



Published in final edited form as:

Nat Neurosci. 2019 March ; 22(3): 447–459. doi:10.1038/s41593-018-0317-8.

State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex

Diego Elgueda^{1,2}, Daniel Duque^{1,3}, Susanne Radtke-Schuller¹, Pingbo Yin¹, Stephen V. David⁴, Shihab A. Shamma^{1,2,5}, and Jonathan B. Fritz^{*,1,2}

¹Institute for Systems Research, Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA

²Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, USA

³Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.

⁴Oregon Hearing Research Center, Oregon Health and Science University, Portland, OR 97239, USA

⁵Laboratoire des Systèmes Perceptifs, École Normale Supérieure, 75005 Paris, France

Abstract

In higher sensory cortices, there is a gradual transformation from sensation to perception and action. In the auditory system, this transformation is revealed by responses in the rostral Ventral Posterior field (VPr), a tertiary area in ferret auditory cortex, which shows long-term learning in trained compared to naïve animals, arising from selectively enhanced responses to behaviorally relevant target stimuli. This enhanced representation is further amplified during active performance of spectral or temporal auditory discrimination tasks. VPr also shows sustained short-term memory activity after target stimulus offset, correlated with task-response timing and action. These task-related changes in auditory filter properties enable VPr neurons to quickly and nimbly switch between different responses to the same acoustic stimuli, reflecting either spectrotemporal properties, timing or behavioral meaning of the sound. Furthermore, they demonstrate an interaction between dynamics of long-term learning and short-term attention as incoming sound is selectively attended, recognized and translated into action.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence: ripple@isr.umd.edu.

Author contributions

JBF, DE and SAS designed experiments; DE, PY, DD, JBF performed experiments; DE, SVD, PY analyzed physiological results; SR-S analyzed neuroanatomical results; DE prepared physiological figures; SR-S and DE prepared the neuroanatomical figures; JBF, DE, SVD, SAS wrote the manuscript.

Code availability

Custom scripts written in MATLAB and Python used in this study are available from the corresponding author upon reasonable request.

Data availability

The data supporting the findings in this study are available from the corresponding author upon reasonable request.

Competing Financial Interests

The authors declare no competing interests.

Introduction:

In order to understand the meaning of sounds, we learn to associate their acoustic features with their behavioral context and link them to appropriate audio-motor responses. Once associative learning has taken place, rapid task-dependent plasticity during active listening may enhance listeners' ability to recognize and respond to relevant incoming sounds by adaptively reshaping auditory cortical filter properties.

Research in visual and somatosensory associative cortices has shown their key role in complex object recognition and perception¹⁻³, formation of learned categorical representations⁴⁻⁶, multisensory integration, memory⁷ and decision-making^{8,9}. However, with a few notable exceptions¹⁰⁻¹³, most neurophysiological studies of auditory cortex in behaving animals have focused on primary auditory cortex (A1) rather than higher order auditory cortical areas.

To investigate the contributions of non-primary auditory cortex to sound processing, we have chosen the ferret, which has become an increasingly valuable animal model to study the neurobiology of auditory behavior and hearing¹⁴. In previous studies, we have described how task engagement induces rapid plasticity in primary auditory cortex (A1) and in tonotopically-organized secondary or "belt" areas in the ferret auditory cortex (PPF and PSF in Figure 1b). The neural representation of sound can be partially transformed in these areas to incorporate behavioral and contextual information^{11,15-17}. We have also characterized a task-dependent, gated representation of behaviorally salient sounds in non-tonotopic dorsolateral frontal cortex (dlFC)¹⁸.

Based on this earlier work, we conjectured (i) there are tertiary auditory cortical areas between secondary areas and frontal cortex where the transformation from sound representation to behavioral meaning is more extensively developed than in lower cortical areas, (ii) long-term task learning permanently shapes neuronal responses in these higher areas, a change that should be evident even during task-free (or "passive") conditions. We also predicted that neurons in higher auditory areas would (iii) display strong attention effects that would amplify long-term changes in the representation of task-relevant stimuli during task performance, (iv) would show response timing linking auditory inputs to reward and motor responses.

Previous studies have shown that ferret auditory cortex is composed of multiple acoustically sensitive adjoining areas in the ectosylvian gyrus of the temporal lobe^{19,20}. Current maps of ferret auditory cortex include nine distinct cortical areas, six of which (A1, AAF, PPF, PSF, ADF, AVF) have been physiologically identified and described previously¹⁹. One field whose function has not been studied previously, VP, lies in a ventral region in the PEG, and its anatomical connectivity makes it a good candidate for a tertiary auditory field²⁰⁻²³.

To test the above hypotheses concerning sound encoding in tertiary auditory cortex, we recorded responses under multiple active task and non-task (passively listening) conditions in the rostral region of VP (VPr in Figure 1a,b)^{11,20,21,23}. Partly because of its extreme lateral location and limited accessibility for surface recordings, VP has remained one of the least studied areas of the ferret auditory cortex. Here, we describe how VPr neurons exhibit

striking state-dependent and context-dependent changes in auditory responses and encode non-acoustical sound features, such as associated behavioral meaning and task timing. These results are consistent with all four of the conjectures above.

Results:

Neurophysiological mapping and neuroanatomical location

Basic tuning properties of VPr were mapped using single-unit activity in 6 animals during passive presentation of pure tone, click train and broadband rippled noise stimuli (see **Methods**).

We marked location of VPr recordings and confirmed that they were ventral and anterior to area PPF (Figure 1b,c, sites labeled with electrolytic lesions, electrolytic deposits of iron, or injections of neuroanatomical tracer). All microelectrode penetrations into VPr followed a 30 degree angle relative to the sagittal plane. The neuronal depths of >90% of our VPr recordings were close to the cortical surface, within the first 500 microns of the first spikes recorded as electrodes entered the brain (Supplementary Figure S-1). The locations of each recording site was also registered with a ferret brain atlas²¹ (Figure 1c). VPr spans an area 1–2 mm below and ventral to high frequency PPF (Figure 1d,e) and ventral to the lower lip of the PSS (pseudosylvian sulcus) and is physiologically characterized by a drastic change in the tonotopic map (Figure 1e) and an increase in response latency (Figure 1f).

Response properties in VPr

Basic auditory tuning properties of single units in VPr are contrasted with previously collected responses from A11 and dPEG (Figure 2). The distribution of tuning properties is consistent with neuroanatomical evidence that VPr is a later processing stage in the auditory pathway, following A1 and dPEG²⁰. Compared to earlier areas, VPr neurons display longer mean latency (Figure 2a; VPr: n=583 neurons, 37.71 ± 1.78 ms, dPEG: n=1125 neurons, 24.45 ± 0.57 ms, A1: n=2309 neurons, 15.57 ± 0.88 ms). We found significant differences in tone response latency (Chi-square=862.83, $p < 0.0001$, $df=2$, Kruskal-Wallis test), where VPr significantly differed from A1 (Tukey's HSD effect size [$\text{mean}_{(A1)} - \text{mean}_{(VPr)}$] = -1265.8 , $CI_{95\%} = [-1388.9, -1142.7]$, $p=0$) and dPEG (effect size [$\text{mean}_{(dPEG)} - \text{mean}_{(VPr)}$] = -341.8 , $CI_{95\%} = [-477.3, -206.3]$, $p=0$).

Neurons in VPr also have broader mean frequency tuning bandwidth (Figure 2b; VPr: n=2594 neurons, 1.77 ± 0.05 octaves; dPEG: n=1202 neurons, 1.4 ± 0.05 octaves; A1: n=2594 neurons, 1.07 ± 0.04 octaves; Chi-square=499.16, $p < 0.001$, $df=2$, Kruskal-Wallis test). Mean VPr bandwidth was significantly greater than A1 (effect size= -1190.9 , $CI_{95\%} = [-1323.6, -1058.2]$, $p=0$) and dPEG (effect size= -630.7 , $CI_{95\%} = [-777.8, -483.7]$, $p=0$).

VPr neurons also display weaker overall following of complex synthetic sounds (Figure 2c; mean SNR, VPr: n=516 neurons, 0.34 ± 0.05 ; dPEG: n=986 neurons, 0.55 ± 0.1 ; A1: n=2399 neurons, 0.73 ± 0.08 ; Chi-square = 291.06, $p < 0.001$, $df=2$, Kruskal-Wallis test). The SNR of VPr responses was significantly lower than A1 (effect size= 816.24 , $CI_{95\%} = [688.15, 944.33]$, $p=0$) and dPEG neurons (effect size= 329.55 , $CI_{95\%} = [186.13, 4472.97]$, $p=0$).

For neurons whose responses did follow the stimulus, spectrotemporal receptive fields (STRFs) were more complex, as indicated by their sparseness index, the peak STRF magnitude, divided by the standard deviation across STRF bins (Figure 2d; VPr: n=180 neurons, 0.44 ± 0.09 ; dPEG: n=472 neurons, 1.35 ± 0.09 ; A1: n=1664 neurons, 2.24 ± 0.13 ; Chi-square=291.06, $p < 0.001$, $df=2$). VPr showed lower STRF sparseness than A1 (effect size=777.47, $CI_{95\%} = [655.48, 899.45]$, $p=0$) and dPEG (effect size=466.29, $CI_{95\%} = [330.18, 602.4]$, $p=0$).

Thus, VPr occupies an intermediate stage in auditory processing, resembling the earlier stages by its tuned responses to tones and occasional phase-locking to modulated stimuli, which allow for STRF measurements in some neurons (in VPr, only 27.3% (180/658) of cells have $SNR > 0.2$, compared to 35.3% in dPEG (472/1337) and 60% in A1 (1644/2740)). However, VPr is also similar to dIFC in its relatively weak auditory responsiveness during passive sound presentation¹⁸, its often poorly-defined tuning, and long response latencies (examples in Figures 3a,c and Supplementary Figures S-2, S-3).

Response Modulation During Task Performance

Responses in VPr changed dramatically during task performance to reflect the behavioral valence of the stimuli as positively (Go) or negative rewarded (NoGo) sounds. A total of 367 single-units were recorded in 4 trained ferrets, before (pre-passive), during and after (post-passive) performance of two distinct conditioned avoidance tasks, learned prior to recordings¹⁵. The tasks were: (1) tone versus noise discrimination task (“Pure Tone-Detection” or PT-D) and (2) Click-Rate Discrimination (or CLR-D) task (Figure 1a). In both tasks, the animals listened to a sequence of reference “Safe” sounds (broadband rippled noise TORCs in PT-D or a range of click-train rates in CLR-D) during which the animal could safely lick a waterspout for reward. The sequence of Safe sounds ended either with a final Safe sound (catch trials) or with a “Warning” Target sound (Tone in PT-D and a different click rate in CLR-D) that alerted the animals to stop licking 400 ms after Target offset to avoid a mild shock. For different CLR-D animals, Warning click rates were either lower or higher than that of the Safe rate. During each recording session, animals often engaged in blocks of two or more tasks with different stimuli.

Examples of single-unit responses in VPr during behavior are shown in Figures 3a, 3d. In the majority of units, engagement in behavior rapidly induced a substantial change in PSTH responses to Warning stimuli, and a lesser change for Safe stimuli (Figures 3a, d). In the extreme, some units were behaviorally gated and showed virtually no response to task-related sounds unless the animal was engaged in behavior (Supplementary Figure S-3). Details of changes varied greatly from cell to cell, reflecting the specific type of response (e.g., onset, sustained, or offset). Nevertheless, the patterns of responses to Warning and Safe stimuli in the population average (Figures 3b,e) remained largely similar for both tasks despite the different stimuli (TORCs/Tones versus TORC-Click-Trains – average responses to 30 different TORCs and multiple click train rates). Population averages (PSTHs) to Safe (Rate 1) and Warning (Rate 2) click rates were averaged across different click rate trains for animals trained with either low or high click rates as Warning stimuli (see Supplementary Figure S-4). Thus, on average, there was a large enhancement in the responses to the class of

NoGo Warning stimuli (e.g., Tone or Target click-train) during behavior, compared to smaller changes in the class of Safe stimulus responses.

Task-dependent response changes were measured by the difference in normalized firing rates ($\Delta nFR_{(B-P)}$) between behaving and passive conditions (Figures 3c,f). This differential change increased the contrast between Safe versus Warning responses, much greater in magnitude but in a similar direction to changes reported earlier in secondary auditory areas¹¹. Behavioral state could alter neuronal responses to a given stimulus from onset to sustained (Figure 3a – lower panel) or even gate VPr neuronal responses so that they only occurred in the active state (Supplementary Figure S-3).

The relationship of VPr responses to behavior is illustrated (Figures 3b, 3e), juxtaposing the population lick probability for Safe and Warning (hits only) sounds to the population neural response for the two tasks. Lick probability for the Safe sounds remains constant during and after these stimuli. However, lick probability for the Warning sound is clearly depressed, not only during the stimulus, but also post-stimulus, until the end of the shock window (shaded area). A comparison with the population neural PSTH for the two tasks, illustrates that the timing of the increased neural responses to the Warning stimuli parallels decreases in the behavioral lick response.

Response transformations from A1 to dPEG, VPr and dlFC

To gain a broader view of VPr in the broader cortical network, we compared population PSTH responses in A1, dPEG, and dlFC during pre-passive and behavior epochs for both PT-D and CLR-D tasks (Figure 4). PT-D data from A1, dPEG, and dlFC of 14 additional ferrets^{11,15,17,18} were re-analyzed and added to PT-D averages to provide a larger sample (see **Methods**). We measured stimulus contrast as the difference between Warning and Safe responses ($\Delta nFR_{(W-S)}$, baseline subtracted, normalized amplitude PSTH, PT-D: 0.1–0.45 s after sound onset; CLR-D: 0.3–1 s after TORC-offset/Click-Train onset) for both passive and behaving conditions (PSTHs in Figures 4a,b, contrasting distributions in Figure 6).

There are several notable findings in the comparison of population average responses across areas and tasks. *First*, the overall pattern of enhanced contrast ($\Delta nFR_{(W-S)}$) between Warning and Safe responses during behavior is similar in both tasks. Overall changes in contrast from one cortical area to another for the two tasks are remarkably similar at a population level, despite considerable differences between their stimuli. This points to the primacy of behavioral meaning of the stimuli in the tasks (as Go or NoGo) rather than their acoustic properties in determining the nature of VPr responses. At a single cell level, many VPr neurons (150/367≈41%) showed similar enhanced target responses in *both* PT-D and CLR-D tasks (Supplementary Figure S-5).

Second, contrast enhancement between Warning vs Safe ($\Delta nFR_{(W-S)}$) gradually increases across areas during behavior compared to the pre-passive state (left versus right columns in Figure 4a,b). We interpret this to indicate a progressively larger weight given to the behavioral distinction between NoGo vs. Go stimuli in higher cortical areas. The overall change in contrast in A1 is much smaller than in dPEG (e.g., CLR-D task of Figure 4b). In fact, the average Warning tone response in A1 during PT-D (Figure 4a) is actually smaller

than the responses to the Safe TORCs. This reversal likely reflects the sensitivity of A1 neurons to tone frequency. In many experiments, recordings were made simultaneously from neurons with different frequency tuning. Hence the Target tone frequency could not be optimized to deliver enhancements as described earlier in studies where the Warning tone was often placed close to the best frequency (BF) to achieve maximal plasticity^{11,15}.

We compared responses in different areas using a 3-way repeated-measures ANOVA (rmANOVA, see **Methods**). In the PT-D task (A1 n=71, dPEG n=199, VPr n=251, dIFC n=138 neurons) the rm-ANOVA for response difference ($nFR_{(w-s)}$) yielded significant main effects for area ($F=8.91$; $p<0.0001$) and task condition (passive or behaving, $F=11.52$; $p=0.0007$). HSD Tukey's test confirmed that response differences are smaller in A1 compared to the other areas (A1-dPEG=2.759, $CI_{95\%}=[1.201, 4.316]$, $p<0.0001$; A1-VPr=2.944, $CI_{95\%}=[1.418, 4.470]$, $p<0.0001$; A1-dIFC=2.202, $CI_{95\%}=[0.580, 3.823]$, $p=0.0028$). A t-test confirmed that response differences are larger when the animal is engaged in the task (Passive-Behavior=0.571, $CI_{95\%}=[0.247, 0.894]$, $t=-3.468$, $p=0.0006$). The analysis also yielded a significant area vs behavior interaction ($F=2.64$; $p<0.0487$), suggesting that the effect of engagement on the response depends on the area. HSD Tukey's post-hoc analysis again confirmed that behavior enhances response contrast ($nFR_{(w-s)}$) in VPr and dIFC (VPr Passive-Behavior=0.903, $CI_{95\%}=[0.135, 1.670]$, $p=0.0089$; dIFC Passive-Behavior=1.120, $CI_{95\%}=[0.171, 2.070]$, $p=0.0086$), but not in A1 or dPEG (A1 Passive-Behavior=0.003, $CI_{95\%}=[-1.356, 1.363]$, $p=1$; dPEG Passive-Behavior=0.213, $CI_{95\%}=[-0.616, 1.044]$, $p=0.994$). Altogether, this analysis suggests that VPr neurons show contrast enhancement that more closely resembles dIFC than dPEG.

In the CLR-D task (A1 n=57, dPEG n=60, VPr n=266, dIFC n=38 neurons), the rmANOVA yielded a significant main effect for task condition ($F=29.47$; $p<0.0001$). A t-test confirmed that response differences are larger during the active behavior condition (Passive-Behavior=1.010, $CI_{95\%}=[0.701, 1.497]$, $t=5.429$, $p<0.0001$). The analysis also yielded a significant area vs task condition interaction ($F=2.74$; $p<0.0429$). Post-hoc HSD Tukey analysis confirmed that behavior enhances the response contrast between Warning and Safe click trains in VPr and dIFC (VPr Passive-Behavior=0.798, $CI_{95\%}=[0.213, 1.382]$, $p=0.001$; dIFC Passive-Behavior=2.146, $CI_{95\%}=[0.551, 3.740]$, $p=0.0013$), but not in A1 or dPEG (A1 Passive-Behavior=0.312, $CI_{95\%}=[-0.979, 1.603]$, $p=0.999$; dPEG Passive-Behavior=1.142, $CI_{95\%}=[-0.097, 2.381]$, $p=0.096$). These findings suggest that a better representation of click trains in A1 and dPEG – when the animal is engaged in a behavioral task – may be used to generate more highly differentiated behavioral percepts in higher order areas of the auditory and frontal cortices.

The relation of VPr responses to motor action (licking) was analyzed by cross-correlating spikes with licks¹⁸ (see **Methods**). Based on this analysis, we found that 37% of VPr neurons ($N=93/251$ neurons tested in the PT-D task) had a significant motor component in their activity. However, behavior induced changes in sound-evoked activity were independent of these motor effects. When we subtracted all lick-predicted spike activity from the 37% of VPr neurons with significant motor-related activity, population mean PSTHs did not change significantly (one-way ANOVA, $F=0.43$, $p=0.5122$, Supplementary Figure S-6). This analysis also highlights an observation about the prevalence of motor-

related activity in VPr. The prevalence of neurons with motor-related activity is 20% in A1 (14/71), 13% in dPEG (26/199), 37% in VPr (93/251) and 20% in dlFC (161/788). Thus, motor-related activity is more common in VPr than in the other auditory cortical areas.

Progressive contrast enhancement exists even in the quiescent state

Previous work on auditory learning in adult animals has shown that auditory cortex undergoes long-term changes that reflect training on behaviorally relevant sound features^{24,25}. Conversely, artificially enhancing neural responses to acoustic stimuli can improve behavioral responses to those stimuli^{26–28}. Thus, we predicted that we would observe enhanced contrast between Warning and Safe responses, not only during behavior, but also during passive listening. We measured Warning versus Safe contrast during the pre-passive epoch in both tasks, and observed that contrast indeed increased from A1 to VPr (Figures 4, 5). We note that because of behavioral gating¹⁸, the dlFC is somewhat different from earlier auditory cortical areas, in that it rarely responds to task stimuli during passive listening in the PT-D or CLR-D tasks. We hypothesize that the significant change in contrast effects (from A1 to VPr) during the pre-passive state may reflect persistent effects of behavioral training. A consequence of this explanation would be that task-naïve animals should *not* exhibit any such effects, as we shall demonstrate and discuss below.

To summarize, there is a gradual shift towards an enhanced representation of behavioral meaning of task stimuli beginning in the early cortical stages (A1 and dPEG) and increasing towards the higher cortical regions where it becomes clearly manifested in dlFC. VPr is similar to the early auditory cortical areas, responding to both contrasted Warning-Safe sounds reflecting their acoustic features such as tone frequency and temporal dynamics. However, VPr responses, on the other hand, also resemble those in dlFC in their state-dependent response changes and selective representation of Warning stimuli during behavior.

Behavioral gating in VPr – comparison with responses in dlFC

There is a subset of neurons in VPr that exhibit behaviorally-gated responses. They are non-responsive to acoustic stimuli during passive listening but show clear responses to the same sounds during behavior (Supplementary Figure S-3). As mentioned, these behaviorally-gated responses in VPr are similar to responses previously observed in dlFC¹⁸. About 28% of VPr neurons (127/453) showed no response to a variety of passively presented acoustic stimuli (i.e. no behavioral task context). However, in active task conditions, only 12% (54/453) were unresponsive. Thus, 16% (73/453) of VPr neurons were behaviorally gated. However, unlike dlFC, a majority (72%) of VPr neurons still do display some broad pre-passive responses. Passive responses are largely absent in the dlFC for either of the two tasks, especially for the CLR-D task (Figure 4). In dlFC, the small pre-passive responses that are observed for PT-D may be largely due to persistent enhancement from previous tasks performed in the same recording session¹⁸.

Contrast between Warning and Safe stimuli is qualitatively different in naïve and trained animals

Since the behavioral meaning of the Warning and Safe stimuli emerges as a result of behavioral training on the Go-NoGo tasks, we conjectured that these two classes of sound might leave a trace in higher cortical sensory regions reflecting their meaning, even when the animal was *not* engaged in performing the task. The strong response contrast between Warning and Safe stimuli ($nFR_{(W,S)}$) during passive listening suggests that this is the case (Figure 4). However, if behavioral training causes these long-term changes, the difference between Warning and Safe responses should be less pronounced and should not increase in higher auditory areas of task-naïve animals. To test this prediction, we recorded responses to task stimuli in A1, dPEG and VPr of a task-naïve animal (Figure 5a, b). For the PT-D stimuli, A1 responses in the naïve and the trained animals are quite similar and clearly discriminate between tones and TORCs. These different responses faithfully reflect differences in the stimuli. For the CLR-D stimuli, there is no difference in the A1 population response to low and high rate click trains, in either naïve or trained animals. Enhanced contrast (compared to the naïve) begins to emerge in the trained animals in dPEG, where the Warning response significantly exceeds the Safe stimuli response. The contrast becomes even clearer in VPr, where the Go-NoGo behavioral meaning that they have acquired during training is clearly manifested in both pre-passive and active behavior conditions in VPr (Figure 6, Supplementary Figure S-7). Figure 6 compares the distribution of Warning-Safe response contrast ($nFR_{(W,S)}$) recorded in the four different cortical regions studied. Consistent with all previous average population findings, response differences reflecting behavioral meaning of the Go-NoGo stimuli increase with training and with active performance.

Responses to TORC stimuli depend on both sensory and behavioral context

Encoding of stimulus meaning in VPr and other cortical fields is also demonstrated by the changes in response to the class of TORC stimuli (the set of 30 modulated noise sounds), which had at least three distinct behavioral meanings for the ferrets, depending on context. TORCs served as (1) “*Safe*” stimuli in the Pure Tone Detection (PT-D) task (Figure 1a). We note that the same sequence of stimuli in the PT-D task was played in passive (listening, but no task – see context 3) and in active task conditions. TORCs were also (2) behaviorally “*Neutral Anticipatory*” stimuli preceding both Warning and Safe click trains in the Click-rate Discrimination (CLR-D) task. In this context, TORCs carried virtually no information about the upcoming click rate, but they *did* provide information as to the onset time of the upcoming click train (Figure 1a). As noted above, the same sequence of stimuli in the CLR-D task was played in both passive listening and active task conditions. (3) “*Behaviorally irrelevant*” TORCs were also regularly employed to measure STRFs, devoid of any other stimulus task sequence or behavioral task context. Likewise, TORCs in the passive presentations PT-D and CLR-D stimuli also played a mostly “behaviorally irrelevant” role (although the context of the stimulus sequence intermixed with Warning sounds might trigger behavioral associations, even in the absence of reward). Therefore, we compared the responses to TORCs in these three contexts in the same cells, and across different cortical regions, to highlight the extent and manner in which responses are shaped both by stimulus context and behavioral meaning (Supplementary Figure S-8). Passive TORC responses were

stable across stimulus contexts in A1 and dPEG, but varied between contexts in VPr, differences that were amplified during active engagement in PT-D versus CLR-D tasks (Figure 5).

Post-stimulus persistence of target responses

In addition to exhibiting a large contrast enhancement between Warning and Safe sounds during the duration of task stimuli, higher cortical areas (especially VPr and dIFC) also showed a persistent response to the Warning stimulus *after* the sound ended. This extended post-stimulus response preserved a short-term (800 ms) “memory” of the contrast after the offset of the Warning stimulus, which persisted through the 400 ms pre-shock window and the 400 ms shock window (during which the animal had to refrain from licking in order to avoid shock – see Figure 1a). This post-stimulus activity is also evident in Figures 3b-c,e-f and in Figure 4, where the response to the Warning stimulus clearly persists in the post-stimulus interval. To quantify this post-Warning activity, we measured the post-stimulus firing rate change from passive to the active state in the silent interval 50–700 ms after target offset (Figure 7 and Supplementary Figure S-9). Post-warning-stimulus response persistence was not observed in A1, and is most apparent in VPr and dIFC regions. The four cortical areas, A1 (PT-D n=71, CLR-D n=57 neurons), dPEG (PT-D n=199, CLR-D n=60 neurons), VPr (PT-D n=251, CLR-D n=266 neurons) and dIFC (PT-D n=138, CLR-D n=38 neurons), had significantly different post-stimulus Warning responses in both tasks (Kruskal-Wallis test, PT-D: chi-square=40.947, $p=6.7\times 10^{-9}$, df=3; CLR-D: chi-square=12.7391, $p=0.0052$, df=3). A post-hoc Tukey’s HSD test revealed significant differences between higher-order areas (VPr, dIFC) and A1 in the CLR-D task (A1/VPr: effect size [$\text{mean}_{(A1)}-\text{mean}_{(VPr)}$]= -60.709, $\text{CI}_{95\%}=[-109.27, -12.14]$, $p=0.0072$, A1/dIFC: effect size=-70.39, $\text{CI}_{95\%}=[-139.36, -1.43]$, $p=0.0433$) and with both earlier auditory areas (A1, dPEG) in the PT-D task (A1/VPr: effect size=-76.51, $\text{CI}_{95\%}=[-143.97, -9.06]$, $p=0.0187$, A1/dIFC: effect size=-90.73, $\text{CI}_{95\%}=[-164.04, -17.42]$, $p=0.008$, dPEG/VPr: effect size=-99.33, $\text{CI}_{95\%}=[-147.42, -51.25]$, $p=0$, dPEG/dIFC: effect size=-113.55, $\text{CI}_{95\%}=[-169.54, -57.55]$, $p=0$). No significant difference was found between VPr and dIFC post-stimulus activity in either task (PT-D: effect size=-14.21, $\text{CI}_{95\%}=[-67.64, 39.21]$, $p=0.9035$; CLR-D: effect size=-9.69, $\text{CI}_{95\%}=[-65.74, 46.367]$, $p=0.9708$).

Discussion:

The present results extend our understanding of neural encoding of sound in a higher “tertiary” auditory cortical region and comprise the first extensive description of neurophysiological responses to acoustic stimuli in area VPr of ferret auditory cortex. The findings reveal a profound transformation of responses between “passive” listening and “active” behavioral context, producing a representation that is consistent with emergent behavioral control signals observed in frontal cortex during the same behaviors. In the quiescent state, VPr responses are distinguishable from those of lower auditory cortical areas (A1, AAF, PPF, PSF) by their significantly longer response latencies, poor phase locking and broader frequency tuning (Figure 3). However, the distinctiveness of VPr responses emerges more vividly *during* active task performance with (1) selective response enhancements to Warning stimuli, (2) unveiling of long-term effects of learning, and (3)

encoding of behavioral meaning of task stimuli not only during a sound, but also *after* it, reflecting reward contingencies and task-action timing, maintained in short-term memory during behavior. These three characteristic features of VPr responses are discussed in more detail below.

Enhanced Warning sound responses during behavior

Although VPr exhibits task-related plasticity in receptive field and response properties as demonstrated previously in A1 and dPEG^{11,15,17}, the greater magnitude, scale and nature of the current neuroplasticity results place VPr at a higher level in the auditory cortical network, intermediate between dPEG fields (PPF and PSF) and dlFC. This position in the auditory cortical pathway is supported by our neurophysiological findings and also by neuroanatomy. The dramatic selective enhancement of VPr responses to Warning sounds during behavior are presumably mediated by the development of new context-dependent neural circuitry during task learning (Figure 4), which in turn also transforms responses to other task-relevant stimuli depending upon behavioral context (Figure 5, S-2,S-3). A remarkable feature of many VPr cells is how quickly they can transition from general auditory responses (in pre-passive conditions) to highly specific responses to Warning sounds (Figures 3a,c, S-2, S-3). Some VPr cells are even more extreme, exhibiting “frontal-cortex-like” properties¹⁸ in that they show very little or no response to Safe or Warning sounds in the passive condition, while selectively responding to Warning stimuli during active behavior (Supplementary Figure S-3). These results illustrate the importance of behavioral state as well as stimulus choice, in shaping sustained responses²⁹. Although the neural mechanisms for such swift attention-driven transformations are currently unknown, similar rapid changes in response properties have been postulated to reflect top-down influences that dynamically switch local network properties associated with each learned task^{30,31}. The top-down effects of task-engagement on receptive field plasticity have been shown to reach A1¹⁵ and subcortically even to the Inferior Colliculus³². Although not evident in the population responses recorded in A1 (Figures 4, 7, S-9) a recent decoding analysis suggests that post-stimulus A1 activity maintains a memory of stimulus behavioral meaning during task-engagement³³. It is possible that this information in A1 is dependent on top-down projections from VPr, dlFC or other higher areas in the auditory attention network.

Choice Probability in VPr

Multiple groups have reported significant choice probability (CP) in auditory cortex, indicating that even sensory neurons carry information about an upcoming decision. Significant CP has been found in A1 for one type of auditory task³⁴, whereas in other tasks, CP was only observed in higher auditory cortical areas^{10,12}. One recent paper¹² highlights the causal role of the auditory belt field AL in the monkey in contributing to perceptual decision making. In light of the present results in VPr, we predicted that VPr would be involved in extracting the behavioral meaning of the acoustic stimuli and forming auditory perceptual decisions. However, our analysis of CP for the two tasks in the present study did not yield significant results in A1 nor in higher auditory areas as might be predicted from earlier work¹². However, it is quite likely that different auditory cortical regions play different roles depending upon species, task design, level of difficulty and context. Further

studies are needed to test CP in VPr in positive reinforcement Go-NoGo or 2-alternative-forced-choice behavioral paradigms.

Long-term effects of learning in VPr

VPr population responses exhibited a systematic and clear contrast between responses to Warning and Safe stimuli even in the passive state, but *only* in trained rather than task-naïve animals (Figures 5, 6). This training-dependent enhanced contrast was weak or absent in lower auditory cortical areas (Figure 5). We ascribe this to the long-term effects of learning that reshape responses in higher cortical areas such as VPr based on their behavioral significance. However, in the dlFC these training effects are *only* evident during behavior, because of the absence of any significant responses in the passive state, reflecting behavioral gating¹⁸. We conjecture that these VPr learning effects may be similar to the experience-dependent malleability of “protocortex” described in visual area IT after extended training⁴.

Sustained post-stimulus responses may track reward and motor timing

VPr responses exhibit another dimension that reveals a similarity with dlFC: sustained post-Warning responses (Figures 3, 4) coding for task timing and the behavioral-response window (in passive listening, but even more clearly during active task performance). In the two Go-NoGo conditioned avoidance tasks in the present study¹⁵ animals learn to cease licking during a 400–800 ms window following Warning stimulus offset. Activity in VPr clearly encodes this timing in the form of post-stimulus responses that, across different single neurons, (a) occur precisely during this narrow temporal window, (b) persist precisely from stimulus offset up to this window, or (c) persist the full 800 ms and beyond (Figures 3a,c, S-2, S-3, S-9). These post-stimulus responses are not present in A1 and begin to appear only in higher auditory areas for both PT-D and CLR-D tasks in VPr and dlFC, as shown in the cartoon representation of population level profiles of passive and active responses in the cortical hierarchy shown in Figure 8.

This encoding of non-acoustic information, such as task decision, motor response or timing, reward, task-correlated visual or somatosensory signals is in general accord with earlier findings^{34–36} that have emphasized the “semantic” processing that occurs in auditory cortex³⁷.

Evidence for VPr as a tertiary region in ferret auditory cortex

Neuroanatomical and neurophysiological studies of the ferret auditory cortex over the past three decades (Figure 1) have revealed the presence of multiple auditory areas, including primary-like areas A1 and AAF, adjacent secondary areas, ADF, PPF and PSF, and still higher auditory areas AVF, aPSSC, pPSSC and VP^{19,20,23,38}. The most recent neuroanatomical connectational data²⁰ support the idea that in the PEG, PPF and PSF may both be secondary or belt areas, since both areas reciprocally interconnect with core areas A1 and AAF. In contrast, while there are reciprocal projections from both PPF and PSF to VP, there do not appear to be projections from core areas to VP²⁰, suggesting that VP may correspond in hierarchical position to a parabelt auditory area in primates.

Author Manuscript

Author Manuscript

Author Manuscript

VP_r can be reliably accessed by carefully mapping tonotopic organization in the MEG and PEG, and determining the position of A1 and PPF, which have mirror tonotopic maps (Figure 1). In recordings lateral to the high frequency (anterolateral) region of PPF, there is a sudden and abrupt change in passive response properties and frequency tuning as summarized in Figure 2, marking the entry into the VP_r region. We recorded in the rostral area of VP in an area up to 2–3 mm lateral to the boundary with PPF, and rostrally up to the pseudosylvian sulcus. These findings are consistent with the only previously published data on tuning in VP^{20,23}. Although future studies will be needed to determine what differences may exist in the responses of the various VP subfields that have been identified²¹ (VP_r, VP_c and VP_v), as well as their multisensory character and spatial tuning²⁰, our current VP_r results reveal many of the passive auditory response properties associated with an auditory “parabelt” area, including broader receptive fields, longer latency and duration responses, low SNR and sparseness. In addition, VP_r displays an impressive array of strong behavioral effects, including rapid short-term (driven by attentive task-engagement) and long-term task-related plasticity and learning.

Comparison of VP_r with the primate parabelt and other tertiary cortical areas

Author Manuscript

Author Manuscript

Author Manuscript

Tertiary sensory cortex is a higher order cortical sensory area at least two synapses up the cortical hierarchy from primary sensory regions. In the monkey auditory system, the primary (core) regions project to multiple, adjacent areas within a secondary (belt) region, which in turn project to areas in a tertiary (parabelt) region^{39,40}. Although the neuroanatomy and connections of parabelt and other regions of primate auditory cortex have been well elucidated^{39–42}, and there is an abundance of insight about processing in lower auditory cortical areas from neurophysiological studies of responses in core and belt, to our knowledge there are only two published studies on the neurophysiological response properties of parabelt in awake but non-behaving macaque monkeys and marmosets^{43,44}. Both studies demonstrate that tone response latencies increase from A1 to belt to parabelt and were longest in rostral parabelt – consistent with our results in the ferret (see Figures 1f, 1a,d). Similar to the primate, belt regions (PPF and PSF) in the ferret receive strong inputs from primary regions (A1 and AAF) and in turn, VP receives inputs from the belt regions – with no (or negligible) A1 or AAF inputs. Our neurophysiological results are also generally consistent with the possibility that the auditory cortical hierarchy, as in the somatosensory and visual systems, not only follows a hierarchical ordering of increasing response latencies but also of increasingly long temporal windows for sensory integration⁴⁵. Hence, using these criteria of response latency and cortical connectivity, ferret VP is a tertiary region that bears similar features to parabelt as defined in primate auditory cortex. Despite these parallels, establishing clear homologies between cortical areas across species is difficult and daunting, especially in higher-order sensory areas⁴⁶. In order to elucidate the relationship between the organization and architecture of auditory cortical areas in carnivores (ferret and cat) and primates, further careful comparative neuroanatomical and neurophysiological studies which will be required. These future studies will be necessary to clarify possible homologies between ferret dorsal dPEG and VP_r, the multiple higher auditory cortical regions in the cat⁴⁷, and belt and parabelt regions in primate auditory cortex.

However, in general, compared to primary sensory regions, secondary and tertiary cortical sensory areas integrate inputs over longer periods of time, show greater context-dependent adaptive plasticity, are more concerned with the associative functions involved in perception, object recognition and object memory, and have also been shown to be closely linked to perceptual decision-making and action. A recent study of human tertiary auditory cortex describes responses that transform from acoustic to perceptual dimensions in the context of the McGurk effect⁴⁸ that illustrates this transformational process in human auditory processing. A comparable tertiary region in the primate visual system may be the inferotemporal cortex, which also plays a key role in object perception and recognition^{3,46} as part of the gradual progression from sensory to task-related processing in the cerebral cortex⁴⁹.

Conclusion – VPr in the Auditory Attention Cortical Network

In conclusion, the physiological response properties of VPr identify a higher field in ferret auditory cortex, distinct from previously characterized areas, situated midway along the auditory cortical network from A1 to dIFC. Responses in VPr are dynamically driven by selective attention during task engagement and markedly reshaped by task conditions and behavioral state (Figures 4, 8). VPr is also distinctive in showing long-term changes in representation of learned task-relevant stimuli (Figure 5). Another feature of VPr is the prominent post-stimulus response to Warning stimuli that is likely to reflect an emergent representation of non-acoustic task-related information such as reward and task timing for action (Figures 3, 4, 5, 8, S-2-S-7). This marks a transition from a nearly veridical acoustic spectrotemporal representation in A1 to a more cognitive representation based on the behavioral meaning associated with incoming sounds in secondary auditory cortical areas (in dPEG)¹¹ and even more strongly in tertiary areas such as VPr. The beginnings of this transition occur as early as A1^{15–17,33–37} and even in IC³², but are most clearly visible in higher auditory cortical areas such as VPr and are influenced by task engagement^{10,12,34}. Our results provide new insights into the transformation from sound to behavioral meaning in the auditory pathway⁵⁰ and also raise new questions as to the neural basis for the differences in task-driven attentional modulation at multiple hierarchical levels of the auditory system, the functional role of VPr in selective auditory attention and task representation, the mechanisms underlying long-term auditory learning, and the role of top-down projections in mediating higher level auditory processing.

Online Methods:

Training and behavioral tasks

We measured auditory response properties of cortical single units in auditory and frontal cortices of awake animals during passive listening to experimental sounds and during performance of auditory tasks. All animals used were female, de-scented, neutered, 1–4 years old adult ferrets (*Mustela putorius furo*) obtained from Marshall BioResources, North Rose, New York. Three additional untrained (task-naïve) ferrets were used for mapping auditory areas and anatomical studies, one of which was also used for naïve passive recordings (Figures 5,6). For comparison of VPr data with other auditory cortical areas such as A1, dPEG and with dIFC, we also re-analyzed data from 14 animals that were trained and

recorded for previous behavior studies of A1, dPEG, dIFC^{11,15,17,18} and from a total of 32 (including responses from 18 additional animals used in previous published studies from our laboratory) for comparing VPr tuning properties with A1 and dPEG during passive listening (Figure 2). Animal subjects were randomly chosen from the colony for experimentation. Some animals were directly implanted with headposts for naïve neurophysiological recordings. Other animals were trained on auditory tasks and were implanted with headposts and used for behavioral neurophysiological recordings. All experimental procedures were approved by the University of Maryland Animal Care and Use Committee and conformed to standards specified by the National Institutes of Health.

We trained four animals on two auditory discrimination tasks with a conditioned avoidance paradigm⁵¹: pure tone detection (PT-D^{11,15,17,18}) and click rate discrimination (CLR-D). In both tasks, animals were presented with a series of 1 to 6 Reference “Safe” sounds on each behavioral trial, followed by a Target “Warning” sound. No Warning sounds were presented on catch trials. Ferrets learned to freely lick water flowing continuously from a spout during presentation of Safe sounds, and to briefly refrain from licking after the Warning stimulus offset (for a minimum of 400 ms in a 400–800 ms post-stimulus window) to avoid receiving a mild tail shock. In the PT-D task, Safe sounds were 1–2 s duration broadband rippled noise (Temporally-Orthogonal Ripple Combinations, or TORCs⁵²), and Warning sounds were pure tones (duration and sound level matched to the TORC references) of equal amplitude. In the CLR-D task, both Safe and Warning sounds were composed of 1.25 s TORCs immediately followed by a 0.75 s click train, but the Warning click rate was fixed to be either higher or lower than the Safe click train. The click train rates used in CLR-D varied from 4 to 48 Hz, and animals (n=4) were trained in one of the two directions of the task, meaning that half of the ferrets (n=2/4) were trained with low-rate Safe click trains (4–24 Hz) and high-rate Warning click trains (16–48 Hz), while the other half (n=2/4) were trained with high-rate Safe click trains and low-rate Warning click trains. The distribution of click rates used is shown in Supplementary Figure S-4 (**Panel A**). The minimum rate difference between the Safe and Warning click trains was 7 Hz, while the maximum was 32 Hz. The distribution of high-to-low click rate ratios used in every behavioral block is shown in Supplementary Figure S-4 (**Panel B**). The mean (\pm SD) ratio used was 3.2 ± 1.09 . All animals trained in the click-rate discrimination task were also trained in the tone detection task (n=4) and recordings were made from all four cortical areas - A1, dPEG, VPr and dIFC. However, the animals from earlier published studies (recordings in A1, dPEG and dIFC but not in VPr) were exclusively trained on the tone detection task (n=14).

The Warning sound frequency in the tone detection task, and Safe and Warning click-train rates in the click discrimination task, were varied among experiments and training sessions, but were held fixed during a single behavioral block. This variability across training days lead the animals to generalize their behavioral responses to *any* Warning tone frequency or Warning click-train rate. Warning tone frequency was chosen after assessing the frequency tuning of the neurons at hand by presenting tone pips (100 ms) of random frequencies (125–32,000 Hz); we chose Warning frequencies that evoked the strongest possible responses in all neurons being recorded. Similarly, click rates for Safe and Warning sounds were chosen after presenting click trains 4 – 60 Hz. Sometimes, VPr neurons broad frequency and click-

rate tuning did not allow to choose a best frequency or click rate, in those cases we chose a frequency or click rate that evoked responses for most neurons being recorded. The sound level (65 – 70 dB SPL), sound durations (1 or 2 s in tone detection, 2s in click-rate discrimination) and inter-stimulus intervals (1.2 or 1.6 s) were kept fixed in a single training or recording session as well. During electrophysiological recording sessions, the task stimulus set was presented before (pre-passive) and after (post-passive) behavior, while the animal passively listened to task sounds and no water was provided (i.e. animal did not engage in the task and displayed no licking behavior). The normal duration of a behavioral block was 10–15 minutes, which depended on the duration of task sounds and on how thirsty animals were at a given time (each trial was initiated only after a lick); in cases where animals were not very thirsty and licked less frequently, the behavioral block duration could extend up to 20 minutes. Passive task blocks were usually 7–12 minutes long, depending on the duration of task sounds.

No blinding was used in the study, as is common in behavioral neuroscience. However, all analyses of neural responses were conducted in the same way for all data acquired over a period of ~15 years in our laboratory, including data from multiple animals and investigators. The neuroanatomical studies were conducted by a researcher who was not directly involved in behavioral or neurophysiological studies, and received brains to process and during histological processing was blind to each animal history.

Surgery

Initially, animals were trained in a freely-moving setting until they reached a consistent and acceptable performance, that is, achieving > 80% hit rate accuracy and < 20% false alarm rate for a discrimination rate > 0.65 in at least two consecutive training sessions^{15,18}. In order to secure stable electrophysiological recordings, ferrets were surgically implanted with a stainless steel headpost that was attached to the sagittal interparietal suture. During surgery, ferrets were anesthetized with a combination of Ketamine (35 mg/Kg IM) and Dexmedetomidine (0.03 mg/Kg SC) for induction, and deep levels of anesthesia were maintained with 1–2% Isoflurane throughout the surgery. Animals were also medicated with atropine sulphate (0.05 mg/kg SC) to control salivation and to increase heart and respiratory rates. During surgery, ECG, pulse and blood oxygenation were monitored, and rectal temperature was maintained at ~38° C. Using sterile procedure, and in order to be able to reach the ventral areas of auditory cortex, the skull was surgically exposed by making a midline incision in the scalp and by dissecting both temporal muscles from their insertion in the sagittal interparietal crest down to the level of the zygomatic arch. The headpost was secured in the skull with titanium screws and polycarboxylate cement (Durelon), and then areas surrounding frontal and auditory cortices are covered with bone cement (Zimmer), leaving small (2 – 3 square mm) cavities for easy access to auditory and frontal cortex in both hemispheres. Following surgery, antibiotics (Cefazolin, 25 mg/Kg SC) and analgesics (Dexamethasone 2 mg/Kg SC and Flunixin meglumine 0.3 mg/Kg SC) were administered.

Animals were allowed to recover for ~2 weeks prior to being habituated to head restraint in a customized Lucite horizontal cylindrical holder for a period of 1–2 weeks, and then retrained to criterion on the tasks for additional 2–3 weeks while restrained in the holder.

Before recording sessions, small craniotomies were made over auditory or frontal cortex and, in simultaneous recording experiments, ipsilateral auditory and frontal cortices. The bone cement (Zimmer) implant securing the headpost, and sterile impression material placed in wells between recording sessions, allowed the craniotomies to be kept well protected from the environment. The wells in the head cap implant containing the craniotomies were kept sealed between experiments with sterile vinyl polysiloxane impression material (Examix NDS) and were cleaned and treated with topical antiseptic drugs (Povidone-iodine) and antibiotics (cefazolin or enrofloxacin, 0.2 ml) at least once per week. The skin surrounding the implant was cleaned 3 times per week with warm saline and treated with povidone-iodine and sulfadiazine cream ointment.

Neurophysiological recording

Training and neurophysiological recording experiments were conducted in a double-walled sound attenuating chamber (IAC). High impedance (2–6 M Ω at 1 kHz) tungsten microelectrodes (FHC) were used for extracellular neurophysiological recordings and stainless steel electrodes (FHC or Microprobes) for iron deposit markings. Electrodes were arranged in a 4-electrode square array and separated by 0.5 mm from their nearest neighbor. In each recording session, four electrodes per recording area (four in auditory and four in frontal cortex) were independently advanced through the dura into cerebral cortex using an Alpha-Omega EPS drive system. Electrodes were slowly and independently advanced until good spike isolation was found in the majority of the electrodes. Data acquisition was performed with an Alpha-Omega AlphaLab data acquisition system, signals recorded at 25,000 samples/s and amplified 15,000 X. Additionally, 18 recording sessions in primary auditory cortex were performed in a subset of 2 animals using a 24 electrode linear array (Plexon U-Probe) with 75 μ m between electrode contacts and impedances between 275–1500 k Ω at 1 kHz, using Plexon and TBSI 1X headstages, amplified with Plexon preamplifiers and acquired using MANTA, an open source data acquisition suite written for MATLAB⁵³. Single units (usually one or two per electrode) were isolated by *k*-means clustering using custom software written in MATLAB or open-source software Klusta (version c1909dd)⁵⁴. Sound stimuli, behavioral operation, on-line analysis and recording triggering were controlled with customized open source software (Behavioral Auditory PHYsiology, BAPHY).

Localization of auditory fields and recording sites

Prior to record activity in non-primary auditory areas, initial recordings were directed to A1 and determined by relative distance to external cranial landmarks. In female ferrets, A1 is located approximately 16 mm anterior to the occipital midline crest and 12 mm lateral to the skull midline. During initial recording sessions in each animal, small craniotomies were placed above A1 using these coordinates, and A1 responses were confirmed by analyzing the tuning properties of the recorded cells in response to 100 ms tone pips of random frequencies spanning 8 octaves, presented at intervals of 1 s. Also, 3 s TORCs were used to compute spectro-temporal receptive fields (STRFs). A1 neurons are known to present sharp tuning to pure tones and clear single-peak, short latency STRFs^{11,15,19}. Determination of neuronal best frequencies allowed us to confirm the location of A1 based on its characteristic tonotopic organization of high to low frequency gradient in a dorsal-ventral

direction¹⁹. Then, by ventrally expanding the existing craniotomy it was possible to gain access to non-primary auditory areas in the Posterior Ectosylvian Gyrus (PEG)^{19,22}. Two subfields in the dorsal PEG (dPEG), Posterior Pseudosylvian Field (PPF) and Posterior Suprasylvian Field (PSF), display a reversal in the tonotopic map, sharing a low frequency area with A1 and displaying higher frequency regions more ventrally. Both fields are also separated by a low frequency border, meaning that PSF displays a low to high frequency tonotopic gradient in an anterodorsal to postero-ventral direction, while PPF displays low frequencies postero-dorsally and high frequencies antero-ventrally^{11,19}. Neurons in these areas display broader tuning, longer latencies and longer sustained responses than A1 (Figure 2), and their STRFs display more complex patterns of excitatory and inhibitory subfields, with more numerous, longer and less compact excitatory and inhibitory subfields in both the spectral and temporal axis¹¹. The locations of dPEG recordings were confirmed by checking the tuning properties of its neurons and their location relative to the tonotopic map.

VPr (Rostral Ventral Posterior field) recordings were directed to a region ventral to the high-frequency region of field PPF. VPr spans an area 1–2 mm below and ventral to high frequency PPF and ventral to the lower lip of the PSS (pseudosylvian sulcus). Partly because of its extreme lateral location and limited accessibility for surface recordings, VP has remained one of the least studied areas of the ferret auditory cortex. Localization of VPr was based upon a tonotopic discontinuity at the high end of the frequency map in adjacent area PPF, which is characterized by a posteromedial-anterolateral tonotopic gradient from low to high frequency tuning^{11,19}. The transition from PPF to VPr is characterized by a sharp transition in frequency tuning from high to lower best frequency and often much broader frequency tuning (Figures 1d,1e). VPr also is characterized by much longer onset response latencies (see Figure 1f). Careful measurements of the location of the recording electrode relative to two reference marks placed in the bone cement surrounding the craniotomy and to a mark at the center of the headpost, were recorded for later reconstruction of electrode penetration sites. Locations of recordings in VPr (see Supplementary Figure S-1) were marked with electrolytic lesions or iron deposits by passing a small current (10 μ A) for 5 minutes using stainless steel electrodes. *Post-mortem* confirmation of iron deposits was determined by histological examination of Prussian blue reactions.

Stimuli

All acoustical stimuli were presented at 65–70 dB SPL, with the exception of a wider range of amplitudes specifically for the tones (which varied from 40–80 dB SPL) used for multi-level tuning assessment and measurement of frequency response curves. Sounds were digitally generated at 40 kHz with custom-made MATLAB functions and National Instruments A/D hardware (PCI-6052E) and presented with a free-field speaker positioned 30 cm in front of the animal's head. Tones (5 ms onset and offset ramps) were used as target stimuli in the tone detection task and, previous to any behavioral testing, to assess frequency tuning by using tone pips of random frequencies spanning eight octaves. Individual clicks in the click-rate discrimination task (occurring after TORCs in sequential TORC-click train stimuli – see Figure 1a), and during passive testing of click tuning with 1 second click trains of randomly varying click rates from 4–60 Hz) were composed of 0.01 s square pulses of

alternating polarity. Thirty distinct temporally-orthogonal ripple combinations (TORCs) were used as task distractor (Safe) sounds and also for computation of STRFs in and out of task context. TORCs were randomly chosen without replacement from a set of 30 TORCs for each TORC set repetition. Each TORC was composed of 5-octave wide broadband noise with a dynamic spectrotemporal profile that is the superposition of the envelopes of six temporally orthogonal ripples (for 4–24 Hz TORCs) or twelve temporally orthogonal ripples (4–48 Hz TORCs). Ripples composing the TORCs had linear sinusoidal spectral profiles, with peaks equally spaced at 0 (flat) to 1.2 cycles-per-octave; the envelope drifted temporally up or down the logarithmic frequency axis at a constant velocity^{52,55}. The envelope of these ripples drifts temporally up or down the logarithmic frequency axis at a constant velocity (4 – 48 or 4 – 24 Hz). The 5-octave spectrum of TORCs could be varied in several ranges and was chosen at each recording session to best span the frequencies of the neurons being recorded.

Data analysis

Offline data analyses were performed with custom-made MATLAB and Python scripts. Figures were made using MATLAB (R2010B) functions, Matplotlib (3.0.2) and Seaborn 0.9.0 Python libraries and Inkscape (0.92.3). No statistical methods were used for pre-determining sample sizes, see Life Sciences Reporting Summary.

Basic tuning properties were determined by analyzing the responses to random frequency tones spanning 6–8 octaves (11 tones per octave), usually ranging from 125 Hz to 32 kHz at 65–70 dB SPL. A Gaussian function was fit to the mean firing rate during a window of 100 ms after tone onset. Best frequency was determined to be the mean of the Gaussian curve, and tuning spectral bandwidth was measured as its width in octaves at half-height. Tones presented had a duration of 100 ms and were presented at 1 s intervals. Response latency was measured from the peristimulus time histogram (PSTH) binned at 1 ms and computed from the responses to all pure frequency tones by measuring the time from tone onset to the peak spike rate in a 100 ms window.

TORCs were also presented to compute spectro-temporal receptive fields (STRFs) by means of reverse correlation⁵² between a time-varying neural response (i.e., spikes, multi-unit activity) and the spectrogram of the TORCs presented during experiments. Positive STRF values indicate time and frequency components of TORCs correlated with increased neural responses (i.e., an excitatory field), whereas negative values indicate components correlated with decreased responses (i.e., a suppressive or inhibitory field). The reliability of responses to TORCs, and the quality of resulting STRFs, was measured by a signal-to-noise ratio (SNR), computed as the ratio of power in the average PSTH response to all TORCs to power in the difference of single-trial responses from the PSTH⁵⁵. Only responses with SNR values equal or higher than 0.2 were used. Response duration was measured as the width at half-height of the STRF positively rectified and averaged over frequency. STRFs varied in their complexity between auditory areas, with clear excitatory/inhibitory fields in a narrow spectrotemporal range in A1 and more complex tuning in higher-order auditory areas. To compare tuning complexity between areas, we measured an STRF sparseness index, computed as the peak magnitude of the STRF, divided by the standard deviation across

STRF bins¹¹. Higher sparseness values are associated with sharply tuned STRFs concentrated over a few contiguous bins. Consistent with the observation of greater complexity in downstream areas, A1 sparseness index values were greater than for higher-order auditory neurons. Both the Kruskal-Wallis test and Tukey's HSD post-hoc test were conducted to examine differences in tuning parameters between areas. We decided to use the Kruskal-Wallis test instead of the one-way ANOVA because all tuning parameter distributions, with the exception of VPr bandwidth ($p=0.268$, Lilliefors test), significantly deviated from normality (Lilliefors test, $p<0.001$).

Single-unit neural responses to task stimuli were measured by computing PSTHs by binning spikes at 30 Hz and obtaining the mean and SEM of the spike rates obtained over all Safe and Warning sound classes. We only analyzed responses to task sounds from neurons that showed significant responses to auditory stimuli. Units were considered auditory-responsive neurons when there were at least 2 bins (33.3 ms bins) significantly modulated from baseline in the PSTH in response to any (Safe or Warning) sound ($p<0.05$, jack-knifed t-test, Bonferroni corrected). Significance of behavioral effects within cells was measured by performing a Wilcoxon sign rank test, corrected for multiple comparisons using false discovery rate⁵⁶, between passive and active PSTHs of each sound class (Safe or Warning). Neural responses were considered to be significantly modulated by behavior (comparing active vs passive responses) if there were at least 2 consecutive, significant (Wilcoxon sign rank test, $p<0.05$) PSTH bins within the responses to any task-sound class (Safe or Warning stimuli). Normalized average PSTHs were computed by subtracting the baseline firing rate (measured during the silent pre-stimulus period) from each neuron and then dividing each neuron firing rates by the peak modulation of the mean population PSTH, thus adjusting the scale to spikes per second above or below spontaneous activity. The mean and SEM of each PSTH bin were computed by a jack-knife procedure⁵⁷. We computed choice probability (CP) from data from all auditory cortical areas in our study following the method described by Sutter and colleagues⁵⁸.

Motor-related lick responses

As in a previous paper¹⁸, we determined significant neural modulation of neuronal activity in VPr by auditory stimuli by stepwise linear regression of time-varying spike activity (binned at 50 ms) against stimulus (safe and warning sounds) and motor (licking) events. The complete regression modeled spiking activity as a function of Safe/Warning sounds and lick events.

$$r(t) = \sum_{\tau=-T}^T h_s(\tau)s_s(t-\tau) + h_w(\tau)s_w(t-\tau) + h_m(\tau)m(t-\tau) \quad \text{Equation (1):}$$

The stimulus functions, $s_s(t)$ and $s_w(t)$, are 0, except at times, t , of Safe or Warning sound onset, respectively, when they have a value of 1. Similarly, the motor function, $m(t)$, has a value of 0 except at times when lick events occur. The regression functions, $h_s(\tau)$, $h_w(\tau)$ and $h_m(\tau)$, then indicate the average firing rate before and after each corresponding event. The

regression functions were fit by normalized reverse correlation, which discounted spurious effects that might arise as a result of correlations between stimulus events and changes in motor activity. Neurons were classified as being significantly modulated by sensory inputs if the occurrence of a stimulus predicted a change in firing rate that could not be explained by a simpler model on the basis of motor activity alone.

$$r(t) = \sum_{\tau = -T}^T h_m(\tau)m(t - \tau) \quad \text{Equation (2):}$$

Thus, a neuron was considered to be modulated by sensory inputs only if the full model predicted spiking activity significantly better than the model based only on licking activity ($P < 0.05$, jackknifed t test).

Statistical analysis

In order to quantify differences in target/reference response contrast in passive and active behavior conditions, we used a 3-way repeated-measures ANOVA (rmANOVA) for response differences between Warning and Safe sounds, with ‘area’ (A1, dPEG, VPr and dlFC) and ‘condition’ (passive and behavior) as fixed factors, and ‘neuron’ nested in ‘area’ as a random intercept. For PT-D task, we calculated the response difference (spikes/s) for each neuron in a time window between 100 ms after stimulus onset (avoiding the purely sensory driven onset transient) and the end of the shock period (800 ms post-stimulus-offset). For the CLR-D task, we calculated the response difference (spikes/s) for the time window between 300 ms following click train onset and the end of the shock period (800 ms post-stimulus-offset). We used Tukey’s range test for post-hoc analysis.

Since areas VPr and dlFC display qualitative differences in their response patterns with areas A1 and dPEG, we measured responses in two different epochs: during the duration of task-relevant sounds and in the silent period after presentation of target sounds. We quantified the contrast between Safe and Warning responses ($nFR_{(W-S)}$) in a time window between 0.1 – 0.45 s after sound onset (PT-D task) or 0.3 – 1.0 s (CLR-D task), in order to avoid onset transient responses and, in CLR-D, the offset response to TORCs preceding click trains. We measured mean values of $nFR_{(W-S)}$ from normalized PSTH responses to compare population responses between areas and between the trained and naïve animals. The responses observed in the silent period after task target sounds were measured as the change in response from passive to active behavior conditions in a time window of 650 ms starting 50 ms after the target offset. Responses recorded during ‘miss’ trials, when the animal failed to report detection of a Warning sound, were discarded to avoid contamination of the data with shock artifacts. Area comparison of post-stimulus Warning responses was performed with a Kruskal-Wallis test and pair-wise area comparisons were performed with Tukey’s HSD range as a post-hoc test. Distributions of $nFR_{(W-S)}$ data computed from A1, dPEG, VPr and dlFC responses in passive and active conditions were tested for normality using Lilliefors test. All distributions of $nFR_{(W-S)}$ were found to significantly deviate from normality ($p < 0.05$) with the exceptions of three distributions collected in the CLR-D task

context; A1 passive $nFR_{(w-s)}$ ($p=0.1586$, $ks=0.1039$), dPEG active $nFR_{(w-s)}$ ($p=0.0595$, $ks=0.192$) and dIFC active $nFR_{(w-s)}$ ($p=0.1198$, $ks=0.1271$).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Celian Bimbard, Kelsey Dutta and Neha Joshi for assistance with neurophysiological recordings and Alex Meredith and Kelsey Dutta for assistance with neuroanatomical experiments, and L. Artemisia for support. This research was funded by grants from the US National Institutes of Health (R0-1 grants to SVD, and SAS and JBF) and by a DARPA grant to JBF. DD held an AGAUR fellowship (Generalitat de Catalunya) in the frame of the EU COFUND Marie-Curie program (2014BP-A00226) and an Erasmus Mundus ACN fellowship. DE held scholarships from CONICYT-PCHA / Becas Chile, Doctorado Convocatoria 2009-folio 72100839, Postdoctorado/ Convocatoria 2016-folio 74170109 and Fulbright-III.

References

1. Afraz A, Yamins DLK & DiCarlo JJ Neural Mechanisms Underlying Visual Object Recognition. *Cold Spring Harb. Symp. Quant. Biol* 79, 99–107 (2014). [PubMed: 26092883]
2. Yau JM, Kim SS, Thakur PH & Bensmaia SJ Feeling form: the neural basis of haptic shape perception. *J. Neurophysiol* 115, 631–642 (2016). [PubMed: 26581869]
3. Kornblith S & Tsao DY How thoughts arise from sights: inferotemporal and prefrontal contributions to vision. *Curr. Opin. Neurobiol* 46, 208–218 (2017). [PubMed: 28942219]
4. Arcaro MJ, Schade PF, Vincent JL, Ponce CR & Livingstone MS Seeing faces is necessary for face-domain formation. *Nat. Neurosci* 20, 1404–1412 (2017). [PubMed: 28869581]
5. Hernández-Pérez R et al. Tactile object categories can be decoded from the parietal and lateral-occipital cortices. *Neuroscience* 352, 226–235 (2017). [PubMed: 28377175]
6. Rossi-Pool R et al. Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *Proc. Natl. Acad. Sci* 113, E7966–E7975 (2016). [PubMed: 27872293]
7. Romo R, Lemus L & de Lafuente V Sense, memory, and decision-making in the somatosensory cortical network. *Curr. Opin. Neurobiol* 22, 914–919 (2012). [PubMed: 22939031]
8. Freedman DJ & Assad JA Neuronal Mechanisms of Visual Categorization: An Abstract View on Decision Making. *Annu. Rev. Neurosci* 39, 129–147 (2016). [PubMed: 27070552]
9. Rojas-Hortelano E, Concha L & de Lafuente V The parietal cortices participate in encoding, short-term memory, and decision-making related to tactile shape. *J. Neurophysiol* 112, 1894–1902 (2014). [PubMed: 24990569]
10. Niwa M, Johnson JS, O'Connor KN & Sutter ML Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. *J. Neurosci* 33, 8378–95 (2013). [PubMed: 23658177]
11. Atiani S et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486–499 (2014). [PubMed: 24742467]
12. Tsunada J, Liu ASK, Gold JI & Cohen YE Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nat. Neurosci* 19, 135–142 (2016). [PubMed: 26656644]
13. Dong C, Qin L, Zhao Z, Zhong R & Sato Y Behavioral Modulation of Neural Encoding of Click-Trains in the Primary and Nonprimary Auditory Cortex of Cats. *J. Neurosci* 33, 13126–13137 (2013). [PubMed: 23926266]
14. Nodal FR & King AJ in *Biology and Diseases of the Ferret*, 3 (eds. Fox JG & Marini RP) 685–710 (John Wiley & Sons, Inc, 2014). doi:10.1002/9781118782699.ch29
15. Fritz JB, Shamma SA, Elhilali M & Klein D Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci* 6, 1216–23 (2003). [PubMed: 14583754]

16. Fritz JB, Elhilali M & Shamma SA Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J. Neurosci* 25, 7623–35 (2005). [PubMed: 16107649]
17. David SV, Fritz JB & Shamma SA Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A* 109, 2144–9 (2012). [PubMed: 22308415]
18. Fritz JB, David SV, Radtke-Schuller S, Yin P & Shamma SA Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat. Neurosci* 13, 1011–1019 (2010). [PubMed: 20622871]
19. Bizley JK, Nodal FR, Nelken I & King AJ Functional organization of ferret auditory cortex. *Cereb. cortex* 15, 1637–53 (2005). [PubMed: 15703254]
20. Bizley JK, Bajo VM, Nodal FR & King AJ Cortico-cortical connectivity within ferret auditory cortex. *J. Comp. Neurol* 523, 2187–2210 (2015). [PubMed: 25845831]
21. Radtke-Schuller S Cyto- and myeloarchitectural brain atlas of the ferret (*Mustela putorius*) in MRI aided stereotaxic coordinates (Springer International Publishing, 2018). doi: 10.1007/978-3-319-76626-3
22. Pallas SL & Sur M Visual projections induced into the auditory pathway of ferrets: II. Corticocortical connections of primary auditory cortex. *J. Comp. Neurol* 337, 317–333 (1993). [PubMed: 8277005]
23. Bajo VM, Nodal FR, Bizley JK, Moore DR & King AJ The ferret auditory cortex: Descending projections to the inferior colliculus. *Cereb. Cortex* 17, 475–491 (2007). [PubMed: 16581982]
24. Recanzone GH, Schreiner CE & Merzenich MM Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci* 13, 87–103 (1993). [PubMed: 8423485]
25. Galván VV & Weinberger NM Long-Term Consolidation and Retention of Learning-Induced Tuning Plasticity in the Auditory Cortex of the Guinea Pig. *Neurobiol. Learn. Mem* 77, 78–108 (2002).
26. Reed A et al. Cortical Map Plasticity Improves Learning but Is Not Necessary for Improved Performance. *Neuron* 70, 121–131 (2011). [PubMed: 21482361]
27. Froemke RC et al. Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci* 16, 79–88 (2013). [PubMed: 23178974]
28. Bieszczad KM, Miasnikov AA & Weinberger NM Remodeling sensory cortical maps implants specific behavioral memory. *Neuroscience* 246, 40–51 (2013). [PubMed: 23639876]
29. Wang X, Lu T, Snider RK & Liang L Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435, 341–6 (2005). [PubMed: 15902257]
30. Gilbert CD & Li W Top-down influences on visual processing. *Nat. Rev. Neurosci* 14, 350–63 (2013). [PubMed: 23595013]
31. Caras ML & Sanes DH Top-down modulation of sensory cortex gates perceptual learning. *Proc. Natl. Acad. Sci. U. S. A* 114, 9972–9977 (2017). [PubMed: 28847938]
32. Slee SJ & David SV Rapid Task-Related Plasticity of Spectrotemporal Receptive Fields in the Auditory Midbrain. *J. Neurosci* 35, 13090–13102 (2015). [PubMed: 26400939]
33. Bagur S et al. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. *Nat. Commun* 9, 2529 (2018). [PubMed: 29955046]
34. Bizley JK, Walker KMM, Nodal FR, King AJ & Schnupp JWH Auditory cortex represents both pitch judgments and the corresponding acoustic cues. *Curr. Biol* 23, 620–5 (2013). [PubMed: 23523247]
35. Brosch M, Selezneva E & Scheich H Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci* 25, 6797–806 (2005). [PubMed: 16033889]
36. Yin P, Mishkin M, Sutter M & Fritz JB Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields A1 and R. *J Neurophysiol* 100, 3009–3029 (2008). [PubMed: 18842950]
37. Scheich H et al. Behavioral semantics of learning and crossmodal processing in auditory cortex: The semantic processor concept. *Hear. Res* 271, 3–15 (2011). [PubMed: 20971178]
38. Kelly JB, Judge PW & Phillips DP Representation of the cochlea in primary auditory cortex of the ferret (*Mustela putorius*). *Hear. Res* 24, 111–115 (1986). [PubMed: 3771373]

39. Kaas JH & Hackett TA Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U. S. A* 97, 11793–9 (2000). [PubMed: 11050211]
40. Hackett TA Information flow in the auditory cortical network. *Hear. Res* 271, 133–146 (2011). [PubMed: 20116421]
41. Hackett TA et al. Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: Refining the hierarchical model. *Front. Neurosci* 8, 72 (2014). [PubMed: 24795550]
42. Plakke B & Romanski LM Auditory connections and functions of prefrontal cortex. *Front. Neurosci* 8, 1–13 (2014). [PubMed: 24478622]
43. Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA & Hackett TA Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc. Natl. Acad. Sci* 109, 18168–18173 (2012). [PubMed: 23074251]
44. Kajikawa Y et al. Auditory properties in the parabelt regions of the superior temporal gyrus in the awake macaque monkey: an initial survey. *J. Neurosci* 35, 4140–50 (2015). [PubMed: 25762661]
45. Murray JD et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci* 17, 1661–1663 (2014). [PubMed: 25383900]
46. Kaas JH The future of mapping sensory cortex in primates: three of many remaining issues. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* 360, 653–64 (2005). [PubMed: 15937006]
47. Winer JA & Lee CC The distributed auditory cortex. *Hear. Res* 229, 3–13 (2007). [PubMed: 17329049]
48. Smith E et al. Seeing Is Believing: Neural Representations of Visual Stimuli in Human Auditory Cortex Correlate with Illusory Auditory Perceptions. *PLoS One* 8, e73148 (2013). [PubMed: 24023823]
49. Brincat SL, Siegel M, von Nicolai C & Miller EK Gradual progression from sensory to task-related processing in cerebral cortex. *Proc. Natl. Acad. Sci* 115, E7202–E7211 (2018). [PubMed: 29991597]
50. Bizley JK & Cohen YE The what, where and how of auditory-object perception. *Nat. Rev. Neurosci* 14, 693–707 (2013). [PubMed: 24052177]

Methods-only References

51. Heffner HE & Heffner RS in *Methods in Comparative Psychoacoustics* (eds. Klump GM, Dooling RJ, Fay RR & Stebbins WC) 79–93 (Birkhäuser Verlag, 1995).
52. Klein DJ, Depireux DA, Simon JZ & Shamma SA Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J. Comput. Neurosci* 9, 85–111 (2000). [PubMed: 10946994]
53. Englitz B, David SV, Sorenson MD & Shamma SA MANTA--an open-source, high density electrophysiology recording suite for MATLAB. *Front. Neural Circuits* 7, 69 (2013). [PubMed: 23653593]
54. Rossant C et al. Spike sorting for large, dense electrode arrays. *Nat. Neurosci* 19, 634–641 (2016). [PubMed: 26974951]
55. Depireux DA, Simon JZ, Klein DJ & Shamma SA Spectro-Temporal Response Field Characterization With Dynamic Ripples in Ferret Primary Auditory Cortex. *J. Neurophysiol* 85, 1220–1234 (2001). [PubMed: 11247991]
56. Benjamini Y & Yekutieli D The control of the false discovery rate in multiple testing under dependency. *Ann. Stat* 29, 1165–1188 (2001).
57. Efron B & Tibshirani R *An introduction to the bootstrap* (Chapman & Hall, 1994).
58. Niwa M, Johnson JS, O'Connor KN & Sutter ML Activity Related to Perceptual Judgment and Action in Primary Auditory Cortex. *J. Neurosci* 32, 3193–3210 (2012). [PubMed: 22378891]

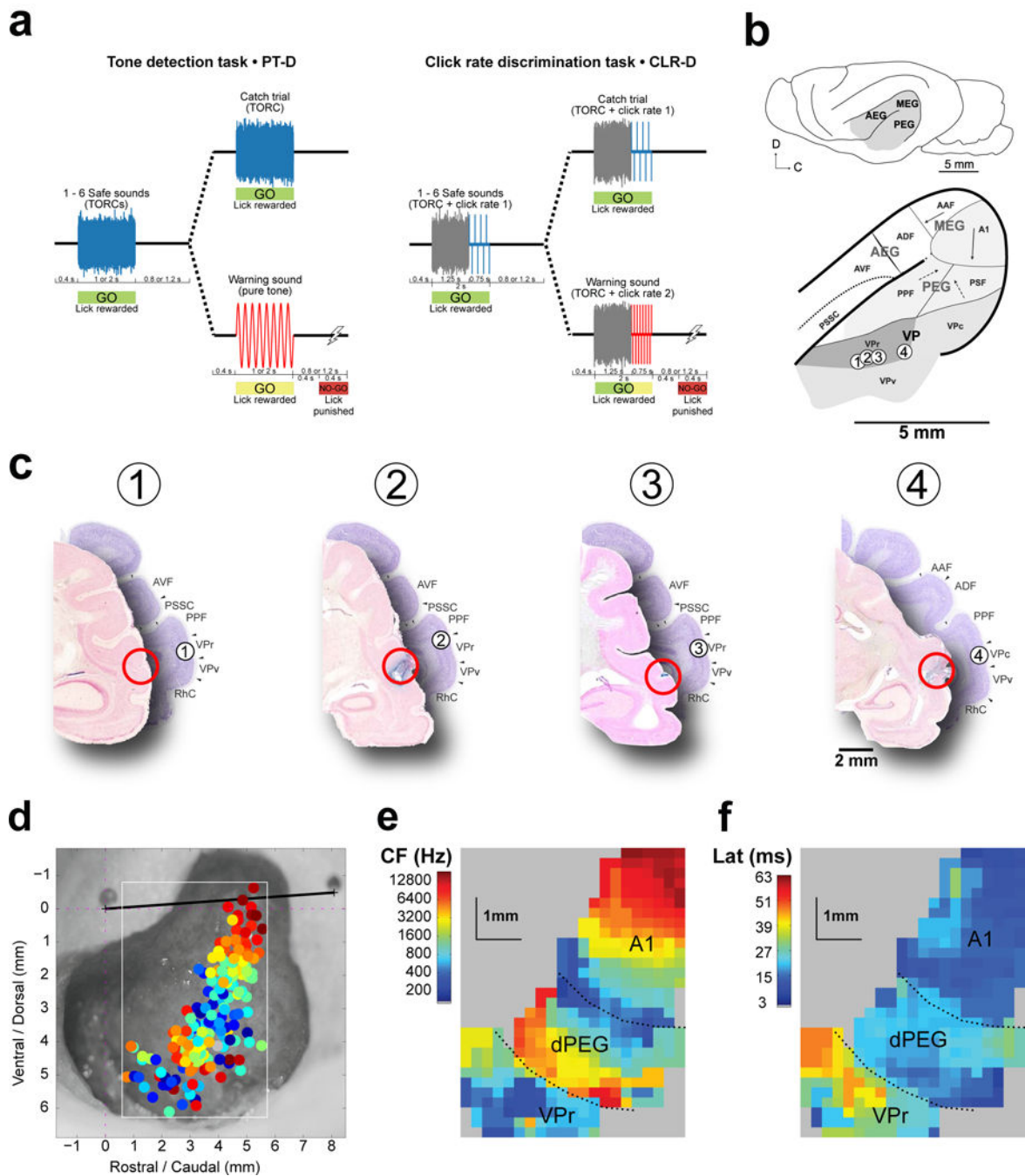


Figure 1. Behavioral task design and location of neurophysiological recordings.

(a) Pure tone detection (PT-D) and click rate (CLR-D) discrimination tasks. Both tasks used a conditioned avoidance paradigm, in which animals were trained to freely lick water from a spout during presentation of Safe sounds, and to refrain from licking for a time window 400–800 ms after offset of a Warning sound in order to avoid a mild tail shock. In the PT-D task, Safe sounds were a class of 30 similar broadband noise-like stimuli (TORCs) and Warning sounds were pure tones. In the CLR-D task, both Safe and Warning sounds were composed of 1.25 s long TORCs followed by a 0.75 s click trains of differing rates, and

animals were trained to discriminate between Safe and Warning click trains of different rates. For both tasks, on a given trial, a random number of Safe sounds (1–6) were followed by a Warning sound. In catch trials, there were no Warning sounds. In each behavioral session (comprised of ~40 trials) the Warning tone frequency (for PT-D), as well as the Safe and Warning click rates (for CLR-D) were varied and chosen after initial characterization of neuronal tuning. **(b)** Location of fields in the ferret auditory cortex. Primary areas A1 and AAF are located in the Medial Ectosylvian Gyrus (MEG) and display a clear tonotopic gradient, shown as a solid arrow. Two secondary areas in the dorsal Posterior Ectosylvian Gyrus (dPEG), Posterior Pseudosylvian Field (PPF) and Posterior Suprasylvian Field (PSF), are lateral-ventrally adjacent to A1 and display coarser and more variable tonotopic gradients (dashed arrow). Tertiary area VP, subdivided into rostral, caudal and ventral fields (VPr, VPc and VPv, respectively), displays broad spectral tuning and no apparent tonotopy. Numbers indicate the location of neuroanatomical markers placed in the vicinity of recording locations in four mapped hemispheres. **(c)** Coronal sections in four hemispheres show recording locations in VPr and their corresponding atlas sections location (Atlas²¹ section positions in mm relative to the occipital crest marking the caudal end of the ferret skull, (1) –18 mm, (2) –17.7 mm, (3) –17.7 mm, (4) –16.5 mm). Projected locations of marks from the map in **b** are depicted with circled numbers **(d)** Characteristic frequencies (CFs) recorded in one ferret, where each dot color corresponds to the mean CF across all neurons recorded in one electrode penetration. These CF measurements were used to generate a map **(e)**, which displays three functionally distinct areas corresponding to A1, dPEG area PPF and VPr. **(f)** Increasing response latencies in the three cortical maps, measured from the same recordings, also suggest three distinct stages in cortical processing from MEG to PEG. (*): TORC/tone durations were either 1 or 2 s. (**): Inter-stimulus silences were either 0.8 or 1.2 s.

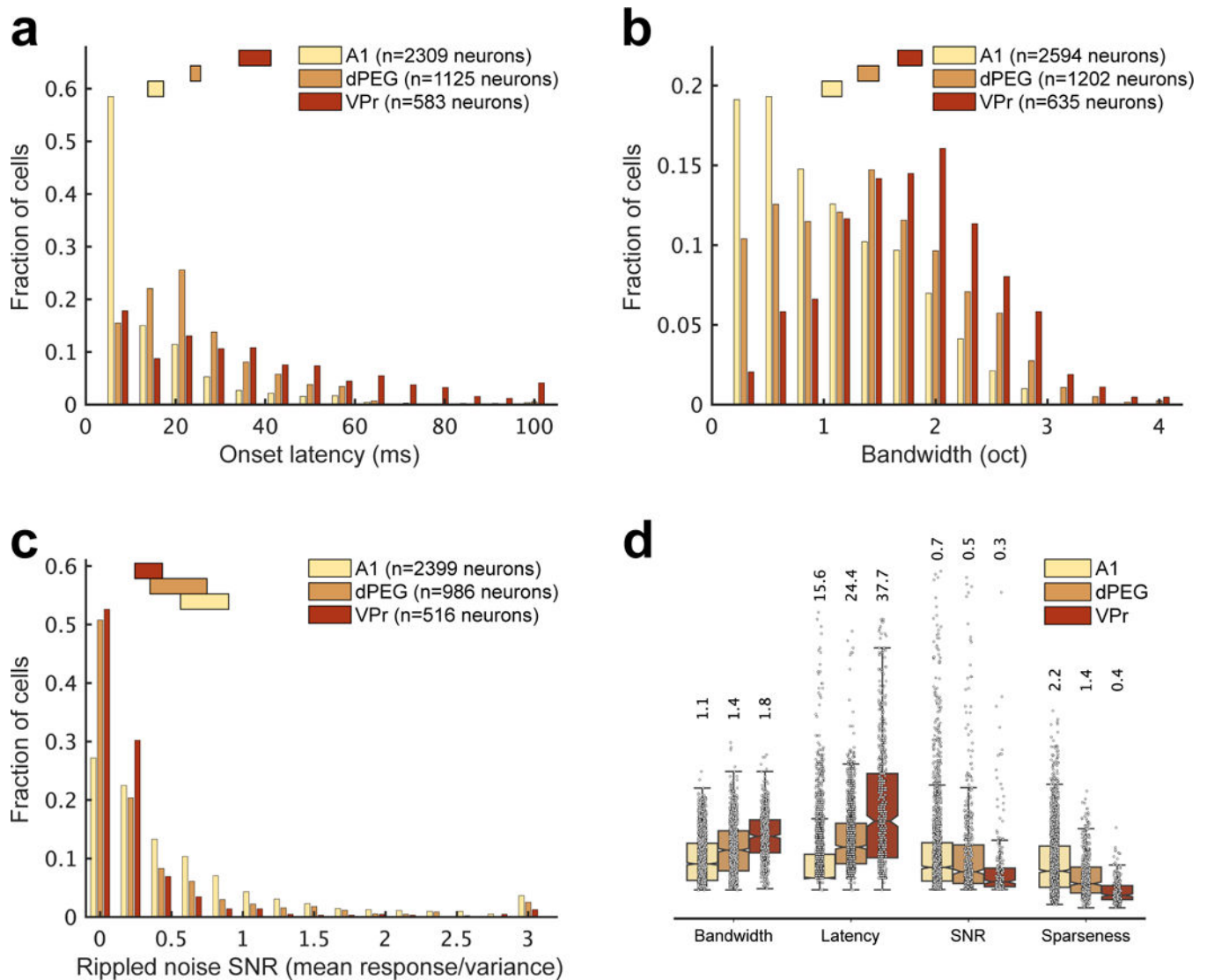


Figure 2. Comparison of tuning properties of areas A1, dPEG and VPr.

Tuning parameters were measured from responses to tones and TORCs and their resultant PSTHs and STRFs during passive listening. Bars at the top of histograms in **a**, **b** and **c** indicate mean \pm SEM in each area (A1: light tan, dPEG: orange, VPr: red). **(a)** Response latency measured from the PSTH response to random tones, measured as the earliest time bin after tone onset significantly modulated from baseline spontaneous activity ($p < 0.05$, two-sided jack-knifed t-test; mean \pm SEM A1: 15.57 ± 0.88 , dPEG: 24.45 ± 0.57 , VPr: 37.71 ± 1.78). All neurons with measurable auditory responses and significant latencies were included (A1: $n=2309/2740$ neurons, dPEG: $n=1125/1337$ neurons, VPr: $n=635/658$ neurons, $p < 0.05$ two-sided jack-knifed t-test). **(b)** Bandwidth of frequency tuning, measured as octaves at half-height of tuning curves constructed by fitting a Gaussian function to average tone responses. Mean \pm SEM A1: 1.07 ± 0.04 , dPEG: 1.4 ± 0.05 , VPr: 1.77 ± 0.05 . All neurons with measurable auditory responses and significant bandwidth values were included (A1: $n=2597/2740$ neurons, dPEG: $1202/1337$ neurons, VPr: $635/658$ neurons, $p < 0.05$ two-sided jack-knifed t-test). **(c)** Signal-to-noise ratio (SNR) of neural responses to TORC

stimuli measured as trial-to-trial phase locking to TORC sounds Mean±SEM A1: 0.73 ± 0.08 , dPEG: 0.55 ± 0.1 , VPr: 0.34 ± 0.05 . All neurons with SNR values > 0 were measured (A1: $n=2399/2740$ neurons, dPEG: $n=986/1337$ neurons, VPr: $n=516/658$ neurons). **(d)** Summary of mean (\pm SEM) tuning parameters measured in A1, dPEG and VPr. Sparseness: Mean STRF sparseness index, measured as the ratio between peak and mean magnitudes measured from STRF estimates (Mean±SEM A1: 2.24 ± 0.13 , dPEG: 1.35 ± 0.09 , VPr: 0.44 ± 0.09). For this measure, only neurons with phase locking (SNR > 0.2 , as measured in **c**) were considered (A1: $n=1664/2740$ neurons, dPEG: $n=472/1337$ neurons, VPr: $n=180/658$ neurons; see **Methods**).

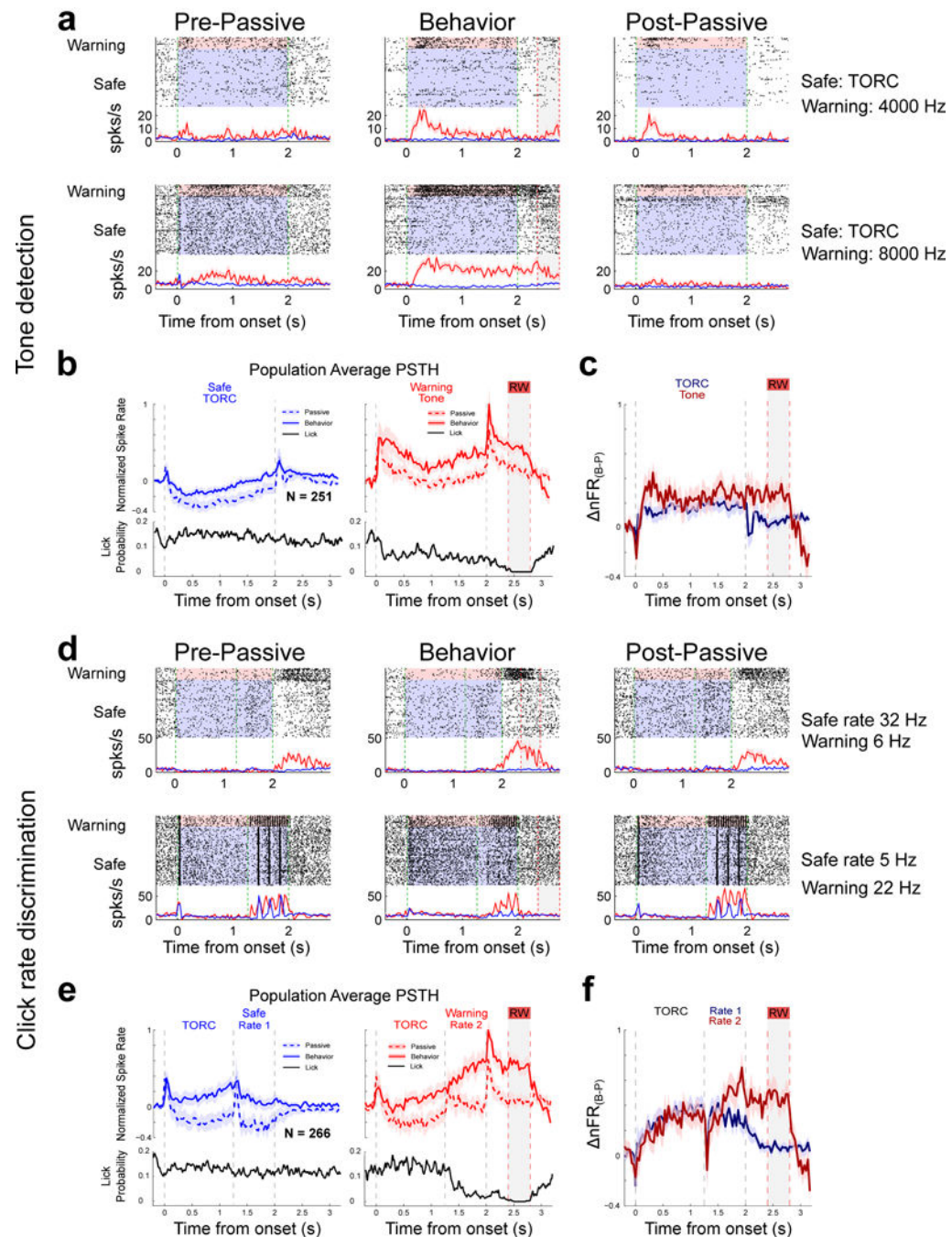


Figure 3. VPr neurons enhance the contrast between responses to Safe and Warning stimuli during behavior.

(a) Raster plots and PSTH (mean firing rate \pm 1 SEM) responses of two VPr neurons to classes of task stimuli before (pre-passive), during (behavior) and after (post-passive) performance of the PT-D task. Shaded areas indicate the duration of Safe (blue) and Warning (red) sounds. Green dashed lines indicate sound onset and offset, Gray shaded areas and red dashed lines indicate the duration of the behavioral response time window. **Left panels (Pre-passive):** Top single unit had a small onset response with a relatively long latency (175 ms).

A completely different response type is seen in the lower unit that has a small onset response but then builds up, reaching a maximum response after 0.5s, and then gradually decays as the tone continues. The responses to TORCs (class of 30 different TORCs) are very weak, and become gradually more suppressed over time. **Middle panels (Behavior):** During behavior, unit responses change dramatically, becoming substantially enhanced especially for Warning stimuli, relative to the much weaker Safe TORC responses. **Right panels (Post-passive):** During the post-passive period, the changes subside across variable time-scales. For the top-unit, Warning responses remain somewhat enhanced compared to the pre-passive state, whereas for the lower unit behaviorally-induced enhanced responses vanish rapidly. **(b)** Population average PSTH (mean firing rate \pm 1 SEM) from all VPr units significantly modulated during **PT-D** behavior (n=251 neurons). Gray dashed lines show sound onset and offset. Gray shaded areas and red dashed lines indicate the duration of the behavioral response time window (RW). **Left panel:** Responses to the Safe TORCs are largely suppressed relative to baseline spontaneous activity during the pre-passive period (blue dashed curve), becoming less so during behavior (blue solid curve). The black curve shows constant lick probability for the Safe sounds. **Right panel:** Excitatory Warning responses in the passive condition (dashed line) become substantially enhanced during behavior (solid line). Note mirror image inverse correlation of population PSTH with the lick probability curve (black) **(c)** Difference in normalized firing rates ($nFR_{(B-P)}$) between active and passive conditions for Safe TORC stimuli (blue curve) and for Warning Tone stimuli (red curve) calculated from data in **(b)** (n=251 neurons). Gray dashed lines show sound onset and offset. Gray shaded areas and red dashed lines indicate the duration of the behavioral response time window (RW). **(d)** Example raster plots and PSTHs of two units before, during and after performance of the **CLR-D** task. Shaded areas indicate the duration of Safe (blue) and Warning (red) sounds. Green dashed lines indicate sound onset and offset, Gray shaded areas and red dashed lines indicate the duration of the behavioral response time window. Cells exhibit little change in responses to the 30 “neutral” TORC stimuli (0–1.25s) between behavioral epochs. Despite the completely different acoustic properties of the click-train versus tone stimuli, behavior-induced response changes are similar to those observed in the **PT-D** task. Thus, Warning responses in the pre-passive (red curves and rasters in **left panels**) become significantly enhanced relative to the Safe responses during behavior (**middle panels**), with varied persistence in the post-passive period (**right panels**). **(e)** Population average PSTH calculated from all VPr units significantly modulated during behavior in the **CLR-D** task (n=266 neurons). Gray dashed lines show sound duration. Gray shaded areas and red dashed lines indicate the duration of the behavioral response window (RW). The large, sustained Warning enhancement contrasts with smaller changes in the Safe responses. The smaller changes in responses to the neutral TORC stimuli (0–1.25s) are identical regardless of their attachment to Safe or Warning stimuli. Note that the population averages for both Safe and Warning click trains includes the full range of different click rates used in the CLR-D task (see Supplementary Figure S-2). Note that as above **(3b)** in the PT-D task, there is no change in licking rate (black curve) during the Safe sound or during the behaviorally neutral TORC component of the Warning sound. However, as soon as the Warning click-train is presented, there is an abrupt decrease in lick rate paralleled by a sharp increase in neural firing rate during active behavior (solid blue curve). Neuronal activity remains high, and lick rate stays low throughout the 800 ms post-stimulus period. The two

mirror image curves come back together after the shock period. **(f)** Difference in normalized firing rates ($nFR_{(B-P)}$) between active and passive conditions for Safe Click Rate 1 stimuli (blue curve) and for Warning Click Rate 2 stimuli (red curve) for population shown in (e) ($n=266$ neurons) Gray dashed lines show sound duration. Gray shaded areas and red dashed lines indicate the duration of the behavioral response window (RW).

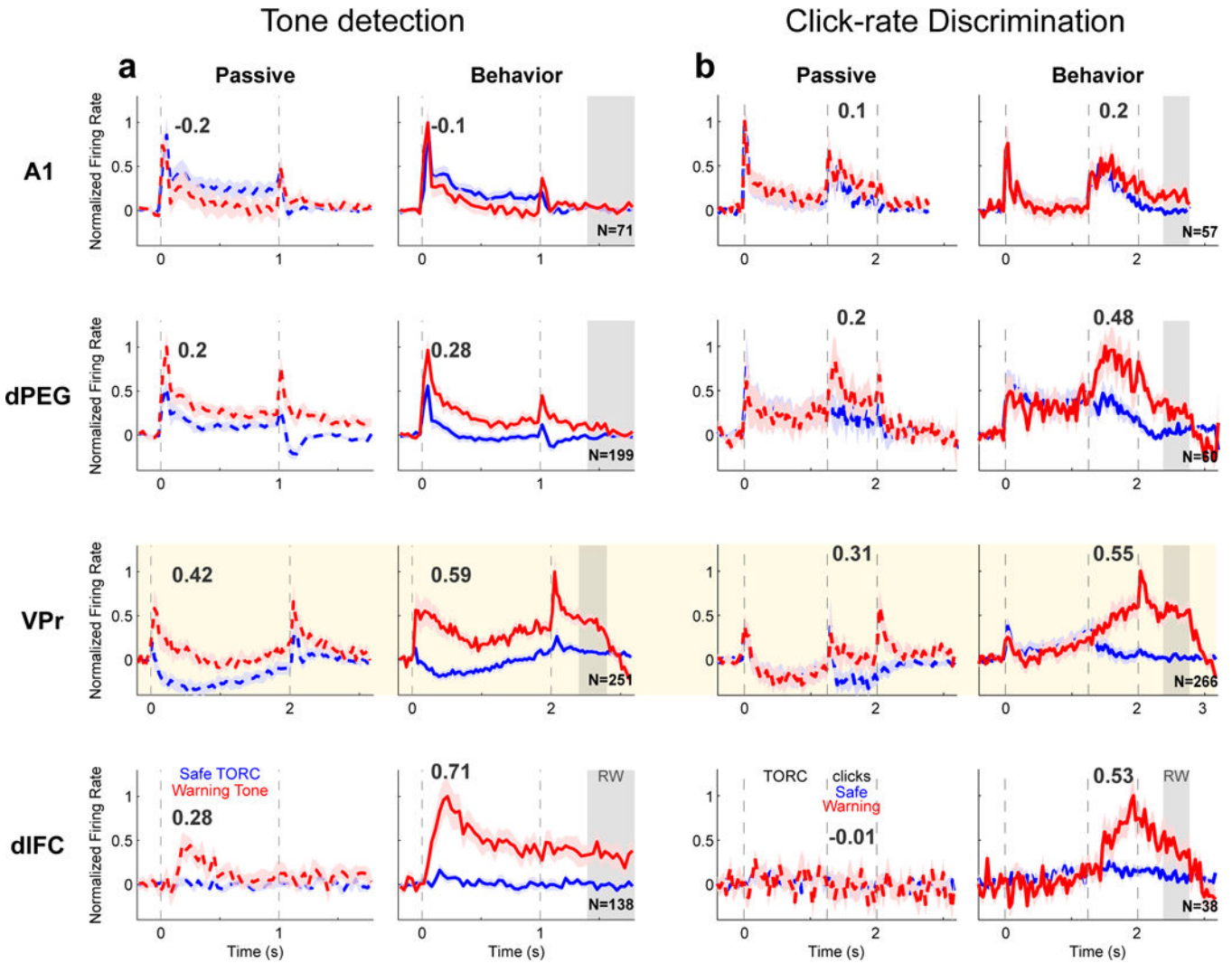


Figure 4.

Average responses to Warning and Safe Stimuli from areas A1, dPEG, VPr and dlFC, during PT-D (a) and CLR-D (b) tasks. In each row, mean \pm SEM time-varying responses to Warning stimulus (red) and Safe stimulus (blue) are compared in pre-passive (left column, dashed curves) and behavior conditions (right column, solid curves). Neural responses were normalized to have the same maximum across behavior conditions before averaging. All cells that exhibited modulated responses during behavior compared to pre-passive were included in the averages. Vertical dashed lines indicate sound onset and offset. Gray shaded area indicates the 400 ms behavioral Response Window (RW) during which animals could receive a shock if they continued licking following Warning sound offset. Cream shaded area indicates data obtained from VPr. Numbers above curves indicate the contrast between Safe and Warning sound responses ($\Delta nFR_{(W-S)}$, see **Methods** and Figure 6), measured as the mean difference in normalized firing rate in response to sounds in a time window 0.1 – 0.45 s from onset (PT-D task) or 0.3 – 1 s after TORC offset (CLR-D task). PT-D: A1 n=71, dPEG n=199, VPr n=251, dlFC n=138 neurons; CLR-D: A1 n=57, dPEG n=60, VPr n=266, dlFC n=38 neurons.

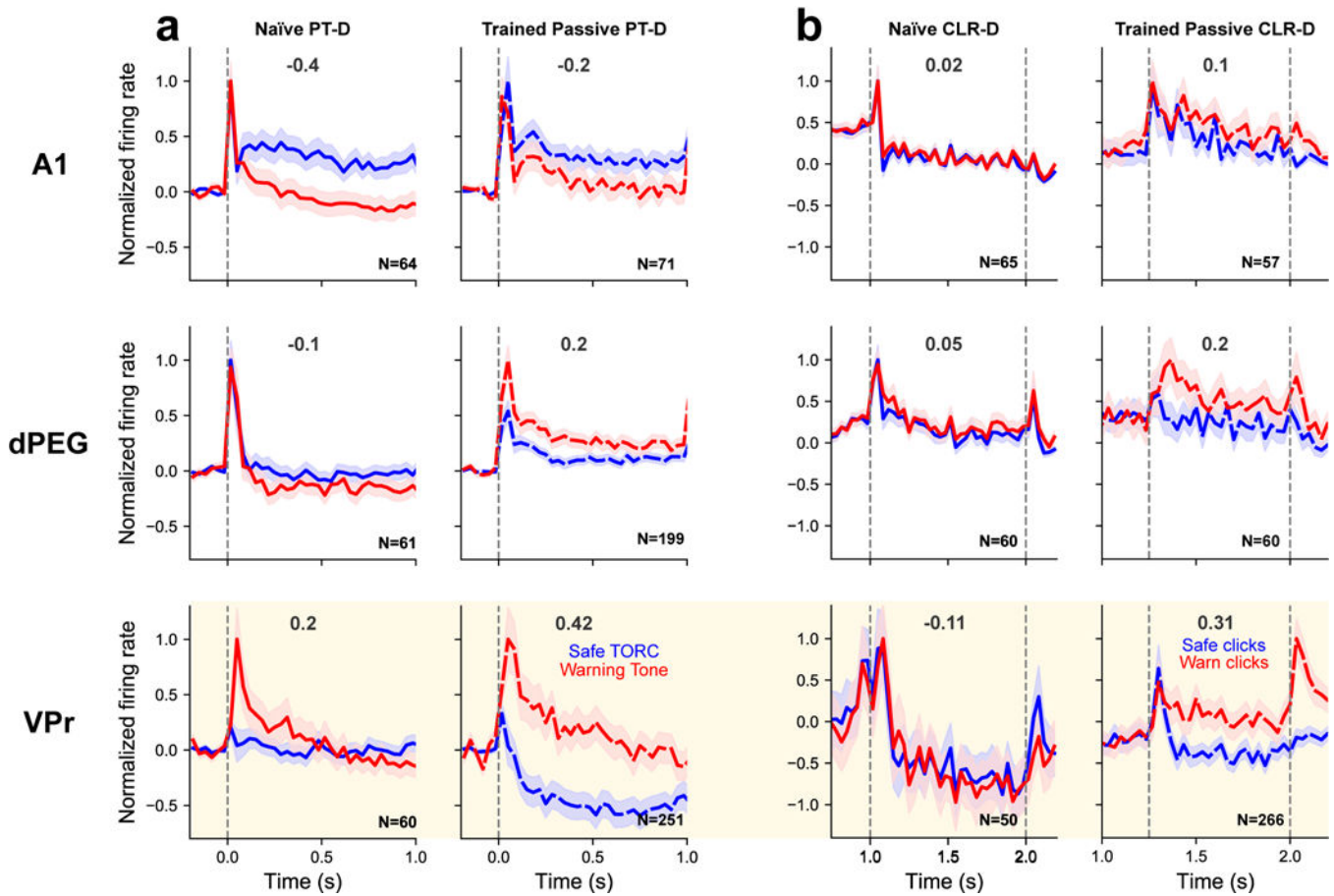


Figure 5. Comparison of contrast between Safe vs Warning sounds in naïve and trained animals in three different auditory cortical areas.

(a) Mean \pm SEM normalized firing rates in response to TORCs (blue) and tones (red) in PT-D. Left column displays the responses to passive presentation of task-sounds to a task-naïve animal (A1 $n=64$, dPEG $n=61$, VPr $n=60$ neurons). Right column displays data acquired during presentation of the PT-D task to trained animals (dashed lines) during the passive state (A1 $n=71$, dPEG $n=199$, VPr $n=251$ neurons). Vertical gray lines indicate sound onset. Cream shaded area indicates VPr responses. (b) Responses to Safe (blue) and Warning (red) click trains recorded while passively presenting the CLR-D task-sounds to a task-naïve ferret (left; A1 $n=65$, dPEG $n=60$, VPr $n=50$ neurons) and trained ferrets (right; A1 $n=57$, dPEG $n=60$, VPr $n=266$ neurons). Even in the behaviorally quiescent listening condition in trained animals (dashed lines), VPr neurons display a greater contrast between Safe and Warning sounds than is observed in a naïve animal. This contrast is further increased during task performance (see Figure 4). Vertical gray lines indicate click train onset and offset. Numbers above curves display the mean Safe and Warning response contrast ($\frac{nFR_{(w-s)}}{nFR_{(w-s)}}$, see **Methods** and Figure 6).

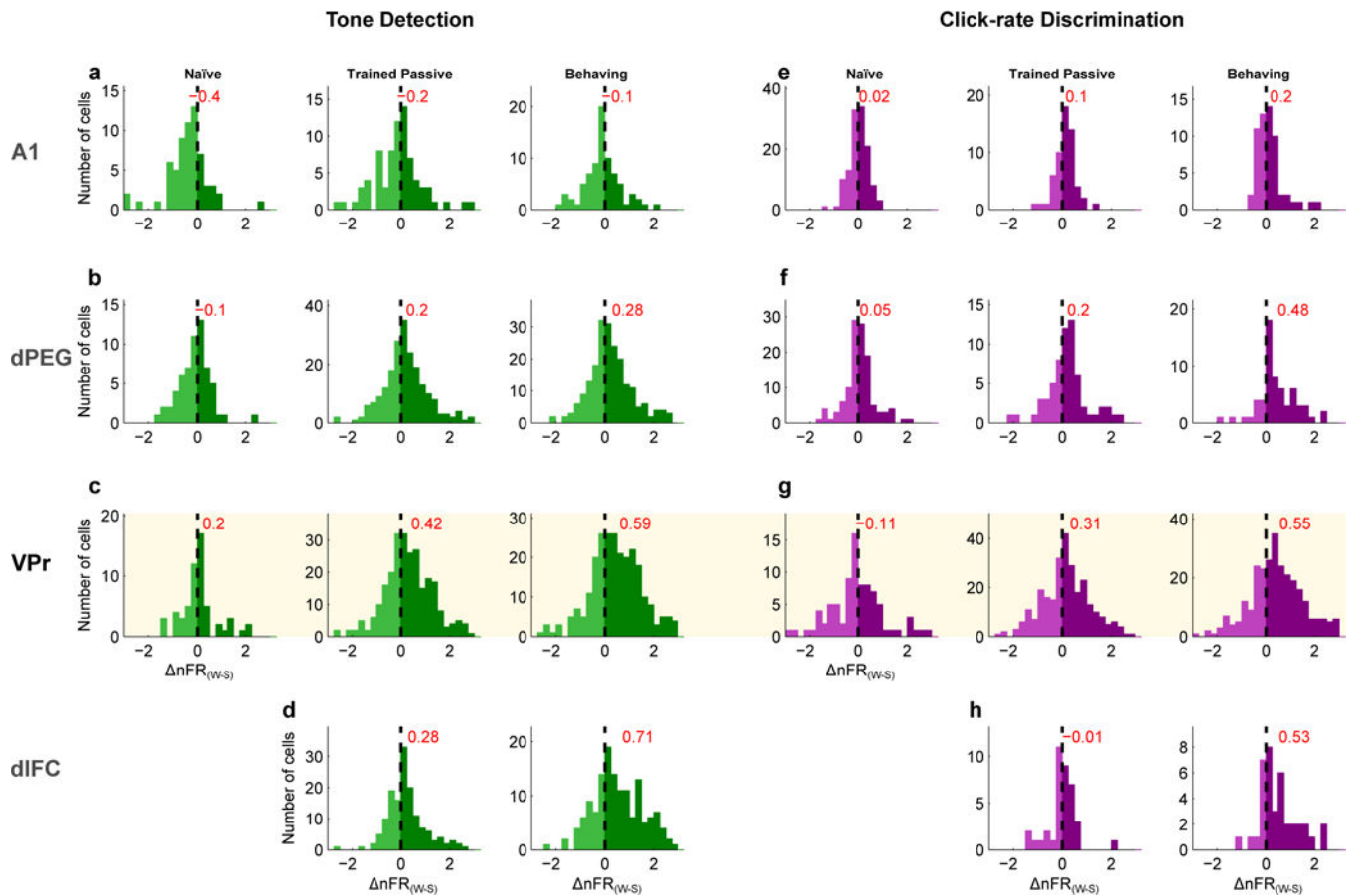


Figure 6. Distributions of Warning-Safe response contrast ($\Delta nFR_{(W-S)}$) recorded from A1, dPEG, VPr and dlFC neurons.

Results are shown for the PT-D (a-d) and CLR-D (e-h) tasks in naïve and trained animals (in both passive and active conditions). $\Delta nFR_{(W-S)}$ was computed as the difference of the mean firing rates (normalized to population maximum) of Warning and Safe sound responses 0.1 – 0.45 seconds after sound onset (PT-D) or 0.3 – 1.0 seconds after TORC offset/click onset (CLR-D). The histograms are arranged in three columns for each task, showing contrast in naïve (left column) and trained animals in passive listening (middle column) and during active behavior (right column). Cream shaded area indicates data obtained from VPr. Histograms are mostly symmetric in the naïve animal in all cortical regions recorded. However, in trained animals, a slight asymmetry towards $\Delta nFR_{(W-S)}$ contrast enhancement shows up in the passive state (middle columns), which is further shifted during behavior (right columns). The distributions become progressively more asymmetric in higher cortical areas. Red values display the mean $\Delta nFR_{(W-S)}$. Naïve PT-D: A1 n=64, dPEG n=61, VPr n=60 neurons; Trained PT-D: A1 n=71, dPEG n=199, VPr n=251, dlFC n=138 neurons. Naïve CLR-D: A1 n=65, dPEG n=60, VPr n=50 neurons; Trained CLR-D: A1 n=57, dPEG n=60, VPr n=266, dlFC n=38 neurons.

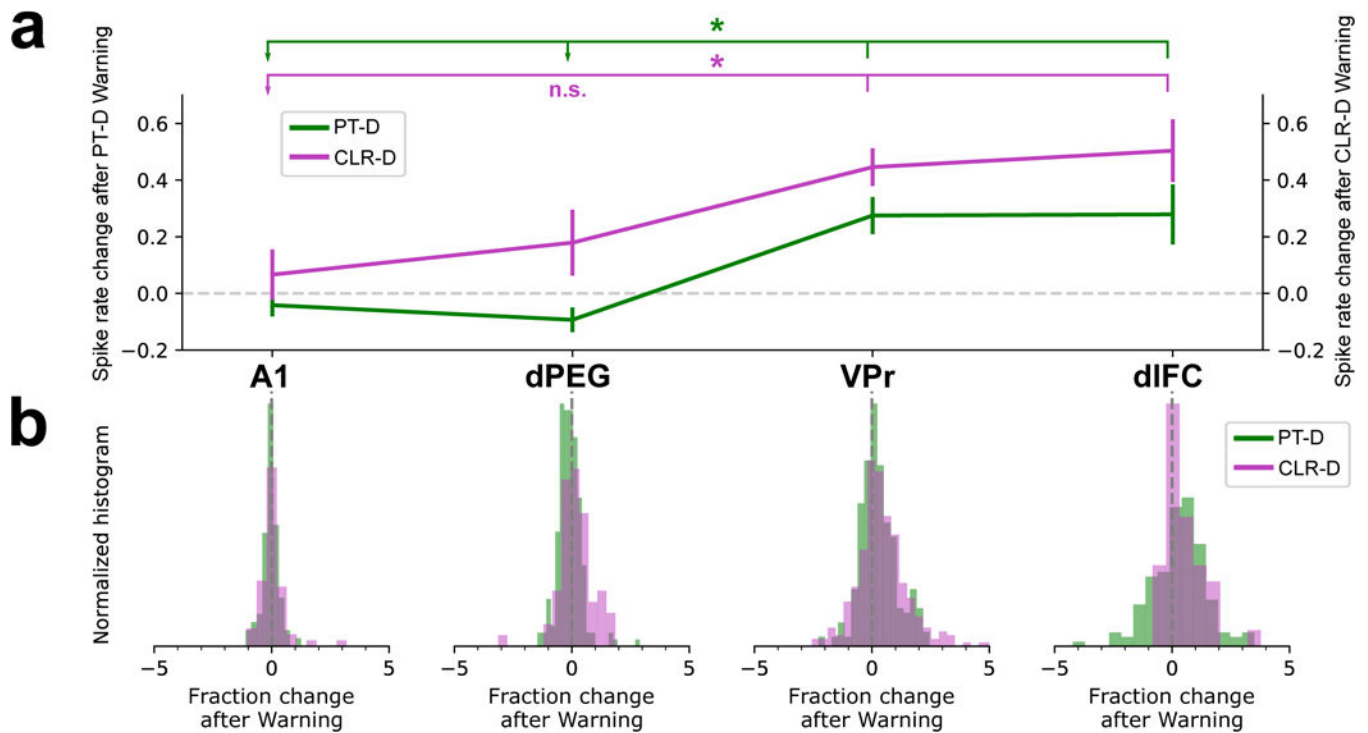


Figure 7. VPr and dIFC neurons display a substantial and sustained response during the silent period following Warning sound presentation during task performance.

This sustained response occurs during a silent period of 800 ms after Warning sound offset and lasts until the end of the 400 ms shock time window (see Figures 1, 3, 4). We measured this response as the change in normalized mean firing rate from the passive to the behaving state in a time window between 50–700 ms after Warning sound offset. **(a)** Behavior-dependent change in normalized after-Warning responses in A1 (PT-D $n=71$, CLR-D $n=57$ neurons), dPEG (PT-D $n=199$, CLR-D $n=60$ neurons), VPr (PT-D $n=251$, CLR-D $n=266$ neurons) and dIFC (PT-D $n=138$, CLR-D $n=38$ neurons) in PT-D (green) and CLR-D (purple) tasks (mean \pm SEM). In both tasks, VPr and dIFC display a significant increase in their responses after Warning sound offset during this silent 650 ms time window. (PT-D: chi-square=13.4, $p=0.0052$, $df=3$; CLR-D: chi-square=40.947, $p=6.7 \times 10^{-9}$, $df=3$; Kruskal-Wallis test). Asterisks and lines above show Tukey's HSD post-hoc pair-wise difference significance between higher-order areas VPr and dIFC with A1 and dPEG (PT-D: A1/VPr $p=0.0072$, dPEG/VPr $p=0$, A1/dIFC $p=0.008$, dPEG/dIFC $p=0$; CLR-D: A1/VPr $p=0.0072$, dPEG/VPr $p=0.3336$, A1/dIFC $p=0.0433$, dPEG/dIFC $p=0.4292$). **(b)** Normalized distributions of the behavior-dependent change in response after Warning sounds in A1, dPEG, VPr and dIFC.

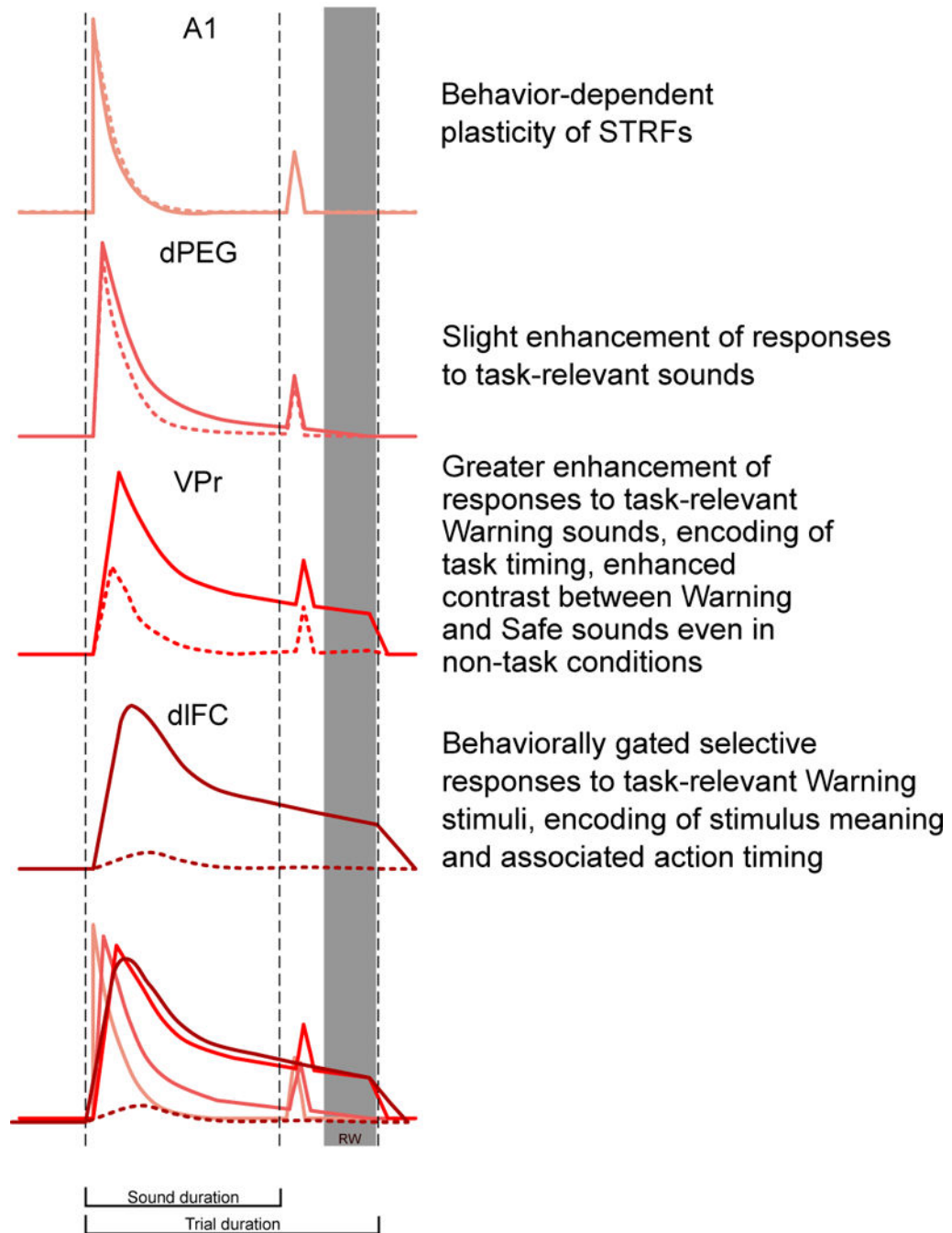


Figure 8. Summary of progression of task-related population responses to the Warning stimulus along the auditory processing hierarchy in the PT-D task.

The overall population-averaged passive response to sound in A1 is slightly suppressed during behavior, particularly for Safe stimuli, though less so for Warning stimuli. However, in contrast, the responses to Warning stimuli are somewhat enhanced in dPEG during behavior, and responses to Warning sounds are even more greatly enhanced in VPr during the active attentive, behavioral state. In the PT-D task, higher-order auditory cortex (VPr) shares common response properties to Warning stimuli with dIFC, suggesting the emergence

of coding for non-acoustical task features, such as timing and sound meaning, may originate in higher auditory cortex.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript