Auditory training alters the cortical representation of both learned and task irrelevant sound features

Huriye Atilgan[1,2], Kerry M Walker[2], Andrew J. King[2], Jan W. Schnupp[2,3] and Jennifer K. Bizley[1,2]

[1] The Ear Institute, University College London, UK

[2] Department of Physiology, Anatomy and Genetics, University of Oxford, UK

[3] City University of Hong Kong, China

Corresponding Author: Jennifer Bizley j.bizley@ucl.ac.uk

UCL Ear Institute, 332 Gray's Inn Road, London, WC1X 8EE.

Abstract

Auditory learning is supported by long-term changes in the neural processing of sound. We mapped neural sensitivity to timbre, pitch and location in animals trained to discriminate the identity of artificial vowels based on their spectral timbre in a two-alternative forced choice (T2AFC, n=3, female ferrets) or to detect changes in fundamental frequency or timbre of repeating artificial vowels in a go/no-go task (n=2 female ferrets). Neural responses were recorded under anaesthesia in two primary cortical fields and two tonotopically organised non-primary fields. Responses were compared these data to that of naïve control animals. We observed that in both groups of trained animals the overall sensitivity to sound timbre was reduced across three cortical fields but enhanced in non-primary field PSF. Neural responses in trained animals were able to discriminate vowels that differed in either their first or second formant frequency unlike control animals whose sensitivity was mostly driven by changes in the second formant. Neural responses in the T2AFC animals, who were required to generalise across pitch when discriminating timbre, became less modulated by fundamental frequency, while those in the go/no-go animals were unchanged relative to controls. Finally, both trained groups showed increased spatial sensitivity and altered tuning. Trained animals showed an enhanced representation of the midline, where the speaker was located in the experimental chamber. Overall, these results demonstrate training elicited widespread changes in the way in which auditory cortical neurons represent complex sounds with changes in how both task relevant and task-irrelevant features were represented.

Introduction

Sensory discrimination tasks are known to drive cortical plasticity, and increases in map area have been proposed as providing the structural substrate for learning in auditory cortex (Rutkowski and Weinberger, 2005; Schreiner and Polley, 2014). However, recent studies have questioned the functional role of map plasticity suggesting that it may be a temporary phenomena associated with learning that does not persist once a task is well-learned (Reed et al., 2011), and demonstrating that auditory learning can occur in the absence of map plasticity (Galindo-Leon et al., 2009; Shepard et al., 2016). Training animals to ‘forage’ for sounds with particular features leads to diverse changes within primary auditory cortex independently of any changes in the tonotopic representation (Bao et al., 2004; Whitton et al., 2014). Since one role for auditory cortex is thought to be in the formation of behaviourally meaningful sound categories (Bizley and Cohen, 2013) a question of great interest is to understand the cortical representation is changed in order to support the categorization of spectrally overlapping naturalistic sounds.

Identifying auditory ‘objects’ requires that animal and human listeners are capable of discriminating sounds along a given perceptual dimension while generalizing across variability in other dimensions (Griffiths and Warren, 2004; Bizley and Cohen, 2013). At the level of the single neuron this requires that neuronal responses are both selective for one sound feature but tolerant (or invariant) across others (Ison and Quiroga, 2008; Bizley et al., 2009). While the consequences of behavioural training on neural invariance is unknown, appropriate environmental exposure during development can shape auditory cortical responses to complex sound features: In the auditory cortical neurons of animals reared in complex acoustic environments fewer neurons respond to any single sound but responses were more selective for particular spectro-temporal features and can tolerate greater acoustic variability(Bao et al., 2013).

In this study we recorded from the auditory cortex of animals trained in one of two behavioural paradigms that required that animals discriminate perceptual features of spectrally overlapping artificial vowels. One set of animals was trained to discriminate the timbre of artificial vowels (Bizley et al., 2013; Town et al., 2018) and did so across a range of fundamental frequencies (F0s). A second set of animals were trained to detect changes in pitch and timbre of an ongoing sound sequence (Walker et al., 2017).

After behavioral training was complete electrophysiological recordings were made from four tonotopic auditory cortical fields. The perceptual features of complex sounds, such as their location in space or spectral timbre, are distributed rather than systematically mapped within or across auditory cortical fields (Bizley and Walker, 2010; Recanzone and Cohen, 2010; King and Middlebrooks, 2011). Given this, and the broadband nature of the stimuli, it seemed unlikely that learning could be supported by an expansion of the tonotopic map. We therefore sought to determine how training altered (i) single neuron response sensitivity and selectivity to both learned and task-irrelevant sound features, and (ii) whether training impacted on the distribution of neural tuning across specific auditory cortical fields. We tested two hypotheses: Firstly, that training would increase the sensitivity to trained features and the tolerance for changes in untrained features. Secondly, consistent with hierarchical encoding theories, we hypothesized that training effects would be strongest in non-primary areas. Our data demonstrate that the representation of both trained and task-irrelevant features was changed by training, but that the direction of these changes varied between cortical fields: increase in selectivity to the trained features occurred specifically in the secondary posterior pseudosylvian field (PPF) whereas neurons in other fields became less sensitive.

Materials & Methods

Animals

All animal procedures were approved by the local animal welfare and ethical review committee at the University of Oxford and performed under license from the UK Home Office in accordance with the Animal (Scientific Procedures) Act 1986. Five adult, female, pigmented ferrets (*Mustela putorius*) were used in this study. Three of these animals experienced 1-2 years of training on a two alternative forced-choice timbre discrimination task which required they report the identity of an artificial vowel(Fig. 1C, T-2AFC; for the details of behavioral training see {Bizley, 2013 #925}). Two animals were trained to perform a Go/No-go change detection task, where animals were presented with a sequence of artificial vowels and had to report (in different sessions) either changes in F0 or changes in vowel identity (Fig.1D, T/P-GNG; see {Walker, 2011 #93}).

Ferrets were housed in groups of either two or three, with free access to high-protein food pellets and water bottles. On the day before behavioural training, water bottles were removed from the home cages and were replaced on the last day of a training run. Training runs lasted for 5 days or less, with at least 2 days between each run. On training days, ferrets received drinking water as positive reinforcement while performing a sound discrimination task. Water consumption during training was measured and supplemented as wet food in home cages at the end of the day to ensure that each ferret received at least 60 ml of water per kilogram of body weight daily. Once behavioural training was complete electrophysiological recordings were made under non-recovery anesthesia. Recording under aneasthsia was necessary for the large scale mapping of neurons across cortical fields, and in order to directly compare the resulting responses with data from control animals previously collected under the same aneasthetic regime. Electrophysiological control data was provided from 5 additional ferrets that formed a previously published dataset {Walker, 2011 #1485;Bizley, 2009 #607}. All animals were routinely monitored to ensure their ears were clean and disease free throughout the study.

Stimuli

Acoustic stimuli for both behavioral testing and electrophysiology were artificial vowel sounds. For electrophysiological testing, sounds were all possible combinations of four F0 values: F0 = 200, 336, 565, and 951 Hz, four spectral timbres: /a/ (F1–F4 at 936, 1551, 2815, and 4290 Hz); /ɛ/ (730, 2058, 2979, and 4294 Hz); /u/ (460, 1105, 2735, and 4115 Hz); and /i/ (437, 2761, 3372, and 4352 Hz) and four spatial locations presented in virtual acoustic space: -45°, -15°, 15°, and 45° azimuth, at 0° elevation; giving a total of 64 sounds, each of which was 150 ms in duration. Additionally, noise bursts and pure tones were used to characterize individual units and to determine tonotopic gradients in order to confirm the cortical field in which any given recording was made {Bizley, 2005 #602}.

For animals trained in the 2AFC timbre identification task stimuli were /u/ and /ɛ/. Animals were trained initially with an F0 of 200 Hz, but tested across a range of values from 150 Hz – 500 Hz. For animals trained in the T/P GNG task reference sounds were the vowel /a/ at 200 Hz with F0 targets being the vowel /a/ with F0 values of 336, 565, and 951 Hz and timbre targets being the vowels /i/, /u/ and /ɛ/ presented with an F0 of 200 Hz.

Electrophysiological Recordings

Experimental methods were identical to those used in Bizley et al., 2009. Recordings were made under medetomidine/ketamine anaesthesia in the left auditory cortex and targeted at primary and non-primary tonotopic areas: primary auditory cortex (A1) and the anterior auditory field, (AAF), on the middle ectosylvian gyrus and the posterior pseudosylvian and posterior suprasylvian fields (PPF and PSF) located on the posterior ectosylvian gyrus (Fig. 2A). Recordings were made with silicon probe electrodes (Neuronexus Technologies, USA) either in a 16 x2 configuration (16 active sites spaced at 100µm intervals on each of two probes), a 32 x 1 configurations (50 µm spacing) or, in one animal, an 8 x 4 configuration (100µm spacing in depth, 200µm between shanks). Voltage signals were bandpass filtered (500-5000Hz), amplified (<20,000 times), and digitized at 25 kHz. Data acquisition and stimulus generation were performed using BrainWare (Tucker-Davis Technologies). A minimum of 80 units were recorded in each field in each dataset (mean = 118 units ± 28.32, Table 1). Data were combined across animals (control or trained) to make composite tonotopic maps (Fig. 1B). The distribution of CFs were not significantly different between trained and naïve animals. (2 way mixed ANOVA (2 groups x 4 fields), group: F (1, 612) = 0.02, p=0.8792; field: F (3,612) = 39.79, p<0.001; group×field: F (3,612) = 4.67, p =0.08).

Behavioural Testing

Full details of the training apparatus and procedure for shaping animals can be obtained in {Bizley, 2013 #925;Walker, 2011 #93}. Briefly, water-restricted ferrets were positively conditioned to report the identity of a vowel sound (either /e/ or /i/) in a two-alternative forced choice task, or trained detect a change in the pitch or timbre of a repeating artificial vowel on a go/no-go task. In each case, the animal initiated each trial by inserting its nose in a poke hole situated at the center of the sound-isolated testing chamber. For the 2AFC task this resulted in the presentation of two repetitions of one of the vowel sounds and animals were rewarded for correctly responding at the side that was associated with that vowel. In the GNG task ferrets heard a sequence of artificial vowels which could change in identity or pitch at the third to seventh vowel in the sequence, and if ferrets withdrew from the nose poke hole during presentation of such a deviant, they were rewarded with water. Failures to withdraw to a deviant (within a 550 ms time window following deviant onset) resulted in a 12 s time out. In both tasks sounds were presented from a speaker located above the central 'go' spout at the animals midline.

Neural Data Analysis

Neural data was analysed using a variance decomposition approach developed in {Bizley, 2009 #607}. We first calculated spike counts for each of the 64 stimuli, averaged over repeated presentations of the same sound and binned with 20 ms resolution over the 300 ms immediately preceding stimulus onset. We then performed a 4-way ANOVA on the spike counts, where the 3 stimulus parameters (azimuth, pitch, and timbre) plus the time bin served as factors. To quantify the relative strength with which one of the three stimulus dimensions influenced the firing of a particular unit, we calculated the proportion of variance explained by each of azimuth, pitch, and timbre, $\mathrm{Var}_{\mathrm{stim}}$, as:

$$\mathrm{Var}_{\mathrm{stim}} = \frac{\mathrm{SS}_{\mathrm{stim\,.bin}} - \mathrm{SS}_{\mathrm{error}} \,.\, \mathrm{df}_{\mathrm{stim\,.bin}}}{\mathrm{SS}_{\mathrm{total}} - \mathrm{SS}_{\mathrm{bin}}}$$

where "stim" refers to the stimulus parameter of interest (pitch, timbre, or azimuth), $\mathrm{SS}_{\mathrm{stim\,.bin}}$ is the sum of squares for the interaction of the stimulus parameter and time bin, $\mathrm{SS}_{\mathrm{error}}$ is the sum of squares of the error term, $\mathrm{df}_{\mathrm{stim\,.bin}}$ refers to the degrees of freedom for the stimulus x time bin interaction, $\mathrm{SS}_{\mathrm{total}}$ is the total sum of squares, and $\mathrm{SS}_{\mathrm{bin}}$ is the sum of squares for the time bin factor. A significant $\mathrm{SS}_{\mathrm{bin}}$ reflects the fact that the response rate was not flat over the duration of the 300 ms response window and by examining the stimulus-by- time-bin interactions, we were able to test the statistical significance of the influence a given stimulus parameter had on the temporal discharge pattern of the response. Subtracting the $\mathrm{SS_{error}} \, . \, \mathrm{df_{stim\,.bin}}$ from the $\mathrm{SS_{stim\,.bin}}$ term allows us to calculate the proportion of response variance attributable to each of the stimuli, taking into account the additional variance explained simply by adding extra parameters to the model. As in our previous work, we considered a main effect or interaction term in the ANOVA to be statistically significant if it exceeded p <0.001.

Experimental Design and Data Analysis

Data from 5 control animals and 5 experimental animals were used in this experiment. This allowed us to have a sufficient sample size of recordings from each cortical field (see Table 1) and to match the frequency tuning distribution of our samples.

For statistical comparison of neural tuning measures (derived using the variance decomposition approach described above) we used Analysis of Variance, with the specific analysis for each test being reported in the Results section. Generalised Linear Mixed Models were fit using maximum likelihood approaches with model selection performed using Akaike Information Criteria values. Where appropriate Bonferoni corrected post-hoc comparisons were applied (p<0.05). The specific analysis for each experimental question is detailed in full in the Results section.

Code and data availability

All code is available on Github: https://github.com/huriyeatg/trainingInducedPlasticity

Data are available on the following repository [link added on publication].

Table 1: Total number of recordings (probe placements and units) in each field for 5 control animals and 5 trained animals.

	A1	*AAF*	*PPF*	*PSF*
Recordings from Control Animals				
Probe Placement(n)	*10*	*7*	*11*	*7*
Units (n)	*189*	*101*	*152*	*96*
Recordings from Trained Animals				
Probe Placement(n)	*14*	*9*	*7*	*11*
Units (n)	*133*	*80*	*115*	*145*

Results

Three ferrets were trained in a two alternative forced choice timbre discrimination (T-2AFC) task to discriminate /u/ from /ɛ/, across a range of F0s (Fig.1C). Two ferrets were trained in a go/no-go task to detect changes in timbre or F0 (T/P-GNG) from a repeating reference vowel (/a/ F0 = 200 Hz, Fig.1D). Once behavioural training and testing was complete we recorded neural activity under medetomidine/ketamine anesthesia, which allowed us to map neural responses across the cortical surface of multiple cortical fields in each animal, and directly compare these data to those obtained in naïve animals in a previous study that constitute the control data for this investigation. The responses of 713 units (459 single neurons, 254 small clusters of units) which were responsive (paired t-test on sound-evoked and spontaneous firing rates, *$p<0.05$*) to vowels were recorded from four tonotopic auditory cortical fields (see Table 1). There were no systematic differences between the response properties of single units and small unit clusters therefore the term ‘units’ will be used to refer to both except when explicitly stated.

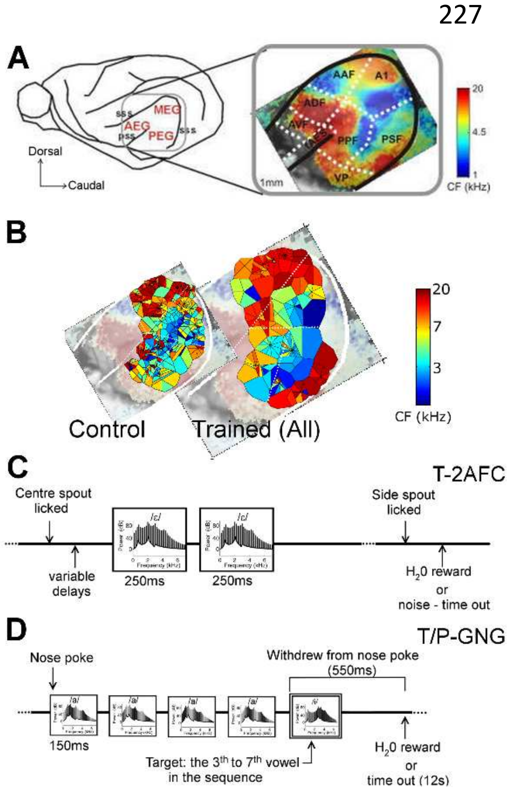

Figure 1:
A, Location of ferret auditory cortical fields and their tonotopic organisation (adapted from (Nelken et al., 2004)). Recordings in this study targeted A1, AAF, PPF and PSF. Field boundaries are marked with dotted lines, and the pseudosylvian sulcus (pss) and suprasylvian sulcus (sss) are drawn as solid lines. **B**, Voronoi tessellation map showing the CFs of all unit recordings made in control animals (478 units, 5 animals) and trained animals (456 units, 5 animals). Tiles represent a recording site and are coloured according to the characteristic frequency (CF) of the unit recorded there. **C**, Schematic illustration of the timbre two-alternative forced choice paradigm (C, T-2AFC) and F0 / timbre change detection paradigm (D, T/P-GNG).

Response modulation by timbre, pitch and azimuth

In order to test our first hypothesis, that we would see an increase in both sensitivity and invariance in control animals, we determined the proportion of units whose responses were significantly modulated by variation in stimulus location, pitch (determined by fundamental frequency, F0) and timbre, using the variance decomposition approach used in Bizley et al. (2009). Since all five animals trained in a timbre discrimination task, we predicted a greater number of neurons might convey timbre information, and that we might observe fewer neurons that were additionally sensitive to untrained stimulus features (pitch in the 2AFC animals and azimuth in 2AFC and T/P GNG animals). However, contrary to these predictions, the number of units that showed timbre sensitivity was equivalent between trained and control groups (391/713 units and 293/538 units respectively, X^2 = 0.005, p=0.95). In both datasets units with joint stimulus sensitivity outnumbered those with sensitivity to only a single parameter, but the distribution of sensitivity to zero, one, two or three stimulus parameters was significantly different between the groups (X^2 =86.9, p<0.001). In the trained dataset many more units showed significant modulation by all three stimulus parameters, or by none of the stimulus parameters, than in the control dataset where most units were modulated by one or two stimulus dimensions. Taken together this suggests that training resulted in a relative increase in the number of units either insensitive to modulations in pitch, timbre or space, or sensitive to all three dimensions. We now consider sensitivity to each feature in turn, asking how training changes the responses of single units, and the cortical distribution of sensitivity.

Table 2: Percentages of recordings modulated by stimuli dimensions (azimuth, F0 and timbre) for naïve, T-2AFC trained and T/P-GNG trained animals

	Control Animals	**T-2AFC trained**	**T/P-GNG trained**
One Stimulus	**24% (120)**	**16% (103)**	**14% (11)**
Only Azimuth	*5% (29)*	*4% (24)*	*3% (3)*
Only F0	*6% (30)*	*7% (47)*	*6% (6)*
Only Timbre	*13% (71)*	*5% (32)*	*5% (5)*
Two Stimuli	**38% (203)**	**16% (92)**	**31% (32)**
Azimuth-F0	*2% (10)*	*6% (30)*	*22% (23)*
Azimuth-Timbre	*7% (35)*	*5% (26)*	*6% (6)*
F0-Timbre	*29% (158)*	*6% (36)*	*3% (3)*
Three Stimuli	**30% (159)**	**44% (272)**	**11% (11)**
Not significant	**13% (70)**	**24% (146)**	**34% (36)**
Total Units	538	607	106

Neural sensitivity to trained stimulus features: Timbre

For each unit we determined what proportion of the response variance was attributable to each of the three stimulus dimensions, and their combinations. We start by considering sensitivity to timbre, as both behavioural tasks required that animals discriminated this feature (Fig. 2A, T-2AFC animals, Fig. 2B, T-GNG). Figure 2C shows the distribution of sensitivity to timbre (i.e. the proportion of variance explained by timbre) across the four cortical fields examined and illustrates that while the proportion of units with significant timbre sensitivity is equivalent across trained and control groups, the magnitude of this sensitivity is substantially *lower* in the trained animals in three out of the four cortical fields examined. This observation was confirmed with a two-way ANOVA (factors: group (control, T-2AFC, T/P-GNG), and cortical field (A1, AAF, PPF, PSF), dependent variable: proportion of variance explained by timbre) which showed a significant effect of group ($F(2, 1224) = 74.48$, $p < 0.001$) and field ($F(3, 1224) = 7.17$, $p < 0.001$) and a significant group × field interaction ($F(6,1224) = 11.17$, $p < 0.001$). In control animals, sensitivity to timbre was highest in the primary fields A1 and AAF. In contrast in the T-2AFC animals timbre sensitivity was substantially higher in PPF than in any of the other three areas. Tukey–Kramer post hoc comparisons between trained and untrained animals ($p < 0.05$) revealed lower timbre sensitivity in the primary areas (A1 and AAF) in T-2AFC and T/P-GNG trained animals than in controls, but higher sensitivity in PPF in T-2AFC trained animals (Fig.2C,D). Therefore, we conclude that long-term training in a spectral timbre discrimination task leads to a field-specific increase in timbre sensitivity in field PPF which is accompanied by a more general decrease timbre sensitivity in A1, AAF and PSF.

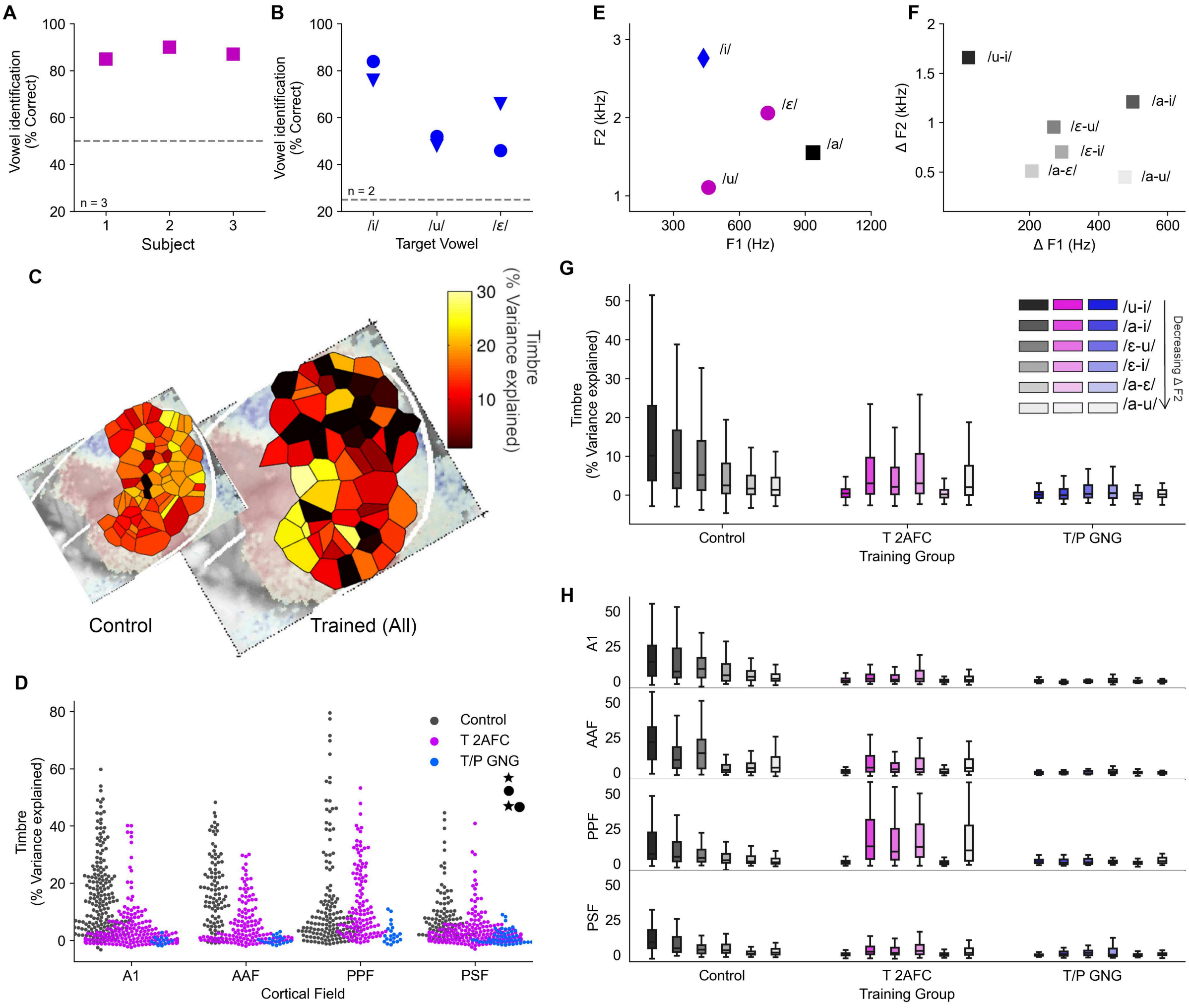

Figure 2: Timbre sensitivity is reduced overall, but enhanced in field PPF, of trained animals.

A, performance of three ferrets discriminating timbre across changes in F0 (chance = 50%). **B** performance of two animals detecting changes in timbre (chance = 25%). **C**, cortical distribution of sensitivity to timbre measured using the Proportion of variance explained metric (see methods). Each tile represents an electrode penetration, with individual sites averaged. **D**, swarm plots showing the distribution of timbre sensitivity across fields and groups. Each datapoint is a unit. The dot indicates a significant main effect of field ($p<0.001$), the star indicates a significant main effect of training ($p<0.001$), and a significant effect of field*training interaction is indicated with both symbols together ($p<0.001$). **E**, first and second formant frequencies for the vowels used in the study, **F**, difference values in F1 or F2 space for given vowel pairs **G,** Neural discriminability for vowel pairs estimated by calculating the proportion of neural response variance attributable to changes in timbre for subsets of the data comprising only those vowels. Boxplots illustrate the median and upper and lower quartile values of the data. **H**, as G, but broken down by cortical field.

Formant cues are reweighted in timbre-trained animals

Contrary to our expectation, long-term training on a spectral timbre discrimination task apparently led to an overall decrease in neural sensitivity to timbre. However, the stimuli with which we recorded neural responses comprised both trained and untrained sounds. To contrast responses to both types of vowels, we therefore repeated the variance decomposition analysis using the neural responses to pairs of vowels (i.e. subsets of 50% of the data comprising the responses to stimulus combinations of two vowels, four F0s and four locations, 32 in total) with the proportion of variance explained by timbre in each case now providing an estimate of how well neuronal responses differentiate a given pair of vowels across variation in F0 and space. We hypothesized that if training led to enhanced selectivity for the target vowels we would observed the highest timbre sensitivity measures for these: for the T2AFC animals /u/ versus /e/ should yield the greatest timbre sensitivity, whereas for the T/P-GNG animals, where /a/ was the reference vowel, we might expect the /a/-/i/, /a/-/e/, and /a/-/u/ contrasts to yield higher sensitivity measures than the pairs of vowels that did not include /a/.

Figure 2E shows the proportion of variance explained by timbre for the six possible vowel combinations, organised according to the magnitude of the difference in second formant frequency (ΔF2). In control animals the neural sensitivity to spectral timbre was well predicted by the size of the difference in F2 frequency. In contrast, the pattern of timbre sensitivity in the T-2AFC animals was neither dominated by the F2 difference, nor by the identity of the trained vowel pair. Instead the pairs with lowest discriminability were those with the smallest difference in the frequency of the first formants (sitting to the left of the space defined by ΔF1 and ΔF2 in Fig.2F).

To better understand how training group, cortical field and change in first and second formant frequencies determined the proportion of the neural response variance was attributed to a change in timbre we ran GLMM which predicted the neural response variance for every pair of vowels with factors field, training group, delta F1 and delta F2 (where delta F1 and F2 were calculated as the difference in first and second formant frequencies for the relevant vowel pair), with unit ID as a random effect. The model was fitted using maximum likelihood estimation and the AIC was used to determine the best model. Within this model both delta F1 and delta F2 were significant predictors, and there was a significant interaction between training group and delta-F1 (but not training group and delta F2) consistent with the idea that training led to an increased integration of both formant frequencies (see Table 1 for full model). From this we conclude that training not only redistributed timbre sensitivity within auditory cortex, but caused a reweighting of spectral integration from being F2 dominated to incorporating differences in both F1 and F2.

Sensitivity to fundamental frequency

T-2AFC animals were able to identify trained vowels across a range of F0 values (Fig 3A) and we therefore predicted that we might see a decreased sensitivity (i.e. increased tolerance) across changes in F0. In contrast, the T/P-GNG animals were never required to discriminate one stimulus dimension across changes in the other, but were trained to detect changes in either F0 (Fig 3B) or spectral timbre (Fig 2B). When sensitivity to F0 is plotted across the cortical surface, or broken down by cortical field it is apparent that compared to control data both trained groups show decreased F0 sensitivity. A two-way ANOVA was performed determine the effect of group and cortical field. There was a significant effect of group (F (2,1224) = 27.76, p<0.001) and a significant training x field interaction (F (6,1224)= 5.02, p<0.001). Tukey–Kramer post hoc comparisons (*p* <0.05) on the responses recorded in each of the cortical field across groups revealed lower F0 sensitivity in A1 of T-2AFC trained animals compared to control animals (Fig.3C,D), with T/P-GNG animals being statistically indistinguishable from controls.

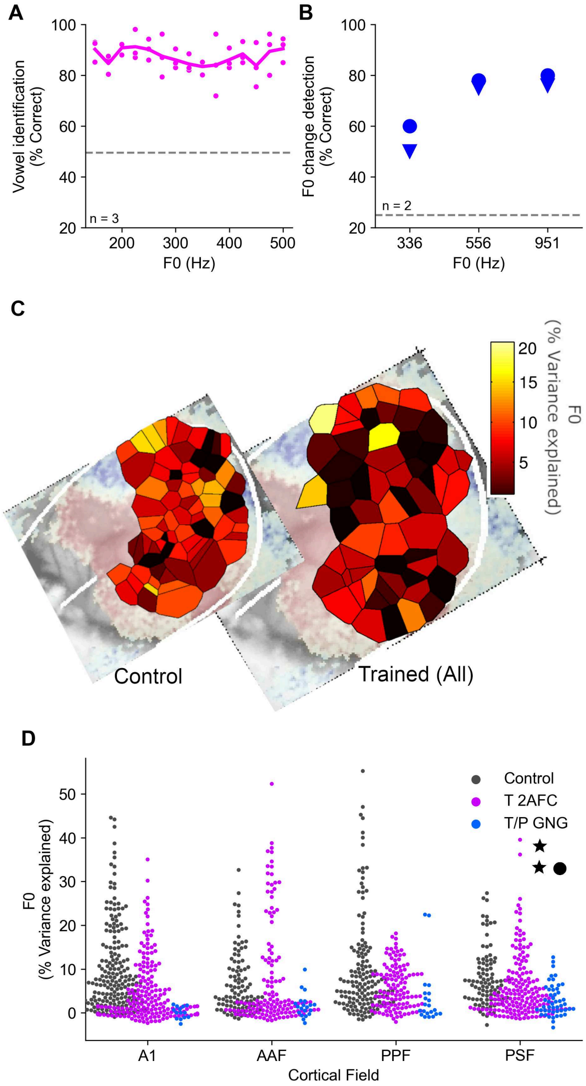

Figure 3: Sensitivity to F0 is decreased in trained animals

A Behavioural performance for discriminating timbre across fundamental frequency (F0). Symbols show individual animals. **B** F0 change detection performance for two animals. **C** Voronoi tessellation maps and box plots plotting the proportion of variance explained by F0 for control and trained animals**.** Conventions as figure 2 C. **D** Swarm plots showing the distribution of F0 sensitivity across fields and training groups. Symbols (black star and dot) show ANOVA results idicating significant main effect of field (p<0.001), training (p<0.001) and a significant effect of field*training interaction.

Sensitivity to task-irrelevant features: Space

Finally we examined how training impacted sensitivity to variation in sound source location. In our control dataset we observed that the spatial tuning elicited by these vowel stimuli presented in virtual acoustic space was modest, and as anticipated, predominantly contralateral (Bizley et al., 2009). When the same neurons were tested with spatially modulated broadband noise (also in VAS) we observed considerably greater spatial modulation leading us to suggest that the low spatial sensitivity was a product of the stimuli rather than the neurons we were recording or the VAS technique. We had speculated that exposure to these sounds might enhance spatial sensitivity measured in auditory cortex using these stimuli. As Figure 4 show spatial sensitivity was indeed higher in trained animals: whether visualized as a cortical map (Fig.4A) or broken down by cortical field (Fig.4B), it is clear that in the non-primary areas spatial sensitivity was higher in the T-2AFC animals than in controls. A two-way ANOVA on the proportion of variance explained by space, with group and cortical field as factors, determined significant effects of group (F(2,1224) = 10.63, p<001) and field (F(3,1224)= 3.45, p=0.01) and a significant training × field interaction (F(6,1224)= 4.22, p<0.001). Tukey–Kramer post hoc comparisons (*p* <0.05) of the trained and control data in specific auditory fields revealed higher azimuth sensitivity in the non-primary areas (PPF and PSF) in trained animals compared to control animals.

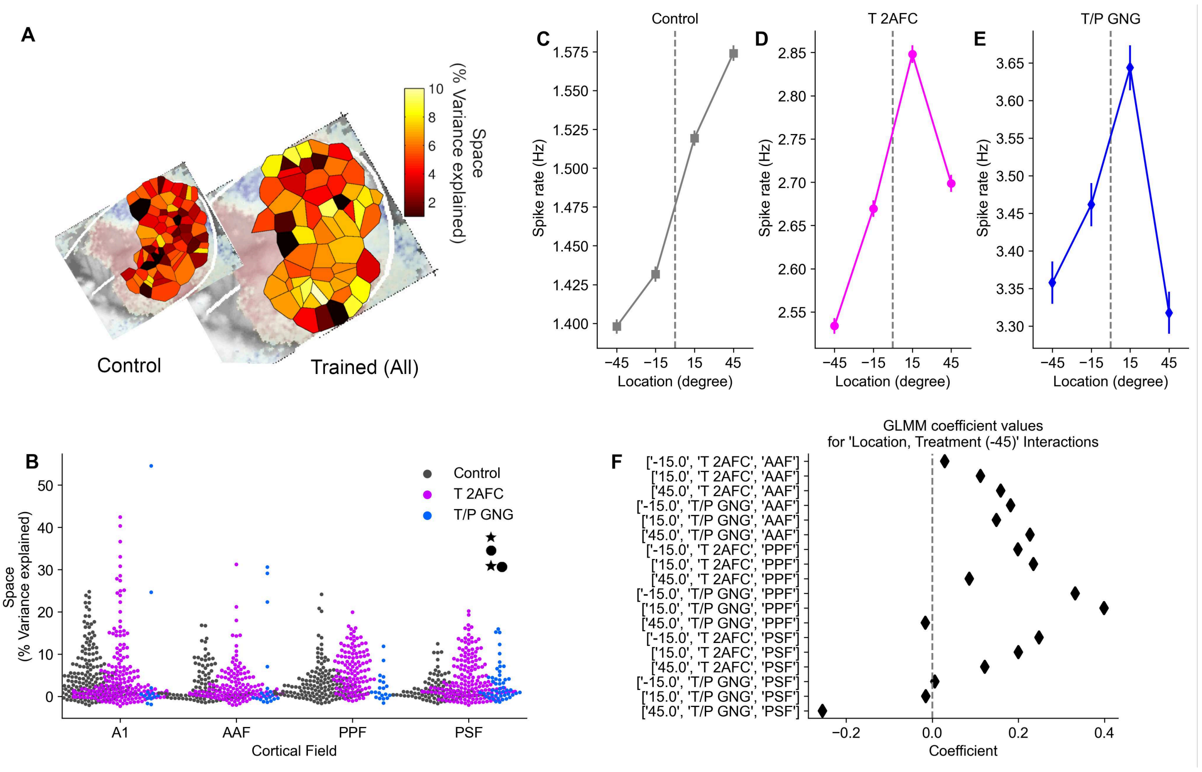

Figure 4: Effect of training on the neural encoding of task-irrelevant features

A, Voronoi tessellation maps plotting the proportion of variance explained by azimuth for control and trained animals. ***B***, Swarm plots showing the proportion of variance explained by azimuth across fields and training groups. The dot indicates a significant main effect of field (p<0.001), the star indicates a significant main effect of training (p<0.001), and a significant effect of field*training interaction is indicated with both symbols together (p<0.001). ***C-E,*** Population spatial tuning functions (mean ± SEM for all responsive units) for control, T2AFC animals and T/P GNG animals. F beta coefficients for the impact of model parameters on spike rate.

We next asked whether the increased spatial sensitivity in trained animals reflected altered spatial tuning, or simply enhanced the gain of spatial receptive fields. To assess this we exploited the typical contralateral spatial preference of most auditory cortical neurons to derive population tuning functions. We calculated a normalised spatial response function for each neuron, taken as the mean sound-evoked spike rate at each location across all pitch and timbre combinations, and then averaged across all recorded units (Fig.4C-E). As expected, in the control animals (Fig.4C) this yielded a monotonically increasing function with the most contralateral (+45°) location eliciting the strongest firing rates. In contrast, in both trained groups of animals, spatial tuning was non-monotonic and showed a peak at +15°. To quantify these effects we ran a GLMM predicting spike rates with group (T-2AFC or T/P-GNG relative to reference category control), field (reference category A1), spatial position (categorical predictor, relative to -45) with unit as a random effect. This confirmed significant main effects of spatial position (-45 and +45 significantly different, reflecting contralateral tuning), and significant field*group*position interactions showing for fields PPF and PSF with both trained groups. The coefficients for these effects (Fig.4F) were negative for +45 and positive for ±15, confirming the changes visible in the spatial tuning functions plotted in (Fig.4C-E). Thus despite the sound source never varying in its spatial position, and not being relevant to the task, the location from which the sounds were presented (i.e. central space) was substantially enhanced in the trained animals.

Discussion

In this study we trained two groups of ferrets to discriminate perceptual attributes of artificial vowels. One group categorized vowels according to their identity across changes in F0 (T-2AFC animals), while the other detected changes in either the F0 or timbre in a sequence of on-going vowel sounds (T/P GNG animals). We predicted that the 2AFC animals would show enhanced sensitivity for timbre and increased tolerance (i.e. decreased sensitivity) to other sound features. In fact, what we observed was more complex: sensitivity to timbre decreased markedly in three cortical areas but was enhanced in field PPF. Moreover, sensitivity to vowel identity became less contingent only on changes in the frequency of the second formant, and instead was dependent on both changes in first and second formant frequency. Sensitivity to F0, which the animals were required to generalize across, decreased. In contrast, sensitivity to space, which had no task-relevance, was enhanced, with spatial receptive fields shifted towards the midline, from where target sounds originated. In the T/P GNG animals our conclusions are somewhat limited by a smaller sample size, but there was also a decrease in timbre sensitivity relative to controls and an increase in spatial sensitivity, with F0 encoding matching that of control animals.

Previous studies investigating the impact of training on neural tuning in auditory cortex have focused on map plasticity in A1. However, higher auditory cortical fields are thought to become increasingly specialized for processing spatial or non-spatial stimulus attributes (Rauschecker and Tian, 2000; Bizley and Cohen, 2013; Elgueda et al., 2019) and neurons in higher cortical fields show enhanced attention related changes during behavior (Mesgarani and Chang, 2012; Atiani et al., 2014; Elgueda et al., 2019). It is currently unknown how these attention-related changes relate to more 'hard-wired' changes in neural sensitivity. Our data bridge this gap by suggesting that the areas that show larger attention related changes may also be those in which receptive fields are optimized through learning to process task-relevant stimuli.

The observation of a marked decrease in sensitivity to task-relevant features in the primary auditory cortical fields A1 and AAF is a potentially surprising finding. However, a number of studies report that engagement in a behavioural task causes suppression of neural responses in auditory cortex (Otazu et al., 2009; Town et al., 2018). One possibility is that training leads to the integration of diverse non-sensory inputs into auditory cortex that ultimately underlie the observation of choice or motor related activity, but that these inputs come at the cost of presumably redundant feed-forward stimulus evoked activity. Cooling primary auditory cortex in ferrets does not lead to an impairment in a vowel discrimination in silence in a task analogous to the one here (Town et al., 2023). Our neural data raise the testable prediction that inactivation of field PPF should cause a timbre identification deficit whereas cooling A1, AAF or PSF would have a more modest effect on behavior.

Calculating discriminability measures for pairs of vowels allowed us to observe that training on a spectral timbre task altered the way in which neuronal responses were sensitive to spectral timbre. In the naïve control animals the most discriminable stimuli were those in which there was a large difference in second formant frequency. This finding is mirrored behaviourally; when first and second formant cues are placed in conflict, animals tend to weight the position of the second formant over the first, with behaviour being best predicted by either F2 position or the position of the spectral centroid (Town et al., 2015). Nonetheless, the animals whose neural data were recorded for this study were also tested behaviourally with single-formant stimuli and these animals accurately classify F1 for

/u/ and F2 for /e/ (Bizley et al., 2013). The finding that neural responses in trained animals are explained by both changes first and second formants for both trained and novel vowels suggest that learning results in an enhanced integration of the cues that define the spectral envelope.

We also considered the impact of training on the neural representation of two task-irrelevant features. In the T-2AFC task animals were required to generalise across F0 while discriminating timbre, and neural responses showed an overall decrease in sensitivity to F0. In contrast, the T/P-GNG animals were required, in separate sessions, to detect changes in timbre or F0 of a sequence of vowels and showed F0 sensitivity that was statistically indistinguishable from controls. It remains possible that, as with timbre in the 2AFC animals, this might reflect a reorganization such that there is a general decrease in sensitivity that is accompanied by a local increase in F0 sensitivity in another cortical area in one or both sets of animals. This possibility requires further investigation, possibly with high-density recordings targeting the low frequency border of primary and posterior fields where specialization for pitch might occur (Walker et al., in prep).

In contrast to the decreased sensitivity to non-spatial sound features we observed a marked increase in spatial sensitivity in trained animals despite the task having no spatial component to it. These changes occurred principally in the non-primary fields PPF and PSF and were not simply gain changes but rather (presumably) adaptive shifts in tuning: the population tuning function shifted from monotonically increasing for more contralateral sounds to peaking 15° contralateral to the midline.

Our spatial receptive fields were very coarsely measured with stimuli at ±15° and ±45° but if the change in tuning observed in the recorded hemisphere were mirrored across the midline we would expect to see that the representation of the midline, where stimuli were presented during behavioural testing, was enhanced. Therefore it seems that repeated exposure to behaviourally relevant sounds from this location led to an enhancement in non-primary auditory cortex (fields PPF and PSF). Engaging in a sound discrimination task has been shown to refine spatial tuning in primary auditory cortex, with changes occurring for both localisation tasks, and – more modestly – for non-spatial tasks (Lee and Middlebrooks, 2011) but to our knowledge this is the first report of enhanced location coding after repeated exposure to behaviourally relevant stimuli. Learning triggers widespread changes in gene expression in auditory cortex (Graham et al., 2023). Future work can seek to unpicking the specific molecular mechanisms which ultimately support the changes in auditory cortical function, and ultimately auditory memory, that we observe here.

The stimuli that the animals were processing were relatively low frequency (<4 kHz). Our sampling yielded balanced samples of neurons across the frequency axis in both groups of animals (Fig.1) but we did not perform the high-resolution mapping that would probably be required to unequivocally argue that there was no difference in the tonotopic organisation in these animals. Nonetheless, we did not see any evidence for map reorganisation in these animals and our sampling of the cortical surface did not yield a dominance of low frequency recording sites. Given evidence that map plasticity may be a temporary phase of learning (Reed et al., 2011), and that learning of natural sounds can occur in the absence of any map plasticity (Shepard et al., 2016) the lack of evidence for tonotopic map plasticity here is perhaps not surprising and likely additionally reflects the broad band nature of the sounds that the animals were exposed to.

A caveat of our study is that it is performed under anaesthesia. This was essential for performing mapping across multiple cortical fields. Methods for recording in awake animals currently allow only high density sampling from a small area of cortex, or sparse sampling across multiple fields. Recording under anaesthesia does, however, allow us to separate out effects of sensitivity to stimulus features from attention and allows us to measure static receptive field features to which attentional and task related effects are likely added during active listening.

In summary, training causes diverse effects in auditory cortex: sensitivity to trained stimulus features which is broadly distributed in naïve animals becomes localised to one distinct auditory field after training. In contrast to control animals which strongly weight second formant frequency, receptive fields integrate information about both first and second formant frequency. Finally, sensitivity to task-orthogonal features – here auditory space – is enhanced when stimuli consistently originated from a single location.

Acknowledgments

This work was supported by a RNID (formally Action on Hearing Loss) PhD studentship to H.A., grants from the Biotechnology and Biological Sciences Research Council (to J.W.H.S., J.K.B., and A.J.K. and to J.K.B. BB/H016813/1), a Royal Society Dorothy Hodgkin Fellowship (J.K.B.), a Sir Henry Dale Fellowship (J.K.B. WT098418MA) and a Wellcome Trust Principal Research Fellowship (A.J.K.).

References

Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, Fritz JB (2014) Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. Neuron 82:486–499.

Bao S, Chang EF, Teng CL, Heiser MA, Merzenich MM (2013) Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment. Neuroscience 248:30–42.

Bao S, Chang EF, Woods J, Merzenich MM (2004) Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nat Neurosci 7:974–981.

Bizley JK, Cohen YE (2013) The what, where and how of auditory-object perception. Nat Rev Neurosci 14:693–707.

Bizley JK, Walker KM (2010) Sensitivity and selectivity of neurons in auditory cortex to the pitch, timbre, and location of sounds. Neuroscientist 16:453–469.

Bizley JK, Walker KM, King AJ, Schnupp JW (2013) Spectral timbre perception in ferrets: discrimination of artificial vowels under different listening conditions. J Acoust Soc Am 133:365–376.

Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. J Neurosci 29:2064–2075.

Elgueda D, Duque D, Radtke-Schuller S, Yin P, David SV, Shamma SA, Fritz JB (2019) State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex. Nat Neurosci.

Galindo-Leon EE, Lin FG, Liu RC (2009) Inhibitory plasticity in a lateral band improves cortical detection of natural vocalizations. Neuron 62:705–716.

Graham G, Chimenti MS, Knudtson KL, Grenard DN, Co L, Sumner M, Tchou T, Bieszczad KM (2023) Learning induces unique transcriptional landscapes in the auditory cortex. Hear Res 438:108878.

Griffiths TD, Warren JD (2004) What is an auditory object? Nat Rev Neurosci 5:887–892.

Ison MJ, Quiroga RQ (2008) Selectivity and invariance for visual object perception. Front Biosci 13:4889–4903.

King AJ, Middlebrooks JC (2011) Cortical Representation of Auditory Space. In: The Auditory Cortex, pp 329–341. Springer.

Lee CC, Middlebrooks JC (2011) Auditory cortex spatial sensitivity sharpens during task performance. Nat Neurosci 14:108–114.

Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485:233–236.

Nelken I, Bizley JK, Nodal FR, Ahmed B, Schnupp JW, King AJ (2004) Large-scale organization of ferret auditory cortex revealed using continuous acquisition of intrinsic optical signals. J Neurophysiol 92:2574–2588.

Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. Nat Neurosci 12:646–654.

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of “what” and “where” in auditory cortex. Proc Natl Acad Sci U S A 97:11800–11806.

Recanzone GH, Cohen YE (2010) Serial and parallel processing in the primate auditory cortex revisited. Behav Brain Res 206:1–7.

Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, Kilgard MP (2011) Cortical map plasticity improves learning but is not necessary for improved performance. Neuron 70:121–131.

Rutkowski RG, Weinberger NM (2005) Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. Proc Natl Acad Sci U S A 102:13664–13669.

Schreiner CE, Polley DB (2014) Auditory map plasticity: diversity in causes and consequences. Curr Opin Neurobiol 24:143–156.

Shepard KN, Chong KK, Liu RC (2016) Contrast Enhancement without Transient Map Expansion for Species-Specific Vocalizations in Core Auditory Cortex during Learning. eNeuro 3.

Town SM, Atilgan H, Wood KC, Bizley JK (2015) The role of spectral cues in timbre discrimination by ferrets and humans. J Acoust Soc Am 137:2870–2883.

Town SM, Poole KC, Wood KC, Bizley JK (2023) Reversible Inactivation of Ferret Auditory Cortex Impairs Spatial and Nonspatial Hearing. J Neurosci 43:749–763.

Town SM, Wood KC, Bizley JK (2018) Sound identity is represented robustly in auditory cortex during perceptual constancy. Nature communications 9:4786.

Walker KM, Davies A, Bizley JK, Schnupp JW, King AJ (2017) Pitch discrimination performance of ferrets and humans on a go/no-go task. Animal Behavior and Cognition. Available at: http://biorxiv.org/lookup/doi/10.1101/165852 [Accessed December 29, 2023].

Whitton JP, Hancock KE, Polley DB (2014) Immersive audiomotor game play enhances neural and perceptual salience of weak signals in noise. Proc Natl Acad Sci U S A 111:E2606-15.