nature neuroscience

Article

https://doi.org/10.1038/s41593-025-01932-3

Spatially clustered neurons in the bat midbrain encode vocalization categories

Received: 16 June 2023

Accepted: 26 February 2025

Published online: 14 April 2025

Check for updates

Jennifer Lawlor[1,2], Melville J. Wohlgemuth[3], Cynthia F. Moss[1,2,4,5] & Kishore V. Kuchibhotla[1,2,4,6]

Rapid categorization of vocalizations enables adaptive behavior across species. While categorical perception is thought to arise in the neocortex, humans and animals could benefit from a functional organization tailored to ethologically relevant sound processing earlier in the auditory pathway. Here we developed two-photon calcium imaging in the awake echolocating bat (*Eptesicus fuscus)* to study the representation of vocalizations in the inferior colliculus, which is as few as two synapses from the inner ear. Echolocating bats rely on frequency-sweep-based vocalizations for social communication and navigation. Auditory playback experiments demonstrated that individual neurons responded selectively to social or navigation calls, enabling robust population-level decoding across categories. When social calls were morphed into navigation calls in equidistant step-wise increments, individual neurons showed switch-like properties and population-level response patterns sharply transitioned at the category boundary. Strikingly, category-selective neurons formed spatial clusters, independent of tonotopy within the dorsal cortex of the inferior colliculus. These findings support a revised view of categorical processing in which specified channels for ethologically relevant sounds are spatially segregated early in the auditory hierarchy, enabling rapid subcortical organization into categorical primitives.

Vocal communication exemplifies the ecological importance of categorical perception. Humans can readily discriminate phonemes that otherwise share many acoustic features. Moreover, categorical discrimination of vocal communication calls appears to be phylogenetically conserved across species, including those that communicate on land (nonhuman primates[1], naked mole rats[2], mice[3], crickets[4] and frogs[5]), sea (dolphins[6] and pinnipeds[7]) and air (songbirds[8] and echolocating bats[9]). Current models of categorical perception propose that sensory cortical networks are the engine underlying category formation[10], binding relevant features across a basis set of lower-level features that are initially decomposed from natural scenes by sensory organs and processed hierarchically in a feedforward manner[11–13]. While the 'perception' of categories may first arise in the cortex, categorical representations at earlier stages in sensory processing would provide notable computational advantages, conferring increased speed and fidelity of categorization.

To what extent do categorical representations emerge early in the sensory hierarchy? Categorical representations may exist in the form of single-neuron selectivity or population selectivity or even in spatially segregated ensembles[14]. We reasoned that two-photon imaging of large neural populations with single-neuron resolution in a mammalian species that exhibits a rich repertoire of vocalization categories would provide high spatial and sufficient temporal resolution to assess functional activity and test these possibilities.

[1]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA. [2]Kavli Neuroscience Discovery Institute, Johns Hopkins University, Baltimore, MD, USA. [3]Department of Neuroscience, University of Arizona, Tucson, AZ, USA. [4]The Solomon Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [5]Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA. [6]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. e-mail: kkuchib1@jhu.edu

The big brown bat, *Eptesicus fuscus*, is an insectivorous bat species that stands out in the animal kingdom for its acoustic repertoire for echolocation-based navigation[15] and for social interactions with conspecifics[16,17]. *E. fuscus* flies in three dimensions through its environment and must rapidly distinguish between vocalizations produced for navigation or social interaction during competitive foraging.

Results

Two-photon calcium imaging in the awake, echolocating bat

One candidate subcortical region for early categorical processing is the inferior colliculus (IC) because of its key position as an obligatory auditory processing station in the central nervous system. The IC has been suggested to occupy a processing level comparable to the primary occipital cortex (V1) in the visual pathway[18], due to the auditory system's multiple subcortical stations that project to the IC. As such, a number of complex feature encoding properties, such as duration tuning[19], and context-dependent responses, such as deviance detection[20–22], are reported to emerge at the level of the IC. The nucleus is organized into multiple subdivisions that receive inputs from the auditory brainstem[23]. One subdivision, the dorsal cortex of the IC (DCIC), is of particular interest for categorical representations as it receives both feedforward input and considerable top-down input from the auditory cortex[24,25]. Here, we developed an approach for two-photon imaging in the IC of the awake, big brown bat, *E. fuscus* (Fig. 1) to study spatially resolved single-neuron and population coding of auditory stimuli. We first sought to characterize the sound-evoked properties of DCIC neurons and validate the use of two-photon calcium imaging in this nontraditional animal model. We injected AAV5-CAMKII-GCaMP6f bilaterally into the DCIC of *E. fuscus* and used a thinned-skull procedure to gain chronic optical access (Fig. 1a,b). We used two-photon imaging to monitor large populations of excitatory neurons (Extended Data Fig. 1a,b) with single-cell resolution (Fig. 1c). We imaged from multiple sites in the left and right IC (LIC and RIC; Fig. 1b), and at multiple depths, from the brain's surface and down to 200 µm. Using this technique, we recorded neural activity from 9,443 spatially resolved neurons across 27 sites and 5 bats. Awake bats were presented with an array of auditory stimuli including pure tones (4–90 kHz), complex stimuli (linear frequency modulated (FM) sweeps and white noise) and a series of bat and mouse vocalizations. Individual neurons exhibited robust sound-evoked responses (Fig. 1c,d) with most of the recorded population responding to at least one of the stimuli presented (Fig. 1e,f, $n = 3$ bats, 79.6 ± 3.7%, $P < 0.05$, paired t-test).

Functional microarchitecture of the DCIC

We next sought to understand the functional microarchitecture of the DCIC in echolocating bats. A fundamental feature of the auditory system is the precise topographic organization of responses to sound frequency, commonly referred to as tonotopy. Tonotopy emerges from the biomechanical properties of the basilar membrane in the inner ear and propagates throughout the central auditory system[26]. The central nucleus of the IC (CNIC) in bats and other mammalian species exhibits a dorsoventral tonotopic axis[27], but less is known about the topographical organization of sound frequency in the DCIC, in part because prior studies in the bat have not consistently identified the precise locations of recording sites[28,29]. We presented pure tones ranging from 4 to 90 kHz, matching *E. fuscus*' hearing range[30], at 70 dB with half-octave spacing and then calculated single-neuron tuning curves to identify a neuron's best frequency (BF), here defined as the sound frequency that elicits the highest-amplitude stimulus-evoked response. Most neurons were significantly tuned to a limited band of sound frequencies (66.4 ± 11%, two-way analysis of variance (ANOVA) for tone and duration, $P < 0.05$; Fig. 2a–c) in the low-to-mid frequency range (Fig. 2c, 4–32 kHz: 93.3 ± 1.7% of tuned neurons). These findings are consistent with those reported in the mouse DCIC using two-photon calcium imaging[24,31]. In addition, these data are consistent with anatomical evidence showing that anteroventral cochlear nucleus neurons that encode lower-frequency sounds project more superficially than dorsal axonal projections[32], suggesting a conserved functional microarchitecture across species.

We then used the spatial resolution of two-photon imaging to reconstruct the precise location of each neuron in three dimensions with its corresponding BF (Fig. 2d). We aligned across imaging sites and bats in the rostrocaudal and the mediolateral extents, based on a common anatomical landmark, the intersection of the IC, superior colliculus and midline. Z depth was measured using the optical sectioning capabilities of two-photon imaging coupled with an estimate of the skull-to-brain transition, based on a change in background fluorescence (Extended Data Fig. 2a–c). This allowed us to test whether a spatial, tonotopic gradient exists in the mediolateral, rostrocaudal or dorsoventral axes. We determined that a positive linear relationship exists for both the mediolateral (ML) and the rostrocaudal (RC) axes but not for the dorsoventral (DV) axis (Fig. 2e, $P_{ML} = 5.02 \times 10^{-95}$ ($t = 23.7$, $R^2 = 0.4$), $P_{RC} = 4.79 \times 10^{-15}$ ($t = 8$, $R^2 = 0.07$) and $P_{DV} = 0.13$ ($t = -1.5$, $R^2 = 0.003$); additional example sites in Extended Data Fig. 3a–c). To capture the direction of this tonotopic axis, we estimated the best-fit angle per reconstructed DCIC left/right side across bats (Supplementary Table 1), yielding an average angle of 47.5°, a value remarkably close to the 50° reported in the mouse DCIC[31]. The tonotopic gradient was computed as the linear fit of the mean frequency along this rostrolateral axis (Fig. 2f,g) per side and per bat. On average, the BF increased by 2.1 octaves mm^{-1} along this axis (Fig. 2g,h). To confirm that there was no observable tonotopic organization in the dorsoventral axis, we compared tuning at three different depths and observed similar frequency tuning distributions (Fig. 2i, Friedman test, $\chi^2(2) = 1.8$, $P_{DV} = 0.4066$, $n = 10$).

Functional selectivity to ethologically relevant vocalizations in the DCIC

We next sought to characterize the response properties of DCIC neurons in response to playbacks of natural vocalizations from conspecifics. Bats communicate to other bats with a frequency-modulated bout[9,33] to claim food when competing with conspecifics (social vocalizations), while also using echolocation to track prey and steer around obstacles (navigation vocalizations)[34,35]. Social and navigation vocalizations exhibit largely overlapping spectrotemporal features with subtle and variable differences, including shallower FM sweep rate and greater low-frequency energy[35] (see Fig. 3a and Extended Data Fig. 4a–d for more information on the acoustic distribution). Most recorded neurons responded to conspecific vocalizations (86 ± 18.4%, one-way ANOVA for responsivity, $P < 0.05$; Extended Data Fig. 5a) whose spectral content spans from 22 to 90 kHz, even though BFs of the DCIC neurons were generally below 32 kHz (Fig. 2c). The population sound-evoked response was highest for conspecific vocalizations and lowest for pure tones (Fig. 3b). These data suggest that ethologically relevant vocalizations drive DCIC neurons more than pure tones or even heterospecific complex stimuli, such as mouse ultrasonic vocalizations (mouse USV; Fig. 3a). This result is consistent with electrophysiological recordings in a number of species showing that conspecific vocalizations elicit larger evoked responses than more artificial stimuli in the IC[36]. We observed a striking selectivity within individual neurons for categories of bat vocalizations (social: 30.2 ± 4.5%; navigation: 19.7 ± 3.7%, Kruskal–Wallis test, $\chi^2(6) = 29.4$, $P = 5.0278 \times 10^{-5}$, $n = 42$, post-hoc Tukey's honest significant difference (HSD), $P_{social/tone} = 0.0001$, $P_{navig.l/tone} = 0.0049$, $P_{social/us} = 0.0146$) but little to no selectivity for mouse USVs (Fig. 3d, $P_{usv/tone} = 0.11$; detailed selectivity in Extended Data Fig. 5b), highlighting the importance of ethological relevance in auditory response selectivity. Given the lack of selectivity to mouse USVs, we focused on the bat vocalizations for further analysis. We calculated the neural selectivity to social versus navigation calls by computing the difference in evoked responses between both categories

https://doi.org/10.1038/s41593-025-01932-3

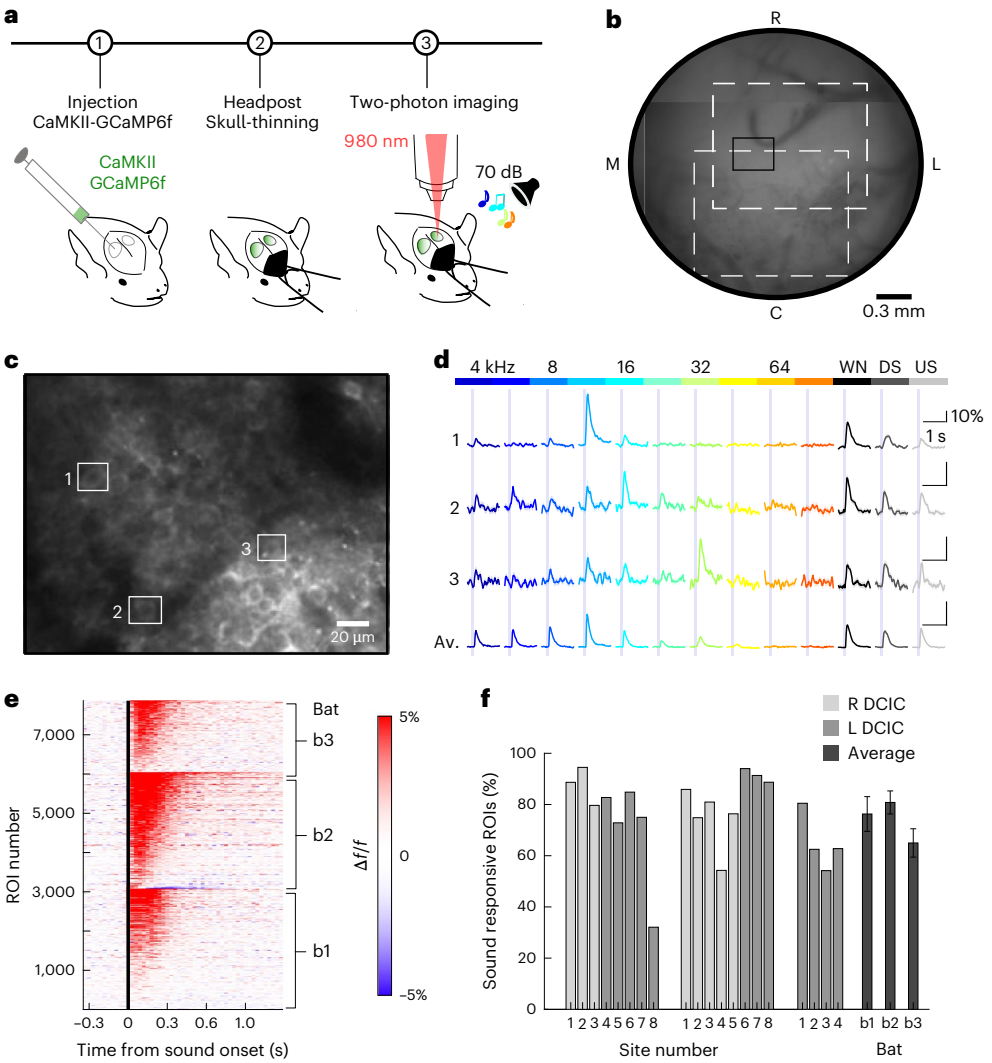


Fig. 1 | Robust sound-evoked responses in the DCIC of awake bats observed using two-photon calcium imaging. a, In vivo two-photon calcium imaging methods and timeline graphical representation. **b**, Composite vasculature map of an example LIC with overlaid imaging sites at 2× magnification (white box; R, rostral; C, caudal; M, medial; L, lateral). The small black box highlights the position of the cells displayed in **c**. **c**, Example site mean image cropped. The white boxes surround examples cells presented in **d**. **d**, Average (Av.) activity ($\Delta f/f$) of three example ROIs to pure tones and complex sounds presentations (color indicates stimulus identity displayed in **c**; shading: mean ± s.e.m; WN, white noise; DS, frequency-modulated downsweep; US, frequency-modulated upsweep). The fourth calcium trace is the average trace for the imaged population of the example site (shading: mean ± s.e.m.). **e**, Sound-evoked heatmap for all identified ROIs across all sites and bats (n_{cells} = 7,868, averaged over all stimulus types presented in this manuscript, bat number: b1, b2 and b3). **f**, Percentage of significantly sound-evoked ROIs per site (light gray, right DCIC; medium gray, left DCIC) and averaged across site (darker gray) for each bat.

(Fig. 3c–e and Extended Data Fig. 5c,d). While there were neurons selective to both bat vocalization categories (social or navigation), the DCIC network exhibited a strong preference for social calls (Fig. 3e and Extended Data Fig. 5d, overall, n = 1,383 social-selective neurons, n = 502 navigation-selective neurons; per bat, 24 ± 5.4% social-selective and 8.4 ± 1.9% navigation-selective, Kruskal–Wallis test, $\chi^2(1) = 16.3$, $P_{all} = 5.3 \times 10^{-5}$, $n = 28$)[35]. These data demonstrate that individual DCIC neurons are selective to conspecific vocalizations with a strong preference for encoding social vocalizations.

Bat vocalizations tend to be grouped in sequences; for example, the social vocalizations presented here contain three to four successive calls. Therefore, categorical information could be carried not only by the spectral content but also by the temporal pattern of call sequences. To determine if the selectivity observed is due to the processing of the spectral information of the first call, the processing of the entire sequence or a mixture of both, we analyzed the time course of the neural selectivity. The activity of selective neurons diverged rapidly after the sound onset (Fig. 3f, top) and before the end of the call sequence. A linear decoder reliably discriminated between conspecific categories on the second imaging frame after sound onset (~32 ms, 31.25 Hz acquisition rate; Fig. 3f, bottom, and Extended Data Fig. 6a; decoding weights and selectivity indices (SI) increase linearly 60 ms after sound presentation), indicating that the selectivity and population-level representations arise within one to two calls.

Categorical representation of ethologically relevant vocalizations in the DCIC

The high proportion of social-selective neurons and rapid population discriminability suggests that the DCIC network may represent social vocalizations with high fidelity. To test this, we shifted from single-neuron analyses to population decoding. We trained a linear decoder for all pairwise stimuli presented to the bats and found that decoding within category performed best for social vocalizations (Fig. 3g,h and Extended Data Fig. 6b; decoding weights and SI increase linearly), suggesting that the higher proportion of social-selective neurons allows the network to encode these vocalizations with higher

https://doi.org/10.1038/s41593-025-01932-3

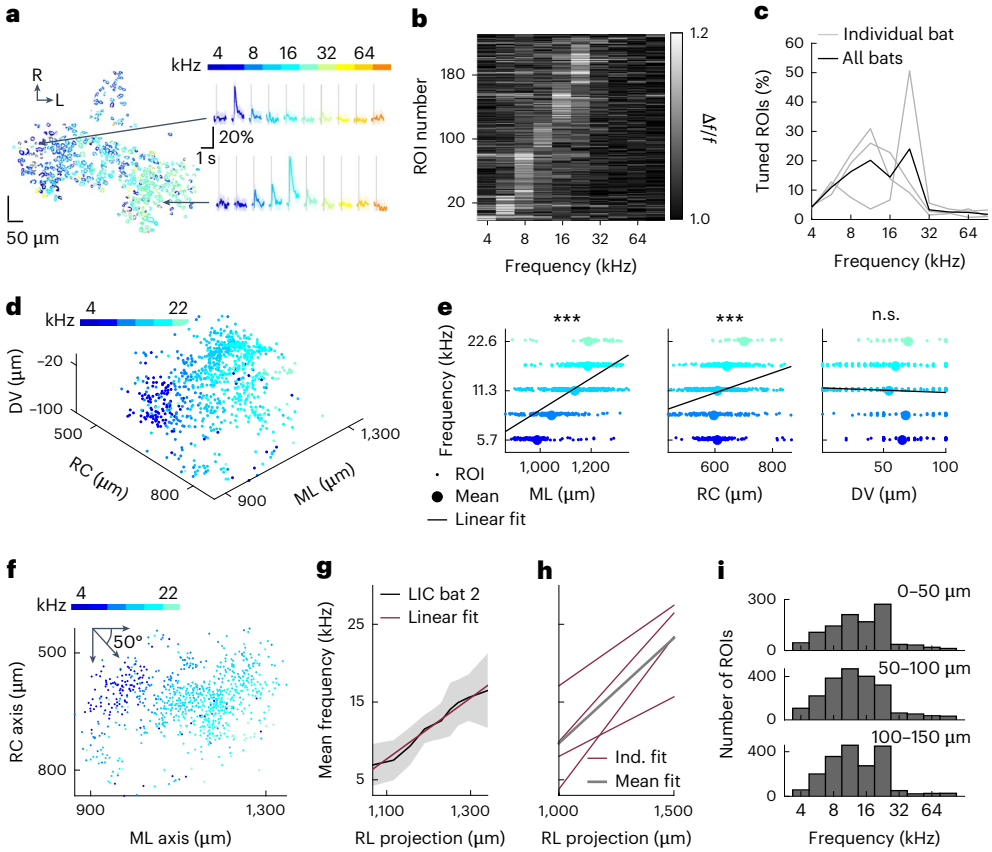


Fig. 2 | Superficial tonotopy in the DCIC. a, Left: example site from the RIC of bat 2, ROIs color-coded by their BF. Right: tone-evoked PSTHs for example cells. The first cell showed clear tuning to 5.7 kHz and the second to 16 kHz. **b**, Tuning curves for all tone-responsive ROIs from the example site presented in **a**. **c**, Distribution of tuning per bat (gray, $n = 3$) and averaged (black) as a percentage of the tuned population. Most of the ROIs' tuning is in the low-to-mid frequency range (93.3 ± 1.7% of tuned neurons' BF between 5.7 kHz and 22.6 kHz). **d**, 3D reconstruction of bat 2 LIC ($n_{site} = 3$, $n_{cells} = 835$). Tuned ROIs are color-coded according to their BF, showing low frequency rostrally and medially and higher frequency caudally and laterally. **e**, ROIs' BF (individual ROI, small dot; average, larger dot) as a function of distance along three anatomically defined axes, for example, LIC in **d**, for each frequency (left, ML; middle, RC; right, DV). Linear fit (black line, one per anatomical axis) is significant for ML and RC but not for DV ($n = 835$ cells; best line fit; $t = 23.7$, $R^2 = 0.4$ and $P = 5.02 \times 10^{-95}$, $t = 8$, $R^2 = 0.07$ and $P = 4.79 \times 10^{-15}$, $t = -1.5$, $R^2 = 0.003$ and $P = 0.13$, respectively; n.s., not significant). **f**, 2D reconstruction of bat 2 LIC. Tuned ROIs are color-coded according to their BF, showing frequency bands oriented rostromedially. **g**, Mean BF (in black, shading: mean ± s.d.) and its linear fit (in red) along the 50° RL axis for LIC in **f**. The mean tuning increases along this axis, confirming the existence of a tonotopic gradient. **h**, Linear fit (red) of tuning along the 50° RL axis for each bat hemisphere (red; Ind., individual fit) and the average (gray). Tuning increases as a function of RL distance by 2.1 octaves mm^{-1} on average. **i**, Distribution of BF across bats along the DV axis (50-µm bins from 0 to 150 µm, $n = 10$). Tuning does not significantly vary with depth (Friedman test, $\chi^2(2) = 1.8$, $P_{DV} = 0.4066$, $n = 10$).

fidelity. We then sought to test the extent to which the DCIC network encoded call category by looking at decoding strength within versus across categories. Interestingly, decoding accuracy was significantly higher across than within categories for all category types (Fig. 3g,h, within social mean: 80 ± 2.3, across social/navigation: 87 ± 1.3, across social/USV: 91.1 ± 1.5, across social/USV: 91.8 ± 1, one-way ANOVA, $F(6,436) = 101.6$, $P = 1.3408 \times 10^{-79}$, post-hoc Tukey's HSD, $P_{S/N/S} = 0.0002$, $P_{S/S/U} < 1 \times 10^{-4}$ and $P_{S/U/N} < 1 \times 10^{-4}$). Despite the lack of single-neuron selectivity to mouse USV's, the DCIC network reliably decoded individual calls within the mouse USV category (70 ± 2.5%; Fig. 3g), but to a lesser extent than across categories of bat vocalizations (91 ± 1.4%; Fig. 3g), which was similar to conspecific categorical decoding (Fig. 3h, 87 ± 1.3%, 91.1 ± 1.5% and 91.8 ± 1% respectively, ANOVA, $P_{N/S/S/U} = 0.58$, $P_{N/S/N/U} = 0.28$). The similarity in cross-category decoding between conspecific vocalizations (social versus navigation) and conspecific versus heterospecific (social versus mouse USV, navigation versus mouse USV), is at odds with the sounds' spectral distances (see Supplementary Table 2 for a summary of decoding accuracy as a function of acoustic distance) and indicates that the DCIC network contains a higher-fidelity representation of conspecific calls, probably relying on species-specific spectrotemporal filtering properties.

To test whether this proposed species-specific filtering is limited to social calls or covers both conspecific vocalization categories, we performed additional experiments and varied the spectral properties of the navigation sequences while preserving the matched temporal structure. We also added three additional social vocalization sequences to tile the acoustic space (Extended Data Fig. 4a–d). Population pairwise decoding in two bats ($n_{cells} = 1,575$) provided similar results with the original set of stimuli (Extended Data Fig. 7a). When spectral variation to the navigation sequences was introduced (Extended Data Fig. 7b), navigation sequences could now be decoded from one another (mean 62.43%, s.e.m. 1.2%). However, this improvement in decoding accuracy did not bridge the gap across categories (mean difference, within navigation versus across: 10%, one-way ANOVA, $F(6,655) = 67.2$, $P = 4.5820 \times 10^{-65}$, post-hoc Tukey's HSD, $P_{N/N/S} = 9.4866 \times 10^{-9}$).

We next sought to understand whether neural selectivity to distinct conspecific vocalizations could be explained by differences in lower-level auditory feature responses. For example, social calls have shallower frequency-modulated component at the lower-frequency tail of the sweep compared with navigation calls, and social-selective neurons may more strongly encode low-frequency stimuli, clustering along the tonotopic axis to reflect their tuning preference. Surprisingly, this

https://doi.org/10.1038/s41593-025-01932-3

was not the case. The ratio of social-selective and navigation-selective neurons was uniformly distributed along the recorded tonotopic axis (Fig. 3i). Importantly, other low-level features, including frequency tuning, bandwidth, the relationship between tuning and bandwidth, duration and complex sound (linear sweeps and white noise bursts) responses, were similar for social- and navigation-selective neurons (Fig. 3j–l, tuning: Friedman test, $\chi^2(2) = 0.06$, $P = 0.9688$; bandwidth: two-way ANOVA, $F(2,76) = 0.2$, $P = 0.7992$; downsweep rate tuning: two-way ANOVA, $F(2,44) = 0.03$, $P = 0.9682$; Extended Data Fig. 8a,b; probability density estimate for social- and navigation-selective cells' tuning and bandwidth: two-sample Kolmogorov–Smirnov test, KS statistic 0.1889, $P = 0.0707$; see Extended Data Fig. 8c–e for other tuning properties). Thus, low-level features could not explain categorical selectivity, indicating that neural selectivity for vocalizations in the DCIC may reflect a higher-order categorical representation. To ascertain the degree to which IC neurons encode species-specific spectrotemporal features, we performed a stimulus-reversal experiment in which calls in each sequence were reversed and presented to two bats. The reversal transforms conspecific downsweep calls into upsweeps, while preserving the overall spectral content and duration (see Extended Data Fig. 7c (left) for a stimulus example). We assessed individual cell sensitivity to the reversal by computing the response difference between forward (natural vocalizations) and reversed stimuli and quantified this using a shuffling procedure. While some individual neurons were insensitive to this reversal, a large percentage of cells was sensitive to the reversal for social calls (31.2%), followed by navigation calls (22.8%), and a smaller fraction showed reversal sensitivity to mouse USVs (15.9%) (Extended Data Fig. 7c). This heightened sensitivity to reversal for conspecific vocalizations mirrors the population decoding results, underscoring a more faithful representation of the vocalization per se rather than underlying spectral features or duration.

To further assess the categorical nature of the encoding at the cell level, we constructed a set of four stimulus continua that morphed a social vocalization into a navigation vocalization in a step-wise fashion (STRAIGHT-TANDEM toolbox[37], modified for bat vocalization parameters; see Fig. 4a for spectrograms of a continuum), similar to speech morphing[38]. Each continuum is composed of a series of stimuli that span the fundamental frequency (f_0) space between social and navigation calls in equidistant steps. There is a total of seven stimuli, progressing in 20% increments, with an additional step at 50%. Following a procedure that revealed learned categories in prefrontal cortical neurons of monkeys[39], and of stimulus categories in the owl optic tectum[40], we reasoned that, if single cells exhibited categorical sensitivity to vocalizations, the response at the boundary between categories would be larger than within categories for equidistant spectral steps. Single neurons responded reliably more to one side of a continuum than the other (Fig. 4b, top: example cell preferring the navigation side of the continuum, bottom: example cell preferring the social side of the continuum). The category sensitivity of each neuron was assessed using a category index (CI)[38–40], which reflects the difference in response between categories and within categories (see Fig. 4a and the Methods for the computation). We found that a substantial portion of the sound-evoked population (294 neurons, 21.75%) was category sensitive to one of the continua. The responses of individual neurons with a significant CI were scaled by dividing by their maximum response to more clearly illustrate the shape of the transition along the continuum (Fig. 4c). Some neurons showed switch-like transition (Fig. 4c, left, example neuron social tuned 1), while others showed more graded responses (Fig. 4c, right, example neuron navigation-tuned 5). Population response patterns transitioned from being dominated by social-tuned neurons to navigation-tuned neurons at the category boundary (Fig. 4d,e, left to right). To determine if the population responses were more similar within categories than across categories, we computed a correlation matrix representing the pairwise similarity between pairs of population responses (Fig. 4f, Pearson correlation coefficient). The response patterns were grouped along the continuum's category boundary and showed an abrupt transition at the boundary (Fig. 4f), consistent with markers of categorization. Interestingly, excluding these neurons from the pairwise decoding analysis did not affect the within-category or across-category decoding accuracy,

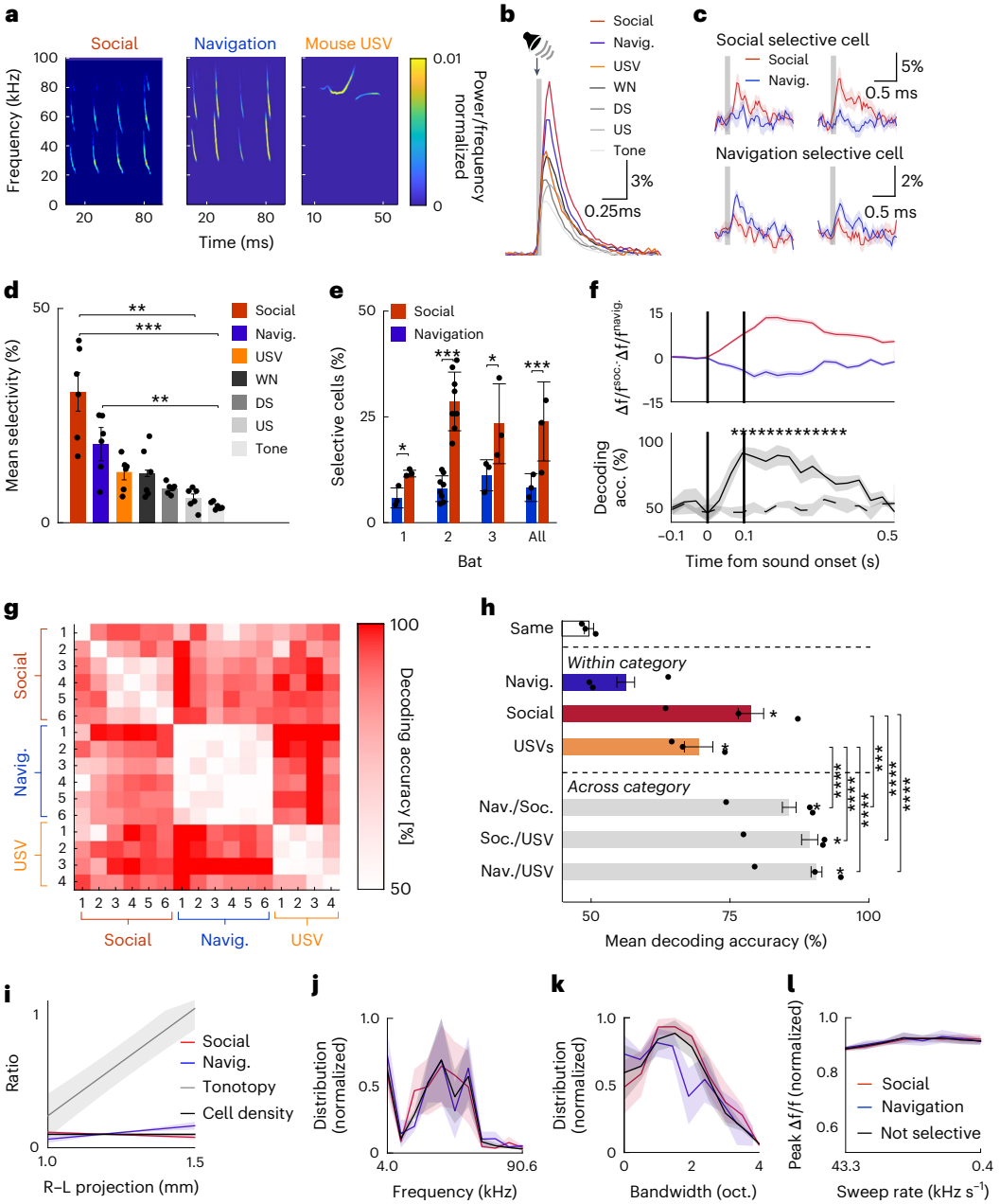
Fig. 3 | Categorical representation of bat vocalizations in the DCIC. a, Example spectrograms for the vocalization categories presented. Left: social call sequence. Center: temporally matched navigation call sequence. Right: mouse ultrasonic vocalization (USV). **b**, Population sound-evoked response as $\Delta f/f$ for all stimulus types presented ($n_{bat} = 2$, $n_{cell} = 2{,}232$). Conspecific bat calls evoked the highest evoked response followed by mouse USVs and white noise, with tones evoking the smallest. All stimuli evoked significantly different responses (Friedman test on peak amplitude, post-hoc Bonferroni, $\chi^2(6) = 5{,}875.7$, $P < 0.05$, $n = 2{,}232$) apart from USVs and navigation calls. **c**, Example cells' sound-evoked responses ($\Delta f/f$, median, shading: mean ± s.e.m.) for two social (red) and corresponding navigation (blue) call sequences. Top: example social-selective cell. Bottom: example navigation-selective cell. **d**, Average population selectivity for stimulus categories (error bar: mean ± s.e.m., black dot: mean selectivity for each pair of stimuli). The imaged population is more selective to the bat calls than other stimulus types (Kruskal–Wallis test, $\chi^2(6) = 29.4$, $P = 5.0278 \times 10^{-5}$, $n = 42$, post-hoc Tukey's HSD, $P_{social/tone} = 0.0001$, $P_{navig.l/tone} = 0.0049$, $P_{social/us} = 0.0146$). For statistical comparison between pairs of stimuli, the numbers of cells, sites and bats are combined ($n_{bats} = 3$, $n_{sites} = 14$, $n_{cells} = 5{,}860$ for tone, social, navigation and USV and $n_{bats} = 2$, $n_{sites} = 6$, $n_{cells} = 2{,}595$ for DS, US and WN). **e**, Average percentage of selective cells per bat (left: black dot percentage of selective cells per site, bat 1: $n_{sites} = 3$, bat 2: $n_{sites} = 8$, $n_{sites} = 3$) and across bats (right: $n_{bats} = 3$, $n_{sites} = 14$, $n_{cells} = 5{,}860$, error bar: mean ± s.d., black dot: mean per bat). The neuronal population shows a bias toward social selectivity (Kruskal–Wallis test; bat 1: $\chi^2(1) = 3.8$, $P_{bat1} = 0.0495$, $n = 6$; bat 2: $\chi^2(1) = 11.3$, $P_{bat2} = 0.0008$, $n = 16$; bat 3: $\chi^2(1) = 3.9$, $P_{bat3} = 0.0495$, $n = 6$; all: $\chi^2(1) = 16.3$, $P_{all} = 5.2683 \times 10^{-5}$, $n = 28$). **f**, Time course of population selectivity. Call sequence onset and offset are indicated by the black lines. Top: difference between mean social-evoked and navigation $\Delta f/f$ for top 5% social-selective cells (red, average across bats, shading: mean ± s.e.m.) and top 5% navigation-selective cells (blue). Note that the difference in activity between social and navigation preferring cells increases rapidly after stimulus onset. Bottom: the time course of decoding accuracy (social versus navigation, mean across bats: plain line, shading: s.e.m.). The star indicates significance compared with chance level decoding (dashed line, permutation test, $\alpha < 0.05$). The bat call category can be decoded from neuronal population 60 ms after sound onset before the end of the call sequence. **g**, Pairwise decoding accuracy averaged across bats for each call (conspecific or heterospecific) exemplar pairs. Decoding accuracy within categories is lower than across category (white, chance; red, 100% accuracy). **h**, Mean decoding accuracy within and across categories (error bar: mean ± s.e.m.). Identity decoding ('same') in black represents chance level. Higher-than-chance decoding accuracy is indicated by a single star above each bar. Decoding accuracy is significantly lower within versus across categories (within social mean: 80 ± 2.3, across social/navigation: 87 ± 1.3, across social/USV: 91.1 ± 1.5, across social/USV: 91.8 ± 1; one-way ANOVA, $F(6,436) = 101.6$, $P = 1.3408 \times 10^{-79}$, post-hoc Tukey's HSD, $P_{S/N/S} = 0.0002$, $P_{S/S/U} < 1 \times 10^{-4}$, $P_{U/N/S} < 1 \times 10^{-4}$ and $P_{U/S/U} < 1 \times 10^{-4}$). Note that decoding accuracy is not significantly different between groups within (one-way ANOVA, $P_{S/U} = 0.1288$) and groups across (ANOVA, $P_{N/S/S/U} = 0.5762$, $P_{N/S/N/U} = 0.2839$, $P_{S/U/N/U} = 0.9996$). For statistical comparison, each pairwise comparison was computed per bat over sites and cells ($n_{bats} = 3$, $n_{sites} = 14$, $n_{cells} = 5{,}860$). For each comparison, the mean decoding accuracy per bat is displayed as a black dot. **i**, Average linear fit of cell ratio along the tonotopic gradient (gray) for both social-selective (red), navigation-selective (blue) and cell density over bats and IC side (shading: mean ± s.e.m.). Only the tonotopy ratio slope is significantly different from the cell density slope ratio (one-way ANOVA, $F(3,12) = 24.2$, $P = 2.2264 \times 10^{-5}$, post-hoc Tukey's HSD, $P = 0.0001$). **j**, Mean BF-normalized distribution for social, navigation and nonselective cells (shading: mean ± s.e.m., Friedman test, $\chi^2(2) = 0.06$, $P = 0.9688$, $n = 10$). **k**, Average bandwidth distribution normalized for social, navigation and nonselective cells (shading: s.e.m., two-way ANOVA, $F(2,76) = 0.2$, $P = 0.7992$). **l**, Mean downsweep rate tuning as peak $\Delta f/f$ normalized for social, navigation and nonselective cells (shading: mean ± s.e.m., two-way ANOVA, $F(2,44) = 0.03$, $P = 0.9682$).

https://doi.org/10.1038/s41593-025-01932-3

yielding a maximum difference of only 1.75% in accuracy for decoding (Friedman test, $\chi^2(2) = 0.09$, $P = 0.7697$, $n = 28$). This indicates that, while these neurons exhibit category sensitivity to f_0 features, they are not the sole contributors to the population's categorical representation. Finally, we computed the CI for all possible boundaries (10%, 30%, 50%, 70% and 90%) and then computed the population average of all cells. For example, for the 10% boundary, we compute a CI for each neuron and each continuum using 10% as the boundary. Importantly, we found that, even when we aimed to maximize the difference at this other boundary (for example, 10%), the population average continued to display a sharp shift centered around 50% (Extended Data Fig. 9a–e). The effective boundary that we obtained using this approach is 54.2 ± 9.4% (Extended Data Fig. 9e,f).

Spatial organization of vocalization category selectivity

Finally, we sought to evaluate whether the neural encoding of vocalizations in the IC is spatially organized by exploiting the high spatial resolution of two-photon imaging. For each category-selective neuron, we computed the percentage of neighboring neurons that were also selective to the same category in 30-µm steps in a site-specific manner. Surprisingly, this revealed clusters of category-selective neurons to either social or navigation calls for all bats in both the left and right DCIC (Fig. 5). The percentage of social-selective neighboring cells is significantly higher than what would be expected by chance from a 'spatial shuffle' for distances of 30 µm and 60 µm (Fig. 5b,c, permutation test two-sided, $P_{30} = 0.01$, $P_{60} = 0.01$, see inset, one-way ANOVA, $F(9,13195) = 39.8$, $P = 1.3669 \times 10^{-70}$, post-hoc Tukey's HSD, $P_{30/90} < 1 \times 10^{-4}$). The same relationship was observed in navigation-selective cells (Fig. 5f–h, permutation test two-sided, $P_{30} = 0.01$, see inset, one-way ANOVA, $F(9,4867) = 39.8$, $P = 1.3199 \times 10^{-5}$, post-hoc Tukey's HSD, $P_{30/90} = 0.001$). To further describe the spatial organization of these selectivity clusters, we identified cluster centers and used a composite map to compare their location with the anatomically defined tonotopic gradient (Fig. 5d,h). Clusters centers are distributed along the tonotopic axis and follow the cell density distribution (see Extended Data Fig. 10a,b for an example site and Extended Data



https://doi.org/10.1038/s41593-025-01932-3

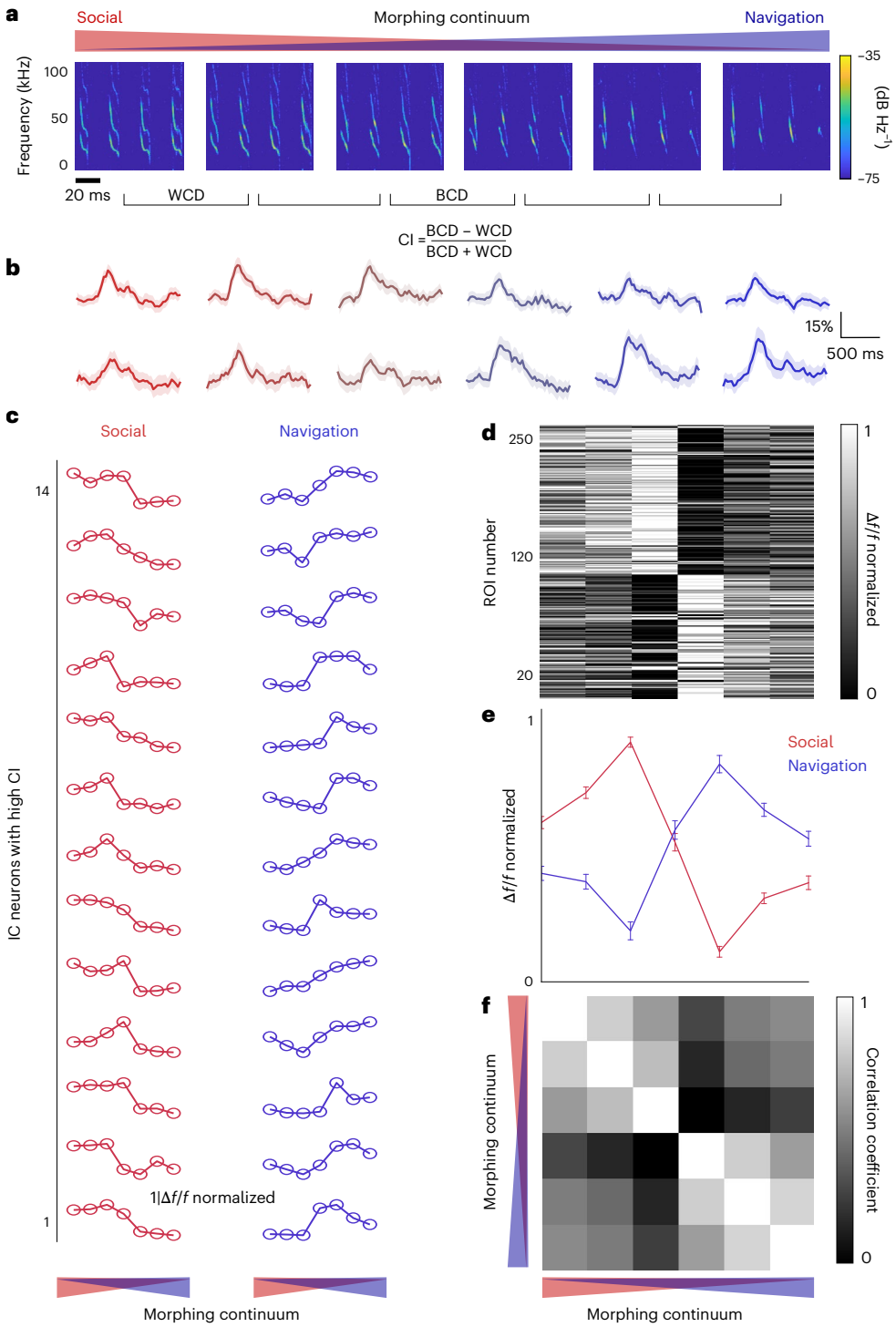


Fig. 4 | Explicit categorical representation at the single-cell level. a, Example social-to-navigation morphing continuum: stimulus spectrogram along the continuum. The reconstructed exemplars on each side (right: social; left: navigation) and the intermediate equidistant steps are shown above. Below, the CI is computed using the BCD and the WCD. **b**, Example cell responses: responses to the stimuli in **a** for a cell showing categorical preference to the navigation part of the continuum (top) and a cell showing categorical preference to the social part of the continuum (bottom). **c**, Response of individual neurons: mean evoked activity of individual neurons (highest CI neurons), normalized by the range of responses for each neuron over the continuum. Left: neurons showing categorical encoding of social stimuli. Right: neurons showing categorical encoding of navigation stimuli. **d**, Response pattern of the categorical population: mean evoked $\Delta f/f$ of 294 cells (21.75% of the recorded population) with significantly positive CI, sorted by continuum preference (social at the top, navigation at the bottom). **e**. Population average response: population average for **d** showing an abrupt transition at the category boundary (error bar: mean ± s.e.m. for the 294 cells with significantly positive CI presented in **d**). **f**, Correlation matrix: pairwise similarity between population responses to different stimuli of the morphing continuum. The white square pattern indicates that population responses are more similar on each side of the continuum.

https://doi.org/10.1038/s41593-025-01932-3

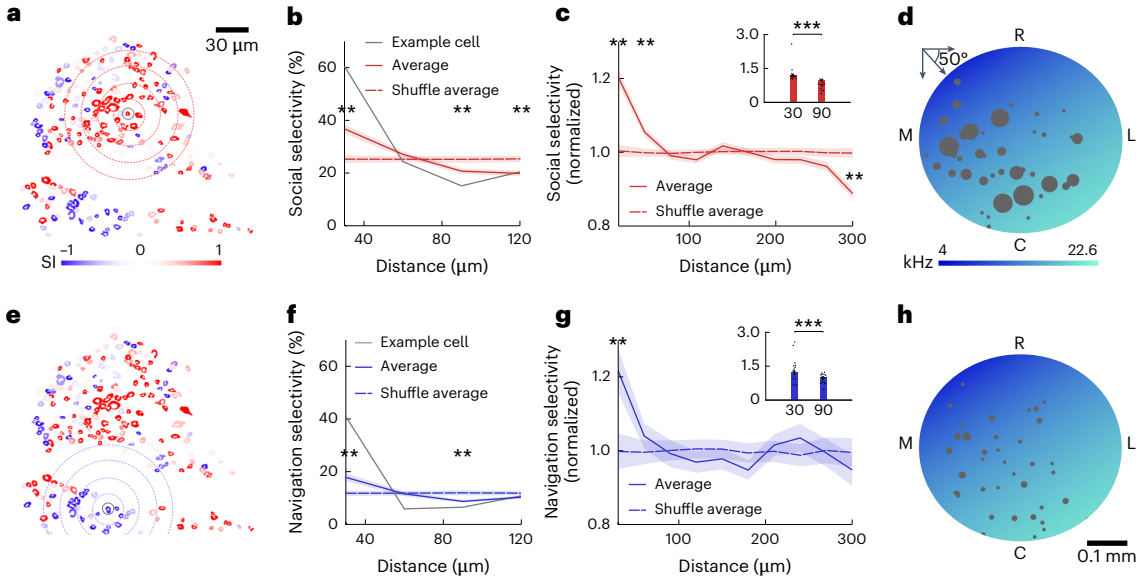


Fig. 5 | Vocalization category selectivity is spatially clustered, independent of tonotopic gradient. a, Example site. The ROI color-coding corresponds to the social (red) versus navigation (blue) selectivity index. The social-selective example cell presented in **b** is highlighted with a gray circle. The dotted red circles indicate areas used to compute the percentage of social-selective neighboring cells (in 30-µm steps). **b**, Percentage of social-selective cells as a function of distance from social-selective cells (n_{social} = 78, shading: mean ± s.e.m.) for the example site in **a**. For the example cell highlighted in **a** (gray line), the site average (plain red line) is significantly higher than the spatial shuffle (dotted red line) at 30 µm (permutation test two-sided, P_{30} = 0.01) and decreases below the shuffle line above 90 µm (permutation test two-sided, P_{90} = 0.01, P_{120} = 0.01). **c**, Percentage of social-selective cells as a function of distance from social-selective cells normalized across all sites and bats (n_{bats} = 3, n_{sites} = 14, n_{cells} = 5,860, n_{social} = 1,383). The average is higher (plain red line) than the spatial shuffle (dotted red line) at distances below 60 µm (permutation test two-sided, P_{30} = 0.01, P_{60} = 0.01) (shading: mean ± s.e.m.) and lower at 300 µm (permutation test two-sided, P_{300} = 0.01). In addition, the percentage (normalized) decreases significantly between 30 µm and 90 µm (one-way ANOVA, $F(9,13195)$ = 39.8, $P = 1.3669 \times 10^{-70}$, post-hoc Tukey's HSD, $P_{30/90} < 10^{-4}$), as displayed in the inset. The average local percentage of social cells (normalized) per recording site is overlaid as black dots. **d**, Social cluster centers (gray dots, dot size is proportional to number of cells in the cluster) are distributed across composite tonotopic gradient (BF color-coded). **e**, Example site in **a**. The gray circle highlights the navigation-selective cell presented in **f**. Here, the concentric dotted blue circles indicate the distance in 30-µm steps to the example cell. **f**, Percentage of navigation-selective cells as a function of distance from navigation-selective cells (shading: mean ± s.e.m.) for the example site in **d**. For the example cell highlighted in **d** (gray line), the site average (plain blue line) is significantly higher than the spatial shuffle (dotted blue line) at 30 µm (permutation test two-sided, P_{30} = 0.01) and decreases below the shuffle line above 90 µm (permutation test two-sided, P_{90} = 0.01). **g**, Percentage of neighboring navigation-selective cells as a function of distance from navigation-selective cells normalized across all sites and bats (n_{bats} = 3, n_{sites} = 14, n_{cells} = 5,860, n_{social} = 502). The average is higher (plain blue line) than spatial shuffle (dotted blue line) below 30 µm (P_{30} = 0.01) (shading: mean ± s.e.m.). In addition, the percentage (normalized) decreases significantly between 30 µm and 90 µm (one-way ANOVA, $F(9,4867)$ = 39.8, $P = 1.3199 \times 10^{-5}$, post-hoc Tukey's HSD, $P_{30/90}$ = 0.001), as displayed in the inset. The average local percentage of navigation cells (normalized) per recording site is overlaid as black dots. **h**, Navigation cluster centers (gray dots, dot size is proportional to number of cells) are distributed across composite tonotopic gradient (BF color-coded).

Fig. 10c for the cluster distribution). Neurons exhibiting call selectivity were, thus, spatially clustered into functional microdomains largely independent of the larger-scale tonotopic gradient.

Discussion

Current models of categorical sound processing rely on the idea that the periphery and midbrain serve primarily a feedforward and filter-bank role[10,41,42]. However, our data demonstrate that categorical representations of vocalization category at the single-neuron and population level emerge earlier in the auditory hierarchy than posited by traditional models and display 'hotspots' of spatially clustered, categorical encoding. The dorsal cortex is one of three subdivisions of the IC. Its role in auditory processing has been less studied compared with the CNIC, mainly due to its unique connectivity patterns and superficial anatomical location. The DCIC has strong feedforward input, exhibits local colliculo-collicular connections and receives prominent projections from the cerebral cortex[23,25]. This interaction between multiple inputs makes it challenging to determine precisely what drives the emergence of categorical representation in the DCIC. The rapid decoding and the similarity of these responses to those reported in the CNIC, however, suggests that local computations play an important role. Top-down signals probably contribute to this—and potentially even more so during active behavior—by further sharpening the categorical nature of these signals or making them context dependent. Future work that leverages projection-specific molecular, optical and physiological tools—techniques that remain challenging in nontraditional species but are becoming increasingly possible[43]—will be required to address this important question.

The categorical representations we observe appear to be independent of the superficial tonotopy in the DCIC. This surprising finding can be contextualized by earlier research on the tuning properties of IC neurons. Several studies investigating the encoding of vocalizations (both conspecific and heterospecific) across various species have highlighted the difficulty in predicting neural responses to vocalizations based on responses to other (simple) auditory stimuli, in both the auditory cortex and the IC[28,44,45]. For example, Machens et al. found that, even using finely characterized spectrotemporal receptive fields, they could predict only 11% of the responses to communication sounds and environmental noise in the rat auditory cortex[46]. One explanation is that BF or excitatory frequency tuning curves alone do not adequately describe neuronal response properties, even at the IC level, which depend on local circuitry and inhibitory processes[28]. IC neurons exhibit complex response properties, such as FM direction selectivity[47], sound duration tuning[19] and multipeak tuning curves[48]. In addition, combination-sensitive neurons, which respond to the specific spectrotemporal features of two distinct acoustic components,

have been reported in the mustached bat[49–51], mouse[52] and anuran IC[53]. Our data suggest that the observed categorical representations could not be explained by a number of low-level features. Future work will be needed to identify the combinatorial set of features that allows rapid discrimination of one particular vocalization category from another.

More broadly, our data support a revised view of categorical sound processing in which ethologically relevant sensory streams are instantiated at the single-neuron and population level in early auditory centers to provide specified and spatially segregated acoustic channels of categorical information to recipient regions. These categorical primitives may serve to increase the speed of transmission (by virtue of being early in the sensory pathway) and reduce the wiring costs (through spatially organized channels) for echolocating bats, and these computational principles could extend to other species that must rapidly extract sound meaning from rich vocal repertoires.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-025-01932-3.

References

1. May, B., Moody, D. B. & Stebbins, W. C. Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *J. Acoust. Soc. Am.* **85**, 837–847 (1989).
2. Barker, A. J. et al. Cultural transmission of vocal dialect in the naked mole-rat. *Science* **371**, 503–507 (2021).
3. Ehret, G. Infant rodent ultrasounds—a gate to the understanding of sound communication. *Behav. Genet* **35**, 19–29 (2005).
4. Wyttenbach, R. A., May, M. L. & Hoy, R. R. Categorical perception of sound frequency by crickets. *Science* **273**, 1542–1544 (1996).
5. Baugh, A. T., Akre, K. L. & Ryan, M. J. Categorical perception of a natural, multivariate signal: mating call recognition in túngara frogs. *Proc. Natl Acad. Sci. USA* **105**, 8985–8988 (2008).
6. Reiss, D., McCowan, B. & Marino, L. Communicative and other cognitive characteristics of bottlenose dolphins. *Trends Cogn. Sci.* **1**, 140–145 (1997).
7. Reichmuth, C. & Casey, C. Vocal learning in seals, sea lions, and walruses. *Curr. Opin. Neurobiol.* **28**, 66–71 (2014).
8. Prather, J., Nowicki, S., Anderson, R., Peters, S. & Mooney, R. Neural correlates of categorical perception in learned vocal communication. *Nat. Neurosci.* **12**, 221–228 (2009).
9. Wright, G. S., Chiu, C., Xian, W., Wilkinson, G. S. & Moss, C. F. Social calls predict foraging success in big brown bats. *Curr. Biol.* **24**, 885–889 (2014).
10. Liu, S. T., Montes-Lourido, P., Wang, X. & Sadagopan, S. Optimal features for auditory categorization. *Nat. Commun.* **10**, 1302 (2019).
11. Theunissen, F. E. & Elie, J. E. Neural processing of natural sounds. *Nat. Rev. Neurosci.* **15**, 355–366 (2014).
12. Tsunada, J. & Cohen, Y. E. Neural mechanisms of auditory categorization: from across brain areas to within local microcircuits. *Front. Neurosci.* **8**, 161 (2014).
13. Yin, P., Strait, D. L., Radtke-Schuller, S., Fritz, J. B. & Shamma, S. A. Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. *Curr. Biol.* **30**, 1649–1663.e5 (2020).
14. Grill-Spector, K. & Weiner, K. S. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* **15**, 536–548 (2014).
15. Surlykke, A. & Moss, C. F. Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory. *J. Acoust. Soc. Am.* **108**, 2419–2429 (2000).
16. Gadziola, M. A., Grimsley, J. M. S., Faure, P. A. & Wenstrup, J. J. Social vocalizations of big brown bats vary with behavioral context. *PLoS ONE* **7**, e44550 (2012).
17. Montoya, J., Lee, Y. & Salles, A. Social communication in big brown bats. *Front. Ecol. Evol.* **10**, 903107 (2022).
18. King, A. J. & Nelken, I. Unraveling the principles of auditory cortical processing: can we learn from the visual system? *Nat. Neurosci.* **12**, 698–701 (2009).
19. Fremouw, T., Faure, P. A., Casseday, J. H. & Covey, E. Duration selectivity of neurons in the inferior colliculus of the big brown bat: tolerance to changes in sound level. *J. Neurophysiol.* **94**, 1869–1878 (2005).
20. Parras, G. G. et al. Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat. Commun.* **8**, 2148 (2017).
21. Carbajal, G. V. & Malmierca, M. S. The neuronal basis of predictive coding along the auditory pathway: from the subcortical roots to cortical deviance detection. *Trends Hear.* **22**, 2331216518784822 (2018).
22. Thomas, J. M. et al. Stimulus-specific adaptation in specialized neurons in the inferior colliculus of the big brown bat, *Eptesicus fuscus*. *Hear. Res.* **291**, 34–40 (2012).
23. Winer, J. A. & Schreiner, C. E. *The Inferior Colliculus* (Springer, 2005).
24. Barnstedt, O., Keating, P., Weissenberger, Y., King, A. J. & Dahmen, J. C. Functional microarchitecture of the mouse dorsal inferior colliculus revealed through in vivo two-photon calcium imaging. *J. Neurosci.* **35**, 10927–10939 (2015).
25. Winer, J. A., Larue, D. T., Diehl, J. J. & Hefti, B. J. Auditory cortical projections to the cat inferior colliculus. *J. Comp. Neurol.* **400**, 147–174 (1998).
26. Guo, W. et al. Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. *J. Neurosci.* **32**, 9159–9172 (2012).
27. Schreiner, C. E. & Langner, G. Laminar fine structure of frequency organization in auditory midbrain. *Nature* **388**, 383–386 (1997).
28. Klug, A. et al. Response selectivity for species-specific calls in the inferior colliculus of mexican free-tailed bats is generated by inhibition. *J. Neurophysiol.* **88**, 1941–1954 (2002).
29. Poon, P. W. F., Sun, X., Kamada, T. & Jen, P. H.-S. Frequency and space representation in the inferior colliculus of the FM bat, *Eptesicus fuscus*. *Exp. Brain Res* **79**, 83–91 (1990).
30. Koay, G., Heffner, H. E. & Heffner, R. S. Audiogram of the big brown bat (*Eptesicus fuscus*). *Hear. Res.* **105**, 202–210 (1997).
31. Wong, A. B. & Borst, J. G. G. Tonotopic and non-auditory organization of the mouse dorsal inferior colliculus revealed by two-photon imaging. *eLife* **8**, e49091 (2019).
32. Oliver, D. L. Projections to the inferior colliculus from the anteroventral cochlear nucleus in the cat: possible substrates for binaural interaction. *J. Comp. Neurol.* **264**, 24–46 (1987).
33. Wright, G., Chiu, C., Xian, W., Moss, C. & Wilkinson, G. Social calls of flying big brown bats (*Eptesicus fuscus*). *Front. Physiol.* **4**, 214 (2013).
34. Chen, C., Xian, W. & Moss, C. F. Flying in silence: echolocating bats cease vocalizing to avoid sonar jamming. *Proc. Natl Acad. Sci. USA* **105**, 13116–13121 (2008).
35. Salles, A. et al. Neural response selectivity to natural sounds in the bat midbrain. *Neuroscience* **434**, 200–211 (2020).
36. Woolley, S. M. N. & Portfors, C. V. Conserved mechanisms of vocalization coding in mammalian and songbird auditory midbrain. *Hear. Res.* **305**, 45–56 (2013).
37. Kawahara, H. & Morise, M. Interactive tools for making vocoder-based signal processing accessible: flexible manipulation of speech attributes for explorational research and education. *Acoust. Sci. Technol.* **45**, 48–51 (2024).

38. Tsunada, J., Lee, J. H. & Cohen, Y. E. Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.* **105**, 2634–2646 (2011).
39. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316 (2001).
40. Mysore, S. P. & Knudsen, E. I. Flexible categorization of relative stimulus strength by the optic tectum. *J. Neurosci.* **31**, 7745–7752 (2011).
41. Kar, M. et al. Vocalization categorization behavior explained by a feature-based auditory categorization model. *eLife* **11**, e78278 (2022).
42. Achutha, A. C., Peremans, H., Firzlaff, U. & Vanderelst, D. Efficient encoding of spectrotemporal information for bat echolocation. *PLoS Comput. Biol.* **17**, e1009052 (2021).
43. Diebold, C. A. et al. Rapid sensorimotor adaptation to auditory midbrain silencing in free-flying bats. *Curr. Biol.* **34**, 5507–5517.e3 (2024).
44. Bar-Yosef, O., Rotman, Y. & Nelken, I. Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J. Neurosci.* **22**, 8619–8632 (2002).
45. Holmstrom, L. A., Eeuwes, L. B. M., Roberts, P. D. & Portfors, C. V. Efficient encoding of vocalizations in the auditory midbrain. *J. Neurosci.* **30**, 802–819 (2010).
46. Machens, C. K., Wehr, M. S. & Zador, A. M. Linearity of cortical receptive fields measured with natural sounds. *J. Neurosci.* **24**, 1089–1100 (2004).
47. Kuo, R. I. & Wu, G. K. The generation of direction selectivity in the auditory system. *Neuron* **73**, 1016–1027 (2012).
48. Palmer, A. R., Shackleton, T. M., Sumner, C. J., Zobay, O. & Rees, A. Classification of frequency response areas in the inferior colliculus reveals continua not discrete classes. *J. Physiol.* **591**, 4003–4025 (2013).
49. Mittmann, D. H. & Wenstrup, J. J. Combination-sensitive neurons in the inferior colliculus. *Hear. Res.* **90**, 185–191 (1995).
50. Leroy, S. A. & Wenstrup, J. J. Spectral integration in the inferior colliculus of the mustached bat. *J. Neurosci.* **20**, 8533–8541 (2000).
51. Portfors, C. V. & Wenstrup, J. J. Excitatory and facilitatory frequency response areas in the inferior colliculus of the mustached bat. *Hear. Res.* **168**, 131–138 (2002).
52. Portfors, C. V. & Felix, R. A. Spectral integration in the inferior colliculus of the CBA/CaJ mouse. *Neuroscience* **136**, 1159–1170 (2005).
53. Lee, N., Schrode, K. M. & Bee, M. A. Nonlinear processing of a multicomponent communication signal by combination-sensitive neurons in the anuran inferior colliculus. *J. Comp. Physiol. A* **203**, 749–772 (2017).

https://doi.org/10.1038/s41593-025-01932-3

Methods

Animals

Data were collected from five adult big brown bats (*E. fuscus*, three females and two males, wild caught) collected from exclusion sites in the State of Maryland. The bats were collected under the permit issued by the Maryland Department of Natural Resources (no. 55440). Epiphyseal–diaphyseal bone fusion was used to determine that all bats included in this manuscript were adults[54]. All procedures were approved by the Institutional Animal Care and Use Committees at Johns Hopkins University (protocol no. BA17A107), where this research was conducted. The number of animals and cells for each result and corresponding figure is presented in Supplementary Table 3.

Anesthesia and pain management. All surgical procedures were performed under isoflurane anesthesia (1.5–3%) with 0.8 oxygen flow in accordance with the protocol no. BA17A107. Thirty minutes before anesthesia induction, animals received half a dose of meloxicam (Metacam) orally for its analgesic effect. The second half dose was delivered once the animal awoke, as well as sulfamide, for antibiotic coverage. Postsurgery recovery lasted 4–5 days with a daily dose of meloxicam and sulfamide.

GCaMP6f viral transfection. We injected AAV5-CamKII-GCaMP6f (Addgene, 100834-AAV5) bilaterally in the IC. We first resected the muscle located on top of the skull (temporalis medius), working from an incision placed above the midline to uncover the IC skull surface bilaterally. To avoid disrupting the recording surface, we entered the IC at an angle of 35° to 50° and traveled ~1 mm to reach an injection site at the center of the DCIC at a depth of −600 µm (see Fig. 1a for a visualization). We used a Hamilton syringe (10 µl, 700 series, 34 gauge) preloaded with the virus (1 µl per side) and injected using a syringe pump (Harvard Apparatus) at 75 nl min^{-1} and a force of 30%. We kept the needle in place for an additional 5 min after injection before slowly removing from the brain. Animals were allowed to recover for at least 1 month before the next procedure (to allow viral expression).

Headpost implantation. We resected the muscle on the rostral part of the skull. A custom-built flat two-pronged headpost (Fig. 1a) was cemented with dental cement (Metabond) to the cranium at a distance of at least 4 mm from the IC center, identified visually. The bat was then put on rest for at least one day before the thinned-skull procedure to allow the cement to cure.

Thinned-skull approach. We adapted the thinned-skull approach developed for mice[55,56]. Using a surgical drill, we carefully thinned the skull above both sides of the IC, without piercing the skull layer. The skull was then dried, and a thin layer of transparent fast-drying adhesive was applied above the thinned area. Bats were allowed to recover for a day before data collection.

Two-photon calcium imaging acquisition. Two-photon calcium imaging was performed using a two-photon resonant microscope (Neurolabware) equipped with a 16× objective (Nikon). Two-photon fluorescence was excited at 980 nm using an Insight X3 laser (SpectraPhysics). Data contained in this manuscript were collected at 31.25 Hz over one plane at either 4× (294 µm × 192 µm, cropped) or 2× (568 µm × 374 µm, cropped). In addition, a volumetric stack (*z* stack) was collected at the end of each session in 2-µm or 5-µm steps from the skull surface up to 200 µm below the surface.

Two-photon calcium imaging preprocessing. We used suite2p[57] to register and align each session. We drew regions of interest (ROIs) on the session mean image using ImageJ (NIH, 2.9; Java 1.8), as automatic ROI identification was perturbed by large differences in brightness related to changes in skull thickness over the field of view (FOV). Fluorescent time traces of individual ROIs were extracted using a custom MATLAB-based toolbox. $\Delta f/f$, the change in fluorescence intensity over the baseline fluorescence, was computed per ROI and sound presentation by dividing the raw fluorescence by the median fluorescence for a baseline period defined as the ten frames (312.5 ms) that immediately preceded sound onset.

Site alignment. Each imaging site position on the surface of the IC was determined by manually matching the site's vasculature with the corresponding widefield vasculature image (larger). All widefield images collected for one side were then tiled using the vasculature as a landmark, creating a composite map per side (for an example of imaging sites overlaid on a composite widefield image, see Fig. 1b). The relative position of each imaging site to the most anterior and medial widefield image was then computed, allowing us to generate ML and RC coordinates within one side. Those coordinates were then placed in the bat's coordinate system by manually matching the vasculature from the obtained composite maps to images covering the full extent of both ICs, obtained during the skull thinning.

ROI coordinates. Mediolateral and caudorostral coordinates for each cell were obtained by aligning recording sites within side and animal ($n_{bat} = 3$, $n_{hemi} = 4$, $n_{sites} = 20$) to an anatomical landmark obtained from surgery images covering both sides: the rostro-medial boundary between SC and IC marked by the confluence of the superior sagittal and lateral sinuses. ROI depth was obtained by estimating the skull thickness above each ROI and subtracting that value from the overall site depth from brain surface identified using the *z* stack. Skull thickness was computed for each FOV as points above a 95% fluorescent boundary based on the FOV raw fluorescence distribution (skull brightness is very high; Extended Data Fig. 2 illustrates this method for an example site).

Auditory stimulation. All auditory stimuli were presented contralaterally to the imaging site using one free-field electrostatic speaker (ES1, Tucker-Davis Technologies (TDT)) placed at 45° angle, 5 cm away and at a 0° azimuth from the bat's head. All auditory stimuli were presented to the awake passively listening bat using a high-frequency auditory signal processor (RZ6, TDT) at a 200 kHz sampling rate. Sounds were generated in TDT (pure tones, white noise, upsweep, downsweep) or were recorded as Waveform Audio File Format (WAV) files and then played back via the TDT. All stimuli were calibrated at 70 dB sound pressure level (SPL) using a GRAS Sound & Vibration microphone (GRAS 46DP), placed at the bat's ear putative position. For each bat call sequence, the amplitude of the waveform was normalized by the peak intensity and set to 70 dB SPL.

Auditory stimuli. A series of stimuli were presented: pure tones ranging from 4 to 90.51 kHz, ½ octave spaced (ten tones, pseudo-random presentation, durations: 2, 4, 6, 10 and 20 ms), white noise, upsweep and downsweep (4 kHz to 90.51 kHz, durations: 2, 4, 6, 10, 20, 50, 100 and 200 ms) and bat and mouse vocalizations (five or six exemplars: food-claiming call sequences[9,33,35], five or six exemplars: temporally matched echolocation call sequence[35], two or four mouse ultrasonic vocalizations[58]). Each sound was typically repeated 10 times (1 out of 23 analyzed sessions had 5 repetitions) and the intersound interval was kept constant at 1.2 s for all stimuli. In two bats, the same series of stimuli was presented with the following modifications: pure tone durations were fixed at 4 ms and were spaced ¼ octave apart within the same range. In addition to the original set of vocalizations, we also presented an expanded and more acoustically varied version of the vocalizations set, consisting of nine exemplars of food-claiming call sequences and nine exemplars of temporally matching but spectrally varying navigation sequences[9,59]. Furthermore, we presented a reversed version of this set of exemplars, which was obtained by reversing the

direction of each call in a sequence using Adobe Audition. Finally, we presented a set of four vocalization morphing continua. Morphed versions of the vocalizations were created using the STRAIGHT toolbox[60]. The fundamental frequency (f_0) of each call in each sequence was extracted, and morphing was accomplished by calculating the shortest trajectory between the fundamental frequencies of two vocalization exemplars. Morphed versions of both a social vocalization sequence and a navigation vocalization sequence were created at intervals of 0% (reconstructed social), 20%, 40%, 50%, 60%, 80% and 100% (reconstructed navigation sequence). In total, four morphing continua were created, derived from two vocalization exemplars and four temporally matched navigation exemplars. Auditory stimuli were presented in blocks and randomized within each block.

Histology. At the end of the experiment, the bats were perfused transcardially with 15 ml of 1% phosphate-buffered saline followed by 20 ml of 4% paraformaldehyde in 1% phosphate-buffered saline. The fixed part of the brain containing the IC was sectioned in 50-µm slices coronally. To ascertain the excitatory expression of GCaMP6f viral expression, immunochemistry was performed following a double-staining procedure for GFP and GAD67 (ref. 61) (Extended Data Fig. 1a,b). For GFP staining, a goat anti-GFP primary antibody (Novus Biologicals, cat. no. NB100-1770) was used at a 1:500 dilution, followed by a donkey anti-goat Alexa Fluor 488 secondary antibody (Jackson ImmunoResearch, cat. no. 705-545-003) at a 1:500 dilution. For GAD67 staining, a rabbit anti-GAD67 primary antibody (ThermoFisher, cat. no. PA5-21397) was used at a 1:1,000 dilution, followed by a donkey anti-rabbit Cy3 secondary antibody (Jackson ImmunoResearch, cat. no. 711-165-152) at a 1:500 dilution. This procedure was performed on brain sections from two experimental bats: one pilot bat (male, adult) and one bat included in this study (B3, female, adult).

Statistics and reproducibility. All data analysis with data collected from five adult big brown bats and statistical testing were performed using custom functions written in MATLAB 2018b or 2021b (The MathWorks). The normality of data was assessed before statistical testing using a one-sample Kolmogorov–Smirnov test. ANOVAs were performed when the distribution was normal, whereas nonparametric tests were applied (Kruskal–Wallis or Friedman) when the data did not follow a normal distribution. The tests used to assess significance are specified in the main text and in the figure legends. The fluorescent example traces depicting single-cell responses (refer to Figs. 1d and 4b) have been smoothed for enhanced visualization. However, it is important to note that the average $\Delta f/f$ curves remain unsmoothed, and all analyses, particularly those involving time course dependencies, were conducted using the unaltered traces. No statistical method was used to predetermine sample size. Imaging sites with lower quality were excluded from the vocalization data analysis (as shown in Extended Data Fig. 5a). This assessment was based on the number of tone-responsive ROIs to prevent contamination of the SI. The investigators were not blinded to allocation during experiments and outcome assessment.

Best frequency identification. Best frequency (BF) was defined as the sound frequency evoking the highest response measured as the median of the peak amplitude across 50 trials (10 frames from stimulus onset, 312.5 ms) at the sound level presented (70 dB). BF analysis was restricted to neurons that exhibited significant tone-evoked activity (paired t-test, $\alpha < 0.05$).

Tonotopic gradient characterization. We first established the relationship between three anatomically defined axes, rostrocaudal, mediolateral and dorsoventral (RC, ML and DV, respectively), and BF tuning (limited to the 4–22.7 kHz range that comprises most of the tuning, Fig. 2c) per side and animal, using a linear regression. We then computed the angle that explained the tuning best by reducing the mean square error per side and bat (Supplementary Table 1). The angle used for the rest of the analysis is the bat-averaged angle 48°. The composite theoretical tonotopic gradient (Fig. 4d,h) was obtained using the averaged angle and averaged linear regression coefficients across bats.

Sound selectivity. An individual cell is referred to as selective if its response shows invariance to members of one category while showing different responses to a second category[35,38]. We computed the selectivity index for each cell as

$$\mathrm{SI} = 1/n\sum_i^n((X_i - Y_i))/((X_i + Y_i)),$$

where X_i is the median activity evoked within the response window (ten frames after sound onset, 312.5 ms) by exemplar i of a stimulus type and Y_i is the corresponding activity evoked by an exemplar i of a second temporally matched stimulus. Because multiple exemplars were presented, SI corresponds to the average of the selectivity index computed for pairs of temporally matched exemplars. SI ranged from 1 to −1, with positive SI indicating selectivity to the first category and negative SI indicating selectivity to the second category. Permutation tests were performed to assess cell selectivity significance (two-sided, $\alpha < 0.05$).

Category decoding. We evaluated the accuracy with which the population activity could be classified as a natural stimulus exemplar by training and testing a linear discriminant classifier[62] in a pairwise fashion, using a leave-one-out cross-validation. Two training population vectors ($\mathbf{c}_{s1}$, $\mathbf{c}_{s2}$) were obtained by averaging the activity for 320 ms (ten frames) after stimulus onset for each stimulus pair and each ROIs for a subset of randomly selected repetitions. The decoding vector $\mathbf{w}$ was defined as

$$\mathbf{w} = \mathbf{c}_{s1} - \mathbf{c}_{s2}$$

and the bias as

$$b = (-(\mathbf{c}_{s2} \times \mathbf{w} + \mathbf{c}_{s1} \times \mathbf{w}))/2.$$

The test population vector (remaining repetition, x) was classified as stimulus 1 or 2, following the rule

$$y(x) = \mathbf{w} \times x + b.$$

If $y(x) > 0$, x is classified as stimulus 1, and as stimulus 2 if $y(x) < 0$.

We repeated this procedure 100 times and averaged the accuracy computed for each stimulus pair per bat. The accuracy presented in Fig. 3g and Extended Data Fig. 6a is averaged across bats. To determine if decoding accuracy is higher than chance level, we compared each group with the identity decoding (same stimulus decoding, different trials, 100 times). The statistical significance for each type of category comparison (within/across) reported in Fig. 3h was assessed using an ANOVA run on the distribution of pairwise comparison over all bats. Similarly, for Fig. 3f we evaluated the accuracy with which the population could discriminate between the two bat call categories using all exemplars of each category per time point (corresponding to the timing of a frame) and the same approach.

The weights shown in Extended Data Fig. 6b correspond to the average vector $\mathbf{w}$ over 100 cross-validations.

Basic auditory feature characterization. To determine if basic functional auditory features could underlie the population selectivity to natural bat calls, we compared a number of auditory functional features for the selective and nonselective populations, including BF tuning, tuning curve bandwidth, amplitude of evoke response to various sounds, and upsweep and downsweep tuning (Fig. 3i–l and Extended Data Fig. 8). Bandwidth is defined as the area under the curve for a

normalized tuning curve (small if the tuning is sharp and large if the tuning is broad). Population tuning to downsweep and upsweep rate and white noise duration was computed as the average of the tuning curves normalized by the peak amplitude.

Category index computation. To determine the categorical nature of the response to different vocalization categories (social versus navigation) at the individual cell level, we computed a category index (CI)[38–40] as described below. For each neuron the within-category difference (WCD) corresponds to the difference between the average of the evoked response (within a 320-ms window after stimulus onset) to morphs 20% and 40% on one side of the continuum and 60% and 80% on the other side of the continuum. We then obtain the between-category difference (BCD) as the difference of the evoked response to morphs of 40% and 60%. The CI is the difference between the BCD and the WCD divided by their sum,

$$\mathrm{CI} = \frac{\mathrm{BCD} - \mathrm{WCD}}{\mathrm{BCD} + \mathrm{WCD}},$$

such that neurons showing a CI >0 present a larger difference in responses between than within categories. The confidence intervals on the CI were created with a shuffling procedure (shuffle of the labels, 100 times). Neurons with a CI higher than the 95th percentiles of the shuffled distribution for each continuum were identified as being category sensitive.

Spatial clustering analysis. The local percentage of selective cells was computed in 30-µm radial increments for each selective cell for either social or navigation selectivity per recording sites. Given the variations in site selectivity, the local percentage was divided by the site overall selectivity to compute a local selectivity ratio as a function of distance per bat. Spatial shuffle distributions were computed per 30-µm radial increment and imaging site using the x and y coordinates (unpaired) to determine significance using a permutation test (two-sided, $\alpha < 0.05$). Selective clusters were identified on a composite map across sites and bats using density-based spatial clustering of applications with noise (DBSCAN, based on the pairwise Euclidean distance between ROIs). The cluster minimum number of cells was set to 3, and epsilon was set to 30 µm.

Reporting summary

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

Data availability

The preprocessed imaging data are publicly available via Zenodo at https://doi.org/10.5281/zenodo.14743696 (ref. 63). The raw data will be made available upon request to the corresponding author.

Code availability

The custom-built MATLAB code used for stimulus generation and analysis will be made available upon request to the corresponding author.

References

54. Brunet-Ross, K. & Wilkinson, G. S. Methods for age estimation and the study of senescence in bats. *Ecological and Behavioral Methods for the Study of Bats* (eds Kunz, T. H. & Parsons, S.) 315–325 (Johns Hopkins University Press, 2009).
55. Yang, G., Pan, F., Parkhurst, C. N., Grutzendler, J. & Gan, W.-B. Thinned-skull cranial window technique for long-term imaging of the cortex in live mice. *Nat. Protoc.* **5**, 201–208 (2010).
56. Shih, A. Y., Mateo, C., Drew, P. J., Tsai, P. S. & Kleinfeld, D. A polished and reinforced thinned-skull window for long-term imaging of the mouse brain. *J. Vis. Exp.* **61**, 3742 (2012).
57. Pachitariu, M. et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at *bioRxiv* https://doi.org/10.1101/061507 (2017).
58. Heckman, J. J. et al. High-precision spatial localization of mouse vocalizations during social interaction. *Sci. Rep.* **7**, 3017 (2017).
59. Wohlgemuth, M. & Moss, C. F. Midbrain auditory selectivity to natural sounds. *Proc. Natl Acad. Sci. USA* **113**, 2508–2513 (2016).
60. Kawahara, H., Masuda-Katsuse, I. & de Cheveigné, A. Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds1. *Speech Commun.* **27**, 187–207 (1999).
61. Ito, T. & Oliver, D. L. The basic circuit of the IC: tectothalamic neurons with different patterns of synaptic organization send different messages to the thalamus. *Front. Neural Circuits* **6**, 48 (2012).
62. Bagur, S. et al. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. *Nat. Commun.* **9**, 2529 (2018).
63. Lawlor, J. et al. Spatially clustered neurons in the bat midbrain encode vocalization categories Dataset. *Zenodo* https://doi.org/10.5281/zenodo.14743696 (2025).

Acknowledgements

We thank A. Salles for providing the conspecific bat calls, B. Englitz for providing the mouse USVs, C. Diebold for animal care and surgical support and Y. Boubenec for help with the spectral distance analysis. We thank C. Drieu, S. Moore and N. Kothari for their thoughtful feedback on our manuscript. This work was supported by an NIH Brain Initiative R34 grant R34NS118462 (K.V.K., C.F.M. and M.J.W.), NSF grant NCS-FO 1734744 and ONR grant N00014-17-1-2736 (C.F.M.).

Author contributions

Conceptualization: J.L., M.J.W., C.F.M. and K.V.K.; experiments: J.L. and M.J.W.; data analysis: J.L. and K.V.K., writing—original draft: J.L. and K.V.K.; writing—review and editing: J.L., M.J.W., C.F.M. and K.V.K.

Competing interests

The authors declare no competing interests.

Additional information

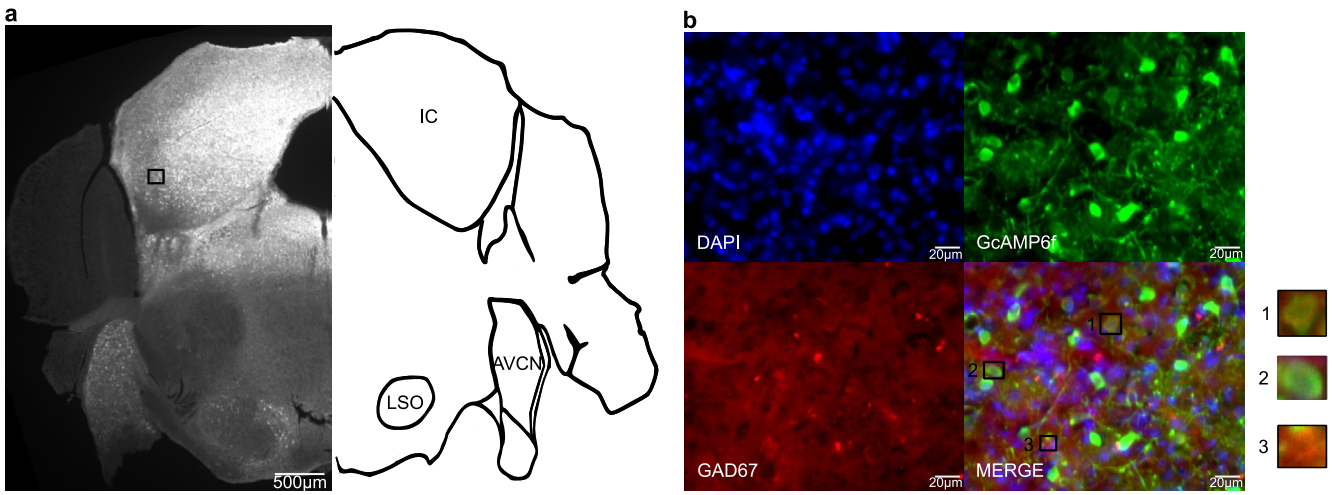
Extended data is available for this paper at https://doi.org/10.1038/s41593-025-01932-3.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-025-01932-3.

Correspondence and requests for materials should be addressed to Kishore V. Kuchibhotla.

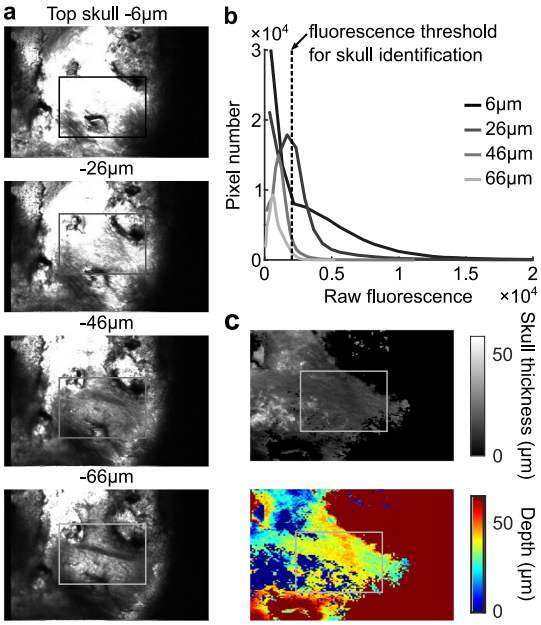
Peer review information *Nature Neuroscience* thanks Manuel Malmierca and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41593-025-01932-3



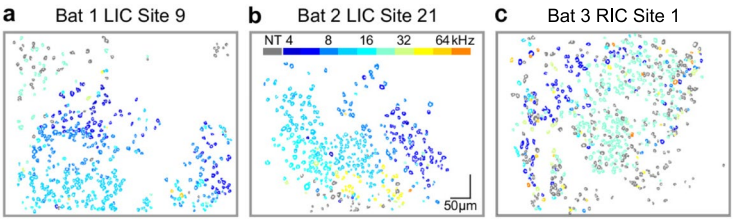
Extended Data Fig. 1 | Histology and immunostaining for GCaMP6f and GAD67. **a**, GFP immunostaining shows expression of GCaMP6f throughout the IC. The black box indicates the location of the site highlighted in **b**. **b**, Immunostaining for IC locus highlighted in **a** for cell nucleus (DAPI, top left), GCaMP6f (GFP, top right), inhibitory neurons (GAD67, bottom left), and merged image (bottom right). Merged image suggests minimal overlap between GCcAMP6f and GAD67 expression. Example neurons numbered on the merged image are displayed on the right: 1) GAD67+ and GcAMP6f+ (yellow), 2) GcAMP6f+ and GAD67- (green), 3) GcAMP6f- and GAD67+ (red). The staining procedure was repeated on the brains of two bats.

https://doi.org/10.1038/s41593-025-01932-3



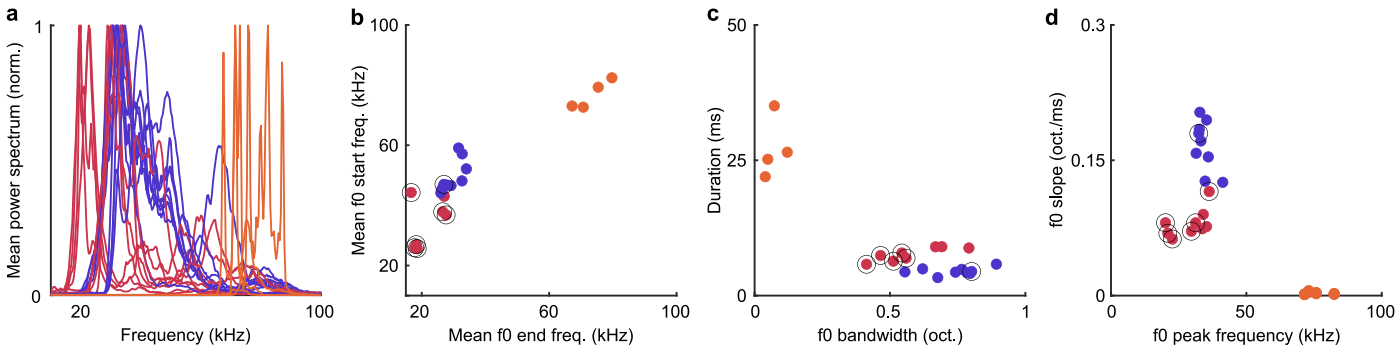
Extended Data Fig. 2 | Site depth estimation corrected by skull thickness. **a**, Mean images (150 frames) from 1X volumetric stack for example 2X imaging site (highlighted by the black to gray box, colors correspond to depth in **b**) every 10 µm until example site imaging depth (bottom image). **b**, Raw fluorescence distribution for images displayed in A. Note that shallower images possess a longer high fluorescence tail indicating a larger skull coverage. The criterion used for skull identification (95th percentile of the overall volumetric stack distribution) is displayed in the dashed line. **c**, Skull thickness and corresponding depth estimation for example volumetric stack in **a** and **b**.

https://doi.org/10.1038/s41593-025-01932-3



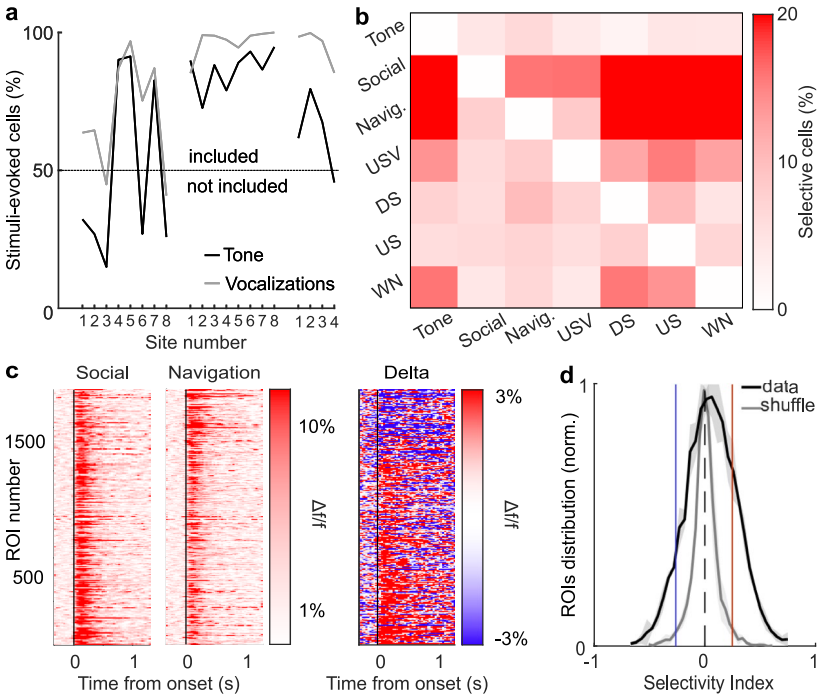
Extended Data Fig. 3 | Tonotopic gradient in example sites for 3 separate bats. a, Example site from the left DCIC (LIC) of bat 1. Cell ROIs are color-coded according to their best frequency (scale in **b**). Cells without an observable tuning are shown in gray ('NT' for not-tuned). **b**, Example site from the left DCIC of bat 2. **c**, Example site from the right DCIC (RIC) of bat 3.

https://doi.org/10.1038/s41593-025-01932-3



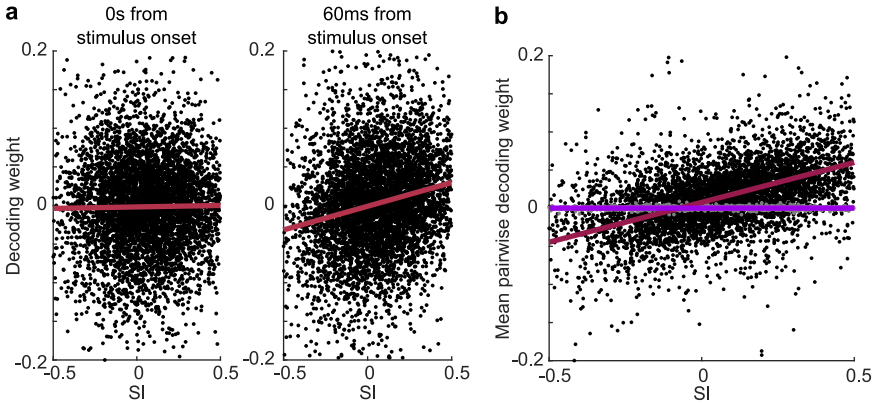
Extended Data Fig. 4 | Acoustic distribution of vocalization. a, Average power spectrum for each vocalization sequences normalized by its max. **b**, Average vocalization sequence f0 end frequency as a function of start frequency. Individual point represents each sequence (social: red, navigation: blue, mouse USVs: orange). Black circles highlight the stimuli presented in Fig. 3, all stimuli were presented for Extended Data Fig. 7a, b. Conspecific sequences end and start frequency are significantly different (mean social: 22.9 and mean navigation: 29.7 kHz KW test, p = 0.012 and p = 9.2e-4, respectively). **c**, Average vocalization duration as a function of duration. Social and navigation sequences can be separated by duration (mean social: 8 ms, mean navigation: 4.5 ms, KW test, p = 4.9e-4). **d**, Vocalization f0 slope as a function of f0 peak frequency. Social vocalizations present shallower slopes (mean social: 0.08 oct./ms, mean navigation: 0.17oct./ms, KW test, p = 3.5e-4).

https://doi.org/10.1038/s41593-025-01932-3



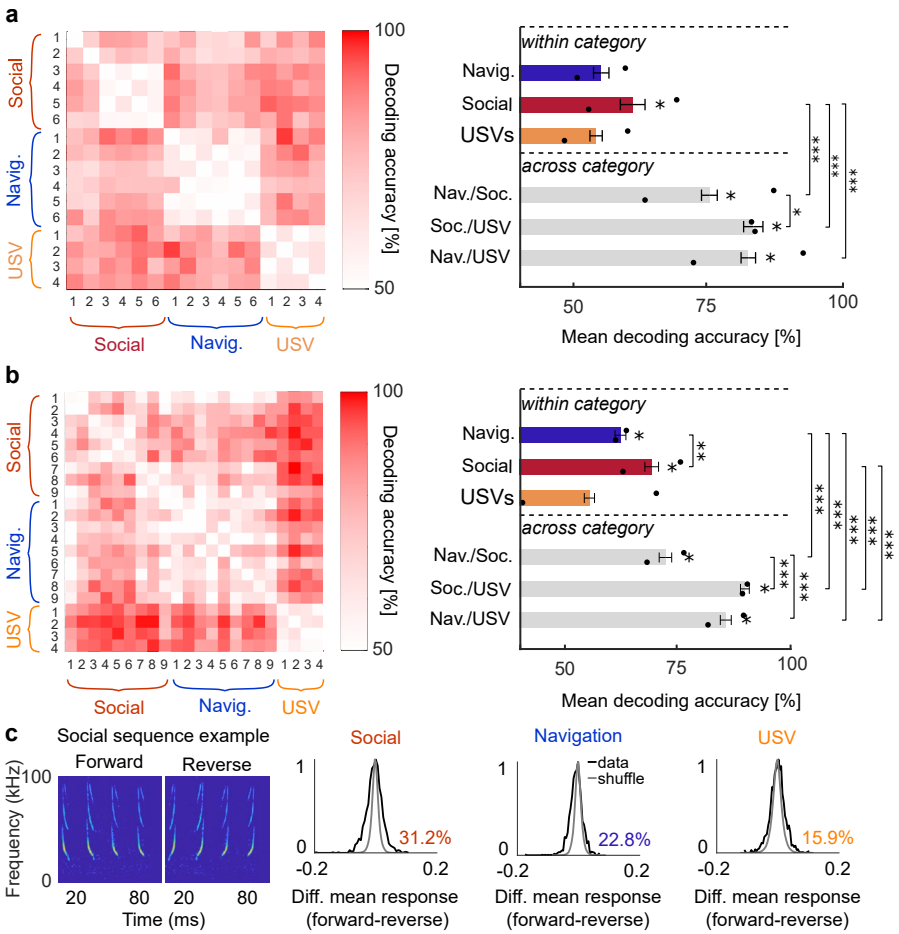
Extended Data Fig. 5 | Selectivity index method. a, Percentage of cells showing significant stimulus-evoked responses (ANOVA, $\alpha < 0.05$) for tones (in black) and vocalizations (in gray) for each recorded site. Sites with fewer than 50% of cells with significant tone-evoked responses were excluded from the vocalization analysis. **b**, Mean percentage of significantly selective cells using SI for each sound category. Note that the percentage of selective cells to conspecific calls is consistently higher than other categories. **c**, Average sound-evoked Δf/f per category for significant selective cells sorted by their selectivity index ($n_{cells} = 1885$). Left: average sound-evoked activity for social calls. Middle: average sound-evoked activity for navigation calls. Right: difference between social and navigation activity. **d**, Average selectivity index distribution across bats (shading: mean+-SEM). Average significance boundaries are displayed in blue (2.5 percentile, $SI_{bound} = -0.2959$) and red (97.5 percentile, $SI_{bound} = 0.3109$).

https://doi.org/10.1038/s41593-025-01932-3



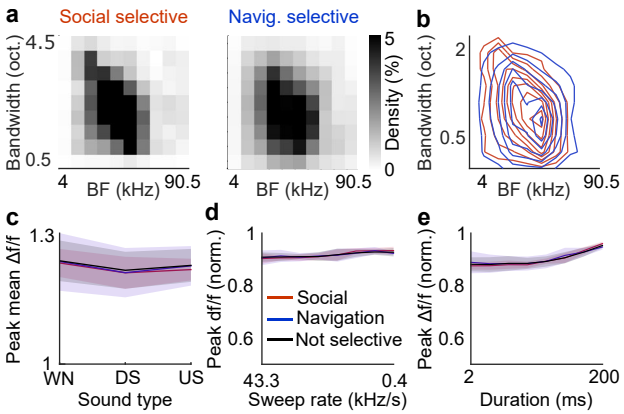
Extended Data Fig. 6 | Linear decoding method. a, Decoding weights for time course decoder (Fig. 3f) as a function of selectivity index per cell (black dot). Left: decoding weights does not increase (linear regression in red, adjusted $R^2 = -0.000129$, $p = 0.623$) with SI at 0 s after stimulus onset. Right: decoding weights increases linearly with SI at 60 ms after stimulus onset (adjusted $R^2 = 0.243$, $p < 10^{-4}$). **b**, Mean decoding weights from pairwise decoding for social vs navigation (Fig. 3g,h) increase with SI (linear regression in red, adjusted $R^2 = 0.0374$, $p = 9.12e\text{-}51$), but not for same stimulus decoding (social only, in grey, linear regression purple, adjusted $R^2 = 2.03e\text{-}05$, $p = 0.29$).

https://doi.org/10.1038/s41593-025-01932-3



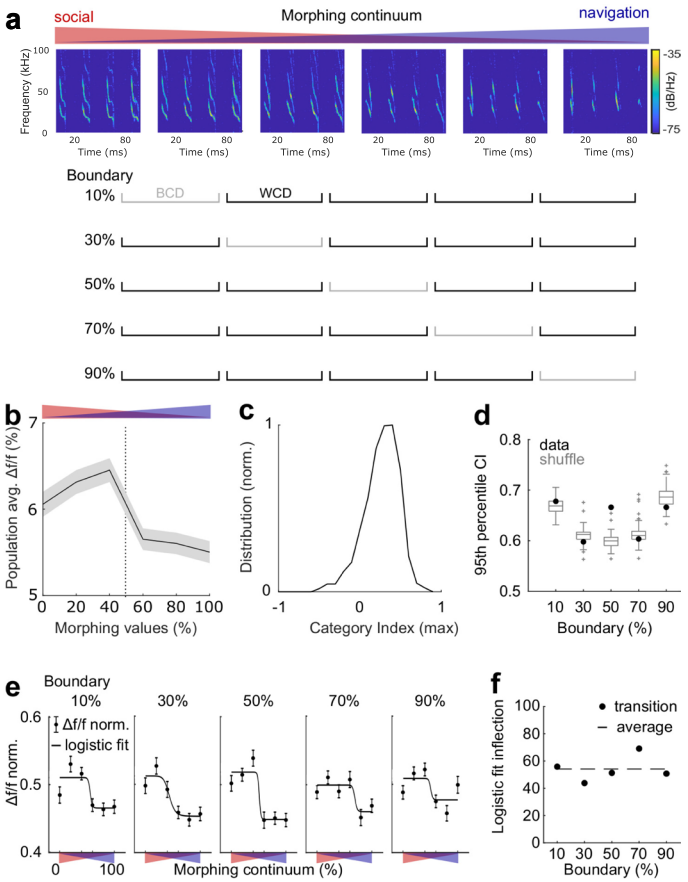
Extended Data Fig. 7 | Population categorical boundaries are maintained with expanded set of stimuli. a, Left: pairwise decoding accuracy matching the set of stimuli in Fig. 3g-h averaged across additional bats (n_{bats}=2, 1 male, 1 female, n_{cells} = 1,575). Similarly, to Fig. 3g-h, decoding accuracy within categories is lower than across category (white = 50% chance, red = 100% accuracy). Right: Mean decoding accuracy within and across categories (error bar: mean+-SEM). Higher than chance decoding accuracy is indicated by a single star above each bar. Decoding accuracy is significantly lower within vs. across categories (one-way ANOVA, $F(6,337) = 33.3$, $p = 1.7728e\text{-}31$, post-hoc Tukey's HSD, $p_{S/N/S} < 10^{-4}$, $p_{S/S/U} < 10^{-4}$, $p_{S/N/U} < 10^{-4}$) in addition decoding accuracy across bat calls is significantly lower than across bat calls and mouse USVs ($p_{N/S/S/U} = 0.0057$). **b**, Left: pairwise decoding accuracy with an extended set of stimuli (n_{bats}=2, same as **a**). Social exemplars extended from 6 to 9, temporally matched navigation exemplars are spectrally varied, USVs remain the same. Decoding accuracy remains significantly lower within vs. across categories (one-way ANOVA, $F(6,655) = 67.2$, $p = 4.5820e\text{-}65$, post-hoc Tukey's HSD, $p_{N/N/S} = 9.4866e\text{-}09$, $p_{N/S/U} < 10^{-4}$, $p_{N/N/U} < 10^{-4}$, $p_{S/S/U} < 10^{-4}$, $p_{S/N/U} = 3.5542e\text{-}15$). In addition, decoding accuracy across bat vocalizations is significantly lower than across bat versus mouse USVs ($p_{N/S/S/U} = 2.5996e\text{-}10$, $p_{N/S/N/U} = 2.1133e\text{-}18$). **c**, Population response to stimuli reversal. Left: example spectrograms for forward (original vocalizations) and corresponding reverse stimulus for a social sequence exemplar. Center: distribution difference between mean evoked response forward and reverse for each category (black, normalized) and corresponding shuffle (100 shuffle, normalized, gray). Numerical value: Percentage of cells sensitive to reversal ($\alpha < 0.05$, two-sided from shuffle) per category. The percentage of cells that changed their response when the stimuli is reversed is highest for social stimuli (31.2%), followed by navigation stimuli (22.8%) and lowest for mouse USV (15.9%).

https://doi.org/10.1038/s41593-025-01932-3



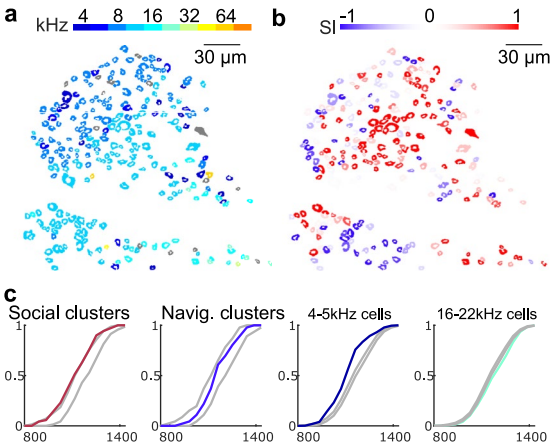
Extended Data Fig. 8 | Additional auditory features for selective and non-selective populations. a, Probability density estimate of tuning and bandwidth for social selective cells (left) and navigation selective cells (right). **b**, Probability density estimate for both cell population represented as contours, showing a large density overlap (Two-sample Kolmogorov-Smirnov test, KS Statistic= 0.1889, p = 0.0707) **c**, Average peak evoked-responses as Δf/f to different type of complex sounds: white noise (WN), downsweep (DS) and upsweep (US), for social, navigation and non-selective cells (shading: mean+-SEM, two-way ANOVA, $F(2,14) = 0.03$, $p = 0.9721$). **d**, Mean upsweep rate tuning as peak Δf/f normalized for social, navigation and non-selective cells (shading: mean+-SEM, two-way ANOVA, $F(2,44) = 0.06$, $p = 0.9459$). **e**. Mean white noise duration tuning as peak df/f normalized for social, navigation and non-selective cells (shading: mean+-SEM, two-way ANOVA, $F(2,44) = 0.04$, $p = 0.9648$).

https://doi.org/10.1038/s41593-025-01932-3



Extended Data Fig. 9 | The 50% continuum boundary is the most salient for the recorded dataset. a, Example continuum and schematic representation of CI computations for different boundaries (between category difference, BCD, in gray and within category difference, WCD, in black for each tested boundary). **b**, Average population activity as Δf/f for continua-responsive cells (n = 2 bats, n = 1122 cells with significant sound-evoked responses). The population average exhibits a shift from high to low activity centered around the 50% boundary, indicating a preference for the social segment of the continua. This pattern is reminiscent of the higher proportion of social-selective cells. **c**, Distribution of the category index for the 50% boundary (selecting for the preferred continuum). The distribution is positively skewed, suggesting a trend toward categorical responses. **d**, 95th percentile for shuffle (gray) and data (black) for each boundary. The 95% percentile for the data lies outside the shuffle distribution only for the 50% boundary, indicating that for other boundaries the CIs are not significantly different than the noise. **e**, Population. average (Δf/f, normalized per cell, error bar: mean+-SEM) along the morphing continuum for each boundary (n = 2 bats, n = 1122 cells with significant sound-evoked responses). A 4-parameter logistic fit of the population average centers around the 50% boundary, regardless of the tested boundary. **f**, Point of inflection for each logistic fit per boundary (black dot) and its average (dotted black line). The population transition is centered around the 50% boundary, regardless of the computed CI boundary (54.2+-9.4%).

https://doi.org/10.1038/s41593-025-01932-3



Extended Data Fig. 10 | Relationship between tuning and selectivity index examples and composite view. a, Tuning map for example site in Fig. 4 color-coded according to the best frequency of the neuron. **b**, Corresponding social/navigation selectivity map for the same example site, color-coded by the neuron's selectivity index. Note that the selectivity 'hotspots' do not follow the tonotopic gradient. **c**, Cumulative distribution for cluster centers and corresponding shuffle distribution (randomly selected from all cells coordinates, in gray) along the composite RL gradient of all imaged sites. Left: social cluster falls just outside of the shuffle distribution (two-sided permutation test, $p < 0.01$), perhaps due to the increase in cluster size with cell density. Center left: navigation cluster centers fall within the 95% confidence interval therefore following cell density independently of the tonotopic gradient. Center right: cells tuned to low frequencies (4 and 5.7 kHz, in dark blue) are located more rostromedially than the corresponding shuffled distribution (two-sided permutation test, $p < 0.01$). Right: cells tuned to high frequencies (16 and 22.7 kHz) are located more caudolaterally than the corresponding shuffled distribution (two-sided permutation test, $p < 0.01$).

nature portfolio

Corresponding author(s): Kishore Kuchibhotla

Last updated by author(s): Jan 30, 2025

Reporting Summary

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

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Imaging : ScanBox (Neurolabware, version: 2.5) Stimuli design: Matlab (MathWorks) R2021b (Custom code) Stimuli presentation: RPvdsEx (Tucker-Davis Technologies, version 90), Matlab R2021b (Custom code)
Data analysis	Custom code for data analysis: Matlab R2021b

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

Data

Policy information about availability of data

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The preprocessed imaging data used for this manuscript is publicly available at 10.5281/zenodo.14743696. Additional data can be made available upon reasonable request to the corresponding author (K.V.K.).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No sample size calculation was performed prior to the study. The number of bats (n = 5) is consistent with the bat electrophysiological literature (Sarel et al., 2022 n = 4, Salles et al. 2020 n = 5). Note that our approach yields a substantially larger number of recorded cells with a total of 9,443 neurons (Sarel et al., 2022, n= 430 single-units, Salles et al. 2020 n = 575 single-units).
Data exclusions	Only recorded imaging sites with visually identifiable cells were analyzed (27 imaging sites over 5 bats). Sites with fewer than 50% of cells with significant tone-evoked responses were excluded from the vocalization analysis (6 excluded sites, so 14 sites for the analysis in original 3 bats, 7 sites for analysis in the 2 new bats for the revision).
Replication	Findings remain similar when performing analyses using neurons, sites, or individual animals as replicates.
Randomization	Sound presentation was randomized within blocks of auditory stimuli.
Blinding	No, this study is an observational study of sound encoding at the cell level in healthy animals. This experimental procedures require hands-on adjustments based on animal condition (e.g., surgery, optical imaging in awake passively listening animal), full blinding is not be feasible.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibody: 1) Goat Anti-GFP (Novus Biologicals, Catalog #: NB100-1770) 2) Rabbit Anti-GAD67 (ThermoFisher, Catalog #: PA5-21397) Secondary antibody: 1)Donkey anti-goat Alexa Fluor 488 (Jackson ImmunoResearch, Catalog #: 705-545-003) 2) Donkey anti-Rabbit Cy3 (Jackson ImmunoResearch, Catalog #: 711-165-152)
Validation	Fluorescence microscopy was used for visual identification. These antibodies are routinely employed in rodent studies. Goat anti-GFP and donkey anti-goat Alexa Fluor 488 were used to enhance GFP fluorescence associated with CaMKII-GCaMP6f. Immunostaining was specific to the infection site (the IC) and corresponded with in vivo imaging results at the IC surface. Double staining with rabbit anti-GAD67 and donkey anti-rabbit Cy3 revealed that GFP-labeled cells (green) occasionally overlapped with GAD67-positive cells (red), as expected.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	No
Wild animals	Big brown bats, Eptesicus fuscus, were collected from exclusion sites in the State of Maryland under a permit issued by the Maryland Department of Natural Resources (no. 55440). Epiphyseal-diaphyseal bone fusion was used to determine that data from all bats included in this manuscript were adults. All procedures were approved by the Institutional Animal Care and Use Committees at Johns Hopkins University (protocol no. BA17A107), where this research was conducted. All animals were euthanized following the experiment in accordance with protocol no. BA17A107.
Reporting on sex	Findings apply to both male and females. Our data set contains n= 5 bats, 3 females and 2 males. We do not report sex-based analyses as we found no differences (median responsivity, KW test, p = 1; % of social selective cells, KW test, p=0.564; % of navigation selective cells, KW test, p=0.248).
Field-collected samples	The study did not involve samples collected in the field.
Ethics oversight	The Institutional Animal Care and Use Committees at Johns Hopkins University approved protocol: no. BA17A107

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA