PLOS Biology

RESEARCH ARTICLE

Category learning drives cortical attentional gain to diagnostic dimensions of perceptual information

Sahil Luthra[1], Raha N. Razin[2], Adam T. Tierney[3], Lori L. Holt[4], Frederic Dick[2]*

1 Department of Psychology, Stony Brook University, Stony Brook, New York, United States of America, **2** Department of Experimental Psychology, Division of Psychological and Language Sciences, University College London, London, United Kingdom, **3** Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom, **4** Department of Psychology, The University of Texas at Austin, Austin, Texas, United States of America

* f.dick@ucl.ac.uk

Check for updates

OPEN ACCESS

Citation: Luthra S, Razin RN, Tierney AT, Holt LL, Dick F (2026) Category learning drives cortical attentional gain to diagnostic dimensions of perceptual information. PLoS Biol 24(11): e3004076. https://doi.org/10.1371/journal.pbio.3004076

Academic Editor: Manuel S. Malmierca, Universidad de Salamanca, SPAIN

Received: December 13, 2025

Accepted: October 12, 2026

Published: November 12, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pbio.3004076

Copyright: © 2026 Luthra et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution,

Abstract

Humans and other animals develop remarkable perceptual and cognitive specializations for identifying, differentiating, and acting on classes of ecologically important signals. This expertise is flexible enough to support diverse perceptual judgments: a voice, for example, simultaneously conveys what a talker says, as well as myriad cues about her identity and state. Expert perception across complex signals thus involves discovering and learning regularities that best inform diverse perceptual judgments, as well as weighting this information flexibly as task demands change. Here, we test whether this process may involve attentional gain. We use two prospective auditory category learning tasks to relate a complex, entirely novel soundscape to four classes of 'alien identity' and two classes of 'alien size.' Identity, but not size, categorization requires discovery and learning of patterned acoustic input situated in one of two simultaneous, non-overlapping frequency bands. This allows us to capitalize on the coarsely segregated, frequency-band-specific channels tiling auditory cortex. We use fMRI to ask whether category-relevant perceptual information present in one frequency band is prioritized relative to simultaneous, uninformative information in the other frequency band. Only among participants expert at alien identity categorization do we observe neural prioritization of the identity-diagnostic frequency band. This prioritization persists even when the diagnostic information becomes irrelevant in the size categorization task. Tellingly, the neural selectivity evoked implicitly in the identity categorization task aligns with that in an independent task, where activation is driven by explicit and sustained selective attention to pure tones in one or the other frequency band. Additionally, the learning trajectories taken to achieve expert-level categorization leave fingerprints on the patterns of neural activity associated with the diagnostic dimension. In all, this indicates that learning categories can drive attentional gain to category-diagnostic input dimensions.

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The summary data used for analysis, and code used for processing and analysis of the current study are available at https://doi.org/10.5522/04/33974425.

Funding: The work was supported by grants from National Institutes of Health (R01DC004674) to LLH AT, and FD, the National Science Foundation (SBE/BCS 2414066, 2420979) to LLH and FD, an individual NRSA from the National Institute on Deafness and Communication Disorders to SL (F32DC020625) under the mentorship of LLH and FD. Neither the National Institutes of Health (www.nih.gov) nor the National Science Foundation (www.nsf.gov) played any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: 2AFC, two-alternative forced choice; FDR, false discovery rate; 4AFC, four-alternative forced-choice; fMRI, functional magnetic resonance imaging; HG, Heschl's Gyrus; LH, left hemisphere; ROIs, regions of interest; RH, right hemisphere; sEEG, stereoelectroencephalography; STG, Superior Temporal Gyrus; STS, Superior Temporal Sulcus; UCL, University College London.

Introduction

Learners of all ages discover structured perceptual input to achieve their goals, and even simple sights and sounds convey rich information that can flexibly support multiple behaviors. A single utterance simultaneously conveys a talker's age [1] and socioeconomic class [2], whether she is a stranger or a loved one [3], and if she is requesting her *bill* or her *pill* [4]. A glimpse of her face communicates her trustworthiness [5], emotion [6], and identity [7]. Acoustic characteristics of her footsteps can convey her gender [8], identity [9], health, and weight [10]. These underlying patterns are carried across multiple perceptual dimensions and – especially in the case of sound – often require integration of information across time. Successful perception thus requires that we discover and learn distinct input patterns that best inform specific judgments. Flexibility in weighting the most diagnostic patterns of perceptual cues or dimensions for the task at hand is crucial for effective face and voice identification, object recognition, and at least in humans, speech comprehension.

Prominent theories of categorization have posited that learning may drive attention to be directed toward diagnostic perceptual dimensions in a task-dependent manner (e.g., [11–13]). These learned attentional biases are proposed to weight perceptual dimensions most relevant to oft-performed tasks and behaviors [14–16], and to language-specific cues [17–19]. For complex skills like spoken language and object recognition, it can take years or even decades to learn which perceptual dimensions are informative for a given behavioral goal, and to optimally weight or attend to them (e.g., Creel et al. [20–25]). Learners also progress along different trajectories to reach equivalent performance outcomes [26,27]. Indeed, distinct perceptual and neural strategies are reflected in developmental neuroimaging studies that demonstrate that there can be quite different weightings of neural activation across brain regions, even among children or adults who perform similarly on a task (e.g., Brown et al. [28]; Stiles et al. [29]; Krishnan et al. [30]; Yates et al. [31,32]). In auditory categorization tasks, both mouse [33,34] and human [35–37] learners exhibit idiosyncratic patterns of behavior over short-term training. These individual differences may in turn reflect where participants are positioned in a longer-term learning trajectory, over which diagnostic input dimensions are discovered and tuned through experience.

Previous functional MRI work has shown that short-term category training can shape distributed visual (Folstein et al. [12,38]) and auditory (Ley et al. [39]; but cf. Kanber et al. [22]) cortical representations. Here, we establish whether learning a new behavior across an informationally rich perceptual space (a) drives distinct patterns of cortical activity that are spatially predictable given the cortical mapping of the category-diagnostic dimension, and (b) does so in a way that is consistent with acquired attention to the perceptual dimensions most diagnostic to the task. By examining cohorts of learners who reach equivalent levels of expertise in this new domain, we also ask whether patterns of cortical activity reflect the distinct trajectories learners take to achieve expertise. To do so, we adopt a prospective learning approach in which naïve adults learn to categorize entirely novel sounds drawn from a complex, multidimensional perceptual space composed of over 36,000 sounds that vary across complex, not-readily-verbalizable acoustic dimensions [40]. Both the

sound exemplars and the categories we have defined across them are entirely novel to listeners, as are the visual icons used during training.

The alien sounds possess information in two simultaneous, non-overlapping frequency bands, with each band having highly similar and complex acoustic information. Each of four aliens' identities is conveyed by spectrotemporal patterns within one or the other of these frequency bands, with the other band conveying simultaneous but non-diagnostic acoustic information. This allows us to take advantage of the tonotopic mapping of auditory cortex, where we expect acoustic information in each frequency band to be preferentially processed in different spatial subregions of the multiple tonotopic maps that lie across the superior temporal lobe.

Within these maps, functional MRI studies have shown that stable, stimulus-driven voxelwise frequency selectivity can be established across much of auditory cortex [41]. Voxels' relative frequency preference can also be mapped by *explicit attention to specific frequency bands* within a complex, dynamic, and full-spectrum acoustic scene [42–45]. For example, the degree to which activity in cortex can be similarly driven by attention to a particular frequency band, and 'bottom up' acoustic stimulation in that same frequency band can be assessed by constructing 'concordance maps' [43]. These maps quantify the similarity of stimulus- and attention-driven tuning of the subset of voxels within small parcellation units within auditory cortex. Such concordance maps allow one to make group-level inferences about *where* there are consistent relationships between stimulus- and attention-driven frequency tuning in auditory cortex, while taking into account the known cross-participant variability in the exact position, size, and orientation of tonotopic maps relative to gyral morphology [41]. We use this analytic approach to ask whether, like explicit attention, the presence of category-diagnostic information in voxels' preferred frequency band drives relatively greater activation compared to when category-diagnostic information is conveyed via the less-preferred frequency band. Such cross-task similarity would support hypotheses that learning and expertise implicitly drive attention to be directed toward category-informative dimensions [46–48].

As noted above, many naturally occurring sounds convey information about multiple types or categories of information, and this information is conveyed along different acoustic dimensions [49]. To model this, we created all sound exemplars to convey both '*alien identity'* via variable, but statistically structured acoustic patterns embedded within one of two frequency bands, and to simultaneously convey *'alien size'* via perceived loudness. Here, lower amplitude sounds are diagnostic of 'small aliens' and higher amplitude sounds are diagnostic of 'big aliens'. Our main group of participants trained over five days to learn alien identity (across spectrotemporal patterns) and also trained to categorize the same sound inventory according to alien size (across amplitude). Another smaller group trained *only* on alien size (conveyed via amplitude) thereby not gaining expertise in classifying alien identity according to the patterned spectrotemporal information embedded in one of the two frequency bands.

This approach allows us to answer several central, yet open, questions: Is category-relevant perceptual information prioritized relative to simultaneously presented category-irrelevant information? If so, does this selectivity persist when the diagnostic dimension is irrelevant to the task at hand? Do emergent patterns of prioritization align closely with neural activity associated with explicit, sustained attention to the dimension as conveyed by different sounds? Finally, do these patterns of neural activity relate to the distinct trajectories that learners take in reaching expert-level categorization?

Results

A group of young adults (N = 95, Fig 1A) trained for 5 days to develop expertise in rapidly and accurately categorizing complex, multidimensional sounds they had never previously encountered (Fig 2). In the main 'alien identity' training task, four sound categories were associated with the identity of four distinct 'alien' images (Fig 1C). Each sound possessed acoustic energy in two non-overlapping acoustic frequency bands (~80-750 Hz and ~1000-9000 Hz), each populated with three 400-ms 'chirps' played one after another and varying in frequency contour (Fig 2A). In the *category-diagnostic* band, these acoustically variable chirps possessed an underlying regularity that defined category membership (alien identity). Specifically, each of the three chirps in the category-diagnostic band was derived from the pitch contours extracted from

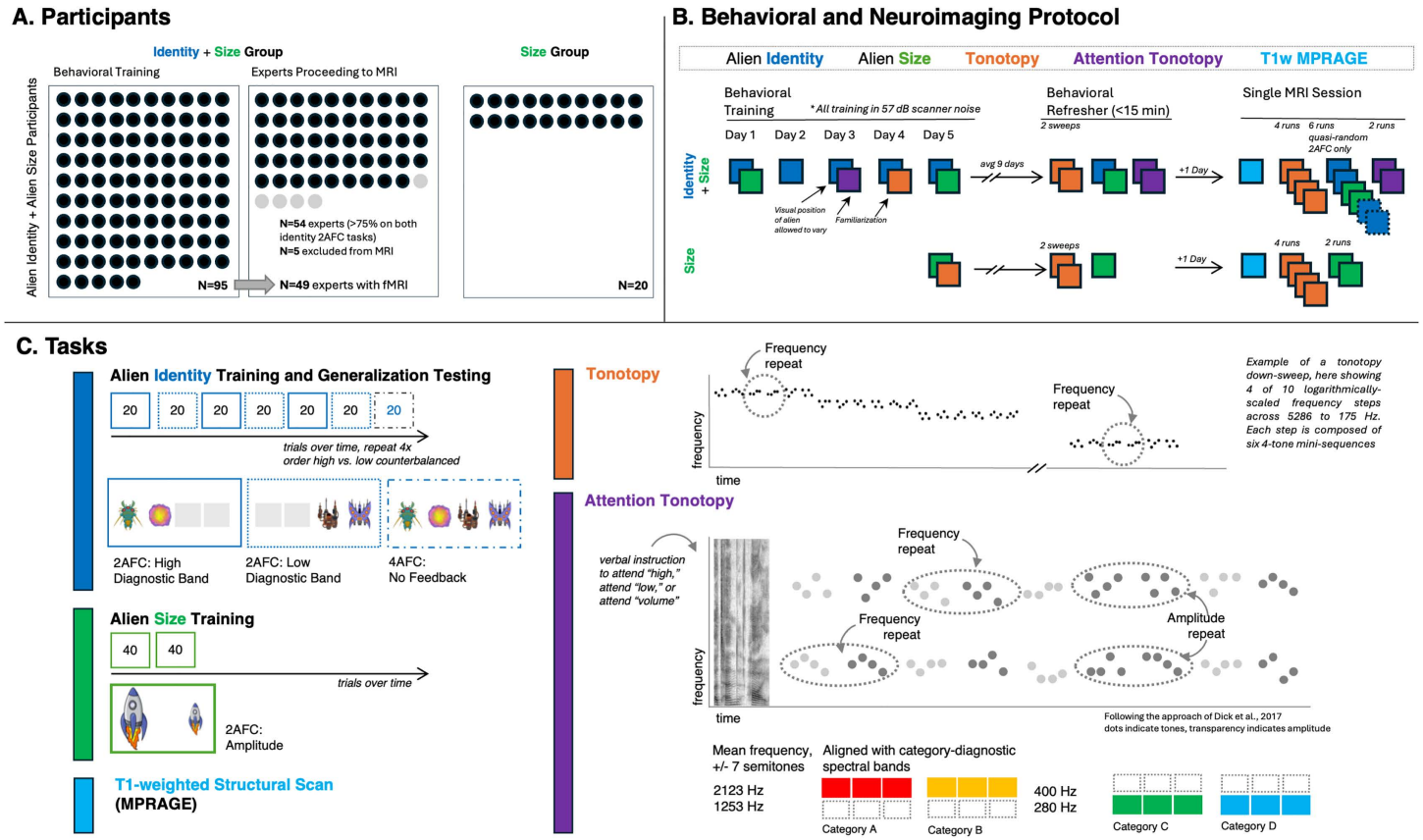


Fig 1. Experimental Approach. (**A**) **Participants**. An *Alien Identity+Size* group of participants (N = 95) trained to achieve expert performance on both alien identity and alien size categorization tasks. Across five days of training with feedback, they learned to associate sound exemplars with four alien identities on a two-alternative forced choice (2AFC) categorization task blocked by high- and low-frequency diagnostic frequency bands (see Fig 2). The N = 54 *Alien Identity+Size* participants who achieved >75% accuracy by Day 5 were recruited as categorization experts for a single fMRI session (5 participants omitted due to motion or noncompliance, for a final fMRI N = 49). A separate *Alien Size* group (N = 20) trained only on the size judgment task *without* training on alien identity, and participated in a single fMRI session. **(B, C) Behavioral and Neuroimaging Protocol, Tasks.** Panel **B** summarizes the behavioral and neuroimaging protocol, with Panel **C** as a color-coded legend that schematizes the component tasks. Tasks are described in detail in **Methods-Procedure**. Stimuli are illustrated in Fig 2 and described in **Methods-Stimuli**.

https://doi.org/10.1371/journal.pbio.3004076.g001

a different instance of the same Mandarin tone category (see **Methods-Stimuli**). Simultaneous and temporally aligned chirps in the other, *non-diagnostic* frequency band were also acoustically variable but possessed no coherent regularity and were not predictive of alien identity (Fig 2B). Each of the three chirps in the non-diagnostic band was derived from pitch contours extracted from three different Mandarin tone categories (see **Methods-Stimuli**).

These stimulus characteristics impose several key demands on learning: (1) because chirp triplets appear simultaneously in two distinct frequency bands but statistically structured patterns diagnostic of alien identity occur in only *one* band, listeners must discover the relevant regularity in the context of competing sounds; (2) chirp triplets within the category-diagnostic frequency band are acoustically variable yet possess an underlying spectrotemporal regularity that maps onto alien identity through feedback; (3) because this regularity differs across aliens, listeners must not only identify the diagnostic frequency band but also learn to map specific spectrotemporal patterns within each frequency band onto alien identity; (4) overall amplitude varies across exemplars, such that learning alien identity requires generalization across identity-irrelevant amplitude variation; (5) learning takes place under adverse listening conditions, with sounds embedded in scanner noise throughout, including during behavioral training outside the scanner.

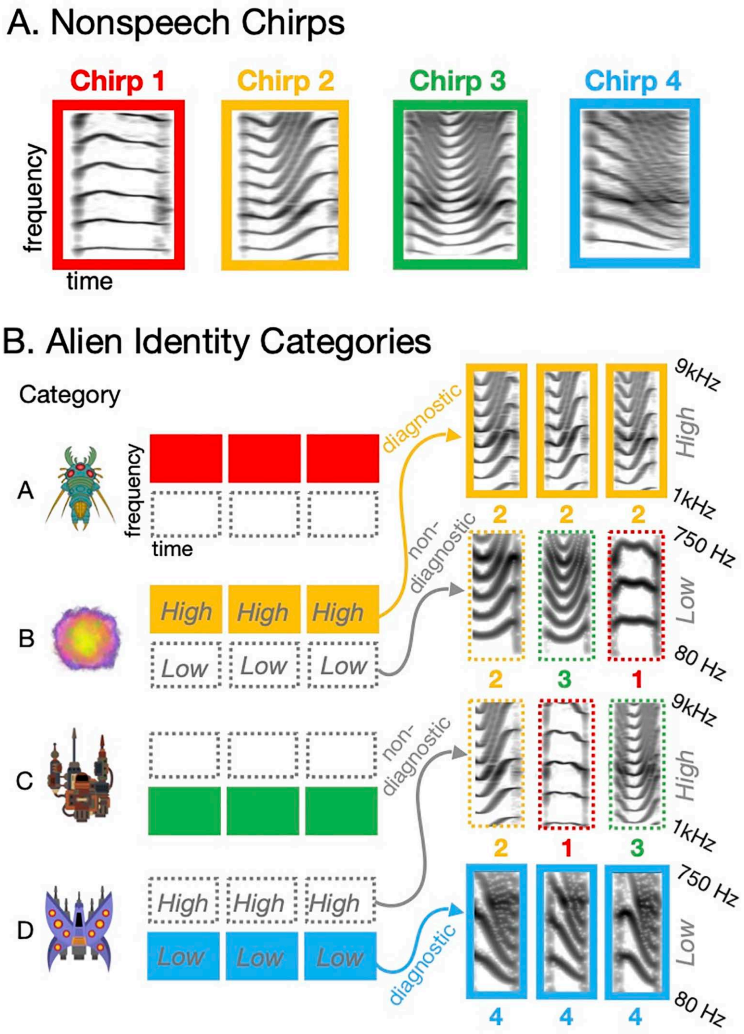


Fig 2. Stimuli. (**A**) **Nonspeech Chirps**. We extracted the fundamental frequency (F0) contour from natural speech recordings of single-syllable words varying in Mandarin lexical tone contour across four (2 female) native Mandarin speakers. This yielded 80 acoustically unique chirps from each of the four classes of Mandarin lexical tone contour (roughly: level, rising, dipping or falling; see Obasih et al. [40]). Spectrograms show one representative chirp from each class (indicated by the colored frame). **(B) Alien Identity Categories.** We next shifted each nonspeech chirp into two non-overlapping frequency bands (low: ~80-750 Hz, high: ~1,000-9,000 Hz), creating two sets of chirps that served as building blocks for constructing novel auditory category exemplars. Each category exemplar was built from 6 unique chirps (3 high, 3 low, presented simultaneously), represented as rectangles. Color indicates the chirp class shown in **(A)**, with three unique chirps that share the same underlying regularity (derived from a single Mandarin tone category). This *category-diagnostic* pattern conveys alien identity through its association with feedback in training. Three simultaneous chirps (uncolored rectangles) from three different randomly selected chirp classes (derived from different Mandarin tone categories) varied such that there was no structured regularity across the triplet; this frequency band was irrelevant to alien identity. For Categories A and B, the high frequency band was diagnostic; for Categories C and D the low frequency band was diagnostic. The insets on the right show a detailed example from Category B (high-band-diagnostic, Chirp 2 structure) and Category D (low-band-diagnostic, Chirp 4 structure) as time x frequency spectrograms with color frames indicating chirp class for visualization. Spectrotemporal modulation characteristics were extremely similar across simultaneously presented low and high frequency bands and between category-diagnostic and non-diagnostic chirp triplets within the bands. (Note that the spectrotemporal characteristics of the chirps displayed in the figure (such as harmonic density and pitch contour) vary considerably across exemplars; for full details, see Obasih et al. [40] and the full set of experimental materials available online.).

https://doi.org/10.1371/journal.pbio.3004076.g002

These demands capture central challenges of learning real-world auditory categories. Speech, for example, is characterized by multiple simultaneous acoustic dimensions that may be relevant to one task (did I hear '*bear*' or '*pear*'?) yet irrelevant to another performed across the same signal (was the speaker frightened?). Speech is also typically learned in naturalistic conditions that include task-irrelevant ambient noise.

These learning demands allow us to test whether, in experienced listeners, category-relevant perceptual information is prioritized over simultaneously presented but uninformative information. We map frequency selectivity across auditory cortex to identify cortical regions sensitive to the frequency bands populated by category-diagnostic versus non-diagnostic chirp triplets. Because spectrotemporal modulation is highly similar across the simultaneously presented low- and high-frequency bands, we expect comparable bottom-up acoustic activation in corresponding frequency-selective regions of auditory cortex. We ask whether learning how frequency-band-specific patterns map onto alien identity in the presence of competing sounds in the opposing band may lead to an exaggerated neural response in the identity-diagnostic band relative to the non-diagnostic band. We further ask whether neural responses elicited by these alien sounds differ between expert listeners and listeners naïve to alien identity (the *Alien Size* group). Finally, we test whether expertise-related modulation of neural activity parallels patterns of frequency-selective attention observed during overt, voluntary attention to a specific band, thereby directly testing the hypothesis that learning and expertise implicitly bias attention toward category-informative dimensions.

Category expertise develops across training, with generalization to novel exemplars

Alien identity categorization training involved a two-alternative forced-choice (2AFC) task, with trials blocked by categories with high- versus low-frequency category-diagnostic information. Trial-by-trial feedback indicated the correct alien identity category (Fig 1C, see **Methods-Procedure**). After each training block, a four-alternative forced-choice (4AFC) task with no feedback assessed generalization of category learning to novel exemplars not experienced in training (Fig 1C).

A separate training task involved categorizing exemplars drawn from all four alien identity categories as 'big' or 'small' according to stimulus amplitude (which varied equiprobably along a continuum from 62−80 dB), *without* regard to the spectrotemporal dimensions associated with alien identity (Fig 1C, see **Methods-Procedure;** all reported dB referenced to $2*10^{-5}$ and measured in Praat; Boersma [50]). Two independent samples of participants completed alien size category training, one learning alien identity and size categories (*Alien Identity+Size*) and the other learning only size categories (*Alien Size*).

This permitted us to test (a) whether identity-training-induced changes in cortical activation associated with the frequency band carrying identity-diagnostic information exist; (b) whether these changes are lacking in identity-naïve listeners; and (c) whether they persist when task demands are redirected to category decisions based on amplitude. The *Alien Size* group is also important for identifying any bottom-up salience differences that may differentiate sounds from specific alien identity categories, independent of category expertise. By design, the stimulus exemplars (and their order) in the identity and size categorization blocks during the fMRI session were identical to avoid potential stimulus confounds.

A subset of expert *Alien Identity+Size* participants (N = 49) who achieved at least 75% 2AFC accuracy in reporting alien identity across both high- and low-frequency diagnostic bands by Day 5 returned an average of 9 days later for a single functional magnetic resonance imaging (fMRI) session (Fig 1A, 1B). Overall, these experts exhibited above-chance 2AFC alien identity training performance on Day 1, $t(48)=15.7$, $p = 7.35 \times 10^{-21}$, with improvements from Day 1 to Day 5, $t(48) = 7.82$, $p = 2.03 \times 10^{-10}$, to reach near-ceiling alien identity categorization (95.5%; SE = 0.7%, Fig 3A**, left panel**). Among these experts, 4AFC generalization to novel exemplars exceeded chance on Day 1, $t(48)=5.64$, $p = 4.41 \times 10^{-7}$; M = 68.5%, SE = 3.3%) and improved significantly from Day 1 to Day 5 (M = 87.7%, SE = 1.9%, $t(48) = 6.56$, $p = 1.73 \times 10^{-8}$; Fig 3A**, middle panel**). Confusion matrices (Fig 3B) from the 4AFC generalization task on the first and last day of training show that category identification sharpens over training. On Day 1, expert listeners made errors in identifying novel exemplars of all categories and tended to confuse novel exemplars that shared a diagnostic band. By Day 5, listeners' categorization

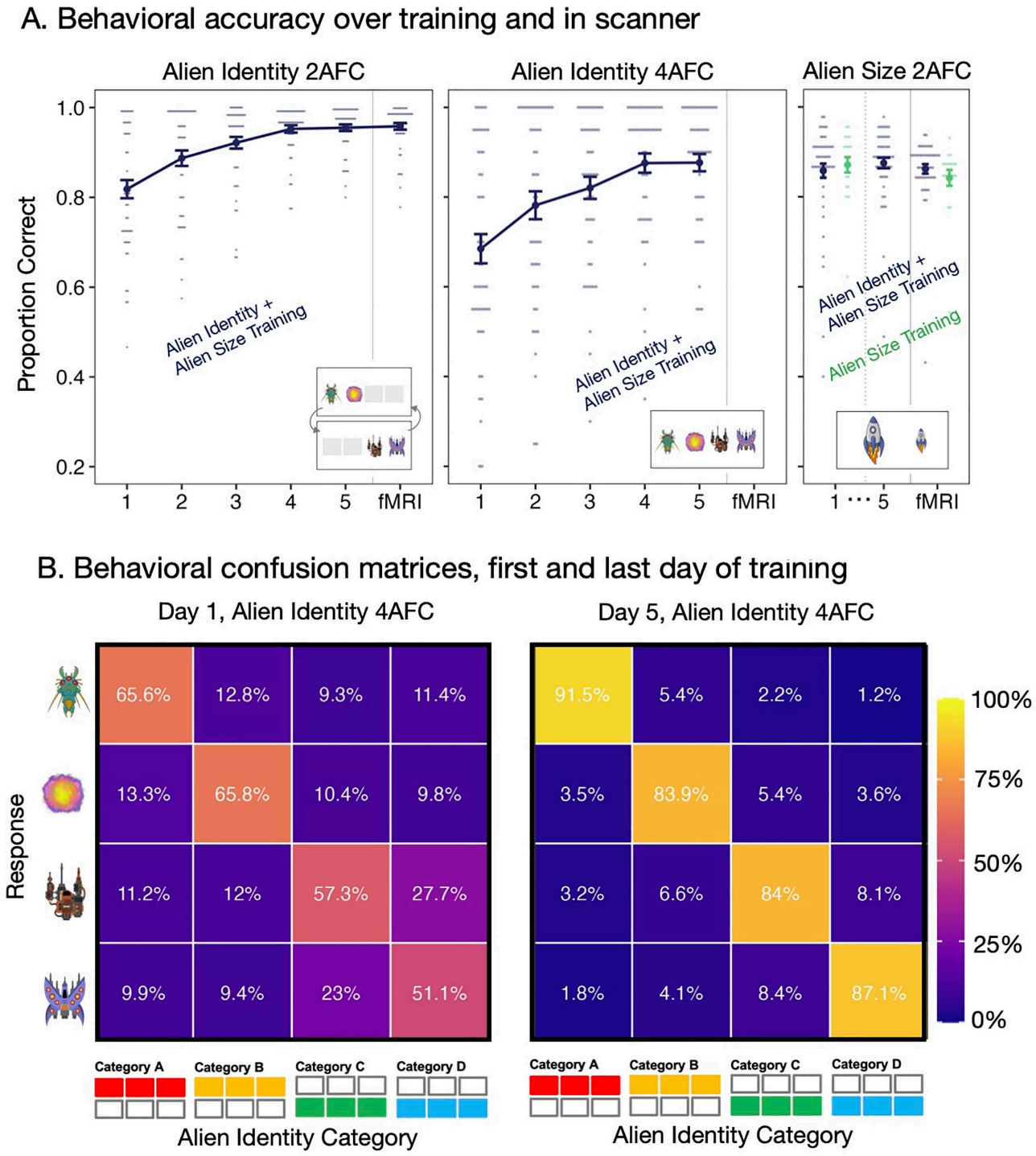


Fig 3. Behavioral Data. (A) Alien Identity Training (2AFC with feedback). Proportion correct 2AFC categorization across Days 1-5 (and in the fMRI scanner, with no feedback) for the N = 49 *Alien Identity+Size* participants who achieved 75% accuracy across all alien identity categories by Day 5. **Alien Identity Generalization (4AFC no feedback).** The same *Alien Identity+Size* participants' 4AFC categorization of novel category exemplars improved across Days 1-5. **Alien Size Training (2AFC with feedback).** A separate training task involved categorizing exemplars drawn from all 4 alien identity categories as 'big' or 'small' according to stimulus amplitude. The *Alien Identity+Size* participants who trained on alien identity trained on alien size on Day 1 and Day 5 of training, and in the scanner without feedback. A separate *Alien Size* group (N = 20) never trained on alien identity; this group trained

on alien size (defined by amplitude) for one day and performed the same task in the scanner with no feedback. **(B) Confusions in 4AFC Alien Identity Generalization.** Confusion matrices comparing category ground truth with response for Day 1 vs. Day 5 reveal increasing category generalization and specificity. Like Fig 2, x-axis rectangles for confusion matrices show diagnostic frequency band for each category (warm colors: high band, cool colors: low band). The data and scripts underlying this figure can be found at https://doi.org/10.5522/04/33974425, and direct links for Fig 3A and 3B are as follows: https://rdr.ucl.ac.uk/ndownloader/files/69528888 (Fig 3A-Panel1-SourceData); https://rdr.ucl.ac.uk/ndownloader/files/69528885 (Fig 3A-Panel2-SourceData); https://rdr.ucl.ac.uk/ndownloader/files/69528894 (Fig 3A-Panel3-SourceData); https://rdr.ucl.ac.uk/ndownloader/files/69528882 (Fig 3B-SourceData).

https://doi.org/10.1371/journal.pbio.3004076.g003

of novel alien exemplars had sharpened considerably across the board. In addition, the learning trajectories across which alien identity experts reached criterion expertise were heterogeneous; we used this variability to subgroup learners for further analysis (see below).

Alien identity experts also trained on the alien size task on Days 1 and 5 to a similar level of accuracy (M = 87.5%, SE = 1.2%; Fig 3A**, right panel**). The *Alien Size* participants (N = 20, sample size matched to the highest-performing 'Early Expert' subgroup of *Alien Identity+Size* group, see below) achieved 87.1% (SE = 1.7%) alien size categorization accuracy at the end of one training session (Fig 3A, **right panel**).

Both groups took part in a single fMRI session following training. In-scanner behavior echoed patterns of training performance (Fig 3A, **right panel)**. Those who trained on both alien identity and alien size successfully categorized alien identity according to patterns evolving in the diagnostic frequency bands (2AFC task, M = 95.8%, SE = 0.7%) and according to stimulus amplitude in the size task (M = 86.2%, SE = 1.1%). Participants trained only on alien size performed with similar accuracy (on alien size) during scanning (M = 84.2%, SE = 1.8%), compared to the group trained on both tasks.

Alien identity learning drives frequency-band-selective activation differences in tonotopic cortex

We first present functional MRI data from the *Alien Identity+Size* participant group trained on both alien identity and alien size tasks. To identify regions likely to differentially respond to category-diagnostic frequency bands, and to test differences in neural responses as a function of categorization task demands, we first characterized each participant's unique tonotopic organization on a voxel-wise basis across bilateral auditory-responsive cortex. Here, participants listened to concatenated series of 4-tone mini-sequences that stepped periodically over a 60-semitone range (175–5,286 Hz, two runs ascending and two descending, see Dick et al. [43] and Fig 1C, **Methods-Procedure**). To maintain task focus, participants performed a 1-back task, reporting infrequent mini-sequence repeats (in-scanner d': M = 3.55, SE = 0.08). From these data we compute voxel-wise activation to mini-sequences grouped into the low and high frequency bands that conveyed alien identity information. The difference in beta coefficients across high-minus-low frequency bands quantifies voxels' frequency selectivity. As expected from prior work [41], we observe coarse alignment of tonotopic organization across individuals (S1 Fig), but cross-individual heterogeneity in the mosaic arrangement of fine-grained frequency selectivity.

The 'concordance map' approach [43] described in the introduction allows us to capture consistent regional differences in responses across participants while retaining frequency sensitivity at a voxel-wise level, despite tonotopic map inter-subject heterogeneity. Here, each individual's auditory-responsive cortices are parcellated into a set of small, roughly equal-area ROIs that are projected via curvature-aligned spherical registration from a participant surface in Dick et al. [43], onto each individual's auditory cortical surface (Fig 4A, see **Methods-Analyses**). (The ROI borders are not designed to align to gyral anatomy or functional fields). Each ROI exhibiting frequency selectivity at the group level (as estimated using the group-average tonotopic map) is sampled into the cortical ribbon in the individual's native volumetric EPI space (Fig 4B). Analogous to the high-minus-low frequency selectivity contrast from the tonotopy runs, voxel-by-voxel beta coefficients relating to activation during alien identity categorization (involving the high- and low-frequency category-diagnostic bands) are computed and a high-minus-low frequency band difference in beta coefficients is calculated for each voxel. These parallel difference scores permit us to fit a regression model across all participants within each ROI (with participant as a random factor, Fig 4C) revealing the ROI-wise concordance between the voxelwise frequency selectivity evoked by

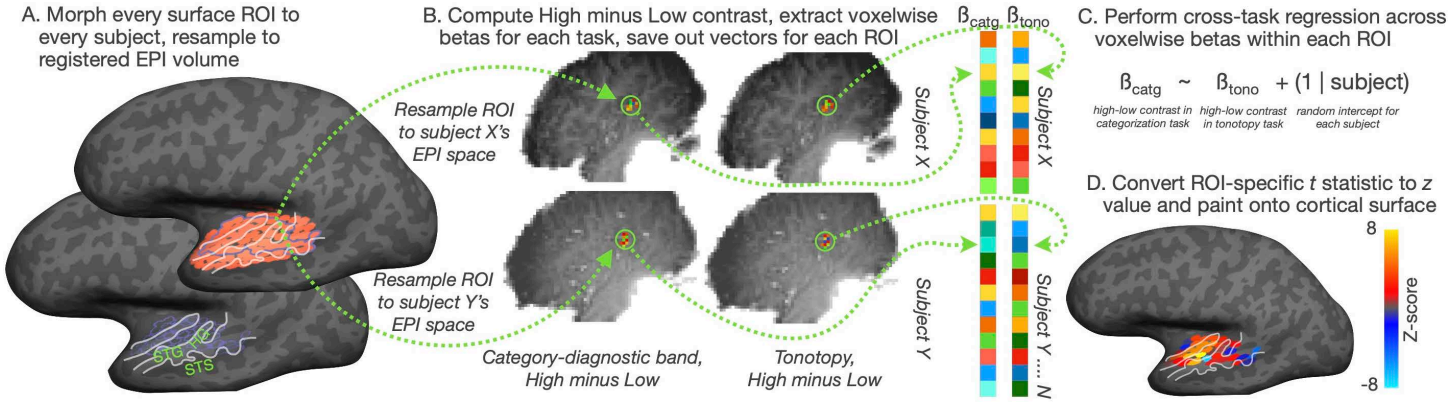


Fig 4. Approach to concordance analyses. To quantify the similarity in activation patterns across tasks, we defined a set of regions of interest (ROIs) and conducted cross-task concordance analyses within each ROI. **(A)** First, we took a set of cortical surface ROIs tiling the superior temporal plane; the ROIs were created to be roughly equivalent in area, with borders quasi-arbitrarily positioned relative to cortical landmarks (ROIs shown in orange; Dick et al. [43]). For the current work, we only considered ROIs that fell within tonotopically organized auditory cortex as determined by a group average map; these ROIs are outlined in purple, with gyral boundaries traced in white and anatomical labels indicated in green (HG: Heschl's Gyrus; STG: Superior Temporal Gyrus; STS: Superior Temporal Sulcus). Each ROI was morphed via cortical-surface-based registration to the individual participant's surface and then sampled to their native volumetric EPI space. **(B)** For the two categorization tasks, we always examined the contrast between the high-band-diagnostic and the low-band-diagnostic identity categories. We used this contrast even when the task was to categorize stimuli based on overall amplitude, because we hypothesized that, for the *Alien Identity+Size* group, relatively intensive training would result in acquired attentional gain for the identity-diagnostic frequency band, even when performing an unrelated task. For the tonotopy task, we used the contrast between bottom-up sensory input from high vs. low frequencies, and for the attention-tonotopy task, we used the contrast between explicit, directed attention to high vs. low frequencies in the context of bottom-up sensory input from both frequency bands. Within each volumetric ROI, we extracted voxel-by-voxel beta coefficients for the high vs. low contrast in each task. **(C)** To assess the similarity of activation across tasks, we submitted voxelwise betas to a regression analysis for each ROI, separately; the goal in each regression analysis was to predict the beta coefficients from one task using the betas from another. Regressions included by-subject random intercepts; note that each participant will contribute slightly different number of voxels to each regression model due to inter-individual anatomical differences. **(D)** The resultant ROI z-statistic, which indicates the strength and direction of the relationship between the two sets of beta coefficients, was painted on a single cortical surface for interpretation and comparison with other maps.

https://doi.org/10.1371/journal.pbio.3004076.g004

tonotopy stimuli, and the hypothesized, category-specific frequency-band selectivity driven by the alien identity learning task (Fig 4D and **Methods-Analyses**). (All of the inferential statistics for each subgroup, as well as cross-subgroup comparisons, are reflected in the statistical maps, with outlines showing FDR-corrected significance in the same way as with the full groups).

In other words, this approach allows us to ask whether categorization decisions guided by chirp patterns embedded in high versus low frequency bands yield activation in corresponding frequency-selective cortical regions. Since each category exemplar possesses rich acoustic energy in each frequency band as well as amplitude variation (signaling size), an exaggerated neural response in the identity-diagnostic versus non-diagnostic frequency band provides evidence for expertise-driven changes in cortical activation to the category exemplars.

We first established that performing category tasks with these novel 'alien' sounds evoked broad activation across much of auditory and auditory-related cortex bilaterally, as well as across a number of other brain networks and regions, including inferior, prefrontal, inferior parietal, and ventral and lateral occipital cortex (Fig 5, **top**). We then turned to the ROI-based regression approach using concordance maps. (Note that all maps show group averages or group comparisons, and do not include any behavior-activation correlation measures). As seen in Fig 5B, we found that, across multiple ROIs, categorization decisions requiring use of information in high versus low frequency bands drives differential activation in voxels exhibiting preferential responses to these acoustic frequencies. Most left hemisphere (LH) auditory cortical ROIs along the swath of the planum temporale – from Heschl's gyrus posteriorly to the anterior-most aspect – show a strong positive concordance between experience-dependent and category-diagnostic versus stimulus-evoked voxelwise

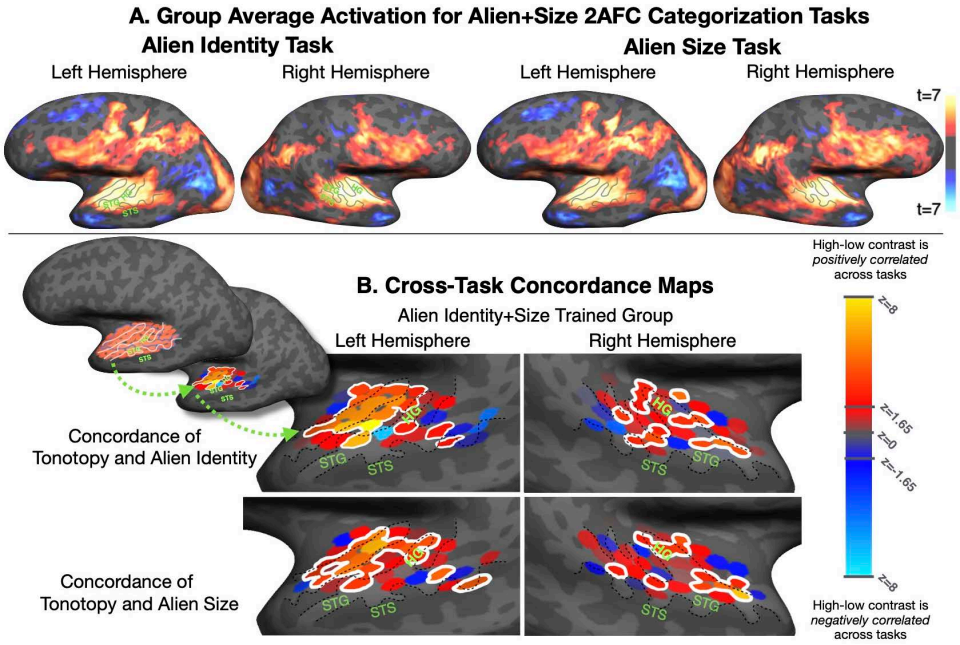


Fig 5. Group average activation and cross-task concordance maps. (A) Cortical-surface-based group averages of BOLD response to the alien identity and alien size 2AFC categorization tasks compared to resting baseline, for *Alien Identity+Size* experts. Average brain-wise responses to both tasks with the auditory category exemplars show extensive bilateral activation of auditory-related cortex in temporal, inferior parietal, and frontal regions, as well as lateral occipital cortical regions typically associated with visual object recognition. The activation maps are leniently thresholded at a $t > 2$ to show full activation patterns. The colorbar shows mapping of t-statistics; the grey tracing shows underlying curvature patterns, and anatomical labels are as in Fig 4. **(B)** Following the approach illustrated in Fig 4, we assessed the concordance among the tonotopy, frequency-diagnostic identity categorization and amplitude-diagnostic size categorization tasks. Each inset figure shows a blow-up of a given concordance map rendered on the left or right temporal lobe. In general, there was a strong positive relationship among betas, indicated by warm-colored ROIs. ROIs outlined in white are statistically significant ($p < 0.05$ after applying FDR correction, see **Methods-Analyses**; dotted black tracing shows underlying curvature patterns). The top row shows that there is greater cortical activation for category-diagnostic, compared to non-diagnostic, frequency bands during alien identity categorization among the *Alien Identity+Size* group. The second row shows that this pattern of activation persists even when categorization according to alien size demands reliance on an orthogonal dimension (amplitude). The bottom row shows the concordance map between alien identity and alien size tasks. Note that all maps show group averages or group comparisons, and do not include any behavior-activation correlation measure. The summary data used for this figure are available at https://doi.org/10.5522/04/33974425.

https://doi.org/10.1371/journal.pbio.3004076.g005

frequency selectivity. Some LH posterolateral ROIs along the superior temporal gyrus (STG) also show positive concordance. In right hemisphere (RH) auditory cortex, we observe significant positive cross-task concordance in multiple clusters of ROIs; in general, FDR-corrected groupings are separated by ROIs showing positive but subthreshold category-tonotopy associations. Across both hemispheres, the few ROIs with negative relationships across tasks tended to be situated in regions exhibiting relatively weak group-level tonotopic selectivity (S1 Fig).

Thus, across expert participants actively engaged in the alien identity category task, we observe that voxels in multiple auditory regions show greater cortical activation associated with the category-diagnostic versus non-diagnostic frequency band in a way that is predicted by the same voxels' stimulus-driven preferences for one or the other frequency band.

Greater cortical activation in voxels preferring category-diagnostic frequency bands persists even in behavioral contexts that do not demand this information

We hypothesized that learning may cause identity-diagnostic dimensions to acquire salience, such that frequency-selective recruitment may persist even when task demands shift to categorization across an orthogonal dimension. To test

this, in different blocks, participants also categorized exemplars from all four alien categories according to whether they were 'big' or 'small' aliens based on amplitude. (Note that sound exemplars varied in amplitude in both alien identity and size tasks, but amplitude was task-relevant only in the size task). As shown in Fig 3A **(right panel)**, in-scanner alien size categorization was accurate, but well below ceiling performance. This indicates that participants successfully relied upon the amplitude dimension for categorization, and that the alien size task was non-trivial.

Although task demands directed participants' overt attention to an orthogonal acoustic dimension (amplitude), concordance between voxels' stimulus-evoked frequency preference and exemplars' frequency-diagnostic band for the over-practiced and more challenging *identity* task persisted (Fig 5B). Across all *Alien Identity+Size* participants, in the LH, we observed significant concordance across most of the ROIs in the planum temporale, as well as two ROIs in the posterior STS. In the RH, we observed concordance in three small groups of ROIs in the middle inferior circular sulcus, the anterolateral STG, and two isolated ROIs along the posterior crown of the STG.

The differential cortical activation across the high versus low frequency bands is consistent with an 'acquired salience' for category-diagnostic dimensions that reflects stable changes in cortical activation elicited by category exemplars that are persistent across changing task demands. Prioritization of category-diagnostic dimensions is present even in contexts that do not demand reliance on them. This effect is challenging to account for via a 'bottom-up' acoustic salience account, as all stimuli possessed complex information in each frequency band, in addition to amplitude variation. Further, results from the repetition suppression fMRI literature would generally suggest that the acoustic difference between the category diagnostic- and non-diagnostic bands - namely the higher-level regularity in pitch contour across the three individual tokens in the diagnostic band – would diminish activation in auditory regions sensitive to that information [51,52], but cf. Segaert et al. [53]). In the present results, we observe the opposite pattern. To examine this question further, in the next two sections we make more direct tests of the impact of alien identity category expertise on activation to the category-diagnostic band.

Cortical activation mirrors individuals' trajectories of learning

All *Alien Identity+Size* participants who completed the fMRI scan (N = 49) had achieved high levels of expertise in a novel domain. Yet their paths to expertise varied (Fig 6A). To understand how these learning trajectory differences might affect neural representations, we grouped participants based on the speed and depth of learning over behavioral training. Early Experts (N = 17) achieved at least 90% accuracy on Day 1 for 2AFC training on both high- and low-frequency diagnostic bands. Late Experts (N = 16) started at lower accuracies but showed an accuracy increase of at least 20% over the course of 2AFC training and achieved accuracy similar to the Early Experts at the end of training.

We also estimated ultimate depth and generalizability of learning by creating separate groups based on generalization performance in the 4AFC task, with the High Achievers (N = 16) and Low Achievers (N = 16) defined across the top and bottom generalization terciles. These paired subgroups capture (partially overlapping) sets of participants with distinct learning trajectories.

We observe the fingerprints of these trajectories on subsequent cortical responses to category exemplars (Fig 6B – note that these maps show concordance maps and do not include any measure of correlation between individual participants' behavioral performance and neural activation). Compared to Late Experts, Early Experts exhibit significantly stronger concordance between alien-identity-categorization-driven and stimulus-driven voxelwise frequency selectivity in ROIs within anterior temporal regions, as well as posterior STG/STS (Fig 6B, **top three panels**). In the RH, Early Experts show significantly greater concordance in ROIs along much of the STG and postero-medially. During the Alien Size task, where spectral band was category-irrelevant, Early Experts also showed significantly stronger concordance in ROIs along much of the medial part of the left STG, with very limited group differences in the right hemisphere.

Similarly, in the Alien Identity task, the High Achievers exhibited significantly greater concordance than the Low Achievers in anterior temporal and STG LH ROIs, with similar group differences in the right hemisphere. (Fig 6B, **bottom three**

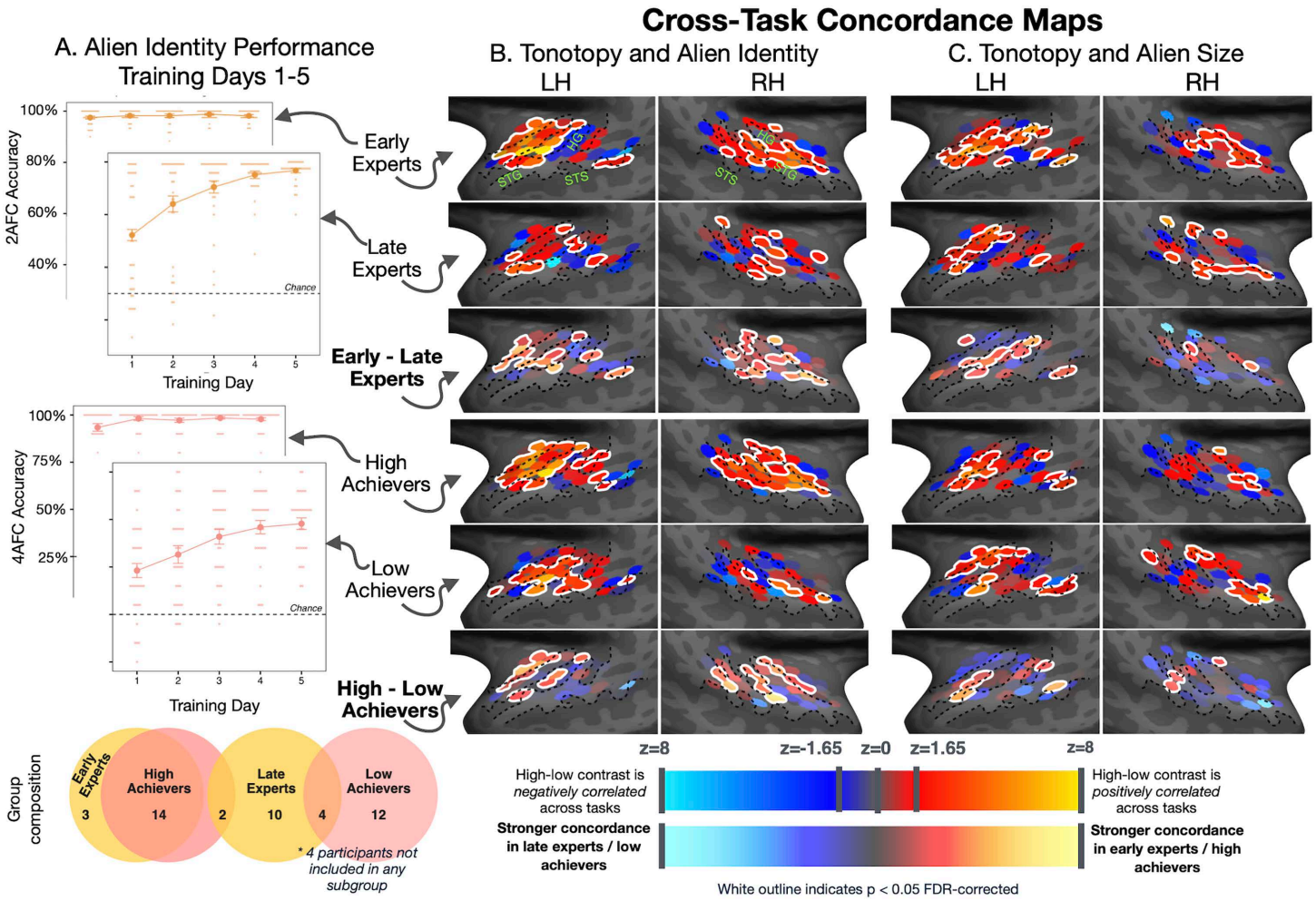


Fig 6. Sub-Group analyses tested for differences in cross-task concordance as a function of frequency-diagnostic categorization performance during the course of training. (A) One set of analyses considered differences as a function of *how quickly participants mastered the 2AFC Alien Identity+Size categorization task*: *Early experts* demonstrated mastery for both high-band-diagnostic and low-band-diagnostic categorization on Day 1, whereas *late experts* demonstrated substantial improvement over the course of training. A second set of analyses considered differences as a function of *generalization ability*, separately analyzing data from the top tercile (*highest achievers*) and bottom tercile (*lowest achievers*) on the 4AFC generalization task. Behavioral performance and Venn diagrams for these sub-groups are shown in the left panel. Light dots indicate individual participants (no datapoints are obscured by graph overlay), solid points indicate mean performance for each day, and error bars indicate standard error. **(B-C).** We assessed cross-task concordance between the tonotopy task and the alien identity (**B**) and alien size task **(C)**; the top two rows show Early and Late Experts concordance maps, with the third row showing the difference between Early vs. Late Experts maps. Parallel concordance maps for High and Low Achievers are shown in the bottom three rows. In the difference maps, warm colors indicate ROIs where stronger cross-task concordance was seen in early experts/highest achievers, and cool colors indicate ROIs where stronger cross-task concordance was seen in late experts/lowest achievers. As in Fig 5, z-statistics are shown in the colorscale, the white outline surrounds ROIs with FDR-corrected p-values at $p < 0.05$ (see also **Methods-Analyses**), and the black dotted outline shows underlying cortical curvature. The data underlying this figure can be found in the repository found at https://doi.org/10.5522/04/33974425; the direct link for the source data in Fig 6A is https://rdr.ucl.ac.uk/ndownloader/files/69528891.

https://doi.org/10.1371/journal.pbio.3004076.g006

panels). In the Alien Size task, High and Low Achievers differed less, with High Achievers showing greater anterior and posterior LH STG concordance, as well as in a few posterior STG and one medial anterior ROI in the right hemisphere.

This indicates that, despite their shared ultimate success in learning and in-scanner task performance, and quite similar overall patterns of concordance (as seen in Fig 6B, rows 1,2, 4,5) *Alien Identity+Size* participants' *degree and rate of*

learning of identity category expertise is associated with specific patterns of brain activation that reflect the degree of tuning to category-diagnostic perceptual dimensions. An important caveat to these findings is that these behaviorally-defined subgroups are relatively small compared to the full sample size, so follow-ups with larger N samples will be helpful for interpreting the strength of these effects.

Learning-driven attentional gain for a diagnostic dimension depends on the demands of learning

As described above, expert-level categorization drove activation of auditory areas tuned to the frequency band diagnostic of alien identity, even when *Alien Identity+Size* experts were successfully categorizing alien size based on stimulus amplitude. To make a critical test of the importance of learning to weight or attend to the identity-diagnostic frequency band in driving these results, we compared the *Alien Size* trained group (N = 20) - which remained naïve to alien identity and *never* trained to learn categories based on frequency-band-delimited acoustic information - to the Early Expert subset (N = 17) of the *Alien Identity+Size* trained group.

As would be expected, alien-size-trained participants exhibited a qualitatively equivalent degree of frequency selectivity and tonotopic organization in auditory cortex (S1 Fig), relative to the Early Expert group that trained on both alien identity and size. They also showed high behavioral accuracy on the tonotopic mapping task (in-scanner d': M = 2.68, SE = 0.13). Having established typical patterns of tonotopic mapping, we next tested whether alien-size-only-trained participants' auditory cortex would show differential activation for stimuli with identity-disambiguating information in high and low frequency bands when they were performing the amplitude-based alien size task. If there is a systematic activation preference for the alien-identity-diagnostic frequency band among participants trained only on alien size – particularly with a similar pattern of sensitivity to alien identity experts – this would suggest that the spectrotemporal regularities present in the identity-diagnostic frequency band are driving at least some of the results we have observed with alien identity experts, potentially via learning-induced attentional capture (Zhao et al. [54]).

Though the alien-size-trained participants demonstrated high behavioral accuracy on amplitude-based size categorization during scanning (M: 84.2%, SE: 1.8%, Fig 3A right panel) they exhibited little sign of activation for the identity-diagnostic frequency band that matched underlying tonotopy, with only a few ROIs showing significant concordance, and only in the most peripheral regions that are inconsistently and weakly frequency-selective across participants (Fig 7A). By comparison, as shown previously in Fig 6C (top right panel), the Early Experts trained on both identity and size tasks showed strong concordance across ROIs between the difference in activation to high- versus low-frequency *identity-diagnostic* band stimuli during alien-size categorization blocks, and high- versus low-frequency tone mapping. A direct statistical comparison between the two groups showed that most LH ROIs just medial to and along the STG, and multiple patches in the RH STG and more medially showed significantly greater frequency-selective modulation in the amplitude-conveyed alien size task among Early *Alien Identity+Size* Experts compared to those trained on alien size only (Fig 7B).

To test the generality of this result, we ran the same comparison between the *Alien-Size-Trained* cohort and the complete *Alien Identity+Size* cohort (N = 49). As seen in S2 Fig, we observe a similar difference in concordance maps, where the Alien Identity+Size cohort shows significantly more concordance in most of the LH ROIs from the middle of the STG extending medially, as well as long the posterior STG. In the RH, the greater concordance in the full *Alien Identity+Size* cohort is again very similar to that observed in the comparison of Early *Alien Identity+Size* experts versus Size experts, with greater concordance in multiple ROIs extending laterally to medially from the middle of the STG, as well as in anterior STG. This adds further support for the hypothesis that learning categories that are dependent upon statistically structured auditory information restricted to one frequency band drives the acquired salience of that structured information [55], such that identity categorization experts (but not identity-naïve participants) exhibit frequency-band-specific recruitment of auditory cortex, even when categorizing stimuli along an orthogonal dimension (stimulus amplitude).

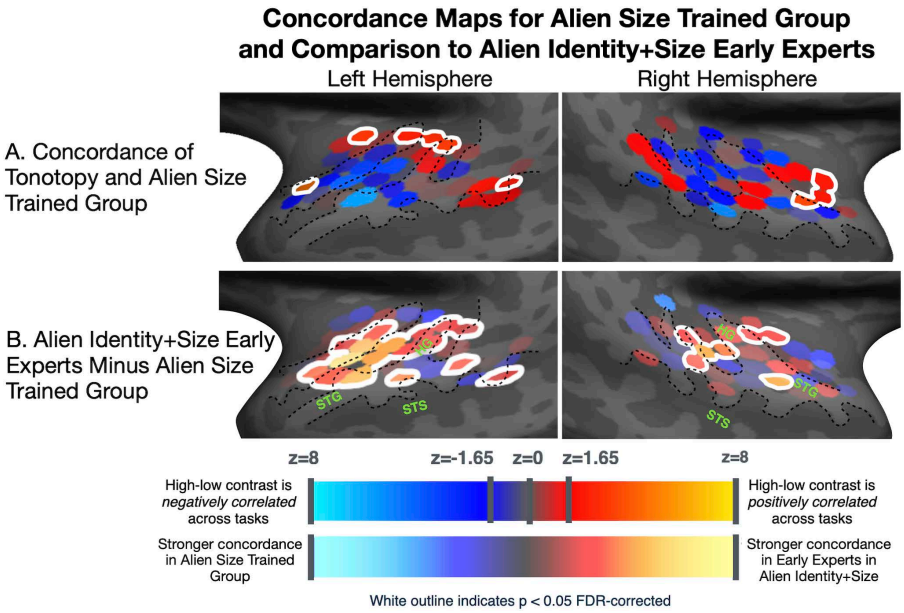


Fig 7. Concordance Maps for *Alien Size* Trained Group and Comparison to Early Experts in *Alien Identity+Size* Trained Group. (**A**) Shows the *Alien Size* trained participants (N = 20) concordance map between high-versus-low-frequency responses from tonotopy, and response differences in the alien size fMRI task when comparing activation for alien sounds with *alien identity* information in high vs. low spectral bands. (Recall that this participant group was naïve to alien identity.) With the exception of a few medial ROIs, participants did not recruit auditory cortex in a frequency-selective way when performing amplitude-based categorization. **(B)** shows the difference between the analogous concordance map for the Early Expert group of *Alien Identity+Size* trained group (top right panel of Fig 6C) vs. the *Alien Size* only trained participants. This difference map suggests that the frequency-selective recruitment of auditory cortex that *Alien Identity+Size* experts exhibited in the amplitude categorization task was not driven by properties of the acoustic signal per se. The summary data used for this figure are available at https://doi.org/10.5522/04/33974425.

https://doi.org/10.1371/journal.pbio.3004076.g007

Frequency-selective activation patterns driven by category learning align with those evoked by overt attention

We have observed that when categorization implicitly hinges on information in a delimited frequency band, voxels more responsive to this band show increased activation compared to when diagnostic information is conveyed via the less preferred frequency band, despite simultaneous acoustic energy in each band. Motivated by theoretical accounts suggesting that perceptual weighting reflects attentional gain, we next asked whether the pattern of increases in activation for category-diagnostic frequency bands is aligned with activation evoked by *explicit* attention to the same frequency band. Among participants trained on *Alien Identity+Size*, we compared the activation difference between high- and low-frequency identity-diagnostic bands with the parallel activation differences evoked by explicit, sustained attention to one of two simultaneously presented streams of sinewave tones situated within the same high- and low-frequency bands.

In addition to the tonotopy and alien categorization tasks, the *Alien Identity+Size* participants also completed a sustained auditory selective attention task based on the approach of Dick et al. [43], who found that overt attention directed at one of two sequences of tones positioned in specific frequency bands elicits an attention-driven map across auditory-responsive cortex. As shown in Fig 1C, participants heard two simultaneously presented sequences of sinewave tones composed of concatenated 4-tone mini-sequences that varied in amplitude; mini-sequences were played simultaneously in higher and lower frequency bands, corresponding to the diagnostic bands of the alien identity task (see **Methods-Procedure**). At the beginning of each task block, explicit instructions directed participants to detect mini-sequence repeats in the high frequency band, the low frequency band, or repeats of the overall relative amplitude of the mini-sequence (1-back task d': M = 2.28, SE = 0.13). As anticipated given previous studies of frequency-selective attention [42–45], voxelwise differences in activation when attending to high versus low frequency bands were strongly associated with stimulus-evoked frequency preference

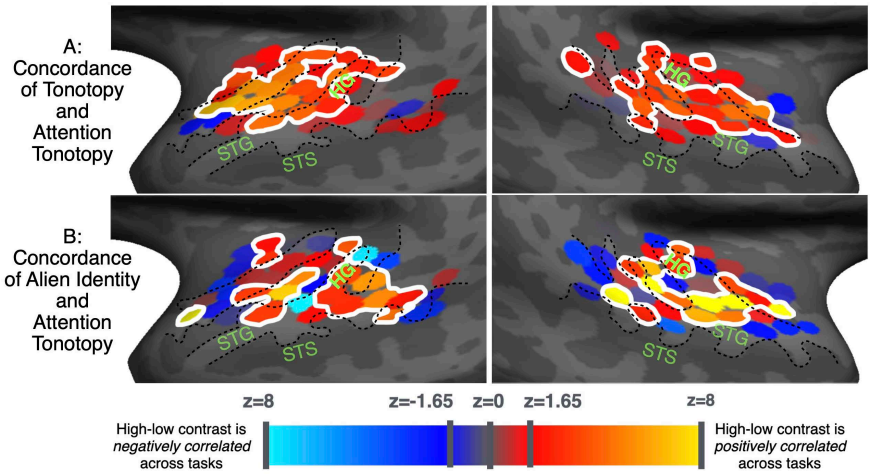


Fig 8. Frequency-selective activation patterns driven by category learning align with those evoked by overt attention. Theoretical accounts suggest that the neural selectivity that emerges with learning is driven by increased attention to diagnostic dimensions. To test this hypothesis, we assessed cross-task concordance with the attention tonotopy task (Fig 1C), where participants were presented with two streams of short tone sequences, and each stream was in a high- or low-frequency band corresponding to the high and low identity-category-diagnostic frequency bands in the *Alien Identity*+-*Size* task. Tone sequences also differed in overall amplitude. Participants were instructed to detect repeated 4-tone sequences in one or the other band or to overall perceived loudness, depending on block. **(A)** Analyses considering all participants showed strong concordance between attention-driven activation to high vs. low frequency bands and the high vs. low frequency band activation difference elicited by the tonotopy task (i.e., stimulus-driven activation). **(B)** Concordance between alien identity categorization-driven activation to high vs. low diagnostic frequency bands, and explicit attention to the same frequency bands populated with simple sequences of tones was observed bilaterally along much of the lateral aspect of tonotopically mappable auditory cortex. The summary data used for this figure are available at https://doi.org/10.5522/04/33974425.

https://doi.org/10.1371/journal.pbio.3004076.g008

(from the tonotopy scans) across most LH auditory cortex ROIs except in the anterior lateral and posterior STG (Fig 8A). In the RH, spectrally directed attention evoked similar activation differences as tonotopic scans along the lateral planum temporale and STG.

We then asked whether patterns of cortical activation elicited by *explicit, directed attention* to the frequency bands align with activation patterns that *arise in category decisions that are implicitly reliant on information* in these same frequency bands. As shown in Fig 8B, cross-task concordance maps showed that analogous differences in activation to high versus low frequency bands were observed laterally in both hemispheres, in a swath of ROIs along the central portions of the STG and abutting ROIs in the planum temporale.

In summary, lateral aspects of auditory cortex that show differential activation when listeners explicitly attend to a delimited region of the frequency spectrum also show correspondingly differential activation when categorization *implicitly* hinges on information situated in these same frequency bands. This holds true despite the considerable acoustic differences between the sequences of pure tone stimuli used for the explicit attention task and the spectrotemporally more complex and structured alien sound category exemplars.

Discussion

How does perception come to weight the most diagnostic patterns of input for the task at hand? We used fMRI to probe the neural consequences of learning to categorize across a novel soundscape defined by multiple, complex acoustic dimensions evolving in time. To do this, we situated category-diagnostic patterns into frequency-delimited bands – thereby coarsely segregating category information into different channels in tonotopically organized auditory cortex. This allowed us to make targeted predictions about: (a) how categorization-driven modulation of cortical activity would intersect with existing tonotopic functional regionalization; (b) how different trajectories of learning to expert-level performance would be reflected

in these representations; and (c) how a neural signature of 'learned salience' and/or attentional gain for the category-diagnostic information might align with overt attention directed toward distinct stimuli in the same frequency bands.

We observe that category learning drives prioritization or gain across a category-relevant dimension. This is reflected in greater cortical activation in voxels whose frequency response is associated with the category-diagnostic frequency band, compared to the simultaneously presented non-diagnostic band. This prioritization persists even when the in-scanner task shifts to categorization based on a completely different auditory dimension, amplitude. Further, it partially aligns with the patterns of cortical activation that emerge when listeners explicitly direct sustained attention to unrelated sounds situated in the same frequency band. This is consistent with attentional gain accounts of advantaged perceptual processing of category-diagnostic dimensions. These neuroimaging results support the notion that learning complex categories can drive the emergence of acquired attentional gain implicitly directed to category-relevant perceptual dimensions.

The question of whether - let alone how - sensory cortical representations are impacted by the acquisition of category expertise has not had a clear answer. Evidence of cortical modulation with category expertise comes predominantly from visual cortex under conditions of active categorization of static images [14,56]. However, these results are mixed, with some studies not observing effects of categorization on visual cortex [57–60] potentially consistent with category representations arising in flexible, more 'amodal' cortical areas with little lasting influence on sensory cortex ([57,61,62], but cf. Hausfeld et al. [63] and Henderson et al. [12]). Jiang et al. [64] report auditory category learning results consistent with this two-stage model, and others have made a case for prefrontal involvement in auditory categorization [65]. In contrast, our results indicate stable, training-induced changes in auditory cortex that persist even when task demands direct attention away from the category-diagnostic dimension (see also Folstein et al. [66], and Van Culick and Cauthier [16] for evidence from repetition suppression across visual categories).

These changes were driven by the specific demands of learning, and not by mere exposure to statistically structured sounds. Participants who trained to use frequency-band-diagnostic information to categorize alien identity exhibited neural selectivity consistent with acquired attentional gain to the diagnostic frequency bands. But those who learned the amplitude-signaled size categories across the very same sounds, thus experiencing the identity-diagnostic spectral information without learning to use it to categorize, did not. One caveat to this conclusion is that *Alien Identity+Size*-trained participants had overall much more exposure to alien sounds than *Alien Size*-trained participants, which could potentially drive differences. However, the fact that neural selectivity differed with learning trajectory and ultimate behavioral expertise within the *Alien Identity+Size*-trained group would argue against this explanation.

Although humans and other organisms exhibit an ability to learn across passive exposure to statistically structured input [67] [68], studies have observed that complex acoustic exemplars that model the complexity of speech are not well learned over passive exposure [69,70]. Yet alignment of these statistically structured sounds with active behavior – even incidentally and without overt feedback – can rescue category learning [35,71,72]. The cross-group differences in the present study demonstrates that the alignment of feedback with a diagnostic dimension played a role in eliciting the cortical selectivity toward those dimensions. These patterns of acquired attentional gain in human auditory cortex connect with nonhuman animal electrophysiological studies that observe that the task-relevance of an acoustic dimension drives changes in receptive fields consistent with attentional gain [73,74].

In this vein, Yin and colleagues [75] present evidence for top-down modulation of auditory cortex by frontal cortex in ferrets, along with induction of intrinsic auditory cortical patterns related to acquired categories that persist even in passive listening. Though we limited our investigation specifically to tonotopically mappable temporal cortex, the regions we observe to be modulated by category learning receive rich top-down feedback inputs, and exhibit modulation by behavioral context that extends beyond basic representation of perceptual dimensions (e.g., Zhong et al. [76]). Such top-down modulation has been well-documented in nonhuman animal electrophysiology (particularly in the visual system, e.g., Freeman et al. [57]; DeCutis and D'Esposito [77]) and may interact with the persistent changes to auditory cortical representation arising from category learning that we observe. However, as Addleman and Jiang [78] note, the neural mechanisms

underlying shifts of experience-driven attention may or may not be supported by the same frontoparietal regions as goal-driven attention. The lack of reliably definable tonotopic maps in frontal and other non-auditory regions using 3T fMRI makes it challenging for the present study to adjudicate this question. However, very recent human intracerebral stereo-electroencephalography (sEEG) reports prefrontal and inferior frontal spectrotemporal selectivity at multiple electrode sites [79] suggesting that ultra-high-resolution fMRI using the present approach may yield fruitful insights.

In the present study, we see that participants - above all the most expert - show selectively greater activation for the category-diagnostic frequency band. Importantly, this effect is widespread across bilateral auditory cortex, particularly in anterior and lateral portions. This degree of auditory cortical modulation by the category-diagnostic dimension - even when task-irrelevant - is striking by comparison with other studies of sensory cortical category sensitivity which tend to show greater category readout in frontal regions (e.g., Jiang et al. [59,64]). In situating category-diagnostic dimensions into discrete frequency bands, our approach may have afforded greater sensitivity by aligning learning-driven cortical changes with the primary representational axis of auditory cortex.

The question of whether such category-learning-related modulation of neural activity is a consequence of learned attentional gain to diagnostic dimensions also has surprisingly little direct support, despite the centrality of this proposition in theories of visual [80–82] and speech [83] categorization. It is clear that explicit attention to stimulus features modulates representations in both visual [84–89] and auditory (Da Costa et al. 2013; [43,45,63,90]) cortical response. Here, we asked whether patterns of cortical gain driven by explicit attention are concordant with patterns that emerge implicitly across category-diagnostic dimensions in categorization.

Consistent with our prior research [43], sustained auditory selective attention across sequences of tones presented simultaneously in two frequency bands results in greater activation across the attended frequency band. Here, in the same listeners, these patterns aligned with the prioritization of frequency-selective cortical activity driven by categorization demands, but primarily in more lateral auditory cortical regions. Of note, this concordance is not driven by task demands; prioritization of the diagnostic frequency band persists even when task demands change and frequency band is no longer task-relevant. This suggests that prioritization of diagnostic information has become a stable, lasting part of how category exemplars are represented in experts that is absent in category-naïve listeners. To our knowledge, this presents the first empirical evidence linking acquired gain to category-diagnostic dimensions of acoustic input and overt, directed selective attention.

We used a novel soundscape that participants would not have previously encountered to minimize interactions with existing auditory representations. This soundscape was designed to model the core learning demands of natural auditory categories, such as speech: listeners must identify the dimensions most diagnostic of category identity and learn their acoustically variable patterns amid competing dimensions, amplitude fluctuations, and background noise. We find that expert listeners who have successfully met these demands display cortical response signatures consistent with implicit, automatic attentional gain for category-diagnostic dimensions. This provides a mechanistic basis for a foundational, yet previously untested, theoretical tenet of speech categorization [83], with implications for understanding and potentially remediating learning-related communication disorders like dyslexia that have been linked to auditory learning [91] and auditory category learning [92,93] challenges. Furthermore, our finding that the specific learning trajectories by which listeners achieve categorization expertise leave a lasting imprint on cortical activation patterns suggests a neural basis for individual differences in speech categorization that persist even in neurotypical adults [94].

We should emphasize that auditory cortical fMRI activation differences driven by categorizing alien identity are decidedly not identical to those evoked by overt attention to pure tones with frequencies in the higher or lower category-diagnostic spectral bands. Although tonotopy concordance with alien identity and attention tonotopy activation is overall quite similar (with both showing particularly strong concordance in left anterior ROIs, and to a lesser extent along the right STG), direct concordance between alien identity and attention tonotopy high-low frequency difference maps is greatest

along bilateral STG, and not the anterior-most left ROIs. While our previous work [43] did show particularly consistent attentional tonotopic modulation along many of these STG regions, it is unclear what is driving this left anterior difference.

One limitation for interpretation lies in attention tonotopy task performance; the attention task was difficult, particularly for some participants, and lapses in sustained attention to one or both frequency bands by a subset of participants could influence the robustness of results. (The task was also run at the end of the long fMRI session, so some participants were somewhat fatigued). Follow-up studies would benefit from a longer training regime for the attention task, similar to that for the alien identity task, to assure consistent performance across runs. A further limitation is that the ROIs were created and morphed on a hemisphere-specific basis, thereby precluding quantitative cross-hemisphere comparisons of qualitative differences, such as the qualitatively higher concordance between attention tonotopy and alien identity tasks in the right compared to left STG.

More globally, the current results shed light on what processes might be leading to expertise-driven changes in neural recruitment [22,35,95–97]. Often, experts show selectively increased sensory cortical activation for the object of expertise (e.g., [98,99]), but the causal change in neural or cognitive processing leading to this activation can be challenging to ascertain [100,101]. Here we show that selective activation changes in experts are systematically related to the locus of category-disambiguating information, even when that information is not immediately task-relevant. This, along with the relative similarity of category- and attention-selective frequency modulation maps (Fig 8), lends support to accounts suggesting learned expertise drives a lasting attentional bias for category-relevant dimensions (Folstein et al. [102]). Alternative accounts emphasize dimensional stretching (Folstein et al. [102], or attention-gated plasticity Sasaki et al. [62]). It is important to note that extensive experience with perceptual categories or objects, such as one's own name or ringtone, can promote automatized overall attentional capture (e.g., [103]). However, the neural signature of such global 'attention-grabbing' effects will differ from those evoked by dimension-selective attention when the category-diagnostic perceptual dimension is topographically mapped in cortex (for instance across retinotopic space, or acoustic frequency as in the present study), and the presence of category-informative information may be differentially distributed within subregions of this space [99,104].

Disambiguation of different accounts could be furthered through explicit modelling of neural response functions that would efficiently partition category exemplars (e.g., Liu et al. [105]). Such modelling may help to shed light on why there is greater category-selective modulation of activation in anterior, versus posterior, auditory cortical regions, particularly in the most expert listeners. Additionally, data that disambiguate signals from upper versus lower cortical layers would potentially reveal more fine-grained information about neural encoding of category-relevant information [106] as well the role of learned attention in such encoding [107,108].

In future work, it will be important to show that the increased activation to identity-category-relevant frequency regions that is evident when they are not task-relevant (e.g., during size judgments) is not resulting from carry-over across experimental blocks. While we believe this is unlikely given the difficulty of the amplitude-conveyed alien size task (where weighting by spectrotemporal cues is counterproductive), it is possible that the activation patterns indicating acquired gain of the identity-diagnostic frequency band may diminish when not regularly reinforced via interleaving with blocks of the identity task.

Use of a more naturalistic or captivating training paradigm [71,109,110] would promote extended training and consolidation, which would be particularly revealing in tracking slower learners like those who did not achieve full mastery in the present study. Greater training time, along with a larger sample for fuller examination of individual differences, would allow for better understanding of potential non-linearities in the relationship between learning and changes in category-related neural sensitivity. As well, since the present study establishes expertise-driven attentional gain to diagnostic dimensions, future studies could observe the emergence of this attentional gain as participants learn categories during scanning. This will reveal how category learning shapes neural activation on a minute-by-minute or session-by-session basis (in the style of Lim et al. [71] and Ittipuripat et al. [111]). We also predict increasing learned salience such that identity category experts

would exhibit further increases in gain for the identity-diagnostic band in the size categorization task, despite the task-irrelevant nature of frequency-band-delimited alien identity information in this task (see also Wang et al. [55]).

In all, learning categories results in prioritization of category-relevant perceptual information across a representational dimension reflected in systematic changes in neural response to complex sounds. This category-selective difference in activation persists even when the dimension is irrelevant to the task at hand and aligns closely with patterns of activation that emerge with explicit, sustained attention to the dimension across unrelated sounds. Thus, acquiring new categories can drive the emergence of acquired attentional salience to category-diagnostic dimensions.

Methods

Ethics statement

This study was approved by the Institutional Review Board of Carnegie Mellon University (approval STUDY2019_00000322), and by the University College London (UCL) Psychological And Language Sciences Ethics Committee (approval number fMRI_2022_002). All participants provided written consent. The study has been conducted according to the principles expressed in the Declaration of Helsinki.

Participants

Our study involved two groups. The *Alien Identity+Size* group trained to categorize sounds according to alien identity (based on complex spectrotemporal patterns) and, separately, to categorize the same exemplars according to alien size (based on amplitude). An *Alien Size* group trained only on alien size. All participants were fluent English speakers with normal hearing, had normal or corrected-to-normal vision and no history of neurological impairment or speech/language disorder, and reported no experience with a tonal language.

Fig 1A shows the participant groups. The *Alien Identity+Size* group included participants from the Pittsburgh, PA, USA and London, UK communities who completed a five-day training protocol (N = 95; 18–40 yrs). Individuals who met criterion levels of expertise in alien identity categorization (75% across both diagnostic frequency bands on the 2AFC alien identity task on Day 5) were invited to participate in a subsequent MRI session, with data collected from N = 54 experts. Five participants' data were excluded (motion, non-compliance with task) resulting a final sample of N = 49 experts (18–37 yrs, M = 23.9 yrs; 29 female, 18 male, 2 non-binary; 59% White; 32 Pittsburgh, 17 London).

The *Alien Size* group (N = 20, 18–43 yrs, M = 28.1 yrs; 12 female, 8 male; 55% White; 20 London) trained only on the alien size task and participated in an abbreviated MRI session.

Stimuli

Four auditory categories were sampled from a highly complex perceptual space (see Obasih et al. [40]). As shown in Fig 2, each stimulus was composed of patterned acoustic energy in two non-overlapping acoustic frequency bands (~80–750 Hz and ~1,000–9,000 Hz). Each band was populated with three 400-ms tonal 'chirps' (100-ms ISI) varying in frequency contour. In the *alien-identity-diagnostic* frequency band (high-frequency for Categories A, B and low-frequency for Categories C, D), the three acoustically variable chirps possessed an underlying regularity that defined alien identity category membership. The three temporally aligned chirps in the non-diagnostic frequency band (low for Categories A, B; high for Categories C, D) were also acoustically variable but possessed no coherent underlying regularity.

To create a large set of naturally varying category exemplars, we derived chirps from four native Mandarin Chinese talkers (two female) uttering single-syllable utterances in each of Mandarin's four lexical tones, which convey variable, but structured, fundamental frequency (F0) contours. We extracted the F0 contour from each and resynthesized the contour to create a non-speech chirp, time-normalizing each to be 400 ms. We pitch-shifted each chirp +33 semitones, then

applied a 1,000 Hz high-pass filter to create a pool of chirps to populate the high frequency band. Equivalently, to create chirps for the low frequency band, we pitch-shifted the F0 contour −1 semitone and applied a 500-Hz low-pass filter.

Acoustic information in the category-diagnostic band consisted of three distinct chirps derived from the same talker and the same Mandarin lexical tone category, thereby conveying an acoustically variable but structured regularity in frequency contour. Acoustic information in the non-diagnostic band consisted of three different chirps drawn from three *different* Mandarin lexical tone categories (from among any of the four) of the same talker such that there was no underlying acoustic regularity across chirps. In all, this created a stimulus pool of 36,000 novel category exemplars [40] from which we sampled 592 unique stimuli.

In addition, we also imposed orthogonal amplitude variation (between 62−80 dB) on copies of each of these 592 sounds. We manipulated amplitude on an exemplar-wise basis after the two spectral bands were combined, so amplitude manipulations scaled equally over the entire spectrum. This met two goals: to increase variability and task difficulty, and to permit creation of two *Alien-Size* categories, with 'Big Alien' exemplars falling between 72−80 dB sound intensity, and 'Small Alien' exemplars falling between 62−70 dB (all reported dB referenced to $2*10^{-5}$ and measured in Praat). Thus, each exemplar could be categorized along identity (spectrotemporal, frequency-band-specific) or size (amplitude) dimensions. Not only were stimuli acoustically identical across Big and Small alien size categories aside from amplitude, they were also identical across the Size and Identity tasks, with the only difference being the task that participants performed on the stimuli.

Collectively, these stimulus features impose several key demands on learning: (1) Because chirp triplets appear in two distinct frequency regions, but statistically structured patterns diagnostic of alien identity occur in only one, listeners must first discover which frequency band carries meaningful regularities; (2) Because two distinct alien identities are conveyed by different patterns within the statistically structured frequency band, listeners must learn how specific regularities map onto identity. Beyond frequency-based demands, identity-category learning requires listeners to generalize across identity-irrelevant amplitude variation, whereas size-category learning requires them to rely on amplitude as the relevant cue. Identity and size category training take place against a background of scanner noise, modelling learning under adverse listening conditions.

Procedure

Online categorization training was implemented using Gorilla software (Anwyl-Irvine et al. [112]), the Google Chrome browser, wired headphones (with headphone compliance screened as in Milne et al. [113] and wired headphone usage reinforced in recruitment instructions), and a computer/laptop (not tablet/phone). Stimuli were created with Matlab version 2021b (The MathWorks), Praat [50] and SoX version 14.4.2 (www.sourceforge.net).

Behavioral Training Prior to fMRI. Fig 1B-1C illustrate the approach to training. *Alien Identity+Size* participants learned to associate 1,400 ms novel sounds with one of four alien images via category-informative feedback identifying the correct alien after each response. Each of 5 daily sessions involved four blocks, each composed of 120 alien identity training trials followed 20 trials to test generalization of learning (Fig 1C). Alien identity categorization training involved 2AFC trials in which participants indicated alien identity with a keypress, and feedback (1,500 ms) indicated the correct alien. In each 120-trial training block, trials were segregated according to the category-diagnostic high or low frequency band (Categories A versus B and Categories C versus D). Every 20 trials, the category pair, and thus diagnostic frequency band, alternated; the other two alien-identity response options were blanked out as response alternatives (Fig 1C), with order counterbalanced across participants. Training stimuli were randomly selected from a pool of 512 exemplars (128/category). Amplitude varied from 62–80 dB across trials and, over the entirety of training, was balanced across this range.

Each 2AFC alien identity training block was followed by a 20-trial 4AFC generalization block with all aliens as response options, and no feedback (Fig 1C). These trials involved a separate pool of 20 exemplars/category that were never presented in training to assess generalization of learning (Fig 1B-1C). Across four 2AFC training and 4AFC generalization

blocks, category exemplars were mixed with a fixed level of scanner noise (57 dB) to familiarize participants with categorization in an MRI environment. The visual position of aliens (and associated keypress) was fixed during the first two days of training, then allowed to vary starting on Day 3. (Note that Obasih et al. [40] found that alien category learning speed and attainment were surprisingly unaffected by different training regimes, including manipulations of exemplar variability, blocking versus interleaving categories in training, and overt instructions to attend to the diagnostic band.)

On Day 1 and Day 5, these participants also trained on alien size categorization (Fig 1B-1C). Across the first eight trials, one exemplar from each of the four categories was presented at the highest amplitude (80 dB), then one exemplar from each alien category was presented at the lowest amplitude (62 dB). The amplitude difference diminished in regularly incremented dB steps every 4 trials for a total of 40 trials, with 71 dB defining the category boundary between 'big' and 'small' aliens. In a second 40-trial block, amplitude values were fully randomized with equiprobable big/small exemplars and an equal number of exemplars from each alien category. A fixed level of scanner noise (again 57 dB) was mixed with stimuli to mimic the MRI sound environment. Additionally, a separate *Alien Size* participant group completed a single session of 2AFC alien size category training over 80 trials (Fig 1B).

Tonotopy. Participants reported infrequent mini-sequence repeats in a 1-back task to characterize voxel-wise frequency selectivity (Fig 1C). Individuals were scanned as they listened to sequences of 178-ms sinewave tones (0 ms ISI, 5 ms ramp at on/offset), organized in 4-tone mini-sequences, each separated by 356 ms. The frequency range of the four tones composing the mini-sequences increased or decreased in frequency incrementally in 10 logarithmically scaled steps to sweep across 175–5286 Hz (60 semitones), each with 6 mini-sequences. Across the 60 mini-sequences comprising each 64-sec sweep, 3 mini-sequence repeats occurred quasi-randomly. Participants reported repeats with a button press. The swept frequency permitted fMRI analysis using a) phase-encoded mapping, where voxels responsive to a particular frequency should respond at a consistent phase delay [43] (Dick et al. [114]; Sereno et al. [115]) and b) multiple regression analysis to identify voxels responsive to the range of the lower (< 500 Hz, lowest three steps) and higher (≥1000 Hz, highest five steps) identity-category-diagnostic frequency bands.

For analysis of training and in-scanner behavior, the response window for each mini-sequence was defined from the onset of the final tone to the onset of the final tone in the subsequent sequence (1068-ms response window). For calculation of d-prime scores, extreme hit and false alarm rates of 0 or 1 were adjusted following the approach of Stanislaw and Todorov [116]. As an introduction to the task, participants were familiarized with one up-sweep and one down-sweep on Day 4 of online training (or at the end of the single training session for the *Alien Size* group), with mini-sequence repeats indicated on-screen (Fig 1B). As with category training, scanner EPI noise (57 dB) was mixed in with the tone stimuli (~80 dB) used to map tonotopy.

Attention tonotopy. Participants who trained on the *Alien Identity+Size* categories completed a task to characterize voxelwise frequency-selective and amplitude-selective overt attentional responses in individual participants. Following the approach of Dick and colleagues [43], participants heard 4-tone mini-sequences evolving simultaneously in two frequency bands with instructions to "attend low" or "attend high" and report mini-sequence repeats from the attended frequency band (Fig 1C). To parallel the amplitude differences signaling alien size categories, mini-sequences alternated between higher and lower amplitude levels, which differed on average by ~10dB. A different spoken instruction directed participants to "attend volume," reporting 1-back matches on the attended amplitude dimension; data from the "attend volume" condition were not used for the present study. Mini-sequences were composed of 180-ms sinewave tones drawn from a pool of four frequencies (1, 3, 5 and 7 semitones above a 'base' frequency). Tones were drawn with replacement, with the constraint that a sequence could never be a single tone played four times. Simultaneous tones from high and low frequency bands were temporally aligned. Mini-sequences were separated by 360 ms silence in short blocks that concluded with 2 sec of silence. Each block included two mini-sequence repeats in the high-frequency band, two repeats in low-frequency band, and two amplitude repeats. The 1080-ms response window for each mini-sequence was defined from the onset of the final tone to the onset of the final tone in the subsequent sequence.

Participants were familiarized with the task on Day 3 of categorization training and first experienced mini-sequences in single frequency bands of fixed amplitude (Fig 1B). Next, amplitude variability was introduced (first only with the task of reporting tone sequence repeats), then finally, participants practiced with dual-frequency-band amplitude-variable stimuli. On-screen messages alerted participants to mini-sequence repeats in the attended band. There was no scanner noise mixed with stimuli during familiarization. Participants then completed 24 attention tonotopy trials (3 blocks/trial, 11–17 mini-sequences/block), with on-screen feedback provided for the first half of the trials. For half of the trials, the dual-band stimuli had base frequencies of 280 and 1,253 Hz, and for the other half, the base frequencies were 400 and 2,123 Hz. "Attend high," "attend low" and "attend volume" trials were equiprobable. The amplitude of the stimuli was set so that the two frequency bands were approximately equated on perceptual loudness (~73 dB average), and a fixed level of scanner noise (57 dB) was presented throughout these trials.

MRI procedure. Images were acquired in a single 90-min session with 3T Siemens Prisma scanners and 32-channel head coils (at the CMU-Pitt BRIDGE Center in Pittsburgh, RRID:SCR_023356 and at the Birkbeck-UCL Centre for Neuroimaging in London). Identical pulse sequences were used across sites, and very similar QA procedures were in place by the respective MRI center teams to monitor scanner performance. Head movement among London participants was minimized with a head stabilization prototype (MR Minimal Motion System). Both sites delivered audio over Sensimetrics earbuds with pre-filtering to accommodate the earbuds' response profile, presented tasks using PsychoPy (v. 2022.2.4), and recorded responses with a button box.

Participants in the *Alien Identity+Size* group who qualified for the MRI session completed a 15-min online refresher one day prior to the scan. This involved the tonotopy task (one 64-sec up-sweep, one 64-sec down-sweep), 80 2AFC categorization trials (48 alien identity, 32 alien size, drawn from the pool of 80 generalization exemplars with no feedback), and two short 9-block attention-tonotopy trials with parameters mirroring the MRI task. Participants in the *Alien Size* group completed a refresher within one day of their scan. It included the tonotopy task (one 64-sec up-sweep, one 64-sec down-sweep) and 32 trials of 2AFC alien size categorization, with no feedback.

At the scanning session, structural images were acquired using a T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence (TR = 2,300 ms, TE = 2.98 ms, FOV = 256 mm, flip angle = 9°) with 1 mm sagittal slices. Functional echo planar images were acquired using a CMRR [117] T2*-weighted sequence (TR = 1.0 s, TE = 30 ms, 44 slices, 2.0 mm thickness, in-plane resolution = 2 mm × 2 mm with 6/8 partial Fourier encoding, FOV = 212 mm, flip angle = 62°, multiband acceleration factor = 4); we included eight dummy volumes at the start of each scan to allow the scanner to reach B1 equilibrium. Functional images were acquired with anterior-to-posterior phase encoding; following each task, we collected two additional volumes with reversed phase encoding to correct for image distortion caused by B0 inhomogeneities.

The scanning session for *Alien Identity+Size* participants started with MPRAGE acquisition. Next, participants completed four runs of the tonotopy task (4 sweeps/run; 1st and 3rd up-sweep, 2nd and 4th down-sweep) during which they fixated on a central cross overlaid on a small landscape image to minimize eye movements associated with the acoustic sweep; images changed over the course each run (5 images/run; random duration to avoid aliasing with the acoustic sweep rate).

Next, for *Alien Identity+Size* participants, there were six runs (10 blocks each, 8 trials/block) of 2AFC categorization. Across runs, 20 blocks involved categorization of alien identity across the high diagnostic frequency band, 20 involved identity categorization across the low frequency band, and 20 involved size categorization across amplitude. Use of the 2AFC (rather than 4AFC) categorization task in the scanner afforded several benefits: (a) increasing power to differentiate across higher and lower frequency bands by minimizing switches within a block; (b) reducing saccades across distinct visual aliens; and (c) constraining each block to two categories making identity and size tasks completely parallel. Our prior behavioral study demonstrated no learning advantage for even explicit instruction indicating the diagnostic frequency band [40]. A pseudorandomized block order common across participants assured that blocks reliant on the same

diagnostic dimension did not occur consecutively. Overall block order in runs 1–3 repeated in runs 4–6, but with trial order shuffled, and with the amplitude set to a different level within the 'big' or 'small' amplitude range. In all, there were 320 alien identity categorization trials (80/category, with equal exemplars from the two size categories) and 160 alien size categorization trials (with equal exemplars from the four identity categories). Note that the 2AFC categorization fMRI stimuli were drawn from the pool of 80 generalization exemplars encountered in training. Prior to the fMRI experiment, participants had received no feedback with these stimuli, and never made a 2AFC decision with them. Each of the exemplars was encountered at two different amplitudes.

To facilitate comparisons between the alien identity and alien size categorization blocks, 20 of the 40 blocks that appeared during the 2AFC alien identity task also appeared in the 2AFC alien size task, with the same fixed trial order within the block. By design, the exemplars (and their order) in the identity and size task blocks were identical to avoid potential stimulus, context, and order confounds. Half of the alien size blocks appeared before their corresponding alien identity blocks, and half appeared after. Each block began with a 3-sec visual display of the task (Alien Identity or Alien Size) followed by 1.5 sec for the audio and a 1.5 sec response window.

Finally, there were two runs of the attention-tonotopy task (one with base frequencies of 280 and 1,253 Hz, one with base frequencies of 400 and 2,123 Hz). Each run involved 15 blocks (22–34 sequence/block), with 12–15 sec of rest after every 3 blocks (except after the final block in a run). Auditory instructions at the beginning of each block indicated the locus of attention ("high," "low," "volume," "rest"). As in the tonotopy task, participants fixated on a central cross overlaid on changing landscape images to avoid high- or low-band-attention-related eye movement (5 images/run).

The *Alien Size* participant group completed an abbreviated MRI protocol that included structural imaging, four runs of the tonotopy task, and two runs of 2AFC size categorization. Critically, the 2AFC categorization task trials followed the same presentation of alien size categorization trials (in the same order) as the *Alien Identity+Size* group.

Analyses

MRI analyses. Cortical surfaces were reconstructed for each participant from the T1-weighted MPRAGE using FreeSurfer (Dale and Sereno, [118]; Dale et al. 1999; Fischl, [119]).

Functional images from the tonotopy task were minimally preprocessed using AFNI (Cox [120]) to account for saturation by discarding the first 8 volumes, to unwarp images using phase-reversed images collected at the end of each task, and to align all volumes to a reference volume from the middle of the first run. Next, the reference volume was aligned to the cortical surface in Freesurfer using boundary-based registration (Greve and Fischl [121]).

To establish the regions of auditory cortex showing frequency selectivity within and across subjects, we calculated tonotopic maps using Fourier-based analyses on the preprocessed functional data from the tonotopy task in csurf (https://pages.ucsd.edu/~msereno/csurf), following standard phase-encoded mapping approaches [43] (Sereno et al. 1995; Dick et al. 2012). Relying on the fact that each frequency step occurs at a consistent time within a sweep, Fourier analysis allowed us to compute a set of F-statistics indicating how each voxel responds at the frequency of stimulus cycling (4 cycles per run, 4 64-sec sweeps per run) relative to other frequencies; the phase lag (i.e., the delay relative to the start of the cycle) of the maximal response can be used to determine the frequency preference of each voxel. We performed Fourier analysis of each run (up-sweeps time-reversed for phase averaging with down-sweeps).

Finally, we painted the phase of the signal (indicating frequency preference) as a color map and visualized the data on each participant's cortical surface reconstruction. Projecting each subject's data onto an icosahedral surface and averaging across participants created a group map that was projected to a single participant's surface reconstruction for visualization. Subsequent region-of-interest (ROI) analyses, described below, were limited to regions in auditory cortex where tonotopic organization was observed at the group level in participants expert in alien-identity categorization.

We next performed a regression-based analysis on the preprocessed functional data from each of the three tasks using 3dREMLfit in AFNI (Chen et al. [122]). For each analysis, we constructed regressors by convolving a vector of the onset

times for each frequency step with a square wave of the appropriate duration. Rigid-body motion parameters (movement in the x-, y- and z-axis directions, as well as pitch, roll, and yaw) were included as regressors of no interest. The resultant beta estimates were used to compute, for each task, the difference in activation observed according to higher versus lower frequency band.

In the categorization tasks, this corresponded to the difference in blocks involving categorization of alien identity across the higher frequency diagnostic band versus those involving the lower frequency diagnostic band. In the attention tonotopy task, this corresponded to the difference between explicitly attending to high- versus low-frequency mini-sequences. Results were mapped to the cortical surface using the mri_vol2surf Freesurfer command. To generate group-level contrast maps, subject-level maps were projected onto an icosahedral mesh and averaged, the result of which was displayed on a single subject's surface.

Although there is a macroscopic cross-participant similarity in large-scale tonotopic gradients, there is considerable cross-participant variability in frequency preference at a finer grain, particularly when voxels are aligned by cortical folding patterns [41,123]. Thus, to measure regional encoding and cross-task concordance in frequency selectivity while taking this variability into account, we calculated the degree of concordance in cross-task voxel-wise frequency selectivity in a set of small cortical surface ROIs defined in a previous study (Dick et al. [43], see Fig 4). Here, we limit our analysis to those ROIs where tonotopic selectivity was observed at the group level in our alien-identity categorization experts. For each participant, the ROIs that tessellate tonotopically organized auditory cortex (37 per hemisphere) were projected from a single participant to the unit icosahedron (sphere_reg), then projected again to each curvature-aligned participant's surface, and finally to the bbregister-aligned native space 2 x 2 x 2 mm EPI space using Freesurfer's mri_label2vol.

For each ROI, cross-task concordance was evaluated via regression analysis, where all subject-wise beta coefficients in one task (e.g., $b_{\text{tonotopy(high-low)}}$) were used to predict the high-low betas in another task (e.g., $b_{\text{attention-tonotopy(high-low)}}$); each regression model included random intercepts for each participant. In other words, there is a regression model fit for each ROI, where for each participant, the number of betas per task equals the number of voxels in the native-space projection of that ROI. The voxelwise beta coefficients for the tonotopy task are from the general linear test (GLT) of high minus low frequency bands; correspondingly, the coefficients for the alien identity task are from the GLT of high minus low category diagnostic bands. Betas are always paired voxelwise for all cross-task concordance maps.

The resultant *t* values from the mixed effect models were transformed to *z* statistics, and these *z* values were painted onto the cortical surface. Significance (alpha < 0.05) was assessed using one-tailed tests, as we had *a priori* hypothesized a positive relationship between betas across tasks. To control for multiple comparisons, we applied a false discovery rate (FDR) correction on our *p* values using the 3dFDR command in AFNI. Note that because 3dFDR performs a two-tailed adjustment, *p* values associated with negative *z* statistics were inverted (i.e., 1-p) both prior to and following FDR correction, yielding FDR-corrected values that reflect one-tailed tests. Similar analyses were performed for different subgroups of participants defined based on behavioral performance during training.

We also conducted a series of analyses to compare cortical responses between subgroups of *Alien Identity+-Size* participants (e.g., in participants who acquired category expertise early in training versus those who acquired expertise over the course of training; see Fig 6) as well as compared to *Alien Size* participants. These regression analyses were akin to the ones described above but also included a main effect of Group, as well as its interaction with Task; this interaction term allows us to assess whether cross-task concordance differed as a function of group. The associated *t* values from the interaction term were converted to *z* values and projected onto a single subject's cortical surface, providing an index of which group demonstrated stronger cross-task concordance. As with other analyses, significance was assessed using one-tailed tests and FDR corrections were applied to control for multiple comparisons.

Supporting information

S1 Fig. Cortical-surface-based group averages of phase-encoded tonotopic maps from the alien identity + size training group (top) and alien size training group (bottom). Color mapping shows phase of maximal fMRI response at 4 cycles/run; this indexes preferred auditory frequency via the high-to-low or low-to-high frequency sweep that repeats 4 times during each tonotopy run. (See Dick et al. 2012 and **Methods-Procedure** for details).
(TIFF)

S2 Fig. Differences in Concordance Maps for *Alien Identity + Size* Trained Group (N = 49) compared to *Alien Size* Trained Group (N = 20). S2 Fig shows the difference between the tonotopy and alien-size-task concordance map for all *Alien Identity+Size* trained experts (shown in bottom panel of Fig 5B) versus that for the *Alien Size* trained participants (shown in Fig 7, panel A). The summary data used for this figure are available at https://doi.org/10.5522/04/33974425.
(TIFF)

Acknowledgments

The authors would like to thank Jenny Bizley, Ayelet Landau, Rob Leech, Federico de Martino, Michelle Moerel, and Jeremy Skipper for comments on previous drafts of this manuscript. We would also like to thank Marty Sereno (https://pages.ucsd.edu/~msereno/csurf/) Bruce Fischl, Anders Dale, and the other developers of Freesurfer (http://surfer.nmr.mgh.harvard.edu), as well as the AFNI team (https://afni.nimh.nih.gov) for their invaluable neuroimaging tools used in this project.

Author contributions

Conceptualization: Sahil Luthra, Adam T. Tierney, Lori L. Holt, Frederic Dick.

Data curation: Sahil Luthra, Raha N. Razin, Lori L. Holt, Frederic Dick.

Formal analysis: Sahil Luthra, Lori L. Holt, Frederic Dick.

Funding acquisition: Adam T. Tierney, Lori L. Holt, Frederic Dick.

Investigation: Sahil Luthra, Raha N. Razin, Adam T. Tierney, Lori L. Holt, Frederic Dick.

Methodology: Sahil Luthra, Adam T. Tierney, Lori L. Holt, Frederic Dick.

Project administration: Raha N. Razin, Lori L. Holt, Frederic Dick.

Resources: Lori L. Holt, Frederic Dick.

Software: Sahil Luthra, Frederic Dick.

Supervision: Lori L. Holt, Frederic Dick.

Validation: Lori L. Holt, Frederic Dick.

Visualization: Sahil Luthra, Lori L. Holt, Frederic Dick.

Writing – original draft: Sahil Luthra, Raha N. Razin, Lori L. Holt, Frederic Dick.

Writing – review & editing: Sahil Luthra, Adam T. Tierney, Lori L. Holt, Frederic Dick.

References

1. Lavan N. The effect of familiarity on within-person age judgements from voices. Br J Psychol. 2022;113(1):287–99. https://doi.org/10.1111/bjop.12526 PMID: 34415575
2. Kraus MW, Torrez B, Park JW, Ghayebi F. Evidence for the reproduction of social class in brief speech. Proc Natl Acad Sci U S A. 2019;116(46):22998–3003. https://doi.org/10.1073/pnas.1900500116 PMID: 31636176
3. Holmes E, Domingo Y, Johnsrude IS. Familiar Voices Are More Intelligible, Even if They Are Not Recognized as Familiar. Psychol Sci. 2018;29(10):1575–83. https://doi.org/10.1177/0956797618779083 PMID: 30096018

4. Murphy TK, Nozari N, Holt LL. Transfer of statistical learning from passive speech perception to speech production. Psychon Bull Rev. 2024;31(3):1193–205. https://doi.org/10.3758/s13423-023-02399-8 PMID: 37884779
5. Todorov A, Pakrashi M, Oosterhof N. Evaluating faces on trustworthiness after minimal time exposure. Soc Cogn. 2009;27:813–33.
6. Krumhuber EG, Skora LI, Hill HC. The role of facial movements in emotion recognition. Nat Rev Psychol. 2023;2:283–96.
7. Young AW, Frühholz S, Schweinberger SR. Face and voice perception: understanding commonalities and differences. Trends Cogn Sci. 2020;24(5):398–410. https://doi.org/10.1016/j.tics.2020.02.001 PMID: 32298625
8. Li X, Logan RJ, Pastore RE. Perception of acoustic source characteristics: Walking sounds. The Journal of the Acoustical Society of America. 2005;90(6):3036–49. https://doi.org/10.1121/1.401778
9. Algermissen S, Hörnlein M. Person Identification by Footstep Sound Using Convolutional Neural Networks. Applied Mechanics. 2021;2(2):257–73. https://doi.org/10.3390/applmech2020016
10. Cunningham S, McGregor I. Audience perceptions of Foley footsteps and 3D realism designed to convey walker characteristics. Personal and Ubiquitous Computing. 2024;28(5):779–99. https://doi.org/10.1007/s00779-024-01819-3
11. Gao M, Turner BM, Sloutsky VM. The role of attention in category representation. Cogn Sci. 2024;48(4):e13438. https://doi.org/10.1111/cogs.13438 PMID: 38605457
12. Henderson MM, Serences JT, Rungratsameetaweemana N. Dynamic categorization rules alter representations in human visual cortex. Nat Commun. 2025;16(1):3459. https://doi.org/10.1038/s41467-025-58707-4 PMID: 40216798
13. Nosofsky RM. Attention, similarity, and the identification-categorization relationship. J Exp Psychol Gen. 1986;115(1):39–61. https://doi.org/10.1037//0096-3445.115.1.39 PMID: 2937873
14. De Baene W, Ons B, Wagemans J, Vogels R. Effects of category learning on the stimulus selectivity of macaque inferior temporal neurons. Learn Mem. 2008;15(9):717–27. https://doi.org/10.1101/lm.1040508 PMID: 18772261
15. O'Bryan SR, Jung S, Mohan AJ, Scolari M. Category learning selectively enhances representations of boundary-adjacent exemplars in early visual cortex. J Neurosci. 2024;44(3):e1039232023. https://doi.org/10.1523/JNEUROSCI.1039-23.2023 PMID: 37968121
16. Van Gulick AE, Gauthier I. The perceptual effects of learning object categories that predict perceptual goals. J Exp Psychol Learn Mem Cogn. 2014;40(5):1307–20. https://doi.org/10.1037/a0036822 PMID: 24820671
17. Chandrasekaran B, Sampath PD, Wong PCM. Individual variability in cue-weighting and lexical tone learning. J Acoust Soc Am. 2010;128(1):456–65. https://doi.org/10.1121/1.3445785 PMID: 20649239
18. Jasmin K, Sun H, Tierney AT. Effects of language experience on domain-general perceptual strategies. Cognition. 2020;206:104481. https://doi.org/10.1016/j.cognition.2020.104481
19. Petrova K, Jasmin K, Saito K, Tierney AT Extensive residence in a second language environment modifies perceptual strategies for suprasegmental categorization. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2023;49(12):1943–55. https://doi.org/10.1037/xlm0001246
20. Creel SC, Mizrahi R, Escobedo AG, Zhao L, Heyman GD. No Heightened Musical Pitch Weighting For Tone Language Speakers in Early Childhood. Music Perception. 2023;40(3):193–201. https://doi.org/10.1525/mp.2023.40.3.193
21. Idemaru K, Holt LL. The developmental trajectory of children's perception and production of English /r/-/l/. J Acoust Soc Am. 2013;133(6):4232–46.
22. Kanber E, Lally C, Razin R, Rosi V, Garrido L, Lavan N, et al. Representations of personally familiar voices are better resolved in the brain. Curr Biol. 2025;35(10):2424-2432.e6. https://doi.org/10.1016/j.cub.2025.03.081 PMID: 40252646
23. McMurray B, Danelz A, Rigler H, Seedorff M. Speech categorization develops slowly through adolescence. Dev Psychol. 2018;54(8):1472–91. https://doi.org/10.1037/dev0000542 PMID: 29952600
24. Yurovsky D, Frank MC. Beyond naïve cue combination: salience and social cues in early word learning. Dev Sci. 2017;20(2):10.1111/desc.12349. https://doi.org/10.1111/desc.12349 PMID: 26575408
25. Yurovsky D, Boyer TW, Smith LB, Yu C. Probabilistic cue combination: less is more. Dev Sci. 2013;16(2):149–58. https://doi.org/10.1111/desc.12011 PMID: 23432826
26. Reetzke R, Xie Z, Llanos F, Chandrasekaran B. Tracing the trajectory of sensory plasticity across different stages of speech learning in adulthood. Curr Biol. 2018;28(9):1419-1427.e4. https://doi.org/10.1016/j.cub.2018.03.026 PMID: 29681473
27. Roark CL, Paulon G, Rebaudo G, McHaney JR, Sarkar A, Chandrasekaran B. Individual differences in working memory impact the trajectory of non-native speech category learning. PLoS One. 2024;19(6):e0297917. https://doi.org/10.1371/journal.pone.0297917 PMID: 38857268
28. Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL. Developmental Changes in Human Cerebral Functional Organization for Word Generation. Cerebral Cortex. 2004;15(3):275–90. https://doi.org/10.1093/cercor/bhh129
29. Stiles, J., Moses, P., Passarotti, A., Dick, F. K., & Buxton, R. Exploring developmental change in the neural bases of higher cognitive functions: the promise of functional magnetic resonance imaging. Developmental Neuropsychology, 2003;24(2–3):641–668
30. Krishnan, S., Leech, R., Mercure, E., Lloyd-Fox, S., & Dick, F. Convergent and divergent fMRI responses in children and adults to increasing language production demands. Cerebral Cortex, 2014;25(10):3261–3277.

31. Yates TS, Ellis CT, Turk-Browne NB. Emergence and organization of adult brain function throughout child development. NeuroImage. 2020;226:117606. https://doi.org/10.1016/j.neuroimage.2020.117606
32. Karmiloff-Smith A. Neuroimaging of the developing brain: taking "developing" seriously. Hum Brain Mapp. 2010;31(6):934–41. https://doi.org/10.1002/hbm.21074 PMID: 20496384
33. Collina JS, Erdil G, Xia M, Angeloni CF, Wood KC, Sheth J, et al. Individual-specific strategies inform category learning. Sci Rep. 2025;15(1):2984. https://doi.org/10.1038/s41598-024-82219-8 PMID: 39848949
34. Drieu C, Zhu Z, Wang Z, Fuller K, Wang A, Elnozahy S, et al. Rapid emergence of latent knowledge in the sensory cortex drives learning. Nature. 2025;641(8064):960–70. https://doi.org/10.1038/s41586-025-08730-8 PMID: 40108473
35. Leech R, Holt LL, Devlin JT, Dick F. Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. J Neurosci. 2009;29(16):5234–9. https://doi.org/10.1523/JNEUROSCI.5758-08.2009 PMID: 19386919
36. Roark CL, Chandrasekaran B. Stable, flexible, common, and distinct behaviors support rule-based and information-integration category learning. NPJ Sci Learn. 2023;8(1):14.
37. Sheth J, Collina JS, Piasini E, Kording KP, Cohen YE, Geffen MN. The interplay of uncertainty, relevance and learning influences auditory categorization. Sci Rep. 2025;15(1):3348. https://doi.org/10.1038/s41598-025-86856-5 PMID: 39870756
38. Folstein JR, Palmeri TJ, Gauthier I. Perceptual advantage for category-relevant perceptual dimensions: the case of shape and motion. Frontiers in Psychology. 2014;5:1394. https://doi.org/10.3389/fpsyg.2014.01394
39. Ley A, Vroomen J, Hausfeld L, Valente G, De Weerd P, Formisano E. Learning of new sound categories shapes neural response patterns in human auditory cortex. J Neurosci. 2012;32(38):13273–80. https://doi.org/10.1523/JNEUROSCI.0584-12.2012 PMID: 22993443
40. Obasih CO, Luthra S, Dick F, Holt LL. Auditory category learning is robust across training regimes. Cognition. 2023;237:105467. https://doi.org/10.1016/j.cognition.2023.105467 PMID: 37148640
41. Moerel M, De Martino F, Formisano E. An anatomical and functional topography of human auditory cortical areas. Front Neurosci. 2014;8:225. https://doi.org/10.3389/fnins.2014.00225 PMID: 25120426
42. Da Costa S, van der Zwaag W, Miller LM, Clarke S, Saenz M. Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. J Neurosci. 2013;33(5):1858–63. https://doi.org/10.1523/JNEUROSCI.4405-12.2013 PMID: 23365225
43. Dick FK, Lehet MI, Callaghan MF, Keller TA, Sereno MI, Holt LL. Extensive tonotopic mapping across auditory cortex is recapitulated by spectrally directed attention and systematically related to cortical myeloarchitecture. J Neurosci. 2017;37(50):12187–201. https://doi.org/10.1523/JNEUROSCI.1436-17.2017 PMID: 29109238
44. De Martino F, Moerel M, Ugurbil K, Goebel R, Yacoub E, Formisano E. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proc Natl Acad Sci U S A. 2015;112(52):16036–41. https://doi.org/10.1073/pnas.1507552112 PMID: 26668397
45. Riecke L, Peters JC, Valente G, Kemper VG, Formisano E, Sorger B. Frequency-selective attention in auditory scenes recruits frequency representations throughout human superior temporal cortex. Cereb Cortex. 2017;27(5):3002–14. https://doi.org/10.1093/cercor/bhw160 PMID: 27230215
46. Fernandez K, Lyu Y. Attentional mechanisms for learning feature combinations. J Neurosci. 2025;45(44):e0784252025. https://doi.org/10.1523/JNEUROSCI.0784-25.2025 PMID: 41162158
47. Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. Neuron. 2017;93(2):451–63. https://doi.org/10.1016/j.neuron.2016.12.040 PMID: 28103483
48. Wang MC, Soltani A. Contributions of attention to learning in multidimensional reward environments. J Neurosci. 2025;45(7):e2300232024. https://doi.org/10.1523/JNEUROSCI.2300-23.2024 PMID: 39681464
49. Giordano BL, Esposito M, Valente G, Formisano E. Intermediate acoustic-to-semantic representations link behavioral and neural responses to natural sounds. Nat Neurosci. 2023;26(4):664–72. https://doi.org/10.1038/s41593-023-01285-9 PMID: 36928634
50. Boersma P a u l. Praat, a system for doing phonetics by computer. Glot International. 2001;5(9/10):341–5.
51. Barron HC, Garvert MM, Behrens TEJ. Repetition suppression: a means to index neural representations using BOLD?. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences. 2016;371:20150355–14. https://doi.org/10.1098/rstb.2015.0355
52. Brown RM, Chen JL, Hollinger A, Penhune VB, Palmer C, Zatorre RJ. Repetition suppression in auditory-motor regions to pitch and temporal structure in music. J Cogn Neurosci. 2013;25(2):313–28. https://doi.org/10.1162/jocn_a_00322 PMID: 23163413
53. Segaert K, Weber K, de Lange FP, Petersson KM, Hagoort P. The suppression of repetition enhancement: a review of fMRI studies. Neuropsychologia. 2013;51(1):59–66. https://doi.org/10.1016/j.neuropsychologia.2012.11.006 PMID: 23159344
54. Zhao J, Al-Aidroos N, Turk-Browne NB. Attention Is Spontaneously Biased Toward Regularities. Psychological Science. 2013;24(5):667–77. https://doi.org/10.1177/0956797612460407
55. Wang M, Jendrichovsky P, Kanold PO. Auditory discrimination learning differentially modulates neural representation in auditory cortex subregions and inter-areal connectivity. Cell Rep. 2024;43(5):114172. https://doi.org/10.1016/j.celrep.2024.114172 PMID: 38703366
56. Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature. 2002;415(6869):318–20. https://doi.org/10.1038/415318a PMID: 11797008
57. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci. 2003;23(12):5235–46. https://doi.org/10.1523/JNEUROSCI.23-12-05235.2003 PMID: 12832548

58. Gillebert CR, Op de Beeck HP, Panis S, Wagemans J. Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. J Cogn Neurosci. 2009;21(6):1054–64. https://doi.org/10.1162/jocn.2009.21089 PMID: 18752400

59. Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M. Categorization training results in shape- and category-selective human neural plasticity. Neuron. 2007;53(6):891–903. https://doi.org/10.1016/j.neuron.2007.02.015 PMID: 17359923

60. van der Linden M, van Turennout M, Indefrey P. Formation of category representations in superior temporal sulcus. J Cogn Neurosci. 2010;22(6):1270–82. https://doi.org/10.1162/jocn.2009.21270 PMID: 19445608

61. Roy JE, Riesenhuber M, Poggio T, Miller EK. Prefrontal cortex activity during flexible categorization. J Neurosci. 2010;30(25):8519–28. https://doi.org/10.1523/JNEUROSCI.4837-09.2010 PMID: 20573899

62. Sasaki Y, Nanez JE, Watanabe T. Advances in visual perceptual learning and plasticity. Nat Rev Neurosci. 2010;11(1):53–60. https://doi.org/10.1038/nrn2737 PMID: 19953104

63. Hausfeld L, Riecke L, Formisano E. Acoustic and higher-level representations of naturalistic auditory scenes in human auditory and frontal cortex. Neuroimage. 2018;173:472–83. https://doi.org/10.1016/j.neuroimage.2018.02.065 PMID: 29518569

64. Jiang X, Chevillet MA, Rauschecker JP, Riesenhuber M. Training humans to categorize monkey calls: auditory feature- and category-selective neural tuning changes. Neuron. 2018;98(2):405-416.e4. https://doi.org/10.1016/j.neuron.2018.03.014 PMID: 29673483

65. Russ BE, Lee Y-S, Cohen YE. Neural and behavioral correlates of auditory categorization. Hear Res. 2007;229(1–2):204–12. https://doi.org/10.1016/j.heares.2006.10.010 PMID: 17208397

66. Folstein JR, Palmeri TJ, Gauthier I. Category learning increases discriminability of relevant object dimensions in visual cortex. Cereb Cortex. 2013;23(4):814–23. https://doi.org/10.1093/cercor/bhs067 PMID: 22490547

67. Saffran JR, Kirkham NZ. Infant statistical learning. Annu Rev Psychol. 2018;69:181–203. https://doi.org/10.1146/annurev-psych-122216-011805 PMID: 28793812

68. Wilson B, Kikuchi Y, Sun L, Hunter D, Dick F, Smith K, et al. Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. Nature Communications. 2015;6(1):8901. https://doi.org/10.1038/ncomms9901

69. Emberson LL, Liu R, Zevin JD. Is statistical learning constrained by lower level perceptual organization?. Cognition. 2013;128(1):82–102. https://doi.org/10.1016/j.cognition.2012.12.006 PMID: 23618755

70. Wade T, Holt LL. Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task. J Acoust Soc Am. 2005;118(4):2618–33. https://doi.org/10.1121/1.2011156 PMID: 16266182

71. Lim S-J, Fiez JA, Holt LL. Role of the striatum in incidental learning of sound categories. Proc Natl Acad Sci U S A. 2019;116(10):4671–80. https://doi.org/10.1073/pnas.1811992116 PMID: 30782817

72. Roark CL, Lehet MI, Dick F, Holt LL. The representational glue for incidental category learning is alignment with task-relevant behavior. J Exp Psychol Learn Mem Cogn. 2022;48(6):769–84. https://doi.org/10.1037/xlm0001078 PMID: 34570548

73. Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, et al. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. Neuron. 2014;82(2):486–99. https://doi.org/10.1016/j.neuron.2014.02.029 PMID: 24742467

74. Bagur S, Averseng M, Elgueda D, David S, Fritz J, Yin P, et al. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. Nat Commun. 2018;9(1):2529. https://doi.org/10.1038/s41467-018-04839-9 PMID: 29955046

75. Yin P, Strait DL, Radtke-Schuller S, Fritz JB, Shamma SA. Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. Curr Biol. 2020;30(9):1649-1663.e5. https://doi.org/10.1016/j.cub.2020.02.047 PMID: 32220317

76. Zhong L, Zhang Y, Duan CA, Deng J, Pan J, Xu N-L. Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. Nat Neurosci. 2019;22(6):963–73. https://doi.org/10.1038/s41593-019-0383-6 PMID: 31036942

77. DeGutis J, D'Esposito M. Distinct mechanisms in visual category learning. Cogn Affect Behav Neurosci. 2007;7(3):251–9. https://doi.org/10.3758/cabn.7.3.251 PMID: 17993211

78. Addleman DA, Jiang YV. Experience-driven auditory attention. Trends Cogn Sci. 2019;23(11):927–37. https://doi.org/10.1016/j.tics.2019.08.002 PMID: 31521482

79. Hullett PW, Leonard MK, Gorno-Tempini ML, Mandelli ML, Chang EF. Parallel encoding of speech in human frontal and temporal lobes. Nat Commun. 2025;17(1):814. https://doi.org/10.1038/s41467-025-67517-7 PMID: 41461641

80. Kruschke JK, Johansen MK. A model of probabilistic category learning. J Exp Psychol Learn Mem Cogn. 1999;25(5):1083–119. https://doi.org/10.1037//0278-7393.25.5.1083 PMID: 10505339

81. Love BC, Medin DL, Gureckis TM. SUSTAIN: a network model of category learning. Psychol Rev. 2004;111(2):309–32. https://doi.org/10.1037/0033-295X.111.2.309 PMID: 15065912

82. McColeman CM, Barnes JI, Chen L, Meier KM, Walshe RC, Blair MR. Learning-induced changes in attentional allocation during categorization: a sizable catalog of attention change as measured by eye movements. PLoS One. 2014;9(1):e83302. https://doi.org/10.1371/journal.pone.0083302 PMID: 24497915

83. Heald SLM, Nusbaum HC. Speech perception as an active cognitive process. Front Syst Neurosci. 2014;8:35. https://doi.org/10.3389/fnsys.2014.00035 PMID: 24672438

84. Foster JJ, Ling S. Feature-based attention multiplicatively scales the fMRI-BOLD contrast-response function. J Neurosci. 2022;42(36):6894–906. https://doi.org/10.1523/JNEUROSCI.0513-22.2022 PMID: 35868860

85. Gundlach C, Wehle S, Müller MM. Early sensory gain control is dominated by obligatory and global feature-based attention in top-down shifts of combined spatial and feature-based attention. Cereb Cortex. 2023;33(19):10286–302. https://doi.org/10.1093/cercor/bhad282 PMID: 37536059

86. Saenz M, Buracas GT, Boynton GM. Global effects of feature-based attention in human visual cortex. Nat Neurosci. 2002;5(7):631–2. https://doi.org/10.1038/nn876 PMID: 12068304

87. Serences JT, Boynton GM. The representation of behavioral choice for motion in human visual cortex. J Neurosci. 2007;27(47):12893–9. https://doi.org/10.1523/JNEUROSCI.4021-07.2007 PMID: 18032662

88. Treue S, Maunsell JH. Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. J Neurosci. 1999;19(17):7591–602. https://doi.org/10.1523/JNEUROSCI.19-17-07591.1999 PMID: 10460265

89. Yoo S-A, Martinez-Trujillo JC, Treue S, Tsotsos JK, Fallah M. Attention to visual motion suppresses neuronal and behavioral sensitivity in nearby feature space. BMC Biol. 2022;20(1):220. https://doi.org/10.1186/s12915-022-01428-7 PMID: 36199136

90. Petkov CI, Kang X, Alho K, Bertrand O, Yund EW, Woods DL. Attentional modulation of human auditory cortex. Nat Neurosci. 2004;7(6):658–63. https://doi.org/10.1038/nn1256 PMID: 15156150

91. Gertsovski A, Ahissar M. Reduced learning of sound categories in dyslexia is associated with reduced regularity-induced auditory cortex adaptation. J Neurosci. 2022;42(7):1328–42. https://doi.org/10.1523/JNEUROSCI.1533-21.2021 PMID: 34969869

92. Gabay Y, Holt LL. Incidental learning of sound categories is impaired in developmental dyslexia. Cortex. 2015;73:131–43. https://doi.org/10.1016/j.cortex.2015.08.008 PMID: 26409017

93. Gabay Y, Roark CL, Holt LL. Impaired and spared auditory category learning in developmental dyslexia. Psychol Sci. 2023;34(4):468–80. https://doi.org/10.1177/09567976231151581 PMID: 36791783

94. Kim H, McMurray B, Sorensen E, Oleson J. The consistency of categorization-consistency in speech perception. Psychon Bull Rev. 2025;32(5):2246–58. https://doi.org/10.3758/s13423-025-02700-x PMID: 40274721

95. Dick F, Lee HL, Nusbaum H, Price CJ. Auditory-motor expertise alters "speech selectivity" in professional musicians and actors. Cereb Cortex. 2011;21(4):938–48. https://doi.org/10.1093/cercor/bhq166 PMID: 20829245

96. Haist F, Lee K, Stiles J. Individuating faces and common objects produces equal responses in putative face-processing areas in the ventral occipitotemporal cortex. Front Hum Neurosci. 2010;4:181. https://doi.org/10.3389/fnhum.2010.00181 PMID: 21206532

97. McGugin RW, Gatenby JC, Gore JC, Gauthier I. High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. Proc Natl Acad Sci U S A. 2012;109(42):17063–8. https://doi.org/10.1073/pnas.1116333109 PMID: 23027970

98. Bilalić M. Revisiting the role of the fusiform face area in expertise. J Cogn Neurosci. 2016;28(9):1345–57. https://doi.org/10.1162/jocn_a_00974 PMID: 27082047

99. Gomez J, Barnett M, Grill-Spector K. Extensive childhood experience with Pokémon suggests eccentricity drives organization of visual cortex. Nat Hum Behav. 2019;3(6):611–24. https://doi.org/10.1038/s41562-019-0592-8 PMID: 31061489

100. Op de Beeck HP, Pillet I, Ritchie JB. Factors determining where category-selective areas emerge in visual cortex. Trends Cogn Sci. 2019;23(9):784–97. https://doi.org/10.1016/j.tics.2019.06.006 PMID: 31327671

101. McKone E, Kanwisher N, Duchaine BC. Can generic expertise explain special processing for faces?. Trends Cogn Sci. 2007;11(1):8–15.

102. Folstein J, Palmeri TJ, Van Gulick AE, Gauthier I. Category learning stretches neural representations in visual cortex. Curr Dir Psychol Sci. 2015;24(1):17–23. https://doi.org/10.1177/0963721414550707 PMID: 25745280

103. Roye A, Schröger E, Jacobsen T, Gruber T. Is my mobile ringing? Evidence for rapid processing of a personally significant sound in humans. J Neurosci. 2010;30(21):7310–3. https://doi.org/10.1523/JNEUROSCI.1113-10.2010 PMID: 20505097

104. Norman-Haignere S, Kanwisher NG, McDermott JH. Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. Neuron. 2015;88(6):1281–96. https://doi.org/10.1016/j.neuron.2015.11.035 PMID: 26687225

105. Liu ST, Montes-Lourido P, Wang X, Sadagopan S. Optimal features for auditory categorization. Nat Commun. 2019;10(1):1302. https://doi.org/10.1038/s41467-019-09115-y PMID: 30899018

106. Montes-Lourido P, Kar M, David SV, Sadagopan S. Neuronal selectivity to complex vocalization features emerges in the superficial layers of primary auditory cortex. PLoS Biol. 2021;19(6):e3001299. https://doi.org/10.1371/journal.pbio.3001299 PMID: 34133413

107. Heynckes M, Lage-Castellanos A, De Weerd P, Formisano E, De Martino F. Layer-specific correlates of detected and undetected auditory targets during attention. Curr Res Neurobiol. 2023;4:100075. https://doi.org/10.1016/j.crneur.2023.100075 PMID: 36755988

108. Moerel M, Yacoub E, Gulban OF, Lage-Castellanos A, De Martino F. Using high spatial resolution fMRI to understand representation in the auditory network. Prog Neurobiol. 2021;207:101887. https://doi.org/10.1016/j.pneurobio.2020.101887 PMID: 32745500

109. Deveau J, Lovcik G, Seitz AR. Broad-based visual benefits from training with an integrated perceptual-learning video game. Vision Res. 2014;99:134–40. https://doi.org/10.1016/j.visres.2013.12.015 PMID: 24406157

110. Whitton JP, Hancock KE, Shannon JM, Polley DB. Audiomotor perceptual training enhances speech intelligibility in background noise. Curr Biol. 2017;27(21):3237-3247.e6. https://doi.org/10.1016/j.cub.2017.09.014 PMID: 29056453

111. Itthipuripat S, Cha K, Byers A, Serences JT. Two different mechanisms support selective attention at different phases of training. PLoS Biol. 2017;15(6):e2001724. https://doi.org/10.1371/journal.pbio.2001724 PMID: 28654635

112. Anwyl-Irvine A, Dalmaijer ES, Hodges N, Evershed JK. Realistic precision and accuracy of online experiment platforms, web browsers, and devices. Behavior Research Methods. 2020;53(4):1407–25. https://doi.org/10.3758/s13428-020-01501-5

113. Milne AE, Bianco R, Poole KC, Zhao S, Oxenham AJ, Billig AJ, et al. An online headphone screening test based on dichotic pitch. Behavior Research Methods. 2020;53(4):1551–62. https://doi.org/10.3758/s13428-020-01514-0

114. Dick F, Taylor Tierney A, Lutti A, Josephs O, Sereno MI, Weiskopf N. In VivoFunctional and Myeloarchitectonic Mapping of Human Primary Auditory Areas. The Journal of Neuroscience. 2012;32(46):16095–105. https://doi.org/10.1523/JNEUROSCI.1712-12.2012

115. Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science (New York, NY), 1995;268(5212):889–893. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=7754376&retmode=ref&cmd=prlinks

116. Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc, 1999;31(1): 137–149.

117. Xu J, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, et al. Evaluation of slice accelerations using multiband echo planar imaging at 3 T. Neuroimage. 2013;83:991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055 PMID: 23899722

118. Dale AM, Sereno MI. Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. Journal of Cognitive Neuroscience. 2008;5(2):162–76. https://doi.org/10.1162/jocn.1993.5.2.162

119. Fischl B. FreeSurfer. NeuroImage. 2012;62(2):774–81. https://doi.org/10.1016/j.neuroimage.2012.01.021

120. Cox RW. AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. Computers and Biomedical Research. 2002;29(3):162–73. https://doi.org/10.1006/cbmr.1996.0014

121. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. NeuroImage. 2009;48(1):63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

122. Chen G, Saad ZS, Nath AR, Beauchamp MS, Cox RW. FMRI group analysis combining effect estimates and their variances. NeuroImage. 2011;60(1):747–65. https://doi.org/10.1016/j.neuroimage.2011.12.060

123. Besle J, Mougin O, Sánchez-Panchuelo R-M, Lanting C, Gowland P, Bowtell R, et al. Is human auditory cortex organization compatible with the monkey model? contrary evidence from ultra-high-field functional and structural MRI. Cereb Cortex. 2019;29(1):410–28. https://doi.org/10.1093/cercor/bhy267 PMID: 30357410