.

ACKNOWLEDGMENTS

This work was supported by the Spanish Ministry of Economy and Knowledge (PSI2015-63664-P-MINECO/FEDER), the Catalan Government (SGR2017- 974), and the Institució Catalana de Recerca i Estudis Avançats (ICREA) Academia Distinguished Professorship awarded to Carles Escera.

FIGURE CAPTIONS

Fig. 1 Schematic illustration of the experimental paradigm. We adopted a frequency roving standard paradigm, where pure tones were arranged in trains of stimuli having the same frequency but different number of repetitions. This way, the first stimulus of a given train always acts as a deviant tone. In our design, one event of interest was represented by a trial of 4 consecutive pure tones spanning a duration of 2000 ms. The first stimulus of a DEV trials is highlighted in red color. The numerical coefficient associated to either a standard (STD) or deviant (DEV) trial indicates the

number of preceding stimuli (e.g., DEV12 means a deviant trial preceded by 12 repeated STD stimuli. STD32 indicates a standard trial preceded by 32 STD stimuli of the same frequency).

Fig. 2 Effect of stimulus repetition on DEV trials processing

A) The unbiased F-test revealed a significant main effect of number of preceding stimuli across DEV trials, as revealed by ROI analysis. Statistical parametric maps are projected over sagittal and axial slices. The STG is shown in orange, while the HG in blue. **B)** Line plot showing the contrast estimates for each DEV trial. Error bars indicate standard error of the mean (S.EM.)

Fig. 3 Brain regions showing significant response to DEV>STD controlling for the number of preceding stimuli

A-C) When matching the number of preceding stimuli across DEV and STD trials, we observed significant responses in the bilateral STG in all the three contrasts , with $DEV_{12}>STD_{12}$ and $DEV_{24}>STD_{24}$ additionally yielding activity in the bilateral HG. The STG is shown in orange, while the HG in blue. **D)** Superimposition of color-coded maps retrieved from **A), B)** and **C)**

Fig. 4 Comparison between deviance detection responses associated to different number of preceding stimuli.

A-C) Statistical parametric maps as revealed by comparing the differential contrasts $[DEV_{12}>STD_{12}]>[DEV_{4}>STD_{4}]$, $[DEV_{24}>STD_{24}]>[DEV_{4}>STD_{4}]$, and $[DEV_{24}>STD_{24}]>[DEV_{12}>STD_{12}]$, respectively. The STG is shown in orange, while the HG in blue. **D-F)** Parametric maps of the effects sizes computed voxel-wise on the whole brain, corresponding to A, B and C), respectively. **G)** Boxplots showing the change in magnitude of the hemodynamic response to distinct contrasts capturing the stimulus history. The lower and upper hinges correspond to the first and third quartile, while dots indicate individual subject data.

Fig. 5 Effect of stimulus repetition on STD trials processing

A) The F-contrast interrogating for any differences spanning across the STD trials revealed a significant main effect of stimulus repetition in the STG (orange) and HG (blue). **B)** Line plot showing the contrast estimate across STD trials in regions shown in A. **C)** Same as in A, assessed only in trains of 36 repetitions. **D)** Same as in B) assessed only in trains of 36 repetitions. Error bars indicate standard error of the mean (S.EM.)

REFERENCES

Anwander, A., Tittgemeyer, M., von Cramon, D.Y., Friederici, A.D., Knosche, T.R., 2007. Connectivity-Based Parcellation of Broca's Area. Cereb Cortex 17, 816-825.

Auksztulewicz, R., Friston, K., 2015. Attentional Enhancement of Auditory Mismatch Responses: a DCM/MEG Study. Cereb Cortex 25, 4273-4283.

Baldeweg, T., 2006. Repetition effects to sounds: evidence for predictive coding in the auditory system. Trends Cogn Sci 10, 93-94.

Baldeweg, T., Klugman, A., Gruzelier, J., Hirsch, S.R., 2004. Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophr Res 69, 203-217.

Bar, M., 2009. The proactive brain: memory for predictions. Philos Trans R Soc Lond B Biol Sci 364, 1235-1243.

Bendixen, A., 2014. Predictability effects in auditory scene analysis: a review. Front Neurosci 8, 60.

Bendixen, A., Schroger, E., Winkler, I., 2009. I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. J Neurosci 29, 8447-8451.

Cacciaglia, R., Escera, C., Slabu, L., Grimm, S., Sanjuan, A., Ventura-Campos, N., Avila, C., 2015. Involvement of the human midbrain and thalamus in auditory deviance detection. Neuropsychologia 68, 51-58.

Costa-Faidella, J., Baldeweg, T., Grimm, S., Escera, C., 2011a. Interactions between "what" and "when" in the auditory system: temporal predictability enhances repetition suppression. J Neurosci 31, 18590-18597.

Costa-Faidella, J., Grimm, S., Slabu, L., Diaz-Santaella, F., Escera, C., 2011b. Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology 48, 774-783.

Cowan, N., Winkler, I., Teder, W., Näätänen, R., 1993. Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). J Exp Psychol Learn Mem Cogn 19, 909-921.

Deouell, L.Y., 2007. The frontal generator of the mismatch negativity revisited. J Psychophysiol 21, 188–203.

Doeller, C.F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., Schröger, E., 2003. Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. Neuroimage 20, 1270-1282.

Dunovan, K.E., Tremel, J.J., Wheeler, M.E., 2014. Prior probability and feature predictability interactively bias perceptual decisions. Neuropsychologia 61, 210-221.

Escera, C., Leung, S., Grimm, S., 2014. Deviance detection based on regularity encoding along the auditory hierarchy: electrophysiological evidence in humans. Brain Topogr 27, 527-538.

Escera, C., Malmierca, M.S., 2014. The auditory novelty system: an attempt to integrate human and animal research. Psychophysiology 51, 111-123.

Fishman, Y.I., 2014. The mechanisms and meaning of the mismatch negativity. Brain Topogr 27, 500-526.

Friston, K., 2005. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360, 815-836.

Gao, P.P., Zhang, J.W., Cheng, J.S., Zhou, I.Y., Wu, E.X., 2014. The inferior colliculus is involved in deviant sound detection as revealed by BOLD fMRI. Neuroimage 91, 220-227.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol 120, 453-463.

Goense, J.B., Logothetis, N.K., 2008. Neurophysiology of the BOLD fMRI signal in awake monkeys. Curr Biol 18, 631-640.

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10, 14-23.

Grimm, S., Escera, C., 2012. Auditory deviance detection revisited: evidence for a hierarchical novelty system. Int J Psychophysiol 85, 88-92.

Grotheer, M., Kovacs, G., 2016. Can predictive coding explain repetition suppression? Cortex 80, 113-124.

Haenschel, C., Vernon, D.J., Dwivedi, P., Gruzelier, J.H., Baldeweg, T., 2005. Event-related brain potential correlates of human auditory sensory memory-trace formation. J Neurosci 25, 10494-10501.

Hindy, N.C., Ng, F.Y., Turk-Browne, N.B., 2016. Linking pattern completion in the hippocampus to predictive coding in visual cortex. Nat Neurosci 19, 665-667.

Hughes, G., Waszak, F., 2014. Predicting faces and houses: category-specific visual action-effect prediction modulates late stages of sensory processing. Neuropsychologia 61, 11-18.

Imada, T., Hari, R., Loveless, N., McEvoy, L., Sams, M., 1993. Determinants of the auditory mismatch response. Electroencephalogr Clin Neurophysiol 87, 144-153.

Javitt, D.C., Grochowski, S., Shelley, A.M., Ritter, W., 1998. Impaired mismatch negativity (MMN) generation in schizophrenia as a function of stimulus deviance, probability, and interstimulus/interdeviant interval. Electroencephalogr Clin Neurophysiol 108, 143-153.

Kanai, R., Komura, Y., Shipp, S., Friston, K., 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. Philos Trans R Soc Lond B Biol Sci 370.

Kiehl, K.A., Stevens, M.C., Laurens, K.R., Pearlson, G., Calhoun, V.D., Liddle, P.F., 2005. An adaptive reflexive processing model of neurocognitive function: supporting evidence from a large scale (n = 100) fMRI study of an auditory oddball task. Neuroimage 25, 899-915.

Kok, P., Failing, M.F., de Lange, F.P., 2014. Prior expectations evoke stimulus templates in the primary visual cortex. J Cogn Neurosci 26, 1546-1554.

Kok, P., Jehee, J.F., de Lange, F.P., 2012. Less is more: expectation sharpens representations in the primary visual cortex. Neuron 75, 265-270.

Kok, P., Mostert, P., de Lange, F.P., 2017. Prior expectations induce prestimulus sensory templates. Proc Natl Acad Sci U S A 114, 10473-10478.

Lecaignard, F., Bertrand, O., Gimenez, G., Mattout, J., Caclin, A., 2015. Implicit learning of predictable sound sequences modulates human brain responses at different levels of the auditory hierarchy. Front Hum Neurosci 9, 505.

May, P.J., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47, 66-122.

Maess, B., Jacobsen, T., Schröger, E., Friederici, A.D., 2007. Localizing pre-attentive auditory memory-based comparison: magnetic mismatch negativity to pitch change. Neuroimage 37, 561-571.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233-1239.

May, P.J., Tiitinen, H., 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. Psychophysiology 47, 66-122.

Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214, 655-667.

Moran, R.J., Symmonds, M., Dolan, R.J., Friston, K.J., 2014. The brain ages optimally to model its environment: evidence from sensory learning over the adult lifespan. PLoS Comput Biol 10, e1003422.

Mühlau, M., Rauschecker, J.P., Oestreicher, E., Gaser, C., Rottinger, M., Wohlschlager, A.M., Simon, F., Etgen, T., Conrad, B., Sander, D., 2006. Structural brain changes in tinnitus. Cereb Cortex 16, 1283-1288.

Näätänen, R., Gaillard, A.W., Mantysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol (Amst) 42, 313-329.

Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin Neurophysiol 118, 2544-2590.

Novitski, N., Tervaniemi, M., Huotilainen, M., Naatanen, R., 2004. Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. Brain Res Cogn Brain Res 20, 26-36.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D.Y., Schröger, E., 2002. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage 15, 167-174.

Opitz, B., Schröger, E., von Cramon, D.Y., 2005. Sensory and cognitive mechanisms for preattentive change detection in auditory cortex. Eur J Neurosci 21, 531-535.

Petrides, M., Tomaiuolo, F., Yeterian, E.H., Pandya, D.N., 2012. The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. Cortex 48, 46-57.

Recasens, M., Grimm, S., Capilla, A., Nowak, R., Escera, C., 2014. Two sequential processes of change detection in hierarchically ordered areas of the human auditory cortex. Cereb Cortex 24, 143-153.

Recasens, M., Leung, S., Grimm, S., Nowak, R., Escera, C., 2015. Repetition suppression and repetition enhancement underlie auditory memory-trace formation in the human brain: an MEG study. Neuroimage 108, 75-86.

Rinne, T., Alho, K., Ilmoniemi, R.J., Virtanen, J., Näätänen, R., 2000. Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage 12, 14-19.

Rinne, T., Kirjavainen, S., Salonen, O., Degerman, A., Kang, X., Woods, D.L., Alho, K., 2007. Distributed cortical networks for focused auditory attention and distraction. Neurosci Lett 416, 247-251.

Sabri, M., Campbell, K.B., 2001. Effects of sequential and temporal probability of deviant occurrence on mismatch negativity. Brain Res Cogn Brain Res 12, 171-180.

Sabri, M., Kareken, D.A., Dzemidzic, M., Lowe, M.J., Melara, R.D., 2004. Neural correlates of auditory sensory memory and automatic change detection. Neuroimage 21, 69-74.

Sabri, M., Liebenthal, E., Waldron, E.J., Medler, D.A., Binder, J.R., 2006. Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. J Cogn Neurosci 18, 689-700.

Sams, M., Alho, K., Näätänen, R., 1983. Sequential effects on the ERP in discriminating two stimuli. Biol Psychol 17, 41-58.

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., Schröger, E., 2013. Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. J Neurosci 33, 8633-8639.

Schall, U., Johnston, P., Todd, J., Ward, P.B., Michie, P.T., 2003. Functional neuroanatomy of auditory mismatch processing: an event-related fMRI study of duration-deviant oddballs. Neuroimage 20, 729-736.

Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. Ear Hear 31, 302-324.

Slabu, L.M., 2010. The effect of slice orientation on auditory FMRI at the level of the brainstem. Brain Topogr 23, 301-310.

Spriggs M.J., Sumner R.L., McMillan R.L., Moran R.J., Kirk I.J., Muthukumaraswamy S.D., 2018. Indexing sensory plasticity: Evidence for distinct Predictive Coding and Hebbian learning mechanisms in the cerebral cortex. Neuroimage 176, 290-300

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat Neurosci 11, 1004-1006.

Szycik, G.R., Stadler, J., Brechmann, A., Munte, T.F., 2013. Preattentive mechanisms of change detection in early auditory cortex: a 7 Tesla fMRI study. Neuroscience 253, 100-109.

Todorovic, A., van Ede, F., Maris, E., de Lange, F.P., 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J Neurosci 31, 9118-9123.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273-289.

Ulanovsky, N., Las, L., Nelken, I., 2003. Processing of low-probability sounds by cortical neurons. Nat Neurosci 6, 391-398.

von Kriegstein, K., Patterson, R.D., Griffiths, T.D., 2008. Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. Curr Biol 18, 1855-1859.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., Dehaene, S., 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A 108, 20754-20759.

Woo, C.W., Krishnan, A., Wager, T.D., 2014. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. Neuroimage 91, 412-419.

Yago, E., Corral, M.J., Escera, C., 2001. Activation of brain mechanisms of attention switching as a function of auditory frequency change. Neuroreport 12, 4093-4097.

Yucel, G., Petty, C., McCarthy, G., Belger, A., 2005. Graded visual attention modulates brain responses evoked by task-irrelevant auditory pitch changes. J Cogn Neurosci 17, 1819-1828.

Figure 1

Fig. 1

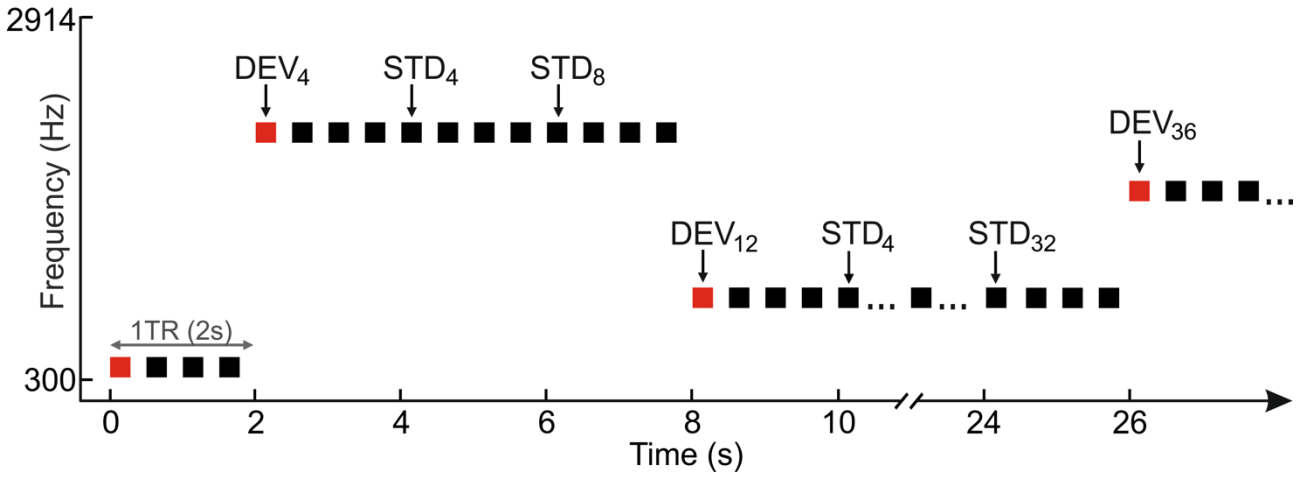


Figure 2

Fig. 2

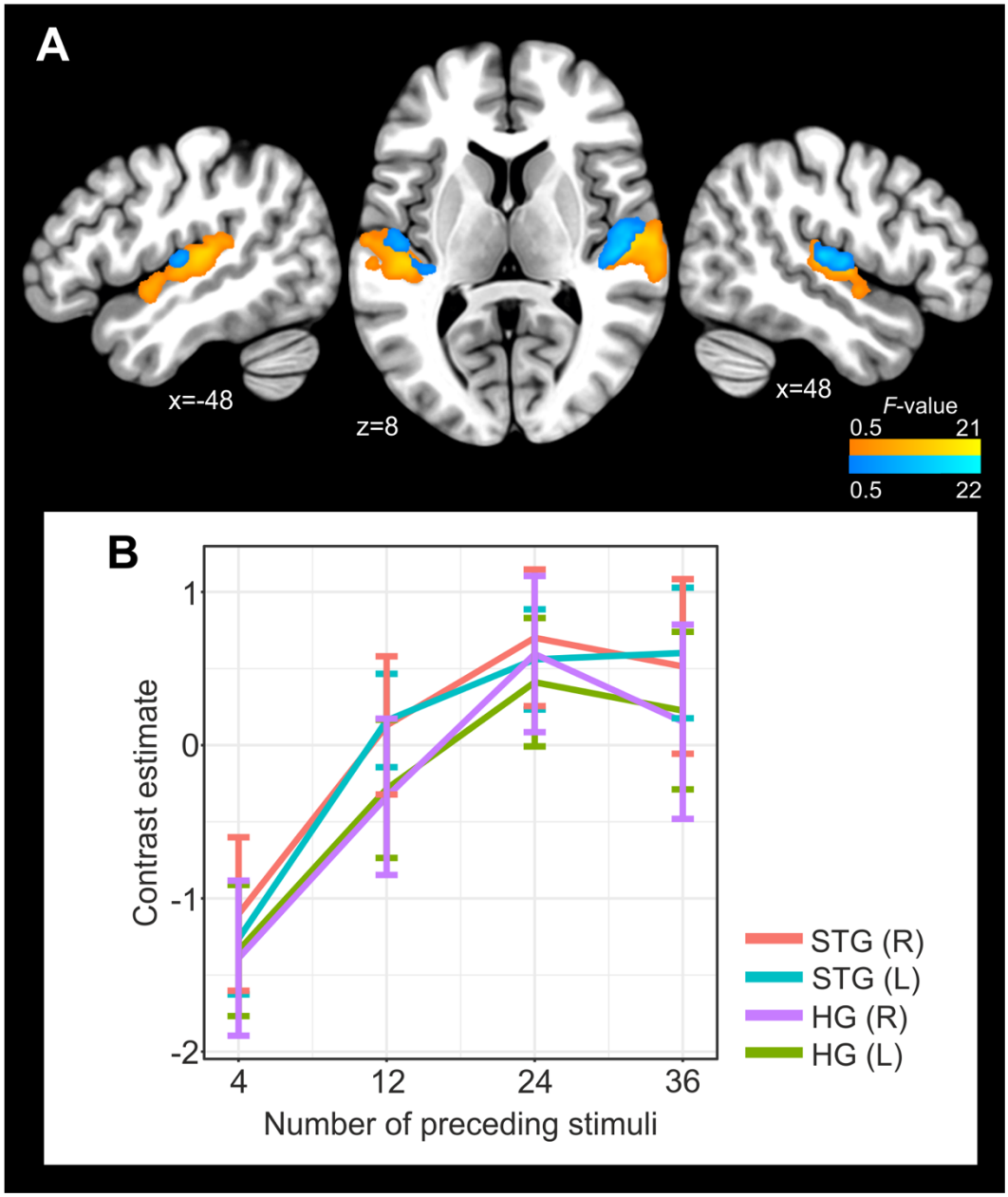


Figure 3

Fig. 3

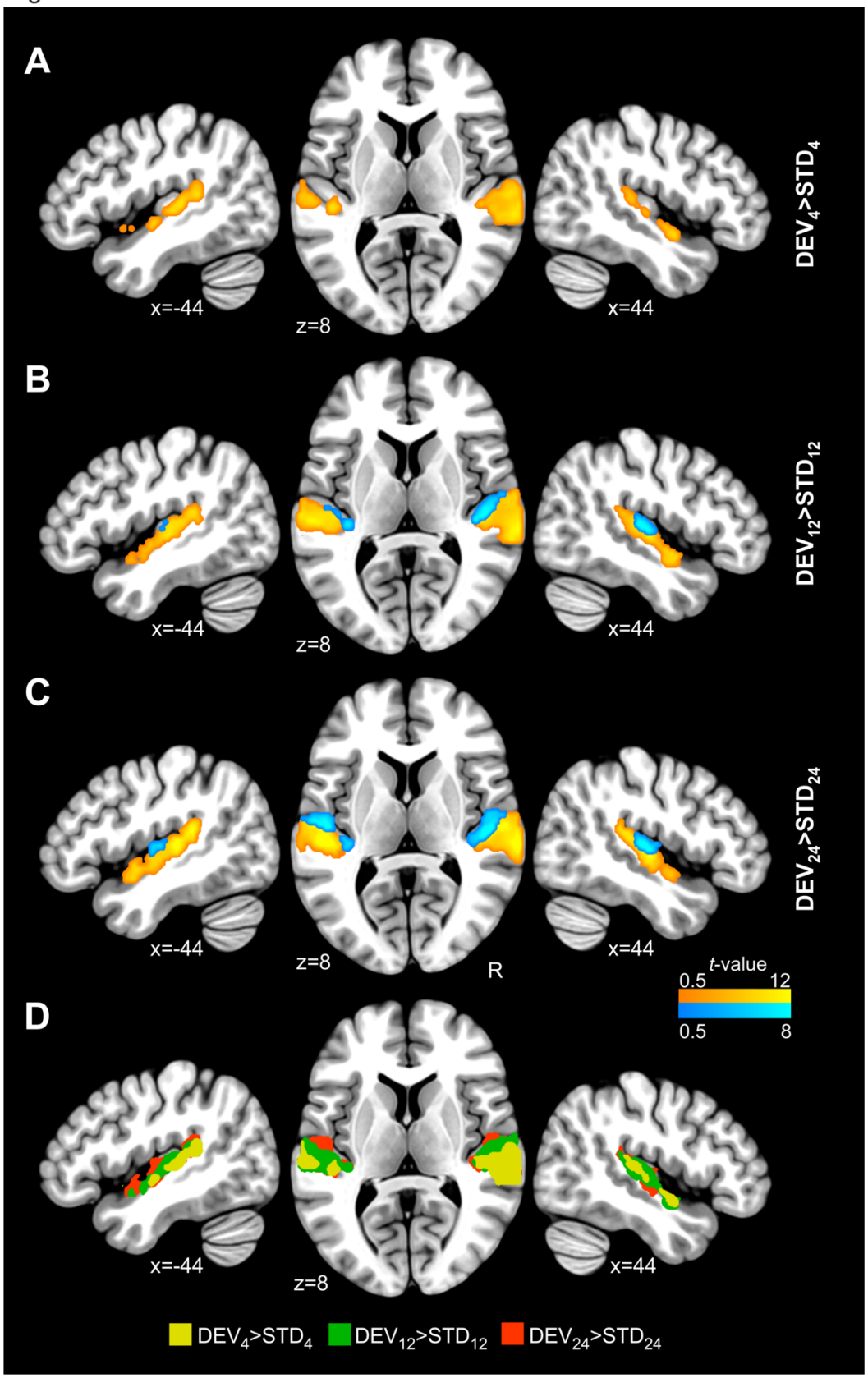


Figure 4

Fig. 4

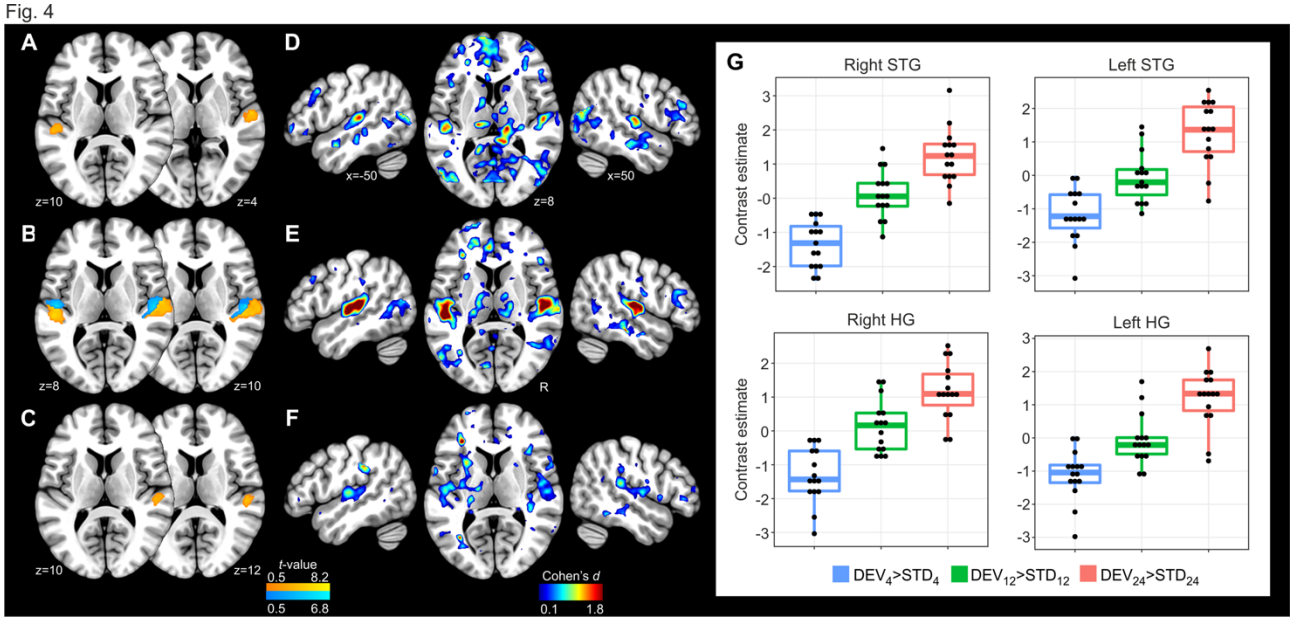
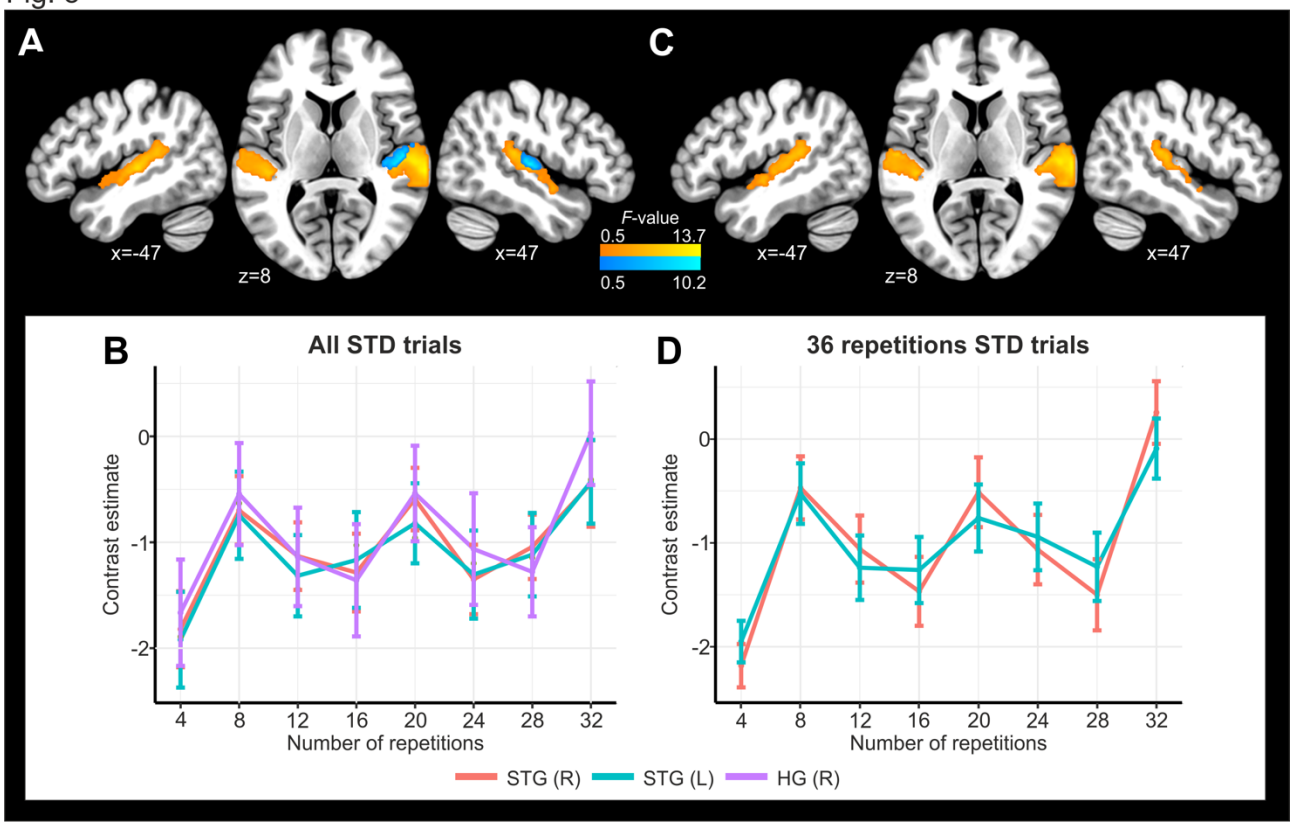
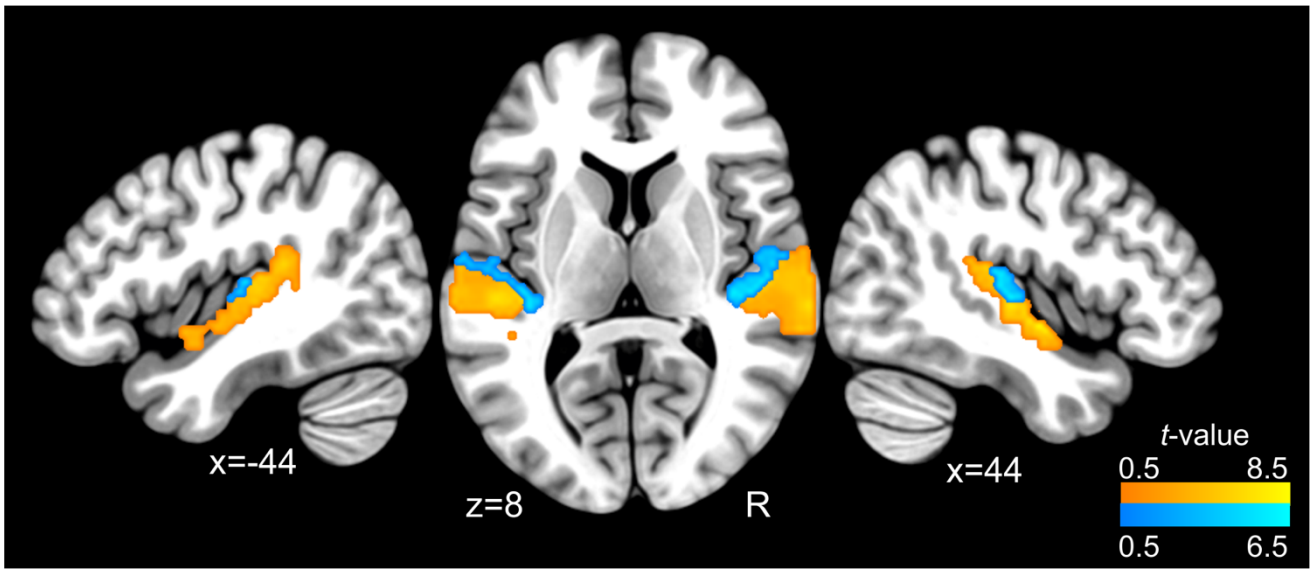


Figure 5

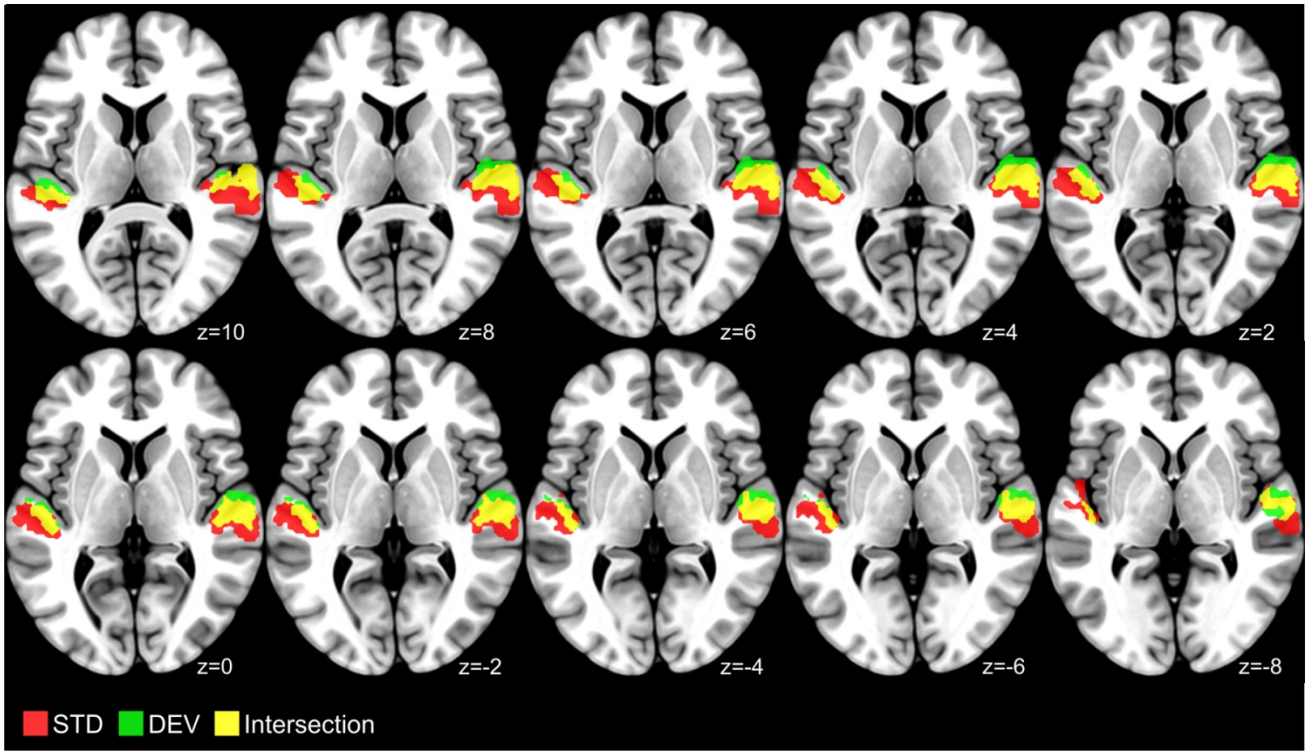
Fig. 5



Supplementary Materials



Supplementary Fig. 1 - Brain regions showing significant response to DEV_{ALL}>STD_{ALL} trials When contrasting all DEV against all STD trials, the ROI analysis revealed significant activations in the bilateral STG, showed in orange (right: $t_{14} = 8.68$, $P_{FWE} < 0.001$, k = 1451, [x =66, y = -28, z = 16]; left: $t_{14} = 7.99$, $P_{FWE} < 0.001$, k = 990, [x = -50, y = -28, z = 10]), and the bilateral HG, showed in blue (right: $t_{14} = 6.06$, $P_{FWE} < 0.001$, k = 146, [x =50, y = -10, z = 4] ; left: $t_{14} = 5.97$, $P_{FWE} = 0.001$, k = 41, [x =-38, y = -28, z = 10]). For visualization purposes t-maps are thresholded at p<0.005 uncorrected for multiple comparisons.



Supplementary Fig. 2 – Stimulus repetition modulated the response to STD and DEV trials in partially segregated clusters of auditory cortex.

Superimposition of significant clusters capturing the responses to any difference among STD and DEV trials separately along stimulus repetition, averaged across subjects. A ROI analysis was performed with a combined mask of the bilateral STG and HG.