nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-04839-9 OPEN

Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex

Sophie Bagur[1], Martin Averseng[2], Diego Elgueda[3], Stephen David[4], Jonathan Fritz[3], Pingbo Yin[3], Shihab Shamma[2,3], Yves Boubenec[2] & Srdjan Ostojic[5]

Primary sensory cortices are classically considered to extract and represent stimulus features, while association and higher-order areas are thought to carry information about stimulus meaning. Here we show that this information can in fact be found in the neuronal population code of the primary auditory cortex (A1). A1 activity was recorded in awake ferrets while they either passively listened or actively discriminated stimuli in a range of Go/No-Go paradigms, with different sounds and reinforcements. Population-level dimensionality reduction techniques reveal that task engagement induces a shift in stimulus encoding from a sensory to a behaviorally driven representation that specifically enhances the target stimulus in all paradigms. This shift partly relies on task-engagement-induced changes in spontaneous activity. Altogether, we show that A1 population activity bears strong similarities to frontal cortex responses. These findings indicate that primary sensory cortices implement a crucial change in the structure of population activity to extract task-relevant information during behavior.

[1] Brain Plasticity Unit, Équipe MOBS, CNRS UMR 8249, École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, 10 rue Vauquelin, 75005 Paris, France. [2] Laboratoire des Systèmes Perceptifs, Département d'Études Cognitives, École Normale Supérieure, PSL Research University, CNRS, 29 rue d'Ulm, 75005 Paris, France. [3] Neural Systems Laboratory, Institute for Systems Research & Electrical and Computer Engineering, University of Maryland in College Park, 2202 A.V. Williams Building, College Park MD 20742 MD, USA. [4] Laboratory of Brain, Hearing and Behavior, Oregon Health & Science University, 3181 S.W. Sam Jackson Park Road, Portland OR 97239-3098 OR, USA. [5] Laboratoire de Neurosciences Cognitives, INSERM U960, École Normale Supérieure, PSL Research University, 29 rue d'Ulm, 75005 Paris, France. These authors jointly supervised this work: Yves Boubenec, Srdjan Ostojic. Correspondence and requests for materials should be addressed to Y.B. (email: boubenec@ens.fr) or to S.O. (email: srdjan.ostojic@ens.fr)

How and where in the brain are sensory representations transformed into abstract percepts? Classical anatomical and physiological studies have suggested that this transformation occurs progressively along a cortical hierarchy. Primary sensory areas are commonly believed to process and extract high-level physical properties of stimuli, such as orientations of visual bars in the primary visual cortex or abstract sound features in the primary auditory cortex[1,2]. These fundamental sensory features are then integrated and interpreted as behaviorally meaningful sensory objects, and relayed to higher cortical areas, which extract increasingly task-relevant abstract information. Prefrontal, parietal, and premotor areas lie at the apex of the hierarchy[3,4]. They integrate inputs from different sensory modalities, transform sensory information into categorical percepts and decisions, and store them in working memory until the time when the appropriate motor action needs to be executed[5,6].

According to this classical feedforward picture, primary sensory areas are often considered as playing a largely static role in extracting and encoding high-level stimulus physical attributes[7–10]. However, a number of recent studies in awake, behaving animals have challenged this view and shown that the information represented in primary areas in fact strongly depends on the behavioral state of the animal. Motor activity, arousal, learning, and task engagement have been found to strongly modulate responses in primary visual, somatosensory, and auditory cortices[11–25]. Effects of task engagement have been particularly investigated in the auditory cortex, where it was found that receptive fields of primary auditory cortex neurons adapt rapidly to behavioral demands when animals engage in various types of auditory discrimination tasks[26–30]. These observations have been interpreted as signatures of highly flexible sensory representations in primary cortical areas, and they raise the possibility that these areas may be performing computations more complex than simple extraction and transmission of stimulus features to higher-order regions.

An important limitation of many previous studies[26–30] is that they relied mostly on a single-cell description, which characterized the selectivity of average individual neurons to sensory stimuli. Here we show that simple population analyses reveal that task engagement induces a shift in the primary auditory cortex from a sensory-driven representation to a representation of the behavioral meaning of stimuli, analogous to the one found in the frontal cortex. We first analyzed the responses during a temporal auditory discrimination task, in which ferrets had to distinguish between Go (Reference) and No-Go (Target) stimuli corresponding to click trains of different rates. The activity of the same neural population was recorded when the animals were engaged in the task and when they passively listened to the same stimuli. Both single-cell and population analyses showed that task engagement decreased the accuracy of encoding the physical attributes of stimuli. Population, but not single-cell, analyses, however, revealed that task engagement induced a shift toward an asymmetric representation of the two stimuli that enhanced target-evoked activity in the subspace of optimal decoding. This shift was in part enabled by a novel mechanism based on the change in the pattern of spontaneous activity during task engagement.

Performing identical analyses on independent datasets collected in A1 during other behavioral discrimination tasks demonstrated that our main finding can be well generalized, independently of the type of stimuli, behavioral paradigm, or reward contingencies. Finally, a comparison between population activity in A1 and single-cell recordings in the frontal cortex revealed strong similarities. Altogether, our results suggest that task-relevant, abstracted information is present in primary sensory cortices and can be read out by neurons in higher-order cortices.

Results

Task engagement impairs A1 encoding of stimulus features. We recorded the activity of 370 units in the primary auditory cortex (A1) of two awake ferrets in response to periodic click trains. The animals were trained using a conditioned avoidance paradigm[26] to lick water from a spout during the presentation of a class of reference stimuli and to stop licking following a target stimulus (Animal 1: 83% hit +/−3% SEM; Animal 2: 69% hit +/−5% SEM) (Fig. 1a; see Methods). Target stimuli thus required a change in the ongoing behavioral output while reference stimuli did not. Each animal was trained to discriminate low vs high click rates, but the precise rates of reference and target click trains changed in every session. The category choice was opposite in the two animals to avoid confounding effects of stimulus rates (low/high) and behavioral category (reference/target). Thus the target for one ferret was high click train rates, and the target for the other ferret was low click train rates. In each session, the activity of the same set of single units was recorded during active behavior (task-engaged condition) and during passive presentations of the same set of auditory stimuli before and after behavior (passive conditions).

We first examined how auditory cortex responses and stimulus encoding depended on the behavioral state of the animal. In agreement with previous studies[14,19], spontaneous activity often increased in the task-engaged condition, while stimulus-evoked activity was often suppressed (Fig. 1b). To quantify the changes in activity over the population, we used a modulation index of mean firing rates between passive and task-engaged conditions, estimated in different epochs (Fig. 1c; see Methods). Spontaneous activity before stimulus presentation increased in the engaged condition ($n = 370$ units, $p < 0.0001$), while baseline-corrected stimulus-evoked activity did not change overall ($n = 370$ units, $p = 0.94$). These changes in average activity suggested that the signal-to-noise ratio between stimulus-evoked and spontaneous activity paradoxically decreased when the animals engaged in the task.

To quantify in a more refined manner the timing of neural responses with respect to click times, we computed the vector strengths (VSs) of individual unit responses, a standard measure of phase-locked activity evoked by click trains[12,31]. VSs quantify the amount of entrainment of the neural response to the clicks and range from 1 for responses fully locked to clicks to 0 for responses independent of click timing. A vast majority of neurons (Passive Ref/Targ: 80%, 81% and Engaged Ref/Targ: 84%, 81%) displayed statistically significant VSs in both conditions. However, VS decreased in the engaged condition compared to the passive condition (Fig. 1c; $n = 574$ (287 units, 2 sounds), $p < 0.0001$), independently of the rate of the click train and the identity of the stimuli (Supplementary Fig. 1). This reduction in stimulus entrainment further suggested that task engagement degraded the encoding of click times in A1.

The change in activity between passive and task-engaged conditions was heterogeneous across the neural population. While stimulus entrainment was on average reduced in the engaged condition, a minority of neurons increased their responses. One possibility is that such changes reflect an increased sparseness of the neural code. Under this hypothesis, the stimuli are represented by smaller pools of neurons in the task-engaged condition but in a more reliable manner. To address this possibility, we built optimal decoders that reconstructed click timings from the activity of all simultaneously recorded neurons, in a trial-by-trial manner (Fig. 1d, Methods). We found that the reconstruction accuracy decreased in the task-engaged condition compared to the passive condition (Fig. 1e–g), confirming that encoding of click times decreased during behavior.

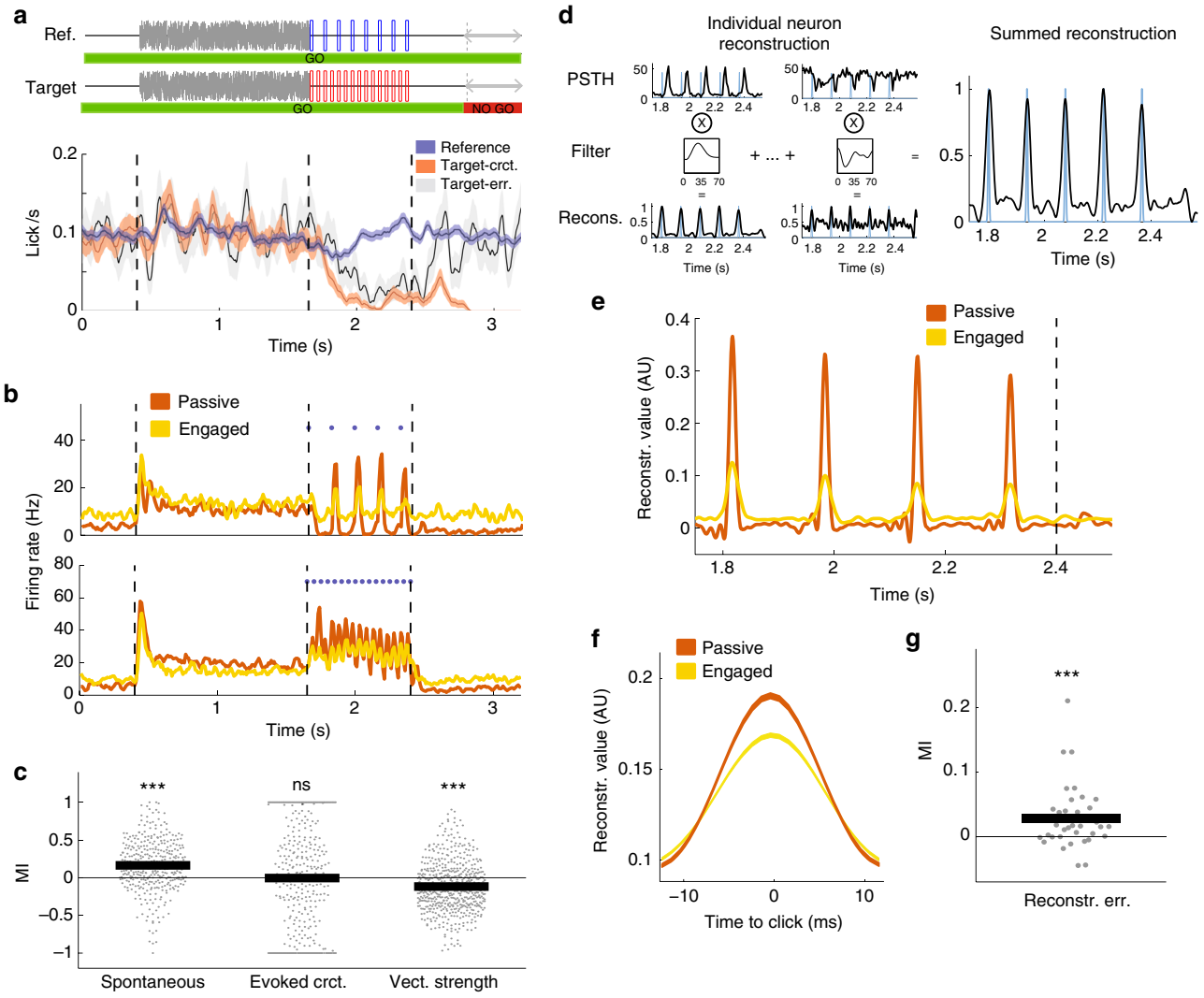


Fig. 1 Task structure and neural encoding of click times in A1. **a** Structure of the click-train discrimination task and average behavior of the two animals. Each sound sequence is composed of 0.4 s silence then a 1.25 s long white noise burst followed by a 0.8 s click train and a 0.8 s silence. On each block, the ferret is presented with a random number (1–7) of reference stimuli (top) followed by a target stimulus (bottom), except on catch trials with no target presentations. On blocks including a target, the animal had to refrain from licking during the final 0.4 s of the trial, the no-go period, to avoid a mild tail shock (error bars are +/− SEM). **b** PSTH of two example units during reference sequences in the passive and engaged state. Note that, in the task-engaged state, the units show enhanced firing during the initial silent period of spontaneous activity and reduced phase locking to the stimulus. **c** Modulation index of each unit for spontaneous firing rate, spontaneous-corrected click-evoked firing rate, and vector strength showing higher spontaneous firing rates and lower vector strength in the task-engaged state. The vector strength was only calculated for units firing above 1 Hz and values for both reference and target are shown. SEM error bars are not shown because not visible at this scale: 0.017, 0.037, and 0.013, respectively (one-sample two-sided Wilcoxon signed-rank test with mean 0, $n = 370$, 574, 370, zval = −8.99, $p = 2.57e\text{-}19$; zval = −0.07, $p = 0.94$; zval = −8.82, $p = 1.16e\text{-}18$; ***$p < 0.001$). **d** Schematic of stimulus reconstruction algorithm. Using PSTHs from half of the trials, a time-lagged filter is fitted to allow optimal reconstruction of the stimulus for each individual unit. Individual reconstructions are summed to obtain a population reconstruction (far right). **e** Stimulus reconstruction from an example session showing degraded reconstruction in the task-engaged state. **f** Mean click reconstruction in passive and engaged states. **g** Modulation index of each session for stimulus reconstruction error. SEM error bar is not shown because it is not visible at this scale: 0.0014 (one-sample two-sided Wilcoxon signed-rank test with mean 0, $n = 36$; zval = −3.4092, $p = 6.51e\text{-}4$; ***$p < 0.001$)

In summary, the fine physical features of the behaviorally relevant stimuli became less faithfully represented by A1 activity when the animals were engaged in this discrimination task.

State-independent discrimination of stimulus category. In the task-engaged condition, the animals were required to determine whether the rate of each presented click train was high or low. They needed to make a categorical decision about the stimuli and correctly associate them with the required actions, before using that information to drive behavior. We therefore asked to what extent the two classes of stimuli could be discriminated based on population responses in A1 in the task-engaged and in the passive conditions.

We first compared the mean firing rates evoked by target and reference click trains. While some units elevated their activity for the target stimulus (Fig. 2a, left), others preferred the reference (Fig. 2a, right). Over the whole population, mean firing rates were not significantly different for target vs reference stimuli (Fig. 2b) or for low vs high rate click trains (Supplementary Fig. 2a). This observation held in both passive and task-engaged conditions. Discriminating between the stimuli was thus not possible on the

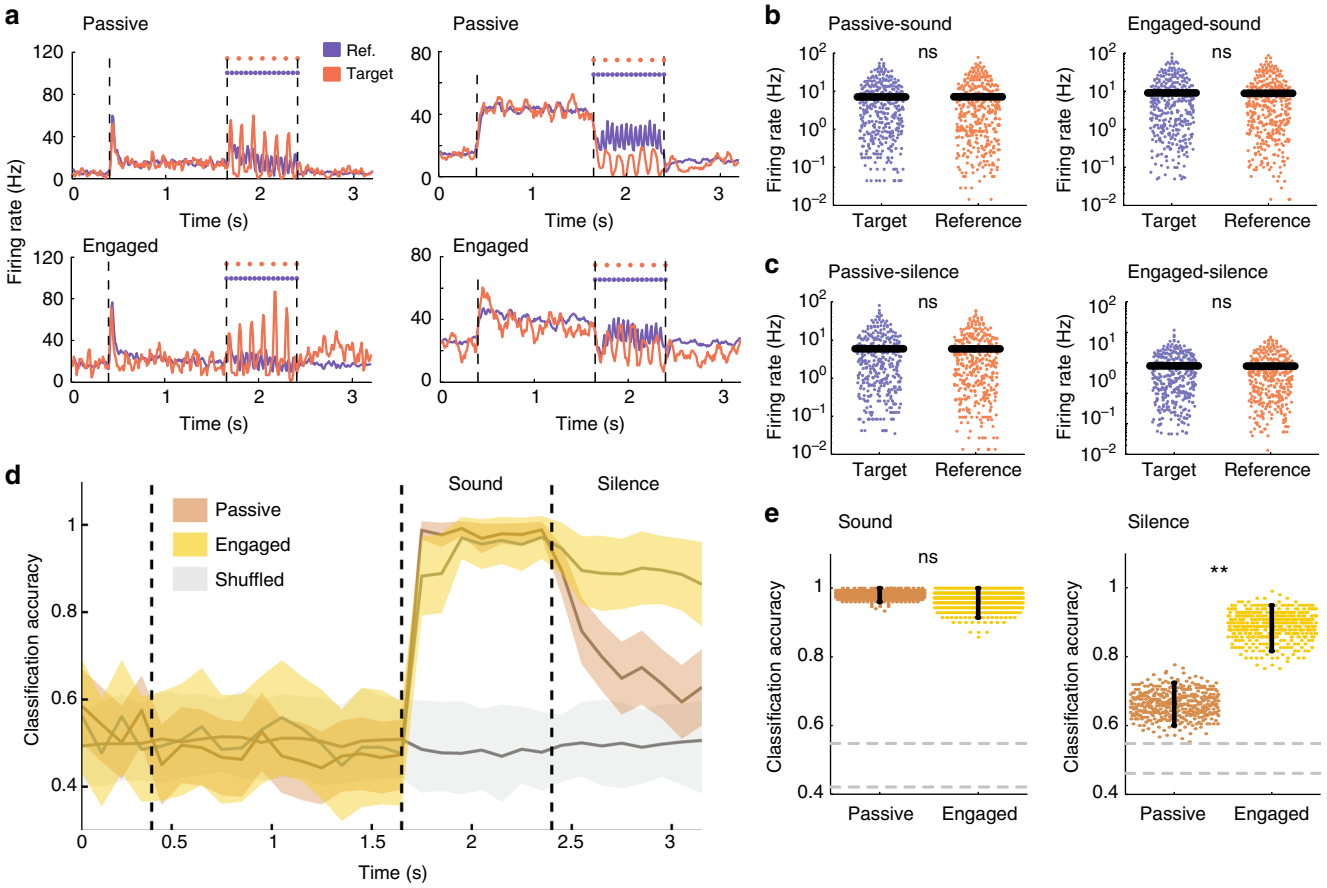


Fig. 2 Discrimination of target and reference stimuli based on A1 activity. **a** PSTHs of two example units during reference (blue) and target (red) trials in the passive (top) and task-engaged (bottom) state. The unit on the left is target-preferring and the unit on the right is reference-preferring. **b**, **c** Comparison of average firing rates on a log scale in passive (left) and engaged (right) between target and reference stimuli during the sound (**b**) and during the post-stimulus silence (**c**) periods. SEM error bars are not shown because it is not visible at this scale. (two-sided Wilcoxon signed rank, $n = 370$; zval = 0.34, $p = 0.73$; zval = 0.35, $p = 0.79$; zval = −0.47, $p = 0.64$; zval = −0.35, $p = 0.73$) (**d**) Accuracy of stimulus classification in passive and engaged states. In gray, chance-level performance evaluated on label-shuffled trials. Error bars represent 1 std calculated over 400 cross-validations. **e** Mean classifier accuracy during the sound (left) and silence period (right) in both conditions. Gray dotted lines give 95% confidence interval of shuffled trials. Error bars represent 95% confidence intervals. ($n = 400$ cross-validations; $p = 0.29$ and $p < 0.0025$; $^{**}p < 0.01$)

basis of population-averaged firing rates (see Supplementary Fig. 2b).

To take into account the heterogeneity of neural responses and quantify the ability of the whole population to discriminate between target and reference stimuli on an individual trial basis, we adopted a population-decoding approach. We used a simple, binary linear classifier that mimics a downstream readout neuron. The classifier takes as inputs the spike counts of all the units in the recorded population, multiplies each input by a weight, and compares the sum to a threshold to determine whether a trial was a reference or a target. The weight of each unit was set based on the difference between the average spike counts evoked by the two stimuli (Supplementary Fig. 3 and Methods). This weight was therefore positive or negative depending on whether it preferred the target or reference stimulus. Different decoder weights were determined at every time-bin in the trial. The width of the time-bins (100 ms) was larger than the interclick intervals (Methods). Shorter time-bins increase the amount of noise but do not affect our main findings (Supplementary Fig. 8A). Training and testing the classifier on separate trials allowed us to determine the cross-validated performance of the classifier and therefore the ability to discriminate between the two stimulus classes based on single-trial activity in A1.

During stimulus presentation, the linear readout could discriminate target and reference stimuli with high accuracy in both passive and task-engaged conditions (Fig. 2d, e). Because the classifier performed at saturation during the sound epoch, it could be that differences between passive and engaged classifiers were masked by the substantial number of neurons provided to the classifiers. Decoders performing with lower numbers of neurons did not reveal any difference between the two behavioral states (Supplementary Fig. 4a). Moreover, this discrimination capability did not appear to be layer-dependent (Supplementary Fig. 4b, c). The primary auditory cortex therefore appeared to robustly represent information about the stimulus class, independently of the decrease in the encoding of precise stimulus properties that occurs during task engagement.

We next examined the discrimination performance during the silence immediately after stimulus offset. This silent period consisted of a 400 ms interval followed by a response window, during which the animal learned to stop licking if the preceding stimulus was a target. As during the sound period, mean firing rates were not significantly different for the two types of stimuli during post-stimulus silence (Fig. 2c). Nevertheless, we found that discrimination performance between target and reference trials remained remarkably high throughout the post-stimulus silence in the task-engaged condition. In the passive condition, the decoding performance decayed during post-stimulus silence but remained above chance level (Fig. 2d, e and Supplementary Fig. 5b). The information about the stimulus class was thus

maintained during the silent period in the neural activity in A1 but more strongly when the animal was actively engaged in the task. Moreover, a comparison between the decoders determined during the sound and after stimulus presentation showed that the encoding of information changed strongly between the two epochs of the trial (Supplementary Fig. 6 and Supplementary Methods).

Shift to target-driven stimulus representation during behavior. We next examined in more detail the neural activity that underlies the classification performance in the two conditions. Target and reference stimuli play highly asymmetric roles in the Go/No-Go task design studied here as their behavioral meaning is totally different. As shown in Fig. 1a, animals continuously licked throughout the task and only target stimuli elicited a change from this ongoing behavioral output, whereas reference stimuli did not. We therefore sought to determine whether target- and reference-induced neural responses play similar or different roles in the discrimination between target and reference stimuli.

We first used dimensionality-reduction techniques to visualize the trajectories of the population activity in three dimensions (Fig. 3a, see Methods for details). The three principal dimensions were determined jointly for the passive and engaged data. This allowed us to visually inspect the difference in population dynamics and decoding axes between the two behavioral conditions. The average neural trajectories on reference and target trials strongly differ in the two behavioral conditions. In the passive condition, reference and target stimuli led to approximately symmetric trajectories around baseline spontaneous activity, suggesting that reference and target stimuli played essentially equivalent roles during the sound (Fig. 3a, c, d). In contrast, in the task-engaged condition, the activity evoked by reference and target stimuli became strongly asymmetric with respect to the decoding axes and the spontaneous activity (Fig. 3b, e, f).

To further characterize the change in information representation between the two conditions, we examined the average inputs from target and reference stimuli to a hypothetical readout neuron corresponding to a previously determined linear classifier. This is equivalent to projecting the trial-averaged population

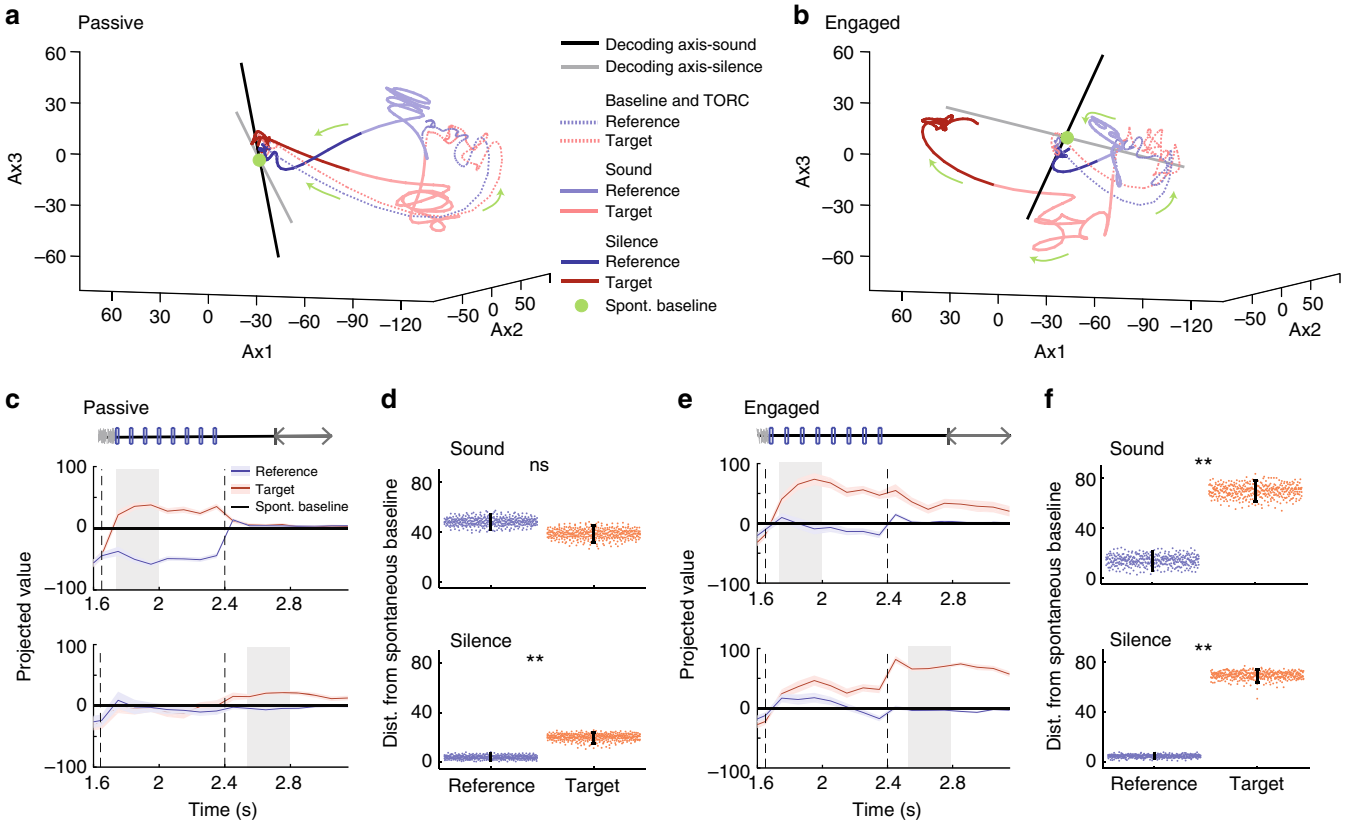
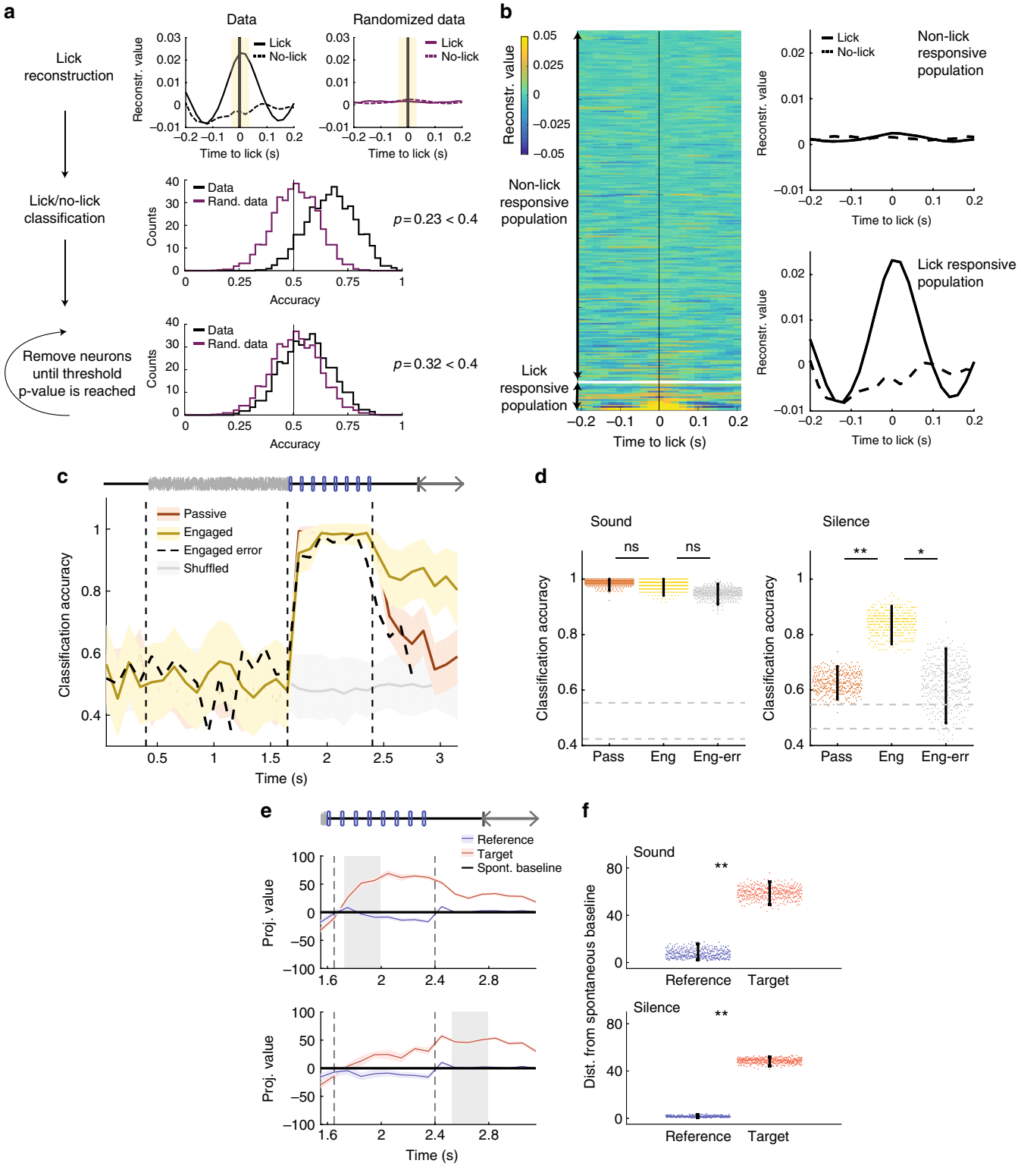


Fig. 3 Task engagement induces shift from symmetric to asymmetric representation of target and reference stimuli. **a** Population response during target and reference stimuli in the passive state along the first three components identified using GPFA (see Methods) on single trial data. The session begins at the baseline (green dot), followed by the TORC presentation, (dotted line) then the click presentation of either the target and the reference sound (light red and blue, respectively) and finally the post-sound silence period (dark red and blue). Note in particular that, in the passive state, the reference and target activities move away symmetrically from the baseline point given by projection of spontaneous activity. **b** As in **a**, for the task-engaged state. Note that, in this state, target activity makes a much larger excursion from the baseline than reference activity. The axes are the same as in **a**, as the GPFA analysis was performed jointly on passive and engaged data. **c** Projection onto the decoding axis of trial-averaged reference- and target-evoked responses for the whole neural population. A baseline value computed from prestimulus spontaneous activity was subtracted for each unit, so that the origin corresponds to the projection of spontaneous activity (shown by black line). Decoding axes determined during sound presentation and post-stimulus silence are, respectively, used for projections in the top and bottom rows. The periods used to construct the decoding axis are shaded in gray. Error bars represent 1 std calculated using decoding vectors from cross-validation. This procedure allows for the visualization of the distance between reference- and target-evoked projections (that corresponds to decoding strength) and the distance of the stimuli-evoked responses from the baseline of spontaneous activity can be interpreted as the contribution of each stimulus to decoding accuracy. **d** Distance of reference and target projections from baseline in each condition during the sound and silence period. Error bars represent 95% confidence intervals ($n = 400$ cross-validations; $p = 0.15$ and $p < 0.0025$; $^{**}p < 0.01$). **e** As in **c** for the engaged state. **f** As in **d** for the engaged state. ($n = 400$ cross-validations; $p < 0.0025$ and $p < 0.0025$; $^{**}p < 0.01$)

activity onto the axis determined by the linear classifier, trained at a given time point in the trial. This procedure sums the neuronal responses after applying an optimal set of weights. It effectively reduces the population dynamics from $N = 370$ dimensions (where each dimension represents the activity of an individual neuron) to a single, information-bearing dimension. The discrimination performance of the classifier is directly related to the distance between reference and target activity after projection, so that the projection allows us to visualize how the classifier extracts the stimulus category from the neuronal responses to the two respective stimuli. Projecting the spontaneous activity along the same axis provides, moreover, a baseline for comparing the changes in activity induced by the target and reference stimuli along the discrimination axis. As the encoding changes strongly between stimulus presentation and the subsequent silence (Supplementary Fig. 6 and Supplementary Note 1), we examined



two projections corresponding to the decoders determined during stimulus and during silence.

As suggested by the three-dimensional visualization, the projections on the decoding axes demonstrated a clear change in the nature of the encoding between the two behavioral conditions. In the passive condition, reference and target stimuli led to approximately symmetric changes around baseline spontaneous activity (Fig. 3c, d). In contrast, in the task-engaged condition, the activity evoked by reference and target stimuli became strongly asymmetric (Fig. 3e, f). In particular, the projection of reference-evoked activity remained remarkably close to spontaneous activity throughout the stimulus presentation and the subsequent silence in the task-engaged condition. The strong asymmetry in the engaged condition and the alignment of reference-evoked activity were found irrespective of whether the projection was performed on decoders determined during stimulus (Fig. 3e, f, top) or during silence (Fig. 3e, f, bottom). The time courses of the two projections were however different, with target-evoked responses rising very rapidly (Fig. 3e, f, top) when projected along the first axis but much more gradually when projected along the second axis (Fig. 3e, f, bottom). In both cases, however, our analysis showed that in the engaged condition the discrimination performance relies on an enhanced detection of the target.

The strong similarity between the projection of reference-evoked activity and the baseline formed by the projection of spontaneous activity is not due to the lack of responses to reference stimuli in the engaged condition. Reference stimuli do evoke strong responses above spontaneous activity in both passive and task-engaged conditions. However, in the task-engaged but not in the passive condition, the population response pattern of the reference stimuli appears to become orthogonal to the axis of the readout unit during behavior. The strong asymmetry between reference- and target-evoked responses is therefore seen only along the decoding axis, but not if the responses are simply averaged over the population, or averaged after sign correction for the preference between target and reference (Supplementary Fig. 7). We verified that these results are robust across a range of time bins (10–200 ms), allowing us to cover timescales both on the order of the click rate and much longer. Both the increase in post-sound decoding accuracy in the engaged state and the increased asymmetry of target/reference representation were observed at all timescales (Supplementary Fig. 8a, b).

Target representation in A1 is independent of motor activity. One simple explanation of the asymmetry between target- and reference-evoked responses could potentially be the motor-evoked neuronal discharge. Indeed, during task engagement, the animals' motor activity was different following target and reference stimuli as the animals refrained from licking before the No-Go window following the target stimulus but not the reference stimulus (Fig. 1a). As neural activity in A1 can be strongly modulated by motor activity[17], such effects could potentially account for the observed differences between target- and reference-evoked population activity.

To assess the role played by motor activity in our findings, we first identified units with lick-related activity. To this end, we used decoding techniques to reconstruct lick timings from the population activity and determined the units that significantly contributed to this reconstruction by progressively removing units until licking events could not anymore be detected from the population activity. We excluded a sufficient number of neurons (10%) such that a binary classifier using the remaining units could no longer classify lick and no-lick time points as compared with random data ($p > 0.4$; Fig. 4a, b, see Methods). We then repeated the previous analyses after removing all of these units. The discrimination performance between target and reference trials remained high and significantly different between the passive and the task-engaged conditions during the post-stimulus silence (Fig. 4c, d), while projection of target- and reference-elicited activity on the updated decoders still showed a strong asymmetry in favor of the target (Fig. 4e, f). This indicated that the information about the behavioral meaning of stimuli was represented independently of any overt motor-related activity. In all subsequent analyses, we excluded all lick-responsive neurons.

Although the information present in A1 during the post-stimulus silent period could not be explained by motor activity, it appeared to be directly related to the behavioral performance of the animal. To show this, we classified population activity on error trials, in which the animal incorrectly licked on target stimuli, using classifiers trained on correct trials. Error trials showed only a slight impairment of accuracy during the sound presentation, but strikingly, the discrimination accuracy of the classifier during the post-stimulus silence on these trials dropped down to the performance level measured during passive sessions (Fig. 4c, e). This analysis therefore demonstrated a clear correlation between the behavioral performance and the information on stimulus category present during the silent period in A1.

Mechanisms underlying task-dependent target representation. The previous analyses of population activity have shown that task engagement induces an asymmetric encoding, in which the

Fig. 4 Relation between A1, motor activity, and behavioral outcome **a** Schematic of the approach used to identify lick-responsive units. First, we reconstructed licks using optimal filters as for click reconstruction (Fig. 1). The filter is applied during licks and also during randomly selected time points with no licks (top left). We evaluated the accuracy of classifying lick and no-lick time events using a linear decoder (black distribution, middle panel). In both cases, the significance was tested using randomized data (top right and purple distribution, middle panel). We iteratively removed the best classification units (bottom plot) until the p value was >0.4 and the two distributions were indistinguishable. **b** Results of reconstruction of lick events and removal of lick units. Left: heatmap of average lick reconstruction for all neurons ordered by classification weight. Right: average reconstruction of lick and no-lick events using units retained for population analysis (non-lick responsive) and units excluded from the population analysis (lick responsive). **c** Accuracy of stimulus classification in passive and engaged states using only non-lick-responsive units. Note that, after removal of lick-responsive units, the discrimination during post-stimulus silence is still enhanced in the task-engaged state on correct trials but is low during error trials. Error bars represent 1 std calculated over 400 cross-validations. **d** Comparison of mean accuracy on passive, task-engaged correct and task-engaged error trials, during sound (left) and post-stimulus silence periods (right). Error bars represent 95% confidence intervals. ($n = 400$ cross-validations; sound: pass/eng $p = 0.22$, eng/err: $p = 0.87$; silence: pass/eng $p < 0.0025$, eng/err: $p = 0.012$; $^{*}p < 0.05$, $^{**}p < 0.01$) **e** Projection onto the decoding axis of baseline-subtracted population vectors during the engaged condition constructed using activity of non-lick-responsive units only for the reference and target stimuli. Projections are shown onto the decoding axes obtained on early sound (top) and silence periods (bottom) (shaded epochs). The origin corresponds to the projection of spontaneous activity (shown by black line). Error bars represent 1 std (cross-validation $n = 400$). **f** Distance of reference and target projections from baseline in the engaged condition during sound and silence periods. Error bars represent 95% confidence intervals ($n = 400$ cross-validations; $p < 0.0025$ and $p < 0.0025$; $^{**}p < 0.01$)

activity elicited by reference stimuli becomes similar to spontaneous background activity when seen through the decoder. Two different mechanisms can potentially contribute to this shift between passive and engaged conditions: (i) the spontaneous activity changes between the two behavioral states such that its projection on the decoding axis becomes more similar to reference-evoked activity; (ii) stimulus-evoked activity changes between the states, inducing a change in the decoding axis and in the projections. In general, both mechanisms can be expected to contribute and their effects can be separated during different epochs of the trial.

To disentangle the effects of the two mechanisms, we chose a fixed decoding axis and projected on the same axis the stimulus-evoked activity from both passive and engaged conditions. We then compared the resulting projections with projections of both passive and engaged spontaneous activity. We performed this procedure separately for decoding axes determined during sound and silence epochs.

Figure 5a (top) illustrates the projections along the decoding axis determined during the sound epoch in the engaged condition. Comparing the passive responses with the passive and engaged spontaneous activity revealed that the projection of passive reference-evoked activity was aligned during sound presentation with the projection of engaged but not passive spontaneous activity (Fig. 5a, top left). A similar observation held for the engaged responses throughout the sound presentation epoch (Fig. 5a, top right). These projections remained similar regardless of whether the decoding axes were determined during the passive or the engaged conditions, as these two axes largely share the same orientation (Supplementary Fig. 6e). Altogether, these results indicate that the change in spontaneous baseline activity during task engagement is sufficient to explain the strongly asymmetric, target-driven response observed early in the trial during sound presentation (Fig. 5b, top).

However, we reached a different conclusion when we examined the activity during the post-stimulus silence (Fig. 5a, bottom). Repeating the same procedure as above but projecting on the decoding axis determined during the post-stimulus silence revealed that the shift in spontaneous activity alone was not able to account for the asymmetry of the projected responses during the post-stimulus silence (Fig. 5b, bottom). The target-driven, asymmetrical projections observed during this trial epoch therefore relied in part on a change in stimulus-evoked responses.

All together, we found that the changes in baseline spontaneous activity induced by the task engagement are key in explaining the enhancement of the target-driven, asymmetric encoding during sound presentation. As described above, the encoding axis during sound presentation is not drastically affected by task engagement. Instead, it is the population spontaneous activity that aligns with the reference-elicited activity with respect to the decoding axis. This observation in particular provides an additional argument against the possibility that the appearance of an asymmetrical representation is due to the asymmetrical motor responses to the two stimuli. Rather, the asymmetry is geometrically explained by baseline changes that precede stimulus presentation and reflects the behavioral state of the animal.

Frontal cortex responses parallel population encoding in A1. The pattern of activity resulting from projecting reference- and

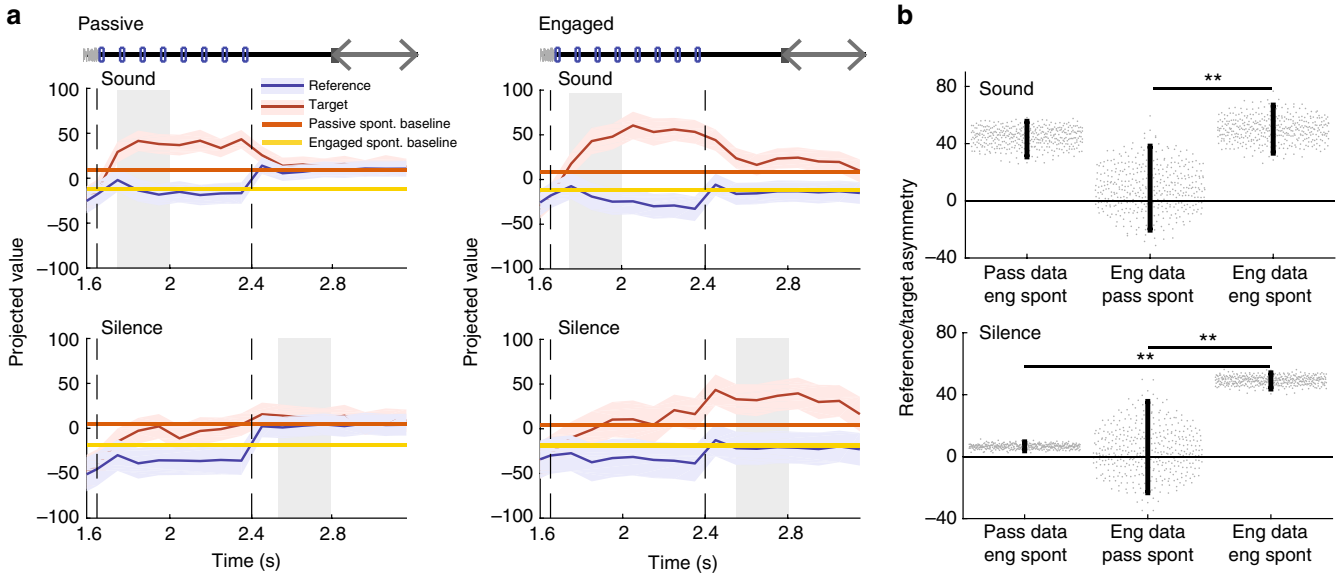


Fig. 5 Shift in spontaneous activity contributes to change in asymmetry. **a** Projection onto the engaged decoding axis of reference- and target-evoked activity in the passive (left column) and engaged state (right column). Decoding axes determined during sound presentation and post-stimulus silence are, respectively, used for projections in the top and bottom rows. This figure differs from Fig. 3c in which the spontaneous activity is subtracted before projection, so 0 corresponds to the null space of the projection. Passive and engaged spontaneous activities after projection are shown by continuous lines. Error bars represent 1 std calculated using decoding vectors from cross-validation ($n = 400$). **b** Comparison of reference/target asymmetry for evoked responses in different states compared to different baselines given by passive or engaged spontaneous activity. Reference/target asymmetry is the difference of the distance of reference and target projected data to a given baseline. We examine three cases: (i) passive evoked responses, distances calculated relative to engaged spontaneous activity; (ii) engaged evoked responses, distances calculated relative to passive spontaneous activity; (iii) engaged evoked responses, distances calculated relative to engaged spontaneous activity. These values are shown during the sound (top) and the silence (bottom). In all three cases, the engaged decoding axis was used for projections. Decoding axes determined during sound presentation and post-stimulus silence are, respectively, used for projections in the top and bottom rows Note that all analysis in this figure is done after excluding lick-responsive units in A1 as described in Fig. 4. Error bars represent 95% confidence intervals ($n = 400$ cross-validations; sound: $p(\text{col1,col3}) = 0.29$ and $p(\text{col2,col3}) < 0.0025$; silence: $p(\text{col1,col3}) < 0.0025$ and $p(\text{col2,col3}) < 0.0025$; $^{**}p < 0.01$)

target-elicited A1 activity on the linear readout is strikingly similar to previously published activity recorded in the dorsolateral frontal cortex (dlFC) of behaving ferrets performing similar Go/No-Go tasks (tone-detect and two-tone discrimination in ref. [32]). We therefore compared in more detail A1 activity with activity recorded in dlFC during the same click-rate discrimination task. When the animal was engaged in the task, single units in dlFC encoded the behavioral meaning of the stimuli by responding only to target stimuli but remaining silent for reference stimuli (Fig. 6a, bottom panel). Target-induced responses were moreover observed well after the end of the stimulus presentation, allowing for a maintained representation of stimulus category. The strong asymmetry of single-unit responses in dlFC clearly resembles the activity extracted from the A1 population by the linear decoder (Figs. 3 and 4). This suggests that the target-selective responses in the dlFC that reflect the cognitive decision process could in part be thought of as a simple readout of information already present in the population code of A1.

To further examine the relationship between dlFC single-unit responses and population activity in A1, we next compared the time course of the projected target-elicited data in A1 (Fig. 3e) and the population-averaged target-elicited neuronal activity in dlFC (Fig. 6a, bottom panel) during engaged sessions. As mentioned above, the optimal decoding axes for A1 activity changes between the stimulus presentation epoch and the silence that follows (Supplementary Fig. 6). The time course of the projected A1 activity depends strongly on the axis used for the projection. When projecting on the axis determined during stimulus presentation, the target-elicited response in A1 was extremely fast (0.08 s +/− 0.009 std) compared to the much longer response latency in the population-averaged response of dlFC neurons (0.48 s +/− 0.12 std) (Fig. 6b). In contrast, when projecting on the axis determined during post-stimulus silence, the target-elicited response in A1 was slower (0.21 s +/− 0.03 std) and closer to the population-averaged response in the dlFC (note that a fraction of individual units in dlFC display very fast responses not reflected in the population average, see Fritz et al.[32]). Our analyses therefore identified two contributions to target-driven population dynamics in A1, a fast component absent in population-averaged dlFC activity and a slower component similar to population-averaged activity in dlFC, thus pointing to a possible contribution of an A1–FC loop that could be engaged during auditory behavior.

Target representation in A1 is a general feature of Go/No-Go tasks. To determine whether the task-related increase in asymmetry between target and reference was a more general feature of primary auditory cortex responses during auditory discrimination, we applied our population analysis to other datasets collected during different tasks. All of these tasks used Go/No-Go paradigms (see Supplementary Fig. 9a, e, I and Methods), in which the animals were presented with a random number of references followed by a target stimulus. In these different datasets, animals were required to discriminate noise bursts vs pure tones (tone-detect tasks) or categorize pure tones drawn from low-, medium-, or high-frequency ranges (frequency range discrimination task). Contrasting datasets were obtained from two groups of ferrets that were separately trained on approach and avoidance versions of the same tone-detect task. These two behavioral paradigms used exactly the same stimuli under two opposite reinforcement conditions[30], requiring nearly opposite motor responses (Supplementary Fig. 9a, e). A crucial feature shared by all these tasks lies in the fact that the behavioral response to the target stimulus always required a behavioral change relative to sustained baseline activity. More specifically, the target was the No-Go stimulus in negative reinforcement tasks and required animals to cease ongoing licking, whereas the target was the Go stimulus in the positive reinforcement task and required animals to begin licking in a non-lick context. In all of the analyses, lick-related neurons were removed using the approach outlined earlier.

Performing the same analyses on all tasks showed that projections of target- and reference-evoked activities in passive conditions contained a variable degree of asymmetry in the sound and silence epochs. However, in all tasks we found that task engagement leads an enhancement of target-driven encoding during sound (Fig. 7a, b; e, f; i, j; m, n). As previously described for the rate discrimination task (Figs. 3 and 4e), target projections more strongly deviated from baseline than projections of reference stimuli in the engaged condition. Moreover, for three of the four tasks we examined, enhancement of target

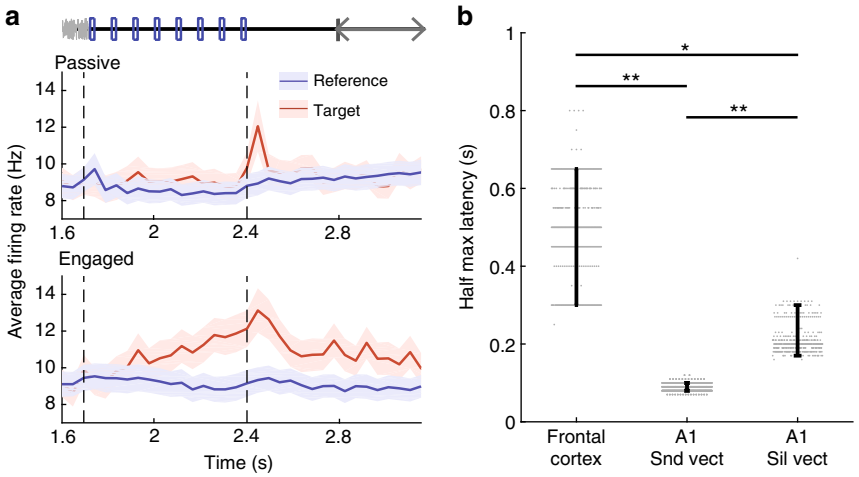
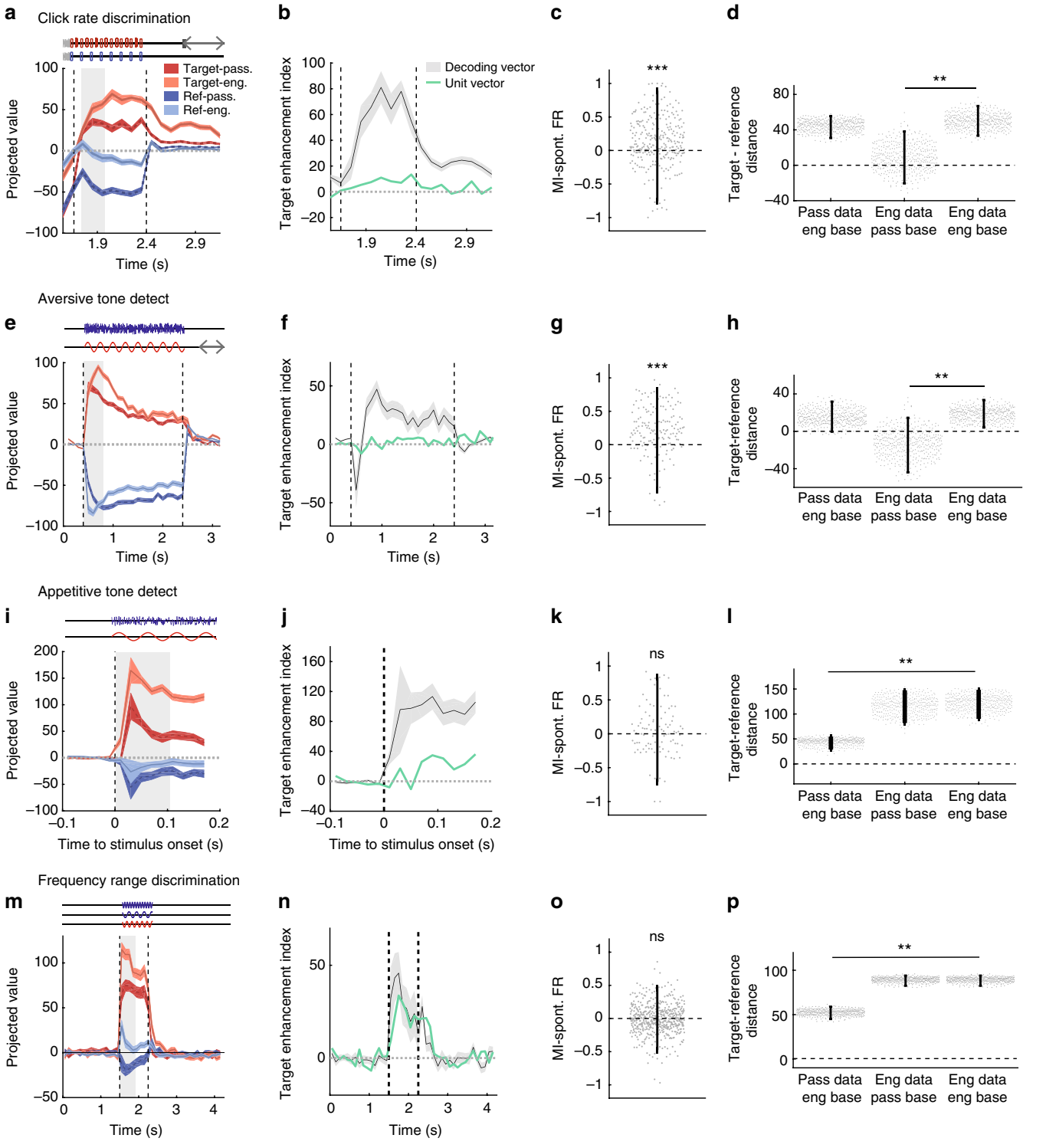


Fig. 6 Persistent, asymmetric response to target and reference stimuli in frontal cortex **a** Average PSTHs of all frontal cortex units in response to target and reference stimuli in both passive and engaged conditions. Note that the response to the target in the task-engaged state is very clear and appears late during the sound. Error bars: SEM over all units ($n = 102$). **b** Latency to half-maximum response for frontal cortex (for average PSTHs) and primary auditory cortex (for projected target-elicited data) in the task-engaged state. For the auditory cortex, data is projected either on the sound decoding vector or the silence decoding vector. Error bars represent 95% confidence intervals. (400 cross-validations. $p \leq 0.0025$, $p \leq 0.0025$ and $p = 0.011$; $^{**}p < 0.01$; $^{*}p < 0.05$). Note that all analysis in this figure is done after excluding lick-responsive units in A1 as described in Fig. 4

representations was not observed at the level of population-averaged responses but only in the direction determined by the decoder (Fig. 7b, f, j, n). During the post-sound silence, decoding accuracy quickly decayed in both passive and engaged states but remained above chance (Supplementary Fig. 9c, g, k). As in the click-train detection task, decoding accuracy relied on a different encoding strategy than the sound period (Supplementary Fig. 9d, h, l), and the asymmetry during the post-sound silence was high both in passive and engaged conditions (Supplementary Fig. 10).

Comparison of appetitive and aversive versions of the same task is particularly revealing as to which type of stimulus was associated with enhanced representation in the engaged state. In the appetitive version of the tone-detect task, ferrets needed to refrain from licking on the reference sounds (No-Go) and started licking the water spout shortly after the target onset (Go) (Supplementary Fig. 9e), whereas in the aversive (conditioned avoidance) paradigm they had to stop licking after the target sound (No-Go) to avoid a shock (Supplementary Fig. 9a). It is important to note that, although the physical stimuli presented to the behaving animals were identical in both tone-detect tasks, the associated motor behaviors of the animals are nearly opposite. Projection of task-engaged A1 population activity reveals a target-driven encoding (compare right panels of Fig. 7f, j with Fig. 7l, j), irrespective of whether the animal needed to refrain from or to



start licking to the target stimulus. This shows that the common feature of stimuli that are enhanced after projection onto the decoding axis is that they are associated with a change of ongoing baseline behavior.

This range of behavioral paradigms provides additional arguments against the described changes in activity being solely due to correlates of licking activity. First, we observed enhanced target-driven encoding in both the appetitive and aversive tone-detect paradigms, even though the licking profiles were diametrically opposite to each other. Second, comparing the projections of the population activity in the approach tone-detect task with the click-rate discrimination task reveals a strong similarity in the temporal pattern of asymmetry observed during task engagement. In <100 ms, projection of target-elicited activity reached its peak in both paradigms (Fig. 7a, i), although the direction and time course of the licking responses were reversed, with a fast decline in lick frequency for the click-rate discrimination task (Fig. 1a), vs a slow increase for the tone detect (Supplementary Fig. 9e, left panel). Last, although the results are more variable partly due to low decoding performance, we observed target-driven encoding during the post-stimulus silence in the passive state (Supplementary Fig. 10) although ferrets were *not* licking during this epoch. The points listed here are again in agreement with a representation of the stimulus' behavioral consequences, independent of the animal motor response.

As pointed out in the case of the click-rate discrimination task, the enhancement of target representation in the engaged condition can rely on two different mechanisms, a shift in the spontaneous activity or a shift in stimulus-evoked activity. We therefore set out to tease apart the respective contributions of the two mechanisms in this novel set of tasks. As in Fig. 5, we compared the distance of target and reference passive and engaged projections to either engaged or passive baseline activities. Out of the three additional datasets, we observed an increase in spontaneous firing rates only in the aversive tone-detect task (Fig. 7g, similar to Fig. 7c). In this latter paradigm, task-induced modulations of spontaneous activity patterns explained the change in asymmetry during sound presentation, similar to what was observed in the click-rate discrimination task (compare Fig. 7d, h). The other two tasks showed no global change of spontaneous firing rate (Fig. 7k, o), and consequently, during the task engagement, the enhancement of the target representation was solely due to the second mechanism, the changes in the target-evoked responses themselves (Fig. 7l, p). During the silence, we observed as previously for the click-rate discrimination that the increase in asymmetry relied only on the second mechanism (Supplementary Fig. 9).

Taken all together, population analysis on four different Go/No-Go tasks revealed an increase of the encoding in favor of the target stimulus as a general consequence of task engagement on A1 neural activity. Viewing activity changes in this light allowed us to interpret the previously observed changes in spontaneous activity as one of two possible mechanisms underlying this task-induced change of stimulus representation in A1 population activity.

Discussion

In this study, we examined population responses in the ferret primary auditory cortex during auditory Go/No-Go discrimination tasks. Comparing responses between sessions in which animals passively listened and sessions in which animals actively discriminated between stimuli, we found that task engagement induced a shift from a sensory driven to an asymmetric, target enhanced, representation of the stimuli, highly similar to the type of activity observed in dlFC during engagement in the same task. This enhanced representation of target stimuli was found in a variety of discrimination tasks that shared the same basic Go/No-Go structure but used a variety of auditory stimuli and reinforcement paradigms.

In the click-rate discrimination task that we analyzed first, the sustained asymmetric stimulus representation in A1 was only observed in the engaged state (Fig. 3). One possible explanation is that this encoding scheme relied on corollary neuronal discharges related to licking activity. However, there are several factors that argue against this interpretation. Firstly we adopted a stringent criterion for the exclusion from the analysis of all units whose activity was correlated with lick events (Fig. 4). After removing lick-responsive units, the results remained unchanged, indicating the absence of a direct link between licking and the observed asymmetry in the encoding. Furthermore, the large differences in the lick profiles between the different tasks were not in line with the remarkably conserved target-driven projections of population activity across tasks and reinforcement types, supporting a non-motor nature of the stimulus encoding in A1 (Fig. 7b, f, j, n). Finally, the role of baseline shifts due to the change in spontaneous activity in two more tasks further argues against a purely motor explanation of the observed asymmetry (Figs. 5 and 7a) since the spontaneous activity occurs during epochs that preceded stimulus presentation and behavioral changes. Altogether, while the different lines of evidence exposed above make an

Fig. 7 Enhanced representation of target stimuli in a range of auditory Go/No-Go tasks. Each line of four panels represent the same analysis for all four tasks; statistics are given in order of appearance in the figure. **a**, **e**, **i**, **m** Projection onto the decoding axis determined during the sound period of trial-averaged reference (blue) and target (ref) activity during the passive (dark colors) and the engaged (light colors) sessions. A baseline value computed from spontaneous activity was subtracted for each neuron, so that the origin corresponds to the projection of spontaneous activity (shown by black line). Note that the target-driven activity is further from the baseline in the engaged state and the reference-driven activity is closer. The periods used to construct the decoding axis are shaded in gray. Error bars represent 1 std (cross-validation $n = 400$). **b**, **f**, **j**, **n** Index of target enhancement induced by task engagement based on projections using the decoding axis determined during the sound. In green, same index computed instead by giving the same weight to all units. The difference between the green and black curved indicates that the change in asymmetry induced by task engagement cannot be detected using the population averaged firing rate alone. Error bars represent 1 std (cross-validation $n = 400$). **c**, **g**, **k**, **o** Modulation index of each unit for spontaneous firing rate after exclusion of lick-related units. Error bars are 95% CI (one-sample two-sided Wilcoxon signed-rank test with mean 0, $n = 277$, zval = 6.35, $p = 2.1e\text{-}10$; $n = 161$, zval = 7.22, $p = 5.4e\text{-}13$; $n = 99$, zval = 1.01, $p = 0.30$; $n = 520$, zval = −0.78, $p = 0.47$; ***$p < 0.001$). **d**, **h**, **l**, **p** Comparison of reference/target asymmetry for evoked responses in different states relative to different baselines given by passive or engaged spontaneous activity. Reference/target asymmetry is the difference of the distance of target and reference projected data to a given baseline. We examine three cases: (i) passive evoked responses, distances calculated relative to engaged spontaneous activity; (ii) engaged evoked responses, distances calculated relative to passive spontaneous activity; (iii) engaged evoked responses, distances calculated relative to engaged spontaneous activity. In all three cases, the engaged decoding axis was used for projections. Error bars represent 95% confidence intervals ($n = 400$ cross-validations; p(col1,col3) = 0.29 and p(col2,col3) < 0.0025; p(col1,col3) = 0.38 and p(col2,col3) < 0.0025; p(col1,col3) < 0.0025 and p(col2,col3) = 0.16; p(col1,col3) < 0.0025 and p(col2,col3) = 0.92; **$p < 0.01$)

interpretation in terms of motor activation unlikely, ultimately a different type of behavioral report, such as the one using similar responses, would help fully rule out this possibility.

Our analyses show that the target-driven representation scheme during task engagement is neither purely sensory nor purely motor but instead argue for a more abstract, cognitive representation of the stimulus behavioral meaning in A1 during task engagement. As the target stimulus was associated with an absence of licking in the tasks under aversive conditioning, one possibility could have been that the A1 encoding scheme was contrasting the only stimulus associated with an absence of licking (No-Go) against all other stimuli (Go). This lick/no-lick encoding was, however, not consistent with the tone-detect task under appetitive reinforcement, in which the target stimulus was a Go signal for the animal. We thus suggest that A1 encodes the behavioral meaning of the stimulus by emphasizing the stimulus requiring the animal to change its behavioral response, i.e., the target stimuli in the different tasks we examined.

Our results critically rely on population-level analyses[33–36], and in particular, on linear decoding of population activity. This is a simple, biologically plausible operation that can be easily implemented by a neuron-like readout unit that performs a weighted sum of its inputs. The summed inputs to this hypothetical readout unit showed that Go and No-Go stimuli elicited inputs symmetrically distributed around spontaneous activity in the passive state. In contrast, in the task-engaged state, only target stimuli, which required an explicit change in ongoing behavior, led to an output different from spontaneous activity, once passed through the readout unit. This switch from a more symmetric, sensory-driven to an increasingly asymmetric, target-driven representation was not clearly apparent if single-neuron responses were simply averaged or normalized (Supplementary Fig. 7, Fig. 7b, f, j, n) but instead relied on a population analysis in which different units were assigned different weights by projecting population activity on the decoding axis. Note that the weights were not optimized to maximize the asymmetry between Go and No-Go stimuli but rather the discrimination between them. The shift toward a more asymmetric representation of the behavioral meaning of stimuli is therefore an unexpected but important by-product of the analysis.

Recordings performed in dlFC in the ferret during tone detection[32] showed that, when the animal is engaged in the task, dlFC single units encode the abstract behavioral meaning of the stimuli by responding only to the target stimuli (that require a change in the ongoing behavioral output) but remain silent for the reference stimuli. Remarkably, projections of reference- and target-elicited A1 activity on the linear readout showed the same type of target-specific patterns of activity. Several possible mechanisms could account for these similarities of representations in A1 and dlFC. Here we propose that, during task engagement, sound-evoked activity in A1 triggers activity in dlFC, which then subsequently feeds back top–down inputs to A1 that may underlie the sustained activity pattern found during post-stimulus silence.

Our analysis suggests a novel population readout mechanism for extracting behaviorally relevant information from A1 while suppressing other, irrelevant sensory information: in the task-engaged state, irrelevant sensory inputs (reference stimuli) elicit changes of activity that are orthogonal to the readout axis and therefore cannot be distinguished from spontaneous activity. This mechanism is reminiscent of the mechanism proposed for movement preparation in motor cortex[37], where preparatory neural activity lies in the null space of the motor readout, i.e., the space orthogonal to the readout of the motor command, and therefore does not generate movements. In our case, the readout is task dependent, as it presumably depends on the performed discrimination task. We showed that the A1 activity in the engaged condition rearranges so that the difference between spontaneous activity and reference-elicited activity lies in the null space of the readout, which is therefore only activated by the target stimuli. This rearrangement can be implemented either by a change of reference-elicited activity or by a change of spontaneous activity. In two of the examined tasks, click discrimination and aversive tone detection, we found that the rearrangement of population activity relied mostly on the change in population spontaneous activity in the engaged condition. Strikingly, these two tasks were performed by the same ferrets, which were trained to switch between the two tasks in the same session. In the two other tasks, reference-elicited activity in the passive condition were already aligned with the passive spontaneous activity when projected on the engaged decoder, suggesting that learning these behavioral tasks may have profoundly reshaped the relation between spontaneous and stimulus-evoked activity. Changes in spontaneous activity have previously been shown to contribute to stimulus responses[33,34,38–40] and task-driven changes have been reported in multiple previous studies[14] but, to our knowledge, have never been given a functional role in stimulus representation[35]. Here we propose that population-level modulations of spontaneous activity act as a mechanism supporting the asymmetric representation of reference and stimuli target in the engaged state.

The simple linear readout mechanism suggested here cannot, however, fully account for the whole set of responses observed in frontal areas as the projections of reference-elicited activity (in A1) during engagement on an aversive task still give rise to a non-null, albeit reduced, output contrary to what is observed in dlFC area recordings. An additional non-linear gating mechanism likely operates between primary auditory cortex and frontal areas, further reducing responses to any stimulus in the passive state and to reference sounds in the engaged state. In particular, neurons in higher-order auditory areas could refine the population-wide, abstracted representation originating in A1 through the proper combinations of synaptic weights. Such a mechanism could also explain why individual single units recorded in belt areas of the ferret auditory cortex show a gradual increase in their selectivity to target stimuli[36].

In summary, we found that task engagement induces a shift from sensory-driven to abstract, behavior-driven representations in the primary auditory cortex. These abstract representations are encoded at a population, but not at a single-neuron level, and strikingly resemble abstract representations observed in higher-level cortices. These results suggest that the role of primary sensory cortices is not limited to encoding sensory features. Instead, primary cortices appear to play an active role in the task-driven transformation of stimuli into their behavioral meaning and the translation of that meaning into task-appropriate motor actions.

Methods

Behavioral training. All experimental procedures conformed to standards specified by the National Institutes of Health and the University of Maryland Institutional Animal Care and Use Committee. Adult female ferrets, housed in pairs in normal light cycle vivarium, were trained during the light period on a variety of different behavioral paradigms in a freely moving training arena. After headpost implantation, the ferrets were retrained while restrained in a head-fixed holder until they reached performance criterion again. Most of the animals in these studies were trained on multiple tasks, including the two ferrets trained both on the click-rate discrimination and the tone-detect tasks. Three out of the four tasks shared the same basic structure of Go/No-Go avoidance paradigms[41], in which ferrets were trained in a conditioned avoidance paradigm to lick water from a spout during the presentation of a class of reference stimuli and to cease licking after the presentation of a different class of target stimuli to avoid a mild shock. The positive reinforcement task is detailed below (see "Tone-detect task—aversive conditioning").

Table 1 Data summary across behavioral tasks

Task	Click-rate discrimination		Tone detect		Frequency range discrimination
Structure	dlFC	A1	A1	A1	A1
Animals	2 ferrets	2 ferrets	2 ferrets	4 ferrets	1 ferret
Conditioning	Aversive	Aversive	Aversive	Appetitive	Aversive
Recorded sessions	—Prepassive —Engaged —Postpassive	—Prepassive —Engaged —Postpassive	—Prepassive —Engaged —Postpassive	—Passive —Engaged	—Prepassive —Engaged —Postpassive
Session num.	25 (17 and 8)	18 (9 and 9)	13 (7 and 6)	56 (8, 37, 2, 9)	149
Recorded units	102 (66 and 36)	370 (188 and 182)	202 (129 and 73)	100 (17, 72, 2, 9)	758

Recordings began once the animals had relearned the task in the holder. Each recording session included epochs of passive sounds presentation without any behavioral response or reinforcement, followed by an active behavioral epoch where the animals could lick. A postpassive epoch was then recorded. This sequence of epochs could be repeated multiple times during a recording session. Table 1 below summarizes the animals and recordings for each task.

Click-rate discrimination task. Two adult female ferrets were trained to discriminate low from high rate click trains in a Go/No-Go avoidance task. A block of trials consisted of a sequence of a random number of reference click train trials followed by a target click train trial (except on catch blocks in which seven reference stimuli were presented with no target). On each trial, the click train was preceded by a 1.25 s neutral noise stimulus (Fig. 1a). Ferrets licked water from a spout throughout trials containing reference click trains until they heard the target sound. They learned to stop licking the spout either during the stimulus or after the target click train ended, in the following 0.4-s time silent response window, in order to avoid a mild shock to the tongue in a subsequent 0.4 s shock window (Fig. 1a). Any lick during this shock window was punished. The ferrets were first trained while freely moving daily in a sound-attenuated test box. Animals were implanted with a headpost when they reached criterion, defined with a discrimination ratio (DR) $>= 0.64$ where $DR = HR \times (1 - FA)$ [hit rate, $HR = 0.8$ and false alarm, $FA = 0.2$]. They were then retrained head fixed with the shocks delivered to the tail. The decision rule was reversed in the two animals, as low rates were Go stimuli for one animal and No-Go for the second one. During each session, rates were kept identical but were changed from day to day.

Tone-detect task—aversive conditioning. The same two ferrets were trained on a tone-detect task previously described[26]. Briefly, a trial consisted of a sequence of 1–6 reference white noise bursts followed by a tonal target (except on catch trials in which 7 reference stimuli were presented with no target). The frequency of the target pure tone was changed every day. The animals learned not to lick the spout in a 0.4 s response window starting 0.4 s after the end of the target. The ferrets were trained until they reached criterion, defined as consistent performance on the detection task for any tonal target for two sessions with >80% hit rate accuracy and >80% safe rate for a DR of >0.65.

Tone-detect task—appetitive conditioning. Four ferrets were on an appetitive version of the tone-detect task previously described[30]. On each trial, the number of references presented before the target varied randomly from one to four. Animals were rewarded with water for licking a water spout in a response window 0.1–1.0 s after target onset. False alarms were punished with a timeout when ferrets licked earlier in the trial before the target window. The average DR during experiments was 0.76. This dataset contained sessions with different trial durations, therefore we analyzed separately data from the first 200 ms after stimulus onset and 200 ms before stimulus offset. For this task, the passive data was not structured in the format of successive reference and target trials as in the engaged session but instead the animal was presented with a block of reference-only trials followed by a block of target-only trials separately. This slight change in the structure of the sound presentation did not affect our results that were highly similar to other tasks but may explain the slightly higher accuracy of decoding during the initial silence in the passive data. Indeed, reference and target trials were systematically preceded by other reference and target trials, possibly allowing the decoder to discriminate using remnant activity from the previous trial.

Frequency range discrimination task. One ferret was trained on a three-frequency-zone discrimination task with a Go/No-Go paradigm. The three frequency zones were defined once and for all and the animal had to learn the corresponding frequency boundaries (Low–Medium: ~500 Hz/Medium–High: ~3400 Hz). Each trial consisted of the presentation of a single pure tone (0.75-s duration) with a frequency in one of the three zones. A trial began when the water pump was turned on and the animal licked a spout for water. The ferret learned to stop licking when it heard a tone falling in the Middle frequency range in order to avoid punishment (mild shock) but to continue licking if the tone frequency fell in either the Low or High range. The shock window started 100 ms after tone offset and lasted 400 ms. The pump was turned off 2 s after the end of the shock window. The learning criterion was defined as DR > 40% in three consecutive sessions of >100 trials.

Acoustic stimuli. All sounds were synthesized using a 44 kHz sampling rate and presented through a free-field speaker that was equalized to achieve a flat gain. Behavior and stimulus presentation were controlled by custom software written in Matlab (MathWorks).

Click-rate discrimination task. Target and reference stimuli were preceded by an initial silence lasting 0.4 s followed by a 1.25 s-long broadband-modulated noise bursts (temporal orthogonal ripple combinations (TORC)[42]) acting as a neutral stimulus, without any behavioral meaning (Fig. 1a). Click trains all had the same duration (0.75 s, 0.8 s interstimulus interval of which the last 0.4 s consisted of the response window) and sound level (70 dB sound pressure level (SPL)). Rates used were comprised between 6 and 36 Hz (ferret A: references [6 7 8 15] Hz, targets [24 26 28 30 32 33 36] Hz/ferret L: references [26 28 30 32 36] Hz, targets [6 8 9 16] Hz).

Tone-detect task. Reference sounds were TORC instances. Targets were comprised of pure tone with frequencies ranging from 125 to 8000 Hz. Target and reference stimuli were preceded by an initial silence lasting 0.4 s. Target and reference stimuli all had the same duration (2 s, 0.8 s interstimulus interval whose last 0.4 s consisted of the response window for the aversive tone-detect task) and sound level (70 dB SPL). In the appetitive version of this paradigm, target and reference duration varied between sessions (0.5–1.0 s, 0.4–0.5-s interstimulus interval).

Frequency range discrimination task. The target frequency region was the Medium range (tone frequencies: 686, 1303, and 2476 Hz) while the reference regions were the Low and High frequency ranges (100, 190, and 361 Hz; 4705, 8939, and 16,884 Hz). Thus the set of tones included 9 frequencies with 90% increment (~0.9 octave) and spanned a ~7.4 octave range. Target and reference stimuli (duration: 0.75 s; level: 70 dB SPL) were preceded by an initial silence lasting 1.5 s and followed by a 2.4 s silence comprising the shock window (400 ms starting 100 ms after the tone offset).

Neurophysiological recordings. To secure stability for electrophysiological recording, a stainless steel headpost was surgically implanted on the skull[26]. Experiments were conducted in a double-walled sound attenuation chamber. Small craniotomies (1–2 mm diameter) were made over primary auditory cortex prior to recording sessions, each of which lasted 6–8 h. The A1 and frontal cortex (dorsolateral FC and rostral anterior sigmoid gyrus) regions were initially located with approximate stereotaxic coordinates and then further identified physiologically. Recordings were verified as being in A1 according to the presence of characteristic physiological features (short latency, localized tuning) and to the position of the neural recording relative to the cortical tonotopic map in A1[43]. Data acquisition was controlled using the MATLAB software MANTA[44]. Neural activity was recorded using a 24 channel Plexon U-Probe (electrode impedance: ~275 kΩ at 1 kHz, 75-µm interelectrode spacing) during the click discrimination task and the aversive version of the tone-detect task. Recordings during the other tasks (frequency range discrimination and appetitive tone-detect task) were done with high-impedance (2–10 MΩ) tungsten electrodes (Alpha-Omega and FHC), using multiple independently moveable electrode drives (Alpha-Omega) to independently direct up to four electrodes. The electrodes were configured in a square pattern with ~800 µm between electrodes. The probes and the electrodes were inserted through the dura, orthogonal to the brain's surface, until the majority of channels displayed spontaneous spiking.

Spike sorting. To measure single-unit spiking activity, we digitized and bandpass filtered the continuous electrophysiological signal between 300 and 6000 Hz. The

Table 2 Summary of time periods used for decoding across behavioral tasks

	Click-rate discrimination	Aversive tone detect	Appetitive tone detect	Frequency range discrimination
Sound	1.7–2 s	0.4–0.8 s	0–0.1 after stim. onset	1.5–1.9 s
Silence	2.5–2.8 s	2.5–2.9 s	0–0.1 after stim. offset	2.4–2.8 s

tail shock for incorrect responses introduced a strong electrical artifact and signals recorded during this period were discarded before processing.

Recordings performed with 24 channel Plextrodes (U-probes) (click discrimination and the tone-detect tasks) were spike sorted using an automatic clustering algorithm (KlustaKwik,[45]), followed by a manual adjustment of the clusters. Clustering quality was assessed with the isolation distance, a metrics developed by Harris et al., 2001, which quantifies the increase in cluster size needed for doubling the number of samples. All clusters showing isolation distance >20 were considered as single units[46,47]. A total of 82 single units and 288 multi-units were isolated. All analyses were reproduced on both pools of units and qualitatively similar results were obtained (see Supplementary Methods). We thus combined all clusters for the analysis. Spike sorting was performed on merged datasets from prepassive, engaged, and postpassive sessions.

For recordings performed with high-impedance tungsten electrodes (frequency range discrimination and relative pitch tasks), single units were classified using principal components analysis (PCA) and *k*-means clustering followed by manual adjustment[26].

Each penetration of the linear electrode array produced a laminar profile of auditory responses in A1 across a 1.8 mm depth. Supragranular and infragranular layers were determined with local field potential responses to 100 ms tones recorded during the passive condition. The border between superficial and middle–deep layer was defined as the inversion point in correlation coefficients between the electrode displaying the shortest response latency and all the other electrodes in the same penetration[48,49].

Click reconstruction from neural data. Optimal prior reconstruction method[50] was used to reconstruct stimulus waveform from click-elicited neural activity. Units with spontaneous firing rate >2 spikes/s in at least one condition were considered for this analysis. Neuronal activity was binned at 10 ms in time with a 1-ms time step. For each trial, we defined $S^k(t)$ the stimulus waveform of trial k ($t \in [1,T]$) and $r_i^k(t)$ the binned firing rate of each neuron $i \in [1,N]$ where $t \in [1,T+\tau]$ with τ the considered delay in the neuronal response. A linear mapping was assumed between the neuronal responses and the stimulus:

$$S^k(t) = \sum_{i=1}^{N} \sum_{\delta=0}^{\tau} g_i(\delta) r_i^k(t+\delta) \quad (1)$$

for unknown coefficients. Equation (1) was rewritten as:

$$S^k = \mathrm{G}R^k \quad (2)$$

with $R^k = \begin{pmatrix} R_1^k \\ R_2^k \\ \vdots \\ R_N^k \end{pmatrix}$ and $R_i^k = \begin{pmatrix} r_i^k(0) & r_i^k(1) & \cdots & r_i^k(T) \\ r_i^k(1) & r_i^k(2) & & r_i^k(T+1) \\ \vdots & & \ddots & \vdots \\ r_i^k(\tau) & r_i^k(1+\tau) & \cdots & r_i^k(T+\tau) \end{pmatrix}$ the lagged neuronal responses, $G = (\,G_1 \quad G_2 \quad \cdots \quad G_N\,)$ and $G_i = (\,g_i(0) \quad g_i(1) \quad \ldots \quad g_i(\tau)\,)$ the corresponding reconstruction filters. The estimate $\hat{G}$ is produced by least-square fitting

$$\hat{G} = S\left(\sum_{k=1}^{K} (R^k)^t\right)\left(\sum_{k=1}^{K} (R^k)^t R^k\right)^{-1} \quad (3)$$

Before the inversion in the previous formula, a single value decomposition was used to eliminate the noisy components of the auto-correlation matrix. The maximal number of components retained was empirically set to 70. Once the values $\hat{G}$ were fitted on all the trials but one, the reconstructed stimulus $\hat{S}^k$ was defined as $\hat{S}^k = \hat{G}R^k$ with the neuronal response R of the remaining run. Each trial was left out in turn. Reconstruction error was quantified with the mean-squared error of the reconstructed stimulus. One passive and engaged reconstruction filters were fitted for each type of stimulus (reference and target) in every session.

Modulation index. To evaluate changes in a given parameter X (firing rate, VS) at the level of the individual unit, we define the modulation index to compare situations 1 and 2 as for each neuron as:

$$\mathrm{MI} = \frac{\mathrm{X}_1 - \mathrm{X}_2}{\mathrm{X}_1 + \mathrm{X}_2}.$$

As a measure of the enhancement of target projection relative to reference projection in the task-engaged state, we used the following index (referred to target enhancement index in the text):

$$\mathrm{MI} = \left(\mathrm{d}\left(\mathrm{Targ}_{\mathrm{eng}}\right) - \mathrm{d}\left(\mathrm{Targ}_{\mathrm{pass}}\right)\right) - \left(\mathrm{d}\left(\mathrm{Ref}_{\mathrm{eng}}\right) - \mathrm{d}\left(\mathrm{Ref}_{\mathrm{pass}}\right)\right)$$

where d is the distance from baseline.

When simply measuring the asymmetry between reference and target in condition X, we used the following index (Figs. 5b, 7d, h, l, and p):

$$\mathrm{Index} = \mathrm{d}\left(\mathrm{Targ}_{\mathrm{X}}\right) - \mathrm{d}\left(\mathrm{Ref}_{\mathrm{X}}\right)$$

Vector strength. VS allows to measure how tightly spiking activity is locked to one phase of a stimulus[51]. If all spikes are at exactly the same phase, VS is 1, whereas if firing is uniformly distributed over phases VS is 0. It is defined in Goldberg and Brown (1969) as:

Significance was assessed using Rayleigh's statistic, $p = e^{nr2}$, where r is the VS and used $p < 0.001$ as the criterion for significant phase locking consistent with previous work[52].

Linear discriminant classifier performance. To evaluate the accuracy with which single-trial population responses could be classified according to the presented stimulus (reference or target), we trained and tested a linear discriminant classifier[53,54] using cross-validation (Supplementary Fig. 3).

Trial-by-trial pseudo-population firing rate vectors were constructed for each 100 ms time bin using units from all sessions and both animals. Training and testing sets were constructed by randomly selecting equal numbers (15) of reference and target trials for each unit. All contribution of noise correlations among neurons are therefore destroyed by this procedure as the pseudo-population vector contains activity of units recorded on different days and on different trials. The classifier was trained for each time bin using the average pseudo-population vectors $c_{R,t}$ and $c_{T,t}$ calculated from a random selection of an equal number of reference and target trials. These vectors define at time bin t the decoding vector w_t given by

$$w_t = c_{T,t} - c_{R,t}$$

and the bias b_t given by

$$b_t = \frac{-\left(c_{R,t} \times w_t + c_{T,t} \times w_t\right)}{2}$$

The decoding vector and and the bias b_t define the decision rule for any population activity vector x:

$$y(x) = w_t^T \times x + b_t$$
$$y(x) > 0,\ x \text{ is classifiedas a target}$$
$$y(x) < 0,\ x \text{ is classifiedas a reference}$$

This rule was applied to an equal number of reference and target testing trials drawn from the remaining trials that were not used to train the classifier. The proportion of correctly classified trials gave the accuracy of the classifier. Cross-validation was performed 400 times by randomly picking training and testing data to estimate the average and variance of accuracy. This allowed comparing the performance of classification in two behavioral states by constructing confidence intervals from the cross-validation. Note that this limits p value estimate to a minimum of $1/400 = 0.0025$.

Random performance. To evaluate whether the classifier performance is higher than chance, the classifier was trained and tested on surrogate datasets constructed

by shuffling the labels ("reference" and "target") of trials. For each of the 100 label permutations, cross-validation was performed 100 times. This allows comparing the performance of classification with chance levels by constructing confidence intervals from the cross-validation and from the random shuffled permutations.

Classifier evolution. When studying the evolution of population encoding (Supplementary Fig. 6), we defined early sound, late sound, and silence periods as 1700–1900, 2200–2400, and 2700–2900 ms (equal duration for comparison) relative to trial onset. The classifier was trained on randomly chosen trials from one time period and then tested on trials at all other 100 ms time bins. We also constructed matrices showing the accuracy of the classifier trained and tested at all 100 ms time bins and evaluated whether these values are higher than chance using surrogate datasets by shuffling labels as described above.

When comparing the classifier during sound and silence periods across tasks (Fig. 7), the time periods summarized in Table 2 were used.

Projection onto decoding vectors. To study the contribution of reference and target trials to classifier performance, we projected population firing vectors at each time bin onto decoding vectors calculated during the sound and silence periods as defined above. Before projection, the mean spontaneous activity of each unit was subtracted from its firing rate throughout the whole trial. Deviations from 0 of the projection show activity deviating from spontaneous activity along the decoding axis.

Controlling for lick-responsive neurons. In order to control for the contribution of units directly linked with task-related motor activity to our results, we combined reconstruction and decoding methods to identify and remove lick-responsive neurons so that linear classification no longer yielded any licking-related information. The approach comprised the following steps:

- Optimal prior reconstruction (described in "Click reconstruction from neural data") was used to reconstruct lick-activity separately for each unit.
- Reconstruction values for each unit were then sampled at the time of licks and at randomly selected times without licking. These values were used to construct population vectors of lick and non-lick activity.
- A linear classifier (described in "Linear discriminant classifier performance") was trained and tested using cross-validation to distinguish lick from non-lick events.
- Reconstruction values and classification was also performed on random data obtained by reconstructing the licking activity of a session with the neural activity of a subsequent session. This made it possible to establish the distribution of accuracy for randomized data.
- The accuracy of classification was compared between the true data and the randomized datasets and a p value was calculated by counting the number of permutations showing better accuracy for the randomized data than the true data.
- We progressively removed units, starting with those with highest classifier weights, which reduced the accuracy of classification, until the p value of population classification rose >0.4. This indicated that the remaining units contained no more information about lick events than randomized data.
- Only the units remaining after this procedure were used to re-analyze the data and verify that reliable classification and difference in projections of reference and tone trials did not rely on the difference in licking activity between the two trials.

For the click-rate discrimination task, only a subset of sessions (15/18) had reliable recordings of all lick events, so the analysis was done on 308 units (not 370), and 277 units were identified as non-lick related. For the appetitive tone task, 99/100 units, for the aversive tone task 161/202 and for the frequency range discrimination 520/758.

Gaussian-process factor analysis. To visualize neural trajectories of the large population of units recorded in A1, we used Gaussian-process factor analysis as described in ref. [55]. This method has the advantage over more traditional methods of dimensionality reduction such as PCA of jointly performing both the binning/smoothing steps and the dimensionality reduction.

Statistics. Statistics on classifier performance relied on p value estimation using cross-validation. For each statistical analysis provided in the manuscript, Kolmogorov–Smirnov normality test was first performed on the data. As the data failed to meet the normality criterion, statistics relied on non-parametric tests. When performing systematic multiple tests, the Bonferroni correction was applied. Data analyses were performed in MATLAB (Mathworks, Natick, MA, USA).

Code availability. Code used in the article can be supplied upon request by writing to the corresponding author.

Data availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 11 March 2018 Accepted: 22 May 2018
Published online: 28 June 2018

References

1. Chechik, G. et al. Reduction of information redundancy in the ascending auditory pathway. *Neuron* **51**, 359–368 (2006).
2. Chechik, G. & Nelken, I. Auditory abstraction from spectro-temporal features to coding auditory entities. *Proc. Natl Acad. Sci. USA* **109**, 18968–18973 (2012).
3. de Lafuente, V. & Romo, R. Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc. Natl Acad. Sci. USA* **103**, 14266–14271 (2006).
4. Siegel, M., Buschman, T. J. & Miller, E. K. Brain processing. Cortical information flow during flexible sensorimotor decisions. *Sci. (80-.).* **348**, 1352–1355 (2015).
5. Vergara, J., Rivera, N., Rossi-Pool, R. & Romo, R. A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* **89**, 54–62 (2016).
6. D'Esposito, M. & Postle, B. R. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* **66**, 115–142 (2015).
7. de Lafuente, V. & Romo, R. Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* **8**, 1698–1703 (2005).
8. Lemus, L., Hernández, A. & Romo, R. Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proc. Natl Acad. Sci. USA* **106**, 9471–9476 (2009).
9. Yildiz, I. B., Mesgarani, N. & Deneve, S. Predictive ensemble decoding of acoustical features explains context-dependent receptive fields. *J. Neurosci.* **36**, 12338–12350 (2016).
10. Sloas, D. C. et al. Interactions across multiple stimulus dimensions in primary auditory cortex. *eNeuro* **3**, 1–7 (2016).
11. Bizley, J. K., Walker, K. M. M., Nodal, F. R., King, A. J. & Schnupp, J. W. H. Auditory cortex represents both pitch judgments and the corresponding acoustic cues. *Curr. Biol.* **23**, 620–625 (2013).
12. Niwa, M., Johnson, J. S., O'Connor, K. N. & Sutter, M. L. Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. *J. Neurosci.* **32**, 9323–9334 (2012).
13. Downer, J. D., Niwa, M. & Sutter, M. L. Task engagement selectively modulates neural correlations in primary auditory cortex. *J. Neurosci.* **35**, 7565–7574 (2015).
14. Otazu, G. H., Tai, L.-H., Yang, Y. & Zador, A. M. Engaging in an auditory task suppresses responses in auditory cortex. *Nat. Neurosci.* **12**, 646–654 (2009).
15. Brosch, M. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.* **25**, 6797–6806 (2005).
16. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
17. Schneider, D. M., Nelson, A. & Mooney, R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–194 (2014).
18. Zhou, M. et al. Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.* **17**, 841–850 (2014).
19. Rodgers, C. C. & DeWeese, M. R. Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. *Neuron* **82**, 1157–1170 (2014).
20. Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y. & Petersen, C. C. H. Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat. Neurosci.* **16**, 1671–1677 (2013).
21. Shuler, M. G. & Bear, M. F. Reward timing in the primary visual cortex. *Science* **311**, 1606–1610 (2006).
22. Petreanu, L. et al. Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489**, 299–303 (2012).
23. Ohl, F. W., Scheich, H. & Freeman, W. J. Change in pattern of ongoing cortical activity with auditory category learning. *Nature* **412**, 733–736 (2001).
24. Quirk, G. J., Armony, J. L. & LeDoux, J. E. Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* **19**, 613–624 (1997).
25. Kuchibhotla, K. V. et al. Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.* **20**, 62–71 (2017).
26. Fritz, J. B., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* **6**, 1216–1223 (2003).
27. Fritz, J. B., Elhilali, M. & Shamma, S. A. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. *J. Neurosci.* **25**, 7623–7635 (2005).
28. Fritz, J. B., Elhilali, M. & Shamma, S. A. Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J. Neurophysiol.* **98**, 2337–2346 (2007).
29. Atiani, S., Elhilali, M., David, S. V., Fritz, J. B. & Shamma, S. A. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* **61**, 467–480 (2009).

30. David, S. V., Fritz, J. B. & Shamma, S. A. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl Acad. Sci. USA* **109**, 2144–2149 (2012).
31. Yin, P., Johnson, J. S. & Sutter, M. L. Coding of amplitude modulation in primary auditory cortex. *J. Neurophysiol.* **105**, 582–600 (2010).
32. Fritz, J. B., David, S. V., Radtke-Schuller, S., Yin, P. & Shamma, S. A. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat. Neurosci.* **13**, 1011–1019 (2010).
33. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).
34. Carcea, A. I., Insanally, M. N. & Froemke, R. C. Dynamics of cortical activity during behavioral engagement and auditory perception. *Nat. Commun.* **8**, 1–12 (2017).
35. Driver, J. & Frith, C. Shifting baselines in attention research. *Nat. Rev. Neurosci.* **1**, 147–148 (2000).
36. Atiani, S. et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* **82**, 486–499 (2014).
37. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
38. Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* **273**, 1868–1871 (1996).
39. Luczak, A., Bartho, P. & Harris, K. D. Gating of sensory input by spontaneous cortical activity. *J. Neurosci.* **33**, 1684–1695 (2013).
40. Tatti, R. & Maffei, A. Synaptic dynamics: How network activity affects neuron communication. *Curr. Biol.* **25**, R278–R280 (2015).
41. Heffner, H. E. & Heffner, R. S. in *Methods in Comparative Psychoacoustics* (eds Klump, G. M., Dooling, R. J., Fay, R. R. & Stebbins, W. C.) 79–93 (Birkhäuser Verlag, Basel, 1995).
42. Klein, D. J., Depireux, D. A., Simon, J. Z. & Shamma, S. A. Robust spectrotemporal and reverse correlation and for the auditory and system: and optimizing stimulus and design. *J. Comput. Neurosci.* **9**, 85–111 (2000).
43. Shamma, S. A., Fleshman, J. W., Wiser, P. R. & Versnel, H. Organization of response areas in ferret primary auditory cortex. *J. Neurophysiol.* **69**, 367–383 (1993).
44. 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).
45. Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H. & Buzsáki, G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* **84**, 401–414 (2000).
46. Belliveau, L. A. C., Lyamzin, D. R. & Lesica, N. A. The neural representation of interaural time differences in gerbils is transformed from midbrain to cortex. *J. Neurosci.* **34**, 16796–16808 (2014).
47. Garcia-Lazaro, J. A., Shepard, K. N., Miranda, J. A., Liu, R. C. & Lesica, N. A. An overrepresentation of high frequencies in the mouse inferior colliculus supports the processing of ultrasonic vocalizations. *PLoS ONE* **10**, e0133251 (2015).
48. Kajikawa, Y. & Schroeder, C. E. How local is the local field potential? *Neuron* **72**, 847–858 (2011).
49. Linden, J. F. & Schreiner, C. E. Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. *Cereb. Cortex* **13**, 83–89 (2003).
50. Mesgarani, N., David, S. V., Fritz, J. B. & Shamma, S. A. Influence of context and behavior on stimulus reconstruction from neural activity in primary auditory cortex. *J. Neurophysiol.* **102**, 3329–3339 (2009).
51. Goldberg, J. M. & Brown, P. B. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* **32**, 613–636 (1969).
52. Gao, X. & Wehr, M. A coding transformation for temporally structured sounds within auditory cortical neurons. *Neuron* **86**, 292–303 (2015).
53. Bishop, C. M. Pattern recognition and machine learning. *Pattern Recognit.* **4**, 180–183 (2006).
54. Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K. & Poggio, T. Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* **100**, 1407–1419 (2008).
55. Yu, B. M. et al. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* **102**, 614–635 (2009).

Author contributions

S.B., M.A., Y.B. and S.O. designed and carried out the analyses and statistical testing. J.F. and S.S. designed the experiments. D.E., S.D., J.F. and P.Y. recorded the data. P.Y. and J. F. provided feedback on the manuscript. S.B., S.S., Y.B. and S.O. wrote all aspects of the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04839-9.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018