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

Diego Elgueda1,2, Daniel Duque1,3, Susanne Radtke-Schuller1, Pingbo Yin1, Stephen V. David3,4, Shihab A. Shamma1,2,5 and Jonathan B. Fritz1,2*

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 auditory field (VPPr), a tertiary area in the 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. VPPr 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 VPPr 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 the dynamics of short-term attention and long-term learning, as incoming sound is selectively attended, recognized, and translated into action.

To understand the meaning of sounds, we learn to associate their acoustic features with their behavioral context and link them to appropriate audiomotor 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-making8–10. However, with a few notable exceptions10–13, most neurophysiological studies of the auditory cortex in behaving animals have focused on the primary auditory cortex (A1) rather than higher-order auditory cortical areas.

To investigate the contributions of nonprimary 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 the primary auditory cortex (A1) and in tonotopically organized secondary or ‘belt’ areas in the ferret auditory cortex (posterior pseudosylvian field (PPF) and posterior suprasylvian field (PSF) in Fig. 1b). The neural representation of sound can be partially transformed in these areas to incorporate behavioral and contextual information11–13,15–17. We have also characterized a task-dependent, gated representation of behaviorally salient sounds in the non-tonotopic dorsolateral frontal cortex (dlFC)14.

Based on this earlier work, we conjectured that (1) there are tertiary auditory cortical areas between secondary areas and the frontal cortex where the transformation from sound representation to behavioral meaning is more extensively developed than in lower cortical areas, (2) 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 (3) display strong attention effects that would amplify long-term changes in the representation of task-relevant stimuli during task performance, (4) would show response timing linking auditory inputs to reward and motor responses.

Previous studies have shown that the ferret auditory cortex is composed of multiple acoustically sensitive adjoining areas in the ectosylvian gyrus of the temporal lobe18–21. Current maps of ferret auditory cortex include nine distinct cortical areas, six of which (A1 anterior auditory field (AAF), PPF, PSF, anterior dorsal field (ADF), anterior ventral field (AVF)) have been physiologically identified and described previously22. One field whose function has not been studied previously, the ventral posterior (VP) area, lies in a ventral region in the posterior ectosylvian gyrus (PEG), and its anatomical connectivity makes it a good candidate for a tertiary auditory field22–23.

To test these hypotheses concerning sound encoding in the tertiary auditory cortex, we recorded responses under multiple active task and non-task (passively listening) conditions in the rostral region of VP (VPPr in Fig. 1a,b)11,12,20,21. 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. In this study, we describe how VPPr neurons exhibit striking state- and context-dependent changes in auditory responses and encode non-acoustical sound features, such as associated behavioral...
**Articles**

**NaTure NeuroscIENce**

| 0 | 7 | 6 | 2 | 0.8 or 1.2 s |
|---|---|---|---|-------------|

**Fig. 1** | Behavioral task design and location of neurophysiological recordings. **a**, PT-D and CLR-D tasks. Both tasks used a conditioned avoidance framework, in which animals were trained to freely lick water from a spout during the presentation of safe sounds and to refrain from licking for a time window of 400–800 ms after the offset of a warning sound to avoid a mild tail shock. In the PT-D task, safe sounds were a class of 30 similar TORCs and warning sounds were pure tones. In the CLR-D task, both safe and warning sounds were composed of 1.25-s TORCs followed by 0.75-s click-trains of differing rates; 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 frequencies (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. TORC and tone durations were either 1 or 2 s. Interstimulus silences were either 0.8 or 1.2 s. **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 by a solid arrow. Two secondary areas in the dorsal posterior ectosylvian gyrus PEG (posterior pseudosylvian field, or PPF, and posterior suprasylvian field, or PSF) are lateroventrally 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. The 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 section locations (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). AEG, anterior ectosylvian gyrus; PSSC, pseudosylvian sulcus cortex; RHC, rhinal cortex. The 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. **e**, These CF measurements were used to generate a map, which displays three functionally distinct areas corresponding to A1, dorsal PEG area PPF, and VPv. **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.
meaning and task timing. These results are consistent with all four of the conjectures above.

Results

Neurophysiological mapping and neuroanatomical location. We mapped the basic tuning properties of VPr using single-unit activity in six animals during passive presentation of pure tone, click-train, and broadband rippled noise stimuli (see Methods).

We marked the location of VPr recordings and confirmed that they were ventral and anterior to area PPF (Fig. 1b,c; sites are 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 μm of the first spikes recorded as electrodes entered the brain (Supplementary Fig. 1). The location of each recording site was also registered with a ferret brain atlas21 for STRF measurements in some neurons (in VPr, only 27.3% of each recording site was also registered with a ferret brain atlas21 for STRF measurements in some neurons (in VPr, only 27.3% of each recording site was also registered with a ferret brain atlas21 for STRF measurements in some neurons (in VPr, only 27.3% of each recording site was also registered with a ferret brain atlas21 for STRF measurements in some neurons (in VPr, only 27.3%

Response properties in VPr. The basic auditory tuning properties of single units in VPr are contrasted with previously collected responses from A1 and dorsal PEG (Fig. 2). The distribution of tuning properties is consistent with neuroanatomical evidence that VPr is a later processing stage in the auditory pathway, after A1 and dorsal PEG20. Compared to earlier areas, VPr neurons display longer mean latency (Fig. 2a; VPr: n = 583 neurons, 37.1 ± 1.78 ms; dorsal PEG: n = 1,125 neurons, 24.45 ± 0.57 ms; A1: n = 2,309 neurons, 15.57 ± 0.88 ms). We found significant differences in tone response latency (χ2 = 862.83, P < 0.0001, d.f. = 2, Kruskal–Wallis test), where VPr significantly differed from A1 (Tukey’s honestly significant difference (HSD) effect size (meanA1 − meanVPr) = −1.265.8, 95% confidence interval (CI) = (−1,388.9, −1,142.7), P = 0.00) and dorsal PEG (effect size (meandorsalPEG − meanVPr) = −341.8, 95% CI = (−477.3, −206.3), P = 0.00).

Neurons in VPr also have broader mean frequency tuning bandwidth (Fig. 2b; VPr: n = 635 neurons, 1.77 ± 0.05 octaves (oct); dorsal PEG: n = 1,202 neurons, 1.4 ± 0.05 oct; A1: n = 2,594 neurons, 1.07 ± 0.04 oct; χ2 = 499.16, P < 0.001, d.f. = 2, Kruskal–Wallis test). Mean VPr bandwidth was significantly greater than A1 (effect size = −1,190.9, 95% CI = (−1,323.6, −1,058.2), P = 0.00) and dorsal PEG (effect size = −630.7, 95% CI = (−777.8, −483.7), P = 0.00).

VPr neurons also display weaker overall following of complex synthetic sounds (Fig. 2c; mean signal-to-noise ratio (SNR); VPr: n = 516 neurons, 0.34 ± 0.05; dorsal PEG: n = 986 neurons, 0.55 ± 0.1; A1: n = 2,399 neurons, 0.73 ± 0.08; χ2 = 291.06, P < 0.001, d.f. = 2, Kruskal–Wallis test). The SNR of the VPr responses was significantly lower than A1 (effect size = 816.24, 95% CI = (688.15, 944.33), P = 0.00) and dorsal PEG neurons (effect size = 329.55, 95% CI = (186.13, 4,724.97), P = 0.00).

For neurons whose responses followed the stimulus, spectrotemporal receptive fields (STRFs) were more complex, as indicated by their sparseness index—the peak STRF magnitude divided by the s.d. across the STRF bins (Fig. 2d; VPr: n = 180 neurons, 0.44 ± 0.09; dorsal PEG: n = 472 neurons, 1.35 ± 0.09; A1: n = 1,664 neurons, 2.24 ± 0.13; χ2 = 291.06, P < 0.001, d.f. = 2). VPr showed lower STRF sparseness than A1 (effect size = 777.47, 95 CI = (655.48, 899.45), P = 0.00) and dorsal PEG (effect size = 466.29, 95% CI = (330.18, 602.4), P = 0.00).

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 dorsal PEG (472/1,337) and 60% in A1 (1,644/2,740)). However, VPr is also similar to dFC in its relatively weak auditory responsiveness during passive sound presentation14, its often poorly defined tuning, and long response latencies (examples in Fig. 3a,c and Supplementary Figs. 2, 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 negatively rewarded (NO-GO) 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 before the recordings43. The tasks were: (1) tone versus noise discrimination task (‘pure tone detection’, PT-D); and (2) click-rate discrimination (CLR-D) task (Fig. 1a).

In both tasks, the animals listened to a sequence of reference ‘safe’ sounds (broadband rippled noise—temporally orthogonal ripple combinations (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 safe rates. 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 Fig. 3a, d. In the majority of units, engagement in behavior rapidly induced a substantial change in peristimulus time histogram (PSTH) responses to warning stimuli, and a lesser change for safe stimuli (Fig. 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 Fig. 3). Details of changes varied greatly from cell to cell, reflecting the specific type of response (for example, onset, sustained, or offset). Nevertheless, the patterns of responses to warning and safe stimuli in the population average (Fig. 3b,c) 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 Fig. 4). Thus, on average, there was a large enhancement in the responses to the class of NO-GO warning stimuli (that is, 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 (ΔFIRSNR) between behaving and passive conditions (Fig. 3c,f). This differential change increased the contrast between safe and 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 (Fig. 3a, lower panel) or even gate VPr neuronal responses so that they only occurred in the active state (Supplementary Fig. 3).

The relationship of VPr responses to behavior is illustrated (Fig. 3b,e), 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 poststimulus, and until the end of the shock window (shaded area). A comparison with the population neural PSTH.
Δ measured stimulus contrast as the difference between warning and the PT-D averages to provide a larger sample (see Methods). We TORC offset/click-train onset) for both passive and behaving conditions (PSTHs in Fig. 4a,b, contrasting distributions in Fig. 6).

To gain a broader view of VPr in the broader cortical network, we compared population PSTH responses in A1, dorsal PEG, and dlFC during pre-passive and behavior epochs for both PT-D and CLR-D tasks (Fig. 4). PT-D data from the A1, dorsal PEG, and dlFC showed similar enhanced target responses in both PT-D and CLR-D tasks (Supplementary Fig. 5).

There are several notable findings in the comparison of population average responses across areas and tasks. First, the overall pattern of enhanced contrast (ΔnFR\text{\(W-S\)}) 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 the 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 NO-GO) rather than their acoustic properties in determining the nature of VPr responses. At a single-cell level, many VPr neurons (150/367) showed similar enhanced target responses in both PT-D and CLR-D tasks (Supplementary Fig. 5).

Second, contrast enhancement between warning and safe (ΔnFR\text{\(W-S\)}) gradually increased across areas during behavior compared to the pre-passive state (left versus right columns in Fig. 4a,b). We interpret this to indicate a progressively larger weight given to for the two tasks illustrates that the timing of the increased neural responses to the warning stimuli parallels decreases in the behavioral lick response.

**Fig. 2** | Comparison of the tuning properties of the A1, dorsal PEG, and VPr areas. Tuning parameters were measured from the responses to tones and TORCs and their resultant PSTHs and STRFs during passive listening. The bars at the top of the histograms in a–c indicate mean ± s.e.m. in each area (A1, light tan; dorsal PEG, 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 ± s.e.m. A1: 15.57 ± 0.88; dorsal PEG: 24.45 ± 0.57; VPr: 37.71 ± 1.78). All neurons with measurable auditory responses and significant latencies were included (A1: n = 2,309/2,740 neurons; dorsal PEG: n = 1,125/1,337 neurons; VPr: 635/658 neurons; P < 0.05, two-sided jack-knifed t test). b, Bandwidth of frequency tuning, measured as octaves at the half-height of tuning curves constructed by fitting a Gaussian function to average tone responses (mean ± s.e.m., A1: 1.07 ± 0.04; dorsal PEG: 1.4 ± 0.05; VPr: 1.77 ± 0.05). All neurons with measurable auditory responses and significant bandwidth values were included (A1: n = 2,597/2,740 neurons; dorsal PEG: 1,202/1,337 neurons; VPr: 635/658 neurons; P < 0.05). All neurons with measurable auditory responses and significant latencies were included (A1: n = 2,399/2,740 neurons; dorsal PEG: 1,202/1,337 neurons; VPr: 635/658 neurons; P < 0.05, two-sided jack-knifed t test). c, SNR of neural responses to TORC stimuli measured as trial-to-trial phase locking to TORC sounds (mean ± s.e.m., A1: 0.73 ± 0.08; dorsal PEG: 0.55 ± 0.01; VPr: 0.34 ± 0.05). All neurons with SNR values > 0 were measured (A1: n = 2,399/2,740 neurons; dorsal PEG: n = 986/1,337 neurons; VPr: n = 516/658 neurons). d, Summary of mean (± s.e.m.) tuning parameters measured in A1, dorsal PEG, and VPr. Sparseness: mean STRF sparseness index, measured as the ratio between peak and mean magnitudes measured from STRF estimates (mean ± s.e.m., A1: 2.24 ± 0.13; dorsal PEG: 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 = 1,664/2,740 neurons; dorsal PEG: n = 472/1,337 neurons; VPr: n = 180/658 neurons; see Methods).
the behavioral distinction between NO-GO and GO stimuli in higher cortical areas. The overall change in contrast in A1 is much smaller than in dorsal PEG (for example, CLR-D task of Fig. 4b). In fact, the average warning tone response in A1 during PT-D (Fig. 4a) is actually smaller than the responses to the safe TORCs. This reversal probably 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 the enhancements described earlier in studies where the warning tone was often placed close to the best frequency to achieve maximal plasticity1,13.

We compared responses in different areas using a three-way repeated-measures analysis of variance (ANOVA) (see Methods). In the PT-D task (A1 n = 71; dorsal PEG n = 199; VPr n = 251; dlFC n = 138 neurons) the repeated-measures ANOVA for the response difference (∆nFR) yielded significant main effects for area (F = 8.91; P < 0.0001) and task condition (passive or behaving, F = 11.52; P = 0.0007). Tukey’s HSD test confirmed that response differences are smaller in A1 compared to the other areas (A1-dorsal PEG = 2.759, 95% CI = (1.201, 4.316), F < 0.0001; A1-VPr = 2.944, 95% CI = (1.418, 4.470), P < 0.0001; A1-dlFC = 2.202, 95% CI = (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, 95% CI = (0.247, 0.894), t = −3.468, P = 0.0006). The analysis also yielded a significant area versus behavior interaction (F = 2.64; P < 0.0487), suggesting that the effect of engagement on the response depends on the area. Tukey’s HSD post hoc analysis again confirmed that behavior enhances response contrast (∆nFR) in VPr and dlFC (VPr passive-behavior = 0.903, 95% CI = (0.135, 1.670), P = 0.0089; dlFC passive-behavior = 1.120, 95% CI = (0.171, 2.070), P = 0.0086), but not in A1 or dorsal PEG (A1 passive-behavior = 0.003, 95% CI = (−1.356, 1.363), P = 1; dorsal PEG passive-behavior = 0.213, 95% CI = (−0.616, 1.044), P = 0.994). Altogether, this analysis suggests that VPr neurons show contrast enhancement that more closely resembles dlFC than dorsal PEG.

In the CLR-D task (A1 n = 57; dorsal PEG n = 60; VPr n = 266; dlFC n = 38 neurons), the repeated-measures ANOVA 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, 95% CI = (0.701, 1.497), t = 5.429, P < 0.0001). The analysis also yielded a significant area versus task condition interaction (F = 2.74; P < 0.0429). Post hoc Tukey’s HSD analysis confirmed that behavior enhances the response contrast between warning and safe click-trains in VPr and dlFC (VPr passive-behavior = 0.798, 95% CI = (0.213, 1.382), P = 0.001; dlFC passive-behavior = 2.146, 95% CI = (0.551, 3.740), P = 0.0013), but not in A1 or dorsal PEG (A1 passive-behavior = 0.312, 95% CI = (−0.979, 1.603), P = 0.999; dorsal PEG passive-behavior = 1.142, 95% CI = (−0.097, 2.381), P = 0.096). These findings suggest that a better representation of click-trains in A1 and dorsal PEG—when the animal is engaged in a behavioral task—may be used to generate more highly differentiated behavioral perceptuals 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 Fig. 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 dorsal PEG (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 the auditory cortex undergoes long-term changes that reflect training on behaviorally relevant sound features. Conversely, artificially enhancing neural responses to acoustic stimuli can improve behavioral responses to those stimuli. 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 (Figs. 4 and 5). We note that because of behavioral gating, 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 the...
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.

To summarize, there is a gradual shift toward an enhanced representation of behavioral meaning of task stimuli beginning in the early cortical stages (A1 and dorsal PEG) and increasing toward the higher cortical regions where it becomes clearly manifested in dlFC. VPr is similar to the early auditory cortical areas, responding to both major and minor acoustic features, such as tone frequency and temporal dynamics. On the other hand, VPr responses 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 Fig. 3). As mentioned, these behaviorally gated responses in VPr are similar to responses previously observed in dlFC.46 About 28% of VPr neurons (127/453) showed no response to a variety of passively presented acoustic stimuli (that is, 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 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 (Fig. 4). In the dlFC, the small pre-passive responses 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/NO-GO 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 ($\Delta nFR_{\text{W-S}}$) during passive listening suggests that this is the case (Fig. 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 the higher auditory areas of task-naive animals. To test this prediction, we recorded the responses to task stimuli in A1, dorsal PEG, and VPr of a task-naive ferret (left; A1 n = 65; dorsal PEG n = 60; VPr n = 251 neurons) and trained ferrets (right; A1 n = 57; dorsal PEG 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 naive animal. This contrast is further increased during task performance (see Fig. 4). The vertical gray lines indicate click-train onset and offset. The numbers above the curves display the mean safe and warning response contrast ($\Delta nFR_{\text{W-S}}$; see Methods and Fig. 6).

**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: (1) TORCs served as ‘safe’ stimuli in the PT-D task (Fig. 1a). We note that the same sequence of stimuli in the PT-D task was played during passive listening but no task—see context (3), and in active task conditions; (2) TORCs were also behaviorally ‘neutral anticipatory’ stimuli preceding both warning and safe click-trains in the CLR-D task. In this context, TORCs carried virtually no information about the upcoming click rate, but they provided information about the onset time of the upcoming click-train (Fig. 1a). As noted earlier, the same sequence of stimuli in the CLR-D task was played during 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

---

**Fig. 5 | Comparison of contrast between safe and warning sounds in naive and trained animals in three different auditory cortical areas.** a. Mean ± s.e.m. normalized firing rates in response to TORCs (blue) and tones (red) in PT-D. The left column displays the responses to the passive presentation of task sounds to a task-naive animal (A1 n = 64; dorsal PEG n = 61; VPr n = 60 neurons). The right column displays the data acquired during presentation of the PT-D task to trained animals (dashed lines) during the passive state (A1 n = 71; dorsal PEG n = 199; VPr n = 251 neurons). The vertical gray lines indicate sound onset. The cream shaded area indicates the VPr responses. b. Responses to safe (blue) and warning (red) click-trains recorded while passively presenting the CLR-D task sounds to a task-naive ferret (left; A1 n = 65; dorsal PEG n = 60; VPr n = 50 neurons) and trained ferrets (right; A1 n = 57; dorsal PEG 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 naive animal. This contrast is further increased during task performance (see Fig. 4). The vertical gray lines indicate click-train onset and offset. The numbers above the curves display the mean safe and warning response contrast ($\Delta nFR_{\text{W-S}}$; see Methods and Fig. 6).
in the passive presentation 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 Fig. 8). Passive TORC responses were stable across stimulus contexts in A1 and dorsal PEG, but varied between contexts in VPr, differences that were amplified during active engagement in PT-D versus CLR-D tasks (Fig. 5).

Poststimulus 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 dlFC) also showed a persistent response to the warning stimulus after the sound ended. This extended poststimulus 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 and 400 ms shock windows (during which the animal had to refrain from licking to avoid shock—see Fig. 1a). This poststimulus activity is also evident in Fig. 3b,c,e,f and in Fig. 4, where the response to the warning stimulus clearly persists in the poststimulus interval. To quantify this post-warning activity, we measured the poststimulus firing rate change from passive to active state in the silent 50–700 ms interval after target offset (Fig. 7 and Supplementary Fig. 9). Post-warning-stimulus response persistence was not observed in A1 and is most apparent in the VPr and dlFC regions. The four cortical areas, A1 (PT-D, n = 71; CLR-D, n = 57 neurons), dorsal PEG (PT-D, n = 199; CLR-D, n = 60 neurons), VPr (PT-D, n = 251; CLR-D, n = 266 neurons), and dlFC (PT-D, n = 138; CLR-D, n = 38 neurons) had significantly different poststimulus warning responses in both tasks (Kruskal–Wallis test, PT-D: X² = 40.947, P = 6.7 × 10⁻⁶, d.f. = 3; CLR-D: X² = 12.7391, P = 0.0052, d.f. = 3). A post hoc Tukey’s HSD test revealed significant differences between higher-order areas (VPr, dlFC) and A1 in the CLR-D task (A1/VPr: effect size (mean(A1)−mean(VPr)) = −60.709, 95% CI = (−109.27, −12.14), P = 0.0072; A1/dlFC: effect size = −70.39, 95% CI = (−139.36, −1.43), P = 0.0433) and with both

![Image of figure 6](https://example.com/figure6.png)

**Fig. 6 | Distributions of warning-safe response contrast (ΔnFR(W-S)) data recorded from A1, dorsal PEG, VPr, and dlFC neurons. a–d.** Results are shown for the PT-D tasks in naïve and trained animals (in both passive and active conditions). e–h. Results are shown for the CLR-D tasks in naïve and trained animals (in both passive and active conditions). ΔnFR(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 s after sound onset (PT-D) or 0.3–1.0 s 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 during passive listening (middle column) and during active behavior (right column). The 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 toward ΔnFR(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 ΔnFR(W-S). Naïve PT-D: A1, n = 64; dorsal PEG, n = 61; VPr, n = 60 neurons. Trained PT-D: A1, n = 71; dorsal PEG, n = 199; VPr, n = 251; dlFC, n = 138 neurons. Naïve CLR-D: A1, n = 65; dorsal PEG, n = 60; VPr, n = 50 neurons. Trained CLR-D: A1, n = 57; dorsal PEG, n = 60; VPr, n = 266; dlFC, n = 38 neurons.
Fig. 7 | VPr and dlFC 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 Figs. 1, 3, and 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 and 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), dorsal PEG (PT-D, n = 199; CLR-D, n = 60 neurons), VPr (PT-D, n = 251; CLR-D, n = 266 neurons), and dlFC (PT-D, n = 138; CLR-D, n = 38 neurons) in PT-D (green) and CLR-D (purple) tasks (mean ± s.e.m.). In both tasks, VPr and dlFC display a significant increase in their responses after warning sound offset during this silent 650-ms time window (PT-D: \( \chi^2 = 13.4, P = 0.0052, d.f. = 3 \); CLR-D: \( \chi^2 = 40.947, P = 6.7 \times 10^{-10}, d.f. = 3 \); Kruskal-Wallis test). The asterisks and lines above the curves show Tukey’s HSD post hoc pair-wise difference significance between higher-order areas VPr and dlFC with A1 and dorsal PEG (PT-D: A1/VPr, \( P = 0.9708 \); dorsal PEG/VPr, \( P = 0.9035 \); CLR-D: A1/VPr, \( P = 0.008 \); dorsal PEG/VPr, \( P = 0.3336 \); A1/dlFC, \( P = 0.0433 \); dorsal PEG/dlFC, \( P = 0.4292 \)). **b**, Normalized distributions of the behavior-dependent change in response after warning sounds in A1, dorsal PEG, VPr, and dlFC.

Enhanced warning sound responses during behavior. Although VPr exhibits task-related plasticity in receptive field and response properties, as demonstrated previously in A1 and dorsal PEG13,14,17, the greater magnitude, scale, and nature of the current neuroplasticity results place VPr at a higher level in the auditory cortical network, at an intermediate level between dorsal PEG 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 (Fig. 4), which in turn also transforms responses to other task-relevant stimuli depending on behavioral context (Fig. 5, and Supplementary Figs. 2 and 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 (Fig. 3a,c and Supplementary Figs. 2 and 3). Some VPr cells are even more extreme, exhibiting ‘frontal-cortex-like’ properties 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 Fig. 3). These results illustrate the importance of behavioral state, as well as stimulus choice, in shaping sustained responses24. 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 task13,15. The top-down effects of task engagement on receptive field plasticity have been shown to reach A115 and subcortically even to the inferior colliculus16. Although not evident in the population earlier auditory areas (A1, dorsal PEG) in the PT-D task (A1/VPr: effect size = -76.51, 95% CI = (-143.97, -9.06), \( P = 0.0187 \); A1/dlFC: effect size = -90.73, 95% CI = (-164.04, -17.42), \( P = 0.008 \); dorsal PEG/VPr: effect size = -99.33, 95% CI = (-147.42, -51.25), \( P = 0.0002 \); dorsal PEG/dlFC: effect size = -113.55, 95% CI = (-169.54, -57.55), \( P = 0.0018 \); CLR-D: A1/VPr, \( P = 0.0052 \); dorsal PEG/VPr, \( P = 0.008 \); dorsal PEG/dlFC, \( P = 0.0002 \); dorsal PEG/VPr, \( P = 0.3336 \); A1/dlFC, \( P = 0.0433 \); dorsal PEG/dlFC, \( P = 0.4292 \)).

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 the VPr area of the 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 the emergent behavioral control signals observed in the 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 (Fig. 2). However, the distinctiveness of VPr responses emerges more vividly during active task performance with (1) selective response enhancements to warning stimuli, (2) the unveiling of the 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 in the next sections.

### Enhanced warning sound responses during behavior.

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 the VPr area of the 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 the emergent behavioral control signals observed in the 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 (Fig. 2). However, the distinctiveness of VPr responses emerges more vividly during active task performance with (1) selective response enhancements to warning stimuli, (2) the unveiling of the 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 in the next sections.

### 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 the VPr area of the 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 the emergent behavioral control signals observed in the 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 (Fig. 2). However, the distinctiveness of VPr responses emerges more vividly during active task performance with (1) selective response enhancements to warning stimuli, (2) the unveiling of the 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 in the next sections.

### Enhanced warning sound responses during behavior.

Although VPr exhibits task-related plasticity in receptive field and response properties, as demonstrated previously in A1 and dorsal PEG13,14,17, the greater magnitude, scale, and nature of the current neuroplasticity results place VPr at a higher level in the auditory cortical network, at an intermediate level between dorsal PEG 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 (Fig. 4), which in turn also transforms responses to other task-relevant stimuli depending on behavioral context (Fig. 5, and Supplementary Figs. 2 and 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 (Fig. 3a,c and Supplementary Figs. 2 and 3). Some VPr cells are even more extreme, exhibiting ‘frontal-cortex-like’ properties 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 Fig. 3). These results illustrate the importance of behavioral state, as well as stimulus choice, in shaping sustained responses24. 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 task13,15. The top-down effects of task engagement on receptive field plasticity have been shown to reach A115 and subcortically even to the inferior colliculus16. Although not evident in the population
Multiple groups have reported significant choice probability in the auditory attention network. 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 in the auditory cortex, indicating that even sensory neurons carry information about an upcoming decision. Significant choice probability has been found in A1 for one type of auditory task, whereas in other tasks, choice probability was only observed in higher auditory cortical areas. One recent paper highlighted the causal role of the auditory belt anterior-lateral field 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 stimulus and forming auditory perceptual decisions. However, our analysis of choice probability 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 work. However, it is quite possible that different auditory cortical regions play different roles depending on species, task design, level of difficulty, and context. Further studies are needed to test choice probability in VPr in positive reinforcement GO/NO-GO or two-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-naive animals (Figs. 5 and 6). This training-dependent enhanced contrast was weak or absent in lower auditory cortical areas (Fig. 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 gating. We conjecture that these VPr learning effects may be similar to the experience-dependent malleability of the 'protocortex' described in the visual area inferotemporal cortex after extended training.

**Sustained poststimulus responses may track reward and motor timing.** VPr responses exhibit another dimension that reveals a similarity with dlFC: sustained post-warning responses (Figs. 3 and 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/NO-GO conditioned avoidance tasks in the present study, animals learned to cease licking during a 400–800 ms window following warning stimulus offset. Activity in the VPr clearly encodes this timing in the form of poststimulus responses that, across different single neurons, (1) occur precisely during this narrow temporal window, (2) persist precisely from stimulus offset up to this window, or (3) persist the full 800 ms and beyond (Fig. 3a,c, and Supplementary Figs. 2, 3, and 9). These poststimulus 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 diagram representation of population-level profiles of passive and active responses in the cortical hierarchy shown in Fig. 8. This encoding of non-acoustic information, such as task decision, motor response or timing, reward, and task-correlated visual or somatosensory signals is in general accord with earlier findings that have emphasized the 'semantic' processing that occurs in the auditory cortex.

**Evidence for VPr as a tertiary region in the ferret auditory cortex.** Neuroanatomical and neurophysiological studies of the ferret auditory cortex over the past three decades (Fig. 1) have revealed the presence of multiple auditory areas, including primary areas such as the A1 and AAF, adjacent secondary areas, such as the ADF, PPF, and PSF, and still higher auditory areas such as the AVE, the anterior and posterior pseudosylvian sulcal cortex (PSSC), and VPr. The most recent neuroanatomical connectional data support the idea that in the PEG, PPF and PSF may both be secondary or belt areas, since they reciprocally interconnect with core areas such as the A1 and AAF. In contrast, while there are reciprocal projections from both the PPF and PSF to VPr, there do not appear to be projections from core areas to VPr, suggesting that VPr may correspond in hierarchical position to a parabelt auditory area in primates.

VPr can be reliably accessed by carefully mapping tonotopic organization in the medial ectosylvian gyrus (MEG) and PEG,
and determining the position of A1 and PPF, which have mirror
tonotopic maps (Fig. 1). In recordings lateral to the high frequency
(anterolateral) region of the PPF, there is a sudden and abrupt
change in passive response properties and frequency tuning as
summarized in Fig. 2, marking the entry into the VPr region. We
recorded in the rostral area of VP in an area up to 2–3 mm lateral
to the boundary with the PPF, and rostrally up to the PSS. These
findings are consistent with the only previously published data on
tuning in VPr4,5,8. Although future studies will be needed to deter-
mine what differences may exist in the responses of the various VP
subfields that have been identified5, (VPr, VP caudal (VPC) and VP
ventral (VPV)), as well as their multisensory character and spatial
tuning5, our current VPr results reveal many of the passive audi-
tory response properties associated with an auditory ‘parabelt’ area,
including broader receptive fields, longer latency and duration
responses, low SNR, and sparseness. In addition, VPr 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 VPr with the primate parabelt and other tertiary
cortical areas.** The tertiary sensory cortex is a higher-order cortical
sensory area at least two synapses up the cortical hierarchy from pri-
mary 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)
region39–41. Although the neuroanatomy and connections of parabelt
and other regions of the primate auditory cortex have been well elu-
cidated39–41, and there is an abundance of insight about processing
in lower auditory cortical areas from neurophysiological studies of
responses in the core and belt, to our knowledge there are only two
published studies on the neurophysiological response properties of
the parabelt in awake but non-behaving macaque monkeys and mar-
mosots36,41. Both studies demonstrate that tone response latencies
increase from A1 to belt to parabelt and were longest in the rostral
parabelt, which is consistent with our results in the ferret (see Figs. 1f
and 2a,d). Similar to primates, belt regions (PPF and PSF) in the
ferret receive strong inputs from primary regions (A1 and AAF); 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 hierar-
chical ordering of increasing response latencies but also of increas-
ingly long temporal windows for sensory integration41. Hence, using
these criteria of response latency and cortical connectivity, ferret
VP is a tertiary region that bears similar features to the parabelt
as defined in the primate auditory cortex. Despite these parallels,
establishing clear homologies between cortical areas across species is
difficult and daunting, especially in higher-order sensory areas42. 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 neurophysiologi-
ical studies are required. These future studies are necessary to clarify
possible homologies between the ferret dorsal PEG and VPr, the
multiple higher auditory cortical regions in the cat42, and the belt
and parabelt regions in the 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 adap-
tive 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 the human tertiary auditory
cortex described responses that transform from acoustic to percep-
tual dimensions in the context of the McGurk effect43, which illus-
trates 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 recognition3,46 as part of the gradual progression
from sensory to task-related processing in the cerebral cortex49.
34. Bizley, J. K., Walker, K. M. M., Nodal, F. R., King, A. J. & Schnupp, J. W. H. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486–499 (2014).

13. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 33, 8378–8395 (2013).

28. Bieszczad, K. M., Miasnikov, A. A. & Weinberger, N. M. Remodeling sensory receptive fields in the auditory midbrain. *J. Neurosci.* 35, 13090–13102 (2015).

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).

27. Froemke, R. C. et al. Long-term modification of cortical synapses improves perceptual learning. *Proc. Natl Acad. Sci. USA* 1011–1019 (2010).

24. Recanzone, G. H., Schreiner, C. E. & Merzenich, M. M. Plasticity in the auditory information in ferret frontal cortex. *J. Neurosci.* 13, 1011–1019 (2010).

19. Bizley, J. K., Bajo, V. M., Nodal, F. R. & King, A. J. Cortico-cortical connectivity of primary auditory cortex. *Nature* 435, 341–346 (2005).

18. Fritz, J. B., David, S. V., Yin, P. & Shamma, S. A. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J. Neurosci.* 25, 7632–7635 (2005).

17. David, S. V., Fritz, J. B. & Shamma, S. A. Task reward structure shapes rapid adaptive persistent encoding of task-relevant sensory information in auditory cortex. *Proc. Natl Acad. Sci. USA* 109, 2144–2149 (2012).

16. Fritz, J. B., Elhilali, M. & Shamma, S. A. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontolateral cortex. *Nat. Neurosci.* 13, 1011–1019 (2010).

15. Fritz, J., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotectal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223 (2003).

14. Nodal, F. R. & King, A. J. A Biology and Diseases of the Ferret. (Wiley-Blackwell, Hoboken, 2014).

13. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 33, 8378–8395 (2013).

12. Tsunada, J., Liu, A. S. K., Gold, J. I. & Cohen, Y. E. Causal contribution of cortical fields A1 and R. to perceptual learning. *Front. Hum. Neurosci.* 8, 72 (2014).

11. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 33, 8378–8395 (2013).

10. Niwa, M., Johnson, J. S., O’Connor, K. N. & Sutter, M. L. Differences between cortical fields A1 and R. stimulus-sequence and task-dependent neuronal activity in monkey auditory cortex. *J. Neurosci.* 25, 9794–9802 (2005).

9. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 33, 8378–8395 (2013).

8. 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).

7. Bajo, V. M., Nodal, F. R., Bizley, J. K., Moore, D. R. & King, A. J. The ferret auditory cortex: descending projections to the inferior colliculus. *Cereb. Cortex* 17, 475–491 (2007).

6. Grothe, B. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 33, 8378–8395 (2013).

5. Galván, V. W. & Weinberger, N. M. Long-term consolidation and retention of learning-induced tuning plasticity in the auditory cortex of the guinea pig. *Neurolbiol. Learn. Mem.* 77, 78–108 (2002).

4. Reed, A. et al. Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron* 70, 121–131 (2011).

3. Bajo, V. M., Nodal, F. R., Bizley, J. K., Moore, D. R. & King, A. J. The ferret auditory cortex: descending projections to the inferior colliculus. *Cereb. Cortex* 17, 475–491 (2007).

2. Recanzone, G. H., Schreiner, C. E. & Merzenich, M. M. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103 (1993).

1. Galván, V. W. & Weinberger, N. M. Long-term consolidation and retention of learning-induced tuning plasticity in the auditory cortex of the guinea pig. *Neurolbiol. Learn. Mem.* 77, 78–108 (2002).
Methods

Training and behavioral tasks. We measured the auditory response properties of cortical single units in the auditory and frontal cortices of awake animals during passive listening to experimental sounds and during performance of auditory tasks. All animals used were the domestic, de-sexed, neutered, 4–5-year-old ferrets (Mustela putorius furo) obtained from Marshall BioResources. Three additional untrained (task-naive) ferrets were used to map the auditory areas and to perform the anatomical studies, one of which was also used for naive passive recordings (Figs. 5 and 6). To compare the VPr data with other auditory cortical areas, such as the A1, dorsal PEG, dIFC, and dIFC(Suprasylv), and from a total of 32 (including responses from 18 additional animals used in previously published studies from our laboratory) to compare the VPr tuning properties with the A1 and dorsal PEG during passive listening (Fig. 2). Animals were randomly chosen from the colony for experimentation. Some 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 the standards specified by the National Institutes of Health. We trained four animals on two auditory discrimination tasks with a conditioned avoidance paradigm: PT-D(Suprasylv) and CLR-D. In both tasks, animals were presented with a series of 1–6 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 the reference sound, then to abruptly stop licking after a warning stimulus offset (for a minimum of 400 ms in a 400–800 ms poststimulus window) to avoid receiving a mild shock. In the PT-D task, safe sounds consisted of 1–2 s duration broadband ripples noise (TORCs), and warning sounds were pure tones (duration and sound level matched to the TORCs 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 Fig. 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 Fig. 4 (panel B). The mean (± s.d.) ratio used was 3.2 ± 1.0. 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, dorsal PEG, VPr, and dIFC. However, the animals from earlier published studies (recordings in A1, dorsal PEG, and dIFC but not in VPr) were exclusively trained on the tone detection task (n = 14).

The warning sound frequencies in the tone detection task, and safe and warning click-train rates in the click discrimination task, were varied among the 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 tones was chosen among those 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, the broadband and click-rate tuning of VPPr neurons did not allow to choose a best frequency, or click rate that evoked responses for most neurons being recorded. The sound level (65–70 dB sound pressure level (SPL)), sound durations (1 or 2 s in tone detection, 2 s in click-rate discrimination), and interstimulus 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 (that is, the animal did not engage in the task and displayed no licking behavior). The normal duration of a behavioral block was 10–15 min, 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 min. Passive task blocks were usually 7–12 min long, depending on the duration of task sounds.

No blinding was used in the study; this is common in behavioral neuroscience. However, all analyses of neural responses were conducted in the same way for all animals. Standardized training sessions’ protocols were stable across all animals (as judged by 15% similar responses from multiple animals and investigators). The neuroanatomical studies were conducted by a researcher who was not directly involved in the behavioral or neurophysiological studies, 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 > 0.63 in at least two consecutive training sessions16. Once 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⁻¹ intramuscularly) and dexmedetomidine (0.03 mg kg⁻¹ subcutaneously) for induction; deep levels of anesthesia were maintained with 1–2% isoflurane throughout the surgery. Animals were also medicated with atropine sulfate (0.05 mg kg⁻¹ subcutaneously) to control salivation and to increase heart and respiratory rates. During surgery, electrophroagdiom, pulse, and blood oxygenation were monitored, and rectal temperature was maintained at ~38°C. Using a sterile procedure, and to be able to reach the ventral areas of the 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 polycarbonate cement; then, areas surrounding the frontal and auditory cortices were covered with bone cement (Zimmer Biomet), leaving small (2–3 mm) cavities for easy access to the auditory and frontal cortices.

Auditory fields and recording sites. Following surgery, the skull was fixed to the stereotaxic frame, and the coordinates for the recording sites were taken from Paxinos and Watson20,21. The auditory–motor–sensory area A1 is located approximately 16 mm anterior to the occipital midline crest and 12 mm lateral to the skull midline. Following the initial recording sessions in each animal, small craniotomies were placed above A1 using these coordinates; 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 oct, presented at intervals of 1 s. Also, 3 s TORCs were used to compute STFRs. A1 neurons are known to present sharp tuning to pure tones and clear single-peak, short latency STFRs (1-2 kHz). Determination of neuronal best frequencies allowed us to confirm the location of A1 based on its characteristic tonotopic organization, including a high-to-low frequency gradient in a dorsolateral direction22. Then, by ventrally expanding the existing craniotomy it was possible to gain access to non-primary auditory areas in the PEG23. Two subfields in the dorsal PEG, PPF, and PSE, 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 posteroventral direction, while PPF displays low frequencies posteroventrally and high frequencies anterodorsally. Neurons in both fields display broader tuning, longer latencies, and longer sustained responses than A1 (Fig. 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 axes. The locations of the dorsal and ventral recordings were confirmed by checking the tuning properties of its neurons and their location relative to the tonotopic map.

VPr recordings were directed to a region ventral to the high-frequency region of the PPF. VPr spans an area 1–2 mm below and ventral to high-frequency PPF and ventral to the lower lip of the PSS. Partly because of its extreme lateral location and limited accessibility for surface recordings, VPr has remained one of the least studied areas of the ferret auditory cortex. Localization of VPr on the basis of a low-to-high tonotopic discontinuity at the high end of the frequency map in the adjacent PPF area, which is characterized by a postero medial-anteralateral tonotopic gradient from low to high frequency tuning. The transition from PPF to VPr is characterized by a sharp transition in frequency tuning from high to lower best frequency and a sharp break in the tuning curves. The transition region is an area of low to high frequency bandwidth, which is characterized by a sharp transition in frequency tuning from high to lower best frequency and a sharp break in the tuning curves.

Matplotlib (version 3.0.2), Seaborn version 0.9.0 Python libraries, and Inkscape Offline data analyses were performed with custom-made MATLAB Data analysis.

Inhibitory subfields, with more numerous, longer, and less compact excitatory and inhibitory subfields in both the spectral and temporal axes. The locations of the dorsal and ventral recordings were confirmed by checking the tuning properties of its neurons and their location relative to the tonotopic map.

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 to 80 dB SPL) used for multilevel tuning assessment and to measure the frequency response curves. Sounds were digitally generated at 40 kHz with custom-made MATLAB functions and A/D hardware (PCI-6052E; National Instruments) 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, prior to any behavioral testing, to assess frequency tuning by using tone pairs of random frequencies spanning eight octaves. Individual clicks in the click-rate discrimination task (occurring after TORCs in sequential TORCs-click-train stimuli; see Fig. 1a) and during passive testing of click tuning with 1 s click-trains of randomly varying click rates from 4 to 60 Hz, were composed of 0.01 s square pulses of alternating polarity. Thirty distinct TORCs were used as task distractor (safe sounds) and also for the 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 a 5-oct-wide broadband noise with a dynamic spectrotemporal profile, that is, the superposition of the envelopes of 6 temporally orthogonal ripples (for 4–24 Hz TORCs) or 12 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. The envelope of these ripples drifted temporally up or down the logarithmic frequency axis at a constant velocity (4–48 or 4–24 Hz). The 5-oct 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 created using MATLAB (R2010b) functions, Matlabplot (version 3.0.2), Seaborn version 0.9.0 Python libraries, and Inkscape (version 0.92.3). No statistical methods were used to predetermined sample sizes (see Life Sciences Reporting Summary).

Basic tuning properties were determined by analyzing the responses to random frequency tones spanning 6–8 oct (11 tones per octave), usually ranging from 125 Hz to 32 kHz. A Gaussian function was fitted 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; tuning spectral bandwidth was measured as its width in octaves at half-height. Tones presented had a duration determined to be the mean of the Gaussian curve; tuning spectral bandwidth was measured as its width in octaves at half-height. Tones presented had a duration determined to be the mean of the Gaussian curve; tuning spectral bandwidth was measured as its width in octaves at half-height.

For each neuron by the peak modulation of the mean population PSTH, thus adjusting the scale to spikes ‘above or below spontaneous activity. The mean and s.e.m. of each PSTH bin were calculated using a jack-knife procedure. We calculated choice probability from data from all auditory cortical areas in our study using the method described by Niwa et al. Motor-related lick responses. As in a previous paper, we determined significant neural modulation of neuronal activity in the VPr by auditory stimuli using a 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-modulated spike activity as a function of safe/warning sounds and lick events is shown in equation (1).

\[
\tau = t + \sum_{r} h_{r}(l)A(l-r) + h_{c}(l)A(l-r) + h_{m}(l)A(l-r)
\]

The stimulus functions, s(l), f(l), and A(l), 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(l), has a value of 0 except at times when lick events occur. The regression functions, h_s(l), h_f(l), and h_m(l), then indicate the average firing rate before and after each corresponding event. t is the time lag (in ms) between either (a) stimulus onset or (b) lick event relative to the spike response. Spiking activity is correlated with events at different time lags which can precede (possible predictive spikes) or follow (possible causal spikes) the stimulus onset (sensory) or lick (motor) events. T indicates the range of possible time lags over which the relevant event (stimulus or lick) might correlate with spiking and was set at 500 ms. The regression functions were fitted using normalized reverse correlation, which discounted the 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 (equation (2)).

\[
\tau = t + \sum_{r} h_{r}(l)A(l-r)
\]

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, jack-knifed t-test).

Statistical analysis. To quantify the differences in target/reference response contrast in passive and active components of the behavior conditions, we used a three-way repeated-measures ANOVA for response differences between warning and safe sounds, with ‘area’ (A1, dorsal PEG, VPr and dIFC) and ‘condition’ (passive and behavior) as fixed factors, and ‘neuron’ nested in ‘area’ as a random intercept. For the PT-D task, we calculated the response difference (spikes−) 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 poststimulus offset). For the CLR-D task, we calculated the response difference (spikes s⁻¹) for the time window between 300 ms following click-train onset and the end of the shock period (800 ms poststimulus offset). We used Tukey's range test for the post hoc analysis.

Since the VPr and dlFC areas display qualitative differences in their response patterns with the A1 and dorsal PEG areas, we measured responses in two different epochs: (1) during the duration of task-relevant sounds; and (2) in the silent period after presentation of target sounds. We quantified the contrast between safe and warning responses (ΔnFR(W-S)) in a time window between 0.1 and 0.45 s after sound onset (PT-D task) or 0.3 and 1.0 s (CLR-D task) to avoid onset transient responses and, in the CLR-D task, the offset response to TORCs preceding click-trains. We measured the mean values of ΔnFR(W-S) from the normalized PSTH responses to compare population responses between areas and between the trained and naive 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 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 poststimulus warning responses was performed with a Kruskal–Wallis test; pairwise area comparisons were performed with Tukey's HSD range test as a post hoc test. The distributions of ΔnFR(W-S) data calculated from the A1, dorsal PEG, VPr and dlFC responses in passive and active conditions were tested for normality with a Lilliefors test. All distributions of ΔnFR(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 ΔnFR(W-S) (P = 0.1586, ks = 0.1039); dorsal PEG active ΔnFR(W-S) (P = 0.0995, ks = 0.192); and dlFC active Δn FR(W-S) (P = 0.1198, ks = 0.1271).

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

**Code availability.** The 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 study findings are available from the corresponding author upon reasonable request.

**References**

51. Heffner, H. E. & Heffner, R. S. in *Methods in Comparative Psychoacoustics* (eds Klump, G. M. et al.) 79–93 (Birkhäuser, Basel, 1995).
52. Klein, D. J., Depireux, D. A., Simon, J. Z. & Shamma, S. A. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. *J. Comput. Neurosci.* 9, 85–111 (2000).
53. Englitz, B., David, S. V., Sorenson, M. D. & Shamma, S. A. MANTA—an open-source, high density electrophysiology recording suite for MATLAB. *Front. Neural Circuits* 7, 69 (2013).
54. Rossant, C. et al. Spike sorting for large, dense electrode arrays. *Nat. Neurosci.* 19, 634–641 (2016).
55. Depireux, D. A., Simon, J. Z., Klein, D. J. & Shamma, S. A. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J. Neurophysiol.* 85, 1220–1234 (2001).
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, Boca Raton, 1994).
58. Niwa, M., Johnson, J. S., O’Connor, K. N. & Sutter, M. L. Activity related to perceptual judgment and action in primary auditory cortex. *J. Neurosci.* 32, 3193–3210 (2012).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

☐ n/a  Confirmed

☐ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement

☐ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly

☐ The statistical test(s) used AND whether they are one- or two-sided

☐ Only common tests should be described solely by name; describe more complex techniques in the Methods section.

☐ A description of all covariates tested

☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons

☐ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)

☐ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted

☐ Give P values as exact values whenever suitable.

☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings

☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes

☐ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

☐ Clearly defined error bars

☐ State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

Most data was collected using AlphaMap 10.10 software from AlphaOmega (Nazareth, Israel). Additional data was collected using open-source programs Baphy (commit E155036, https://bitbucket.org/lbhb/baphy) and MANTA 1.0 (https://code.google.com/archive/p/manta-system/), both coded in MATLAB R2010B (from MathWorks, Natick, MA).

Data analysis

Data was analyzed using custom code written in MATLAB R2010B (MathWorks, Natick, MA) and its integrated functions. Figures 5, 7 and S9 were prepared with Python 3.7, using the Matplotlib 3.0.2 library and Seaborn 0.9.0. Spike sorting was performed using custom code written in MATLAB R2010B and some data was sorted with open-source program Klusta (version c1909dd, Rossant, C. et al. Spike sorting for large, dense electrode arrays. Nat. Neurosci. 19, 634–641 (2016) doi:10.1038/nn.4268)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

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

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications (References: Fritz et al. (2003) Nat Neurosci 6:1216–1223. doi:10.1038/nn1141, Fritz et al. (2010) Nat Neurosci 6:1216–1223. doi:10.1038/nn.2598, Atiani et al. (2014) Neuron 82:486–499. doi:10.1016/j.neuron.2014.02.029).

Data exclusions

In population PSTH averages we excluded neurons that had no significant auditory responses in any presentation of task sounds, in both passive and active conditions. This criterion was pre-established.

Replication

In multiple recording sessions (n=713) in multiple animals (n=32), we placed electrodes acutely in the brain and, by measuring their position with respect to permanent landmarks in the implant, stereotaxic coordinates, and neurophysiological markers such as tonotopic map position, we were able to ascertain the cortical areas (A1, dorsal PEG, VPr, dorsolateral frontal cortex) in which we were recording. We were subsequently able to confirm these cortical areas with neuroanatomical and histological analysis of the electrode paths and electrolytic marks made in the recording sites. We recorded from multiple neurons in each area (see Methods and Figures 2 and 4) and found that the responses of these neurons were consistent within area. As shown in Figure 2, we recorded many different cells from each area (n=2740 in A1, n=1337 in dPEG, n=658 in VPr) in passive listening conditions and found highly consistent and replicable responses that were distinctive for each area. We also recorded responses in the behaving animal, as shown in Figure 4 from multiple areas in many animals and recorded from A1 (n=71), dPEG (n=199), VPr (n=266), dlFC (n=138) and also found consistent responses from all neurons in these different areas, that were distinct from responses of neurons in the different areas. In this way, we were able to replicate successfully our conclusions about response properties of neurons in each of the four different cortical areas that we studied.

Randomization

Animal subjects (female ferrets) 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.

Blinding

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.

Reporting for specific materials, systems and methods
### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ✗   | Unique biological materials |
| ✗   | Antibodies |
| ✗   | Eukaryotic cell lines |
| ✗   | Palaeontology |
| ✗   | Animals and other organisms |
| ✗   | Human research participants |

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
| ✗   | ChIP-seq |
| ✗   | Flow cytometry |
| ✗   | MRI-based neuroimaging |

### Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

| Laboratory animals | All animals used in this study were adult (1-4 years old) female ferrets (Mustela putorius furo) obtained from Marshall BioResources. Animals were spayed and desecrated. During behavioral studies, animals were placed on a water schedule to motivate them to perform the tasks, in which they received water as reward. |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Wild animals       | The study did not involve wild animals.                                                                                                                                                           |
| Field-collected samples | The study did not involve samples collected from the field.                                                                                                                                   |