Sensory and Motor Systems

Enhanced Stability of Complex Sound Representations Relative to Simple Sounds in the Auditory Cortex

Harini Suri[1] and Gideon Rothschild[1,2]

[1]Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109 and [2]Kresge Hearing Research Institute, Department of Otolaryngology-Head and Neck Surgery, University of Michigan, Ann Arbor, Michigan 48109

Abstract

Typical everyday sounds, such as those of speech or running water, are spectrotemporally complex. The ability to recognize complex sounds (CxSs) and their associated meaning is presumed to rely on their stable neural representations across time. The auditory cortex is critical for the processing of CxSs, yet little is known of the degree of stability of auditory cortical representations of CxSs across days. Previous studies have shown that the auditory cortex represents CxS identity with a substantial degree of invariance to basic sound attributes such as frequency. We therefore hypothesized that auditory cortical representations of CxSs are more stable across days than those of sounds that lack spectrotemporal structure such as pure tones (PTs). To test this hypothesis, we recorded responses of identified layer 2/3 auditory cortical excitatory neurons to both PTs and CxSs across days using two-photon calcium imaging in awake mice. Auditory cortical neurons showed significant daily changes of responses to both types of sounds, yet responses to CxSs exhibited significantly lower rates of daily change than those of PTs. Furthermore, daily changes in response profiles to PTs tended to be more stimulus-specific, reflecting changes in sound selectivity, compared with changes of CxS responses. Last, the enhanced stability of responses to CxSs was evident across longer time intervals as well. Together, these results suggest that spectrotemporally CxSs are more stably represented in the auditory cortex across time than PTs. These findings support a role of the auditory cortex in representing CxS identity across time.

Key words: auditory cortex; plasticity; stability

Significance Statement

The ability to recognize everyday complex sounds such as those of speech or running water is presumed to rely on their stable neural representations. Yet, little is known of the degree of stability of single-neuron sound responses across days. As the auditory cortex is critical for complex sound perception, we hypothesized that the auditory cortical representations of complex sounds are relatively stable across days. To test this, we recorded sound responses of identified auditory cortical neurons across days in awake mice. We found that auditory cortical responses to complex sounds are significantly more stable across days compared with those of simple pure tones. These findings support a role of the auditory cortex in representing complex sound identity across time.

Introduction

Everyday sounds such as human speech, animal vocalizations, the sound of running water or rustling of leaves, are spectrotemporally complex (Ehret and Haack, 1982; Doupe and Kuhl, 1998; Gygi et al., 2007). A key brain region involved in the perception of spectrotemporally complex sounds is the auditory cortex (AC; Rauschecker, 1998; Griffiths et al., 2004; Nelken, 2004, 2008; Nelken and Bar-Yosef, 2008; Bizley et al., 2009; King et al., 2018; Maor et al., 2020). For example, AC lesions result in a more profound impairment in processing complex sounds (CxSs) in comparison with pure tones (PTs) and other simple sounds in both humans (Kaga et al., 1997; Griffiths, 2012) and animal models (Ohl et al., 1999; Harrington et al., 2001; Rybalko et al., 2006). Responses of AC neurons to CxSs can often not be predicted from a linear combination of responses to the PT components of the CxS (Nelken et al., 1999; Barbour and Wang, 2003; Wang et al., 2005; Atencio et al., 2008; Sadagopan and Wang, 2009; Schreiner et al., 2011; Mizrahi et al., 2014; Harper et al., 2016; Angeloni and Geffen, 2018; Schwartz et al., 2020). Furthermore, studies using a range of approaches have shown that AC responses to CxSs represent sound "identity" with a substantial invariance to its frequency components and other acoustic parameters (Chechik and Nelken, 2012; Nelken et al., 2003, 2014; Carruthers et al., 2015; Blackwell et al., 2016; Town et al., 2018; Harpaz et al., 2021). While these studies suggest an important role of the AC in representing the identity and meaning of CxSs, to what degree these representations are stable across time remains unknown.

To support the ability to recognize sensory stimuli and their associated meaning, the neural representations of the stimuli are expected to be stable across time (Lütcke et al., 2013; Schoonover et al., 2021). At the large-scale spatial resolution, the representation of tone frequency across the AC tonotopic map is indeed generally stable in adulthood in the absence of instructive learning or manipulation of the acoustic environment (Merzenich et al., 1976; Guo et al., 2012). At the single-cell level, receptive fields of most auditory cortical neurons have been found to be stable across up to 2 h of recording, though a minority of neurons exhibited significant changes within this time frame (Elhilali et al., 2007). However, whether AC sound representations are stable across days and whether the representations of CxSs and PTs are similarly stable, remains unknown. Given the suggested involvement of AC in

representing CxS identity, we hypothesized that CxSs would be more stably represented in the AC across time compared with PTs. Here, we tested this hypothesis by recording the responses of identified layer 2/3 (L2/3) AC excitatory neurons to both PTs and CxSs across days in awake mice using two-photon calcium imaging.

Materials and Methods

All animal procedures were performed in accordance with the regulations of the University of Michigan animal care committee.

Animals

We used 13 Thy1-GCaMP6f mice [C57BL/6J-Tg (Thy1-GCaMP6f) GP5.17Dkim/J; catalog #025393, The Jackson Laboratory; 10 males, 3 females; age, 8–15 weeks], which express the GCaMP6f calcium indicator in excitatory pyramidal neurons (Dana et al., 2014). Mice were housed under a reverse 12 h light/dark cycle, with lights on at 8:30 P.M. Experiments were conducted between 11:00 A.M. and 4:00 P.M., and each animal was imaged around the same time of day across all days of data collection so that the time gap between consecutive imaging days was ~24 h.

Surgical procedure

All surgeries were performed on mice anesthetized using ketamine (100 mg/kg, i.p.) and xylazine (10 mg/kg, i.p.). Anesthetized mice were placed in a stereotaxic frame (catalog #514, Kopf Instruments), and injections of an anti-inflammatory drug (carprofen, 5 mg/kg, s.c.) and a local anesthetic (lidocaine, s.c.) were administered. A craniotomy was performed over the right primary AC (anteroposterior, −3.1 mm; mediolateral, 4.6 mm; lateral from midline; Extended Data Fig. 1-1) using a 3 mm biopsy punch (Integra), and a 3-mm-diameter round glass cranial window was secured over this craniotomy. A custom-made lightweight (<1 g) titanium head bar was attached to the left side of the skull using dental cement and cyanoacrylate glue to allow for head-fixed imaging. During the surgery, body temperature was maintained at 38°C, and the depth of anesthesia was regularly assessed by checking the pinch withdrawal reflex. Mice were treated with carprofen for 48 h postsurgically and allowed to recover for a week.

Two-photon calcium imaging

Mice were first habituated to the imaging setup and the sound protocols for 3 d. During the 3 d habituation period, the animals were exposed to the same PT and CxS stimuli as during imaging days 1–5 while being head-fixed in the same setup under the two-photon microscope while being positioned on a circular treadmill (without imaging). Each stimulus was presented 30–35 times in total across the 3 d habituation period.

During imaging, the objective of the microscope was placed perpendicular to the surface of the cranial window to access the AC. Imaging was conducted using a two-photon microscope (model Ultima IV, Bruker)

through water-immersion objectives [40×: numerical aperture (NA) = 0.65 (n = 2 mice); 16×: NA = 0.8 (n = 11 mice); Nikon], and a pulsed laser was used to provide excitation at 940 nm (MaiTai eHP DeepSee, Spectra Physics). Data were collected using galvanometric ("galvo") scanning of 256 × 256 pixel images at 3 frames/s. We conducted a separate set of recordings from the same neurons using galvo scanning and faster resonant scanning at 60 frames/s (averaging every 4 frames to yield 15 frames/s) and found that responsiveness, response magnitude, and trial-to-trial consistency were not underestimated by the slower galvo imaging sample rate (Extended Data Fig. 1-2). During the period of habituation, focal planes with a high yield of neurons were determined in L2/3 (imaged at depths of 150–330 μm; Meng et al., 2017). The overlying blood vessel patterns and position with respect to the cortical surface were noted and were used to image the same focal planes across 5 consecutive days of the experiment.

Auditory stimuli

Stimuli were generated at a sampling rate of 97.6 kHz using MATLAB and presented to the animal using an SA1 speaker amplifier, an ED1 speaker driver, and a multifield magnetic speaker (MF1) positioned ~10 cm in front of the animal, all by Tucker-Davis Technologies. Acoustic stimuli consisted of the following two protocols: PTs consisted of eight pure tone stimuli at 2–32 kHz (Fig. 1D), while CxSs consisted of four animal vocalizations (cricket, macaque, chiffchaff, and water shrew) and four environmental sounds (glass, thump, scratch, and water; Fig. 1D). The CxSs had significantly higher-frequency bandwidth, spectral entropy, and spectrotemporal modulation compared with the PTs (Fig. 1C). The duration of each sound was 500 ms (padded with silence for some of the CxSs) and sound intensity was 65–70 dB SPL. In a given imaging session for each focal plane, each sound within a protocol was repeated 10 times in a pseudorandom order with an interstimulus interval of 1.5 ± 0.3 s. The order of the sound protocols was shuffled across experiments.

Frequency bandwidth, spectral entropy, and spectrotemporal modulation were quantified for all sounds as attributes of sound complexity. Occupied frequency bandwidth quantifies the range of frequencies a sound is composed of and was calculated as the difference in frequency between the points where the integrated power crosses 0.5% and 99.5% of the total power in the spectrum. Spectral entropy of a sound quantifies how distributed its frequency content is and was calculated as the Shannon entropy of the normalized power distribution of the sound. Spectrogram autocorrelation of each sound measures the similarity of the frequency content of a sound across time bins and was calculated by temporally binning each spectrogram into 20 equally sized time bins (excluding brief periods of silence at the end of some sounds), resulting in column vectors that represent the power distribution of the sound at every time bin. We then calculated the Pearson correlations between all vectors and averaged the values of these correlations. Thus, the spectrogram autocorrelation of each sound inversely represents the degree of

spectrotemporal modulation. The "spectrotemporal modulation index" was defined as one minus spectrogram autocorrelation.

Data analysis
Preprocessing

Imaging data were run through the open-source Suite2p software package (Pachitariu et al., 2016) to correct for movement and neuropil signal, and to select neuronal regions of interest. To ensure reliable physiological measurements, we required that in any given imaging session, detected cell bodies show a compactness >0.8 and that the trace of their relative change in fluorescence (ΔF/F) shows a skewness >1.1 and clear transients (the experimenter was blind to sound responsiveness during the cell inclusion phase). A small minority of responses occurring during locomotion were excluded from all analyses (Extended Data Fig. 1-3). All further analysis was performed on the data preprocessed and output from Suite2p using custom-written MATLAB scripts (MathWorks, 2019a).

To identify the same neurons across imaging sessions, the average across-frames fluorescence image (with Suite2p median-filtering image enhancement) of each focal plane was used. The average fluorescence images of the same focal plane were then manually matched for the same neurons across days. We confirmed cell matching using fully automated image registration (MATLAB command: imregcorr) and calculation of structural similarity index (MATLAB command: ssim) of the cell bodies across days and found >95% agreement (Extended Data Fig. 1-4).

Two-photon imaging data analysis

The ΔF/F was defined for each neuron in a given imaging session as $(F(t) - F_0)/F_0$, where $F(t)$ is the raw fluorescence signal of the cell at time t, and F_0 is the median of the raw fluorescence signal across the session. The response magnitude of a given neuron to a sound was defined as the across-trials average ΔF/F within 0–1.5 s from sound onset. The responsiveness of a given neuron to each stimulus was determined using a bootstrap analysis. Specifically, the difference between the sound response magnitude across trials and the mean prestimulus (prestim) response magnitude [mean ΔF/F during in the prestim windows (−1.5 to 0 s) of all sounds in the protocol] was compared with a distribution of similar differences resulting from 1000 random shuffles of the sound responses and prestim responses. The neuron was considered responsive to a given stimulus if the difference between the real sound response and mean prestim magnitude was >97.5% of the shuffled differences and if the sound response magnitude was at least 10% greater than the prestim magnitude. On a given day, a neuron was considered sound responsive if it was responsive to at least one stimulus on that day (with Bonferroni's correction for the number of stimuli).

To allow pooling changes in daily responses across neurons with different response magnitudes, the responses of each neuron to all stimuli across the 2 d of
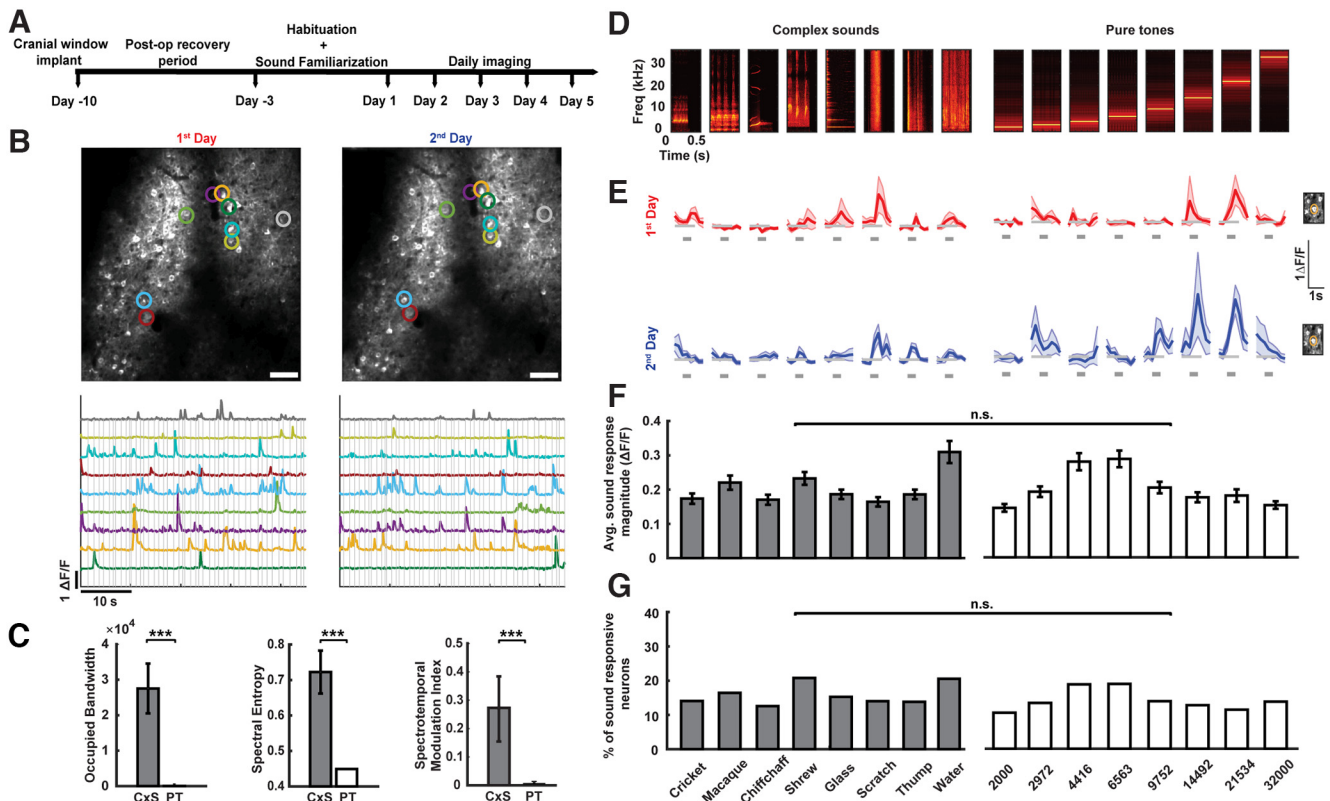
Figure 1. Imaging responses of identified L2/3 auditory cortical excitatory neurons to pure tones and complex sounds across days. **A**, An illustration of the experimental timeline. **B**, Top, Example two-photon micrographs from L2/3 of the auditory cortex in an awake mouse from 2 consecutive days. Scale bar, 60 μm. Colored circles identify example individual neurons matched across 2 consecutive days. Extended Data Figure 1-1 shows example histological verification of the imaging location. See Extended Data Figure 1-4 for validation of neuron matching across days. Bottom, ΔF/F traces of the neurons marked in the micrographs during a pure tone protocol. Gray lines indicate the stimuli. Calibration: 10 s, 1 ΔF/F. **C**, Comparison of spectrotemporal features of the CxS and PT. From left to right: occupied bandwidth (***$p = 0.00015$, Mann–Whitney U test), spectral entropy (***$p = 0.00015$ Mann–Whitney U test), and spectrotemporal modulation index (***$p = 0.00015$, Mann–Whitney U test). **D**, Spectrograms of the sound stimuli presented in the complex sounds (left) and pure tone (right) protocols. The identity of the complex sounds and the frequency of the pure tones are labeled at the bottom of **G**. **E**, Responses of a representative neuron to complex sounds and pure tones (the stimuli correspond to **D** column-wise) across 2 consecutive days (first day in red, second day in blue). Shaded area marks the mean ± SEM across trials. The gray bar below each response indicates the stimulus time (0.5 s). Calibration: 1 s, 1 ΔF/F. The cell body of the neuron as imaged during the 2 d is highlighted on the right. Extended Data Figure 1-2 validates that responses included in the study were not underestimated because of sampling rate, and Extended Data Figure 1-3 shows the distribution of the number of trials included in the data. **F**, Average sound response magnitude (mean ΔF/F across all trials over the stimulus window) across all days of imaging in response to complex sounds and pure tones (the stimuli correspond to **D**–**G** column-wise). Error bars indicate the mean ± SEM. Number of neurons across all pairs of consecutive days (with repetitions): CxSs, 557; PTs, 587; $p = 0.98$ (two-way ANOVA). **G**, Percentage of sound-responsive neurons recorded across all days of imaging in response to complex sounds and pure tones (the stimuli correspond to **D**–**G** column wise). Number of neurons across all pairs of consecutive days (with repetitions): CxSs, 557; PTs, 587, $p = 0.875$ (bootstrap test, see Materials and Methods; Table 1). Sound stimuli for each are indicated below: eight complex sounds and eight pure-tone frequencies (in hertz).

comparison were z scored before further analysis and statistical testing. For each comparison, a neuronal response to a given stimulus was included if the neuron was sound responsive on at least one of the days of comparison.

The significance of a change in response magnitude of a given neuron to a specific sound was quantified using a shuffle test. Specifically, the difference in mean response magnitudes between days was determined to be significant if the difference was >95% of the simulated differences generated from the random

shuffling of trials across the days of comparison (nShuffles = 1000) and in addition the magnitude of change was at least 10%. Using this method, we computed the significance of changes across four 1 d intervals (day 1 → day 2; day 2 → day 3; day 3 → day 4; day 4 → day 5), three 2 d intervals (day 1 → day 3; day 2 → 4; day 3 → day 5), two 3 d intervals (day 1 → day 4; day 2 → day 5), and one 4 d interval (day 1 → day 5).

The percentage of significant change in daily neuronal responses to a stimulus was calculated as follows:

Table 1: Summary of statistical analysis used in this study

Figure	Data structure	Type of test	Statistical data
1C	Non-normal	Mann–Whitney U test	First panel: rank sum = 36; $p = 0.00015$ Second panel: rank sum = 36; $p = 0.00015$ Third panel: rank sum = 100; $p = 0.00015$
1E	Normal distribution with two factors	Two-way ANOVA	F1: $F = 0$; df = 1; $p = 0.98$ F2: $F = 6.86$; df = 7; $p < 0.0001$
1F	Non-normal	Bootstrap test (describedabove)	10,000 randomly simulated proportions for one group (PT), given the probability of the other group (CxS). $p = 0.875$
2B,C	Non-normal	χ^2 test for proportions and confirmed by Fisher's exact test	Figure 2B: CXS = 12.15% (66 of 543); PTs = 22.01% (114 of 518); χ^2 statistic = 18.27; $p = 1.91e^{-5}$, Fisher's exact test: odds ratio = 49.03%; CI = 35.22%, 68.26% Figure 2C: CXS = 0.14 (34 of 241); PTs = 0.23 (59 of 256); χ^2 statistic = 6.52; $p = 0.011$ Fisher's exact test: odds ratio = 0.55; CI = 0.3445, 0.8730
2D	Normal distribution	Two-sided t test	$t = -2.0023$; $p = 0.046$; df = 495; SD = 0.51; CI = −0.1809, −0.0017
3B	Non-normal	z test for proportions	z value = −1.7811; $p = 0.037$
3C	Non-normal	Mann–Whitney U test	z value = −2.8857; $p = 0.0039$
4A,B	Non-normal	Bootstrap test (described above)	10,000 randomly simulated proportions for one group (PT), given the probability of the other group (CxS) computed for each sub category Figure 4A: $p = 0.0015$ Figure 4B: $p = 0.1084$
4C	Normal distribution with two factors	Two-way ANOVA with interaction	F1: $F = 9.52$; df = 3; $p < 0.001$ F2: $F = 5.89$; df = 1; $p = 0.015$ F1 * F2: $F = 0.1$; df = 3; $p = 0.961$

$$\frac{(\text{Number of sound responses that showed a significant change}) * 100}{\text{Total number of significant responses.}}$$

A neuron was determined to show significant change across days if it showed a significant change in response to at least one stimulus (after Bonferroni's correction for the number of stimuli).

For a given neuron, we computed the average Euclidean distance between its response profiles (magnitude of responses across stimuli) across pairs of days using the following equation:

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2},$$

where X_i equals the response of the neuron to stimulus i on the first day and y_i equals the response of the neuron to stimulus i on the second day.

To test for the stimulus specificity of response change, we tested whether the day of recording (1 or 2) significantly interacted with the stimulus identity in determining response magnitude using a two-way ANOVA with interaction. The ANOVA output was used to compute the effect size (ω^2) of the interaction term.

To test whether there is a significant difference between multiday or multistimuli proportions across CxSs and PTs (Fig. 1F; see also Fig. 4), we used a bootstrap analysis. Specifically, for each category across CxSs and PTs (see Fig. 4A, "1-day"), we derived a distribution of 10,000 randomly simulated PT proportions given the probability of the corresponding CxS category. The p-value was calculated as the fraction of "CxS-simulated" PT probabilities that were equal to or higher than the real PT probabilities across categories.

Statistical tests

We used statistical tests at a $p < 0.05$ significance level and $\alpha = 0.05$ for all comparisons unless otherwise indicated (Table 1).

Results

To quantify the degree of stability of auditory cortical representations of PTs and CxSs, we conducted two-photon calcium imaging of identified excitatory neuronal ensembles in L2/3 of the AC (Extended Data Fig. 1-1) in 10 awake head-fixed Thy1-GCaMP6f mice across days. As the degree of sound novelty influences response magnitude in AC (Ulanovsky et al., 2003, 2004; Nelken, 2014; Kato et al., 2015; Parras et al., 2017; Heilbron and Chait, 2018), we familiarized the mice to the experimental sound protocols for 3 consecutive days while being head fixed under the two-photon microscope before data acquisition commenced (Fig. 1A). During this habituation period, in each animal, three optical focal planes were chosen and registered with respect to the overlying blood vessel pattern to allow for repeated imaging of the same neurons across days (Fig. 1B).

From day 1 to day 5 of the experiment, we imaged the daily responses of identified neuronal ensembles to eight PTs of varying frequencies and eight CxSs. The CxSs consisted of animal vocalizations and environmental sounds that broadly overlapped in frequency content with the PTs, while having significantly higher frequency bandwidth, spectral entropy, and spectrotemporal modulation (Fig. 1C,D; see Materials and Methods). As expected, AC
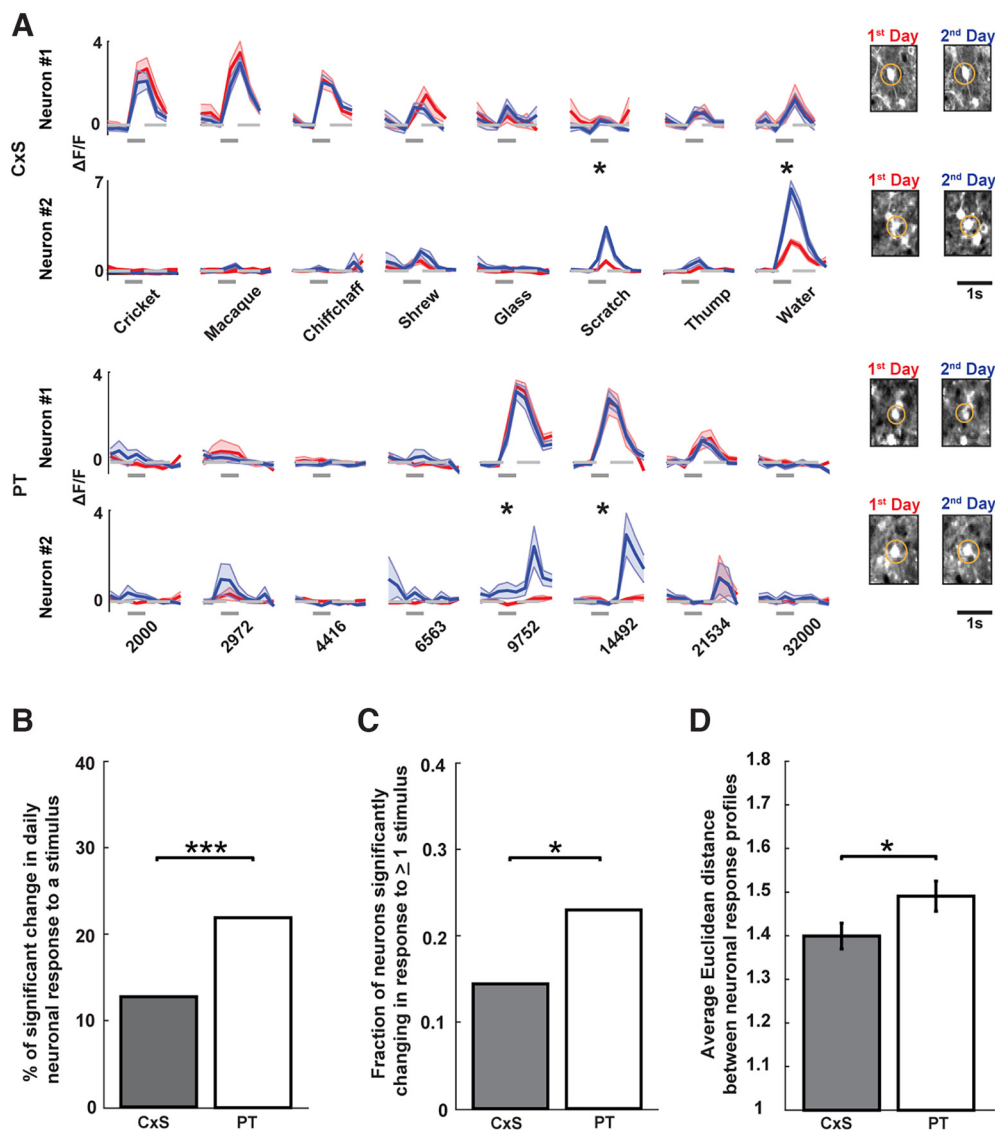
Figure 2. Auditory cortical responses to complex sounds are more stable than those to pure tones across days. ***A***, Responses of two representative neurons to complex sounds (rows 1 and 2) and pure tones (rows 3 and 4) across 2 consecutive days (first day in red, second day in blue). Shaded area marks the mean ± SEM across trials. Gray bars indicate the stimulus time. Stars indicate significant response changes (see Materials and Methods). Calibration: 1 s. The cell bodies of the imaged neurons are shown on the right. Rows 1 and 3 show neuronal responses that were stable from one day to the next, and rows 2 and 4 show neuronal responses that changed from one day to the next. ***B***, Percentage of significant changes in response to complex sounds and pure tones across pairs of consecutive days. CxSs: 12.15% (66 of 543); PTs: 22.01% (114 of 518); ***$p = 1.91 \times 10^{-5}$ (χ^2 test). ***C***, Fraction of neurons that show a significant change in response to at least one stimulus across pairs of consecutive days. CxSs: 0.14 (34 of 241); PTs: 0.23 (59 of 256); *$p = 0.011$ (χ^2 test). ***D***, Average Euclidean distance between the response profile of a neuron (to either CxSs or PTs) from one day to the next day; *$p = 0.046$ (two-sided t test). Extended Data Figure 2-1 shows the relationship between the changes in responsiveness of a neuron to CxS and PT.

neurons responded to both PTs and CxSs with sound-triggered transients in $\Delta F/F$ (Fig. 1E). We first compared the degree of overall sound-evoked responsiveness to CxSs and PTs across the population. We found that response magnitudes to PTs and CxSs were not significantly different (Fig. 1F) and that the rate of responsive neurons to PTs and CxSs were also not significantly different (Fig. 1G). Thus, our chosen set of PTs and CxSs evoked similar magnitudes and rates of responses among L2/3 AC excitatory neurons. Responsiveness, response magnitude, and trial-to-trial

consistency were not underestimated by our imaging sample rate (Extended Data Fig. 1-2).

We next quantified the degree of stability of these neuronal responses by comparing the responses of identified neurons across pairs of consecutive days. The identical variation in daily experimental and physiological conditions for PTs and CxSs allowed us to compare the relative degrees of change in responses between the two sound protocols. We observed that while most responses of individual AC neurons showed stability across days, some
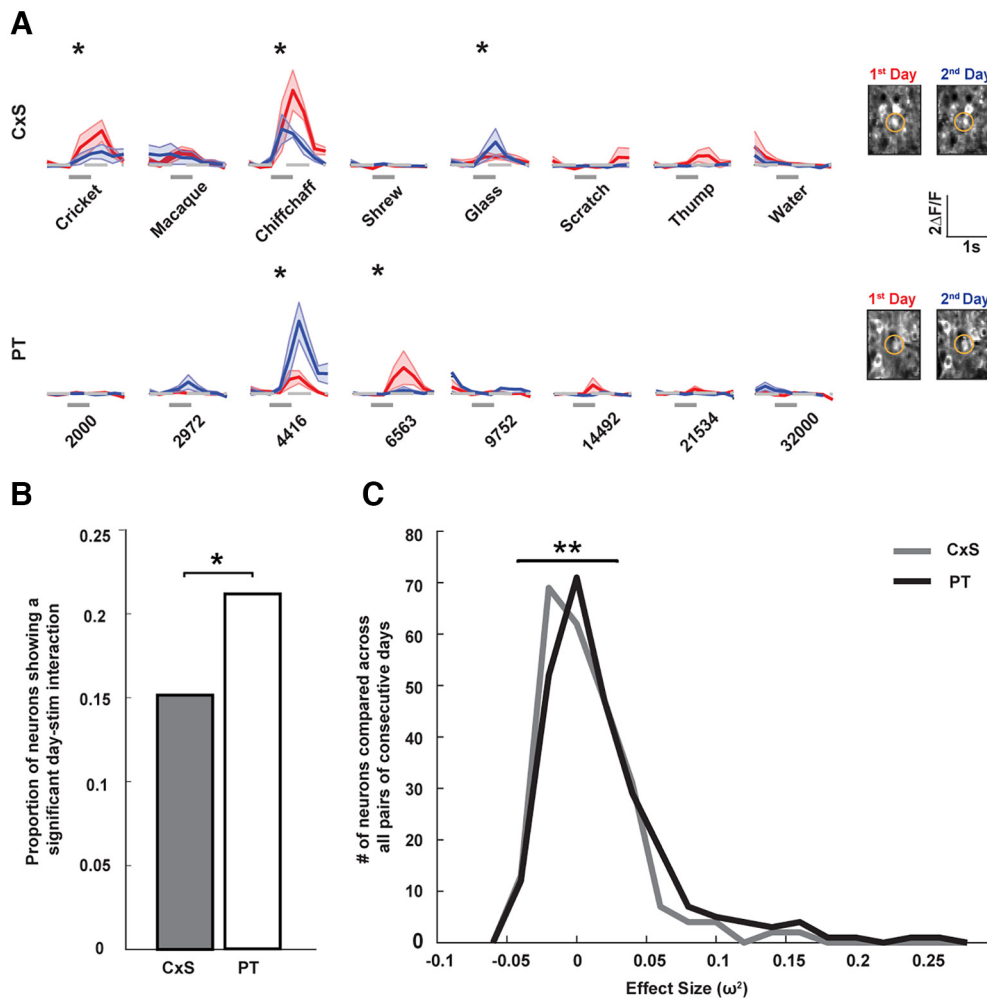
Figure 3. Daily plasticity in responses to pure tones is more stimulus-specific than to complex sounds. ***A***, Responses of two representative neurons to CxSs (row 1) and PTs (row 2) across 2 consecutive days, showing stimulus-specific changes. Shaded area marks the mean ± SEM across trials. Gray bars indicate stimulus timing. Stars indicate significant response changes. Calibration: 1 s, 2 $\Delta F/F$. The cell bodies of the imaged neurons are shown on the right. ***B***. Proportion of neurons across all pairs of consecutive days showing a significant day–stimulus interaction, computed via two-way ANOVA. CxSs: 14.9% of neurons (36 of 241); PTs: 21.1% of neurons (54 of 256); *$p = 0.037$ (z test for proportions). ***C***, Distributions of the effect size (ω^2) indicating the strength of the interaction between day and stimulus for CxSs (gray) and PTs (black) for each neuron across all consecutive days. **$p = 0.0039$ (Mann–Whitney U test).

displayed significant daily variation (Fig. 2A). To measure changes in sound responses, we first focused on the responses of individual neurons to individual stimuli across pairs of consecutive days and restricted our analyses to responses that were significant in at least one of the two days. Across this population, we found that while the majority of responses were stable across days, 22.01% (114 of 518) of significant responses to PTs showed a significant change in response magnitude across successive days (Fig. 2B). These results suggest that, underlying a generally stable representation, responses of AC neurons to PTs show a moderate degree of daily dynamics. Interestingly, however, only 12.15% (66 of 543) of significant responses to CxSs showed a significant change in magnitude across the same time interval (Fig. 2B). This proportion of daily response change to CxSs was significantly lower than that of PTs (Fig. 2B), suggesting that AC responses to CxSs

are more stable than responses to PTs across days. The degree of stability of CxSs with well-defined spectral centroids at <10 kHz (Cricket, Chiffchaff, and Macaque) did not significantly differ from those of more distributed spectra (Glass, Shrew, Thump, Scratch, and Water; 12.36% vs 12.05%, respectively; $p = 0.92$, χ^2 test for proportions).

As a complementary approach, we quantified a similar measure at the single-neuron level rather than the single-stimulus level. To this end, we calculated the fraction of sound-responsive neurons that exhibited a significant change in response magnitude to at least one of the eight PTs or CxSs for each pair of consecutive days. Consistent with our findings at the single-stimulus level, we found that the fraction of neurons showing a significant change in response to CxSs was significantly lower than that to PTs (Fig. 2C).
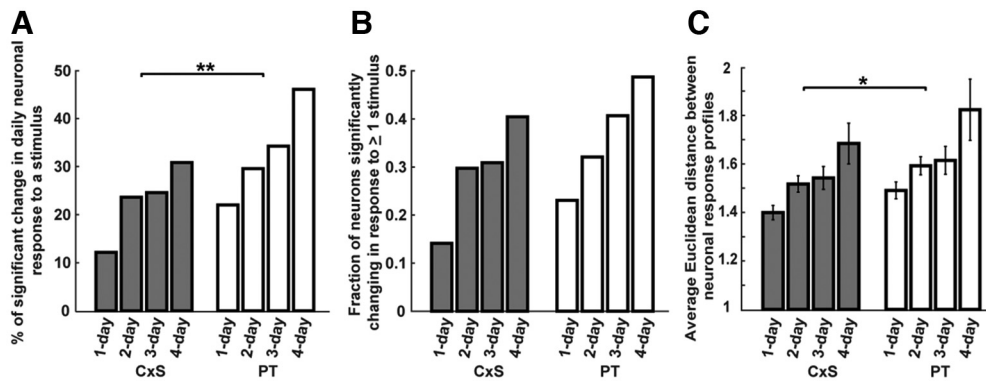
Figure 4. Auditory cortical responses to complex sounds are more stable than those to pure tones across multiple days. ***A***, Percentage of significant change in daily neuronal response to a given stimulus in the CxSs (gray bars) and PTs (white bars) protocol across varying daily intervals. **$p = 0.0015$ (bootstrap test; see Materials and Methods). ***B***, Fraction of neurons significantly changing in response to at least one stimulus in the protocol of CxSs (gray bars) and PTs (white bars) across varying daily intervals ($p = 0.1084$, bootstrap test; see Materials and Methods). ***C***, Average Euclidean distance between response profiles of a neuron for CxSs (gray bars) and PTs (white bars) across varying daily intervals. *$p = 0.015$ (two-way ANOVA).

To quantify the stability/plasticity of sound responses at the level of response profiles across stimuli, we computed for each neuron the Euclidean distance between its response profile (to either PTs or CxSs) on one day and that of the next day. A larger Euclidean distance reflected a higher degree of response change across stimuli. Consistent with the findings above, we found that the Euclidean distance between daily response profiles to PTs was significantly higher than those to CxSs (Fig. 2D). There was no significant correlation between the Euclidean distance of the same neurons to PTs and CxSs (Extended Data Fig. 2-1A) and changes in responses to CxSs were not significantly more strongly correlated with changes in frequency-overlapping PTs compared with frequency-nonoverlapping PTs (Extended Data Fig. 2-1B). Together, these findings across varying quantification methods indicate that AC neuronal responses to CxSs are more stable than those to PTs across consecutive days.

A change in the response profile of a neuron across days may include a change in response gain, manifesting as similar changes in response magnitude across stimuli, or it may be stimulus-specific, reflecting a change in the neuronal sound selectivity (Fig. 3A). To test whether changes in responses to PTs and CxSs differed in the nature of change, we compared the degree of stimulus specificity of response change for each of the stimuli classes. We tested for the responses of each neuron, whether there was a significant interaction between the day of recording and the different stimuli. A significant day–stimulus interaction indicates that responses to the different stimuli were differentially modulated across days, reflecting stimulus specificity in response change. We found that a significantly higher proportion of neurons showed stimulus specificity in daily changes in responsiveness to PTs compared with CxSs (Fig. 3B). Further, the strength of the day–stimulus interaction was significantly higher for PTs than for CxSs (Fig. 3C). These findings indicate that in addition to showing higher overall rates of daily change in responsiveness, the changes

in responses to PTs were more stimulus-specific, and therefore reflected a higher degree of change in sound selectivity, compared with CxSs.

Finally, we investigated how the rates of change across pairs of days relate to rates of change across longer durations. To this end, we quantified the changes in responsiveness in a similar manner across intervals of 1–4 d. We found that the degree of response plasticity increased with increasing time interval between days for both CxSs and PTs (Fig. 4A,B). Moreover, the elevated rates of change in responses to PTs compared with CxSs that were observed across pairs of days also manifested across these intervals (Fig. 4A). The fraction of neurons showing a significant change to at least one stimulus showed a similar trend, though it did not reach significance (Fig. 4B). Last, the Euclidean distance between the PT response profiles was significantly higher than that of CxSs across these intervals (Fig. 4C). Consistent with our previous results, this suggests that AC representations of CxSs are more stable compared with PTs over a range of daily time intervals.

Discussion

In this study, we used two-photon calcium imaging to record the degree of stability and plasticity of sound-evoked responses of L2/3 AC excitatory neurons to PTs and CxSs across days. We found that most responses to both PTs and CxSs were stable, with a moderate but significant degree of change across pairs of consecutive days. Importantly, we report that responses to CxSs exhibited significantly enhanced stability across days compared with PTs. Furthermore, the structure of response profiles to PTs exhibited larger degrees of change than to CxSs across days, as evidenced by a higher degree of stimulus-specific changes. Finally, we found that the enhanced degree of stability in CxS representations generalizes to longer daily time intervals.

Our findings of a significant degree of ongoing daily changes in auditory cortical representations of both CxSs

and PTs add to a number of recent studies describing "representational drift" in other sensory modalities (Peron et al., 2015; Ranson, 2017; Rule et al., 2019; Deitch et al., 2021; Pérez-Ortega et al., 2021; Schoonover et al., 2021). Together, these studies point to a potential common principle, by which, despite the well established link between perception and cortical function (Bergman, 1990; Chait et al., 2010; Chapuis and Wilson, 2011; Leopold, 2012; Frégnac and Bathellier, 2015; Kuchibhotla and Bathellier, 2018; Ceballo et al., 2019; Lee and Rothschild, 2021), a stable sensory perception does not rely on fixed cortical sensory representations. Instead, representational dynamics may reflect a general principle of cortical function. Indeed, the locally heterogeneous organization of AC L2/3 ensembles has been suggested to be well suited to support rapid synaptic reorganization in response to changing environmental conditions (Bandyopadhyay et al., 2010; Rothschild et al., 2010, 2013; Bathellier et al., 2012; Kanold et al., 2014; Kato et al., 2015; Rothschild and Mizrahi, 2015; Maor et al., 2016; Francis et al., 2018; Liu et al., 2019; Liu and Kanold, 2021). The question of whether sound representations in the thalamorecipient L4 are more stable than those in L2/3 remains for future studies.

While auditory cortical representations of both classes of sounds exhibited significant degrees of daily change, representations of CxSs were significantly more stable compared with PTs across quantification methods. These findings likely result from the differences in the acoustic properties of these stimuli. In particular, CxSs are decomposed into narrow frequency channels at the cochlea, and reconstructing their wideband frequency contents throughout the auditory pathway requires reintegration across frequency channels. In contrast, a pure tone evokes responses in a narrower channel throughout the auditory system. If daily variation in responses is at least partly independent in different frequency channels, integration across frequency bands as needed to represent CxSs may "average out" some of this variation compared with that of PTs. Thus, spectrotemporal integration may give rise to enhanced longitudinal stability of CxSs in the AC. Future studies could directly test this possibility by, for example, measuring the stability of representations of noise with systematically varying bandwidths. An alternative acoustic property that may determine the degree of AC stability is based on temporal rather than spectral integration. In particular, temporal modulations in the complex sounds may "reset" neuronal responses multiple times within a stimulus, such that the enhanced degree of overall stability is because of temporal averaging of per modulation fluctuations. This possibility could be tested using sequences of amplitude-modulated tones, which have temporal modulation without spectral bandwidth.

Beyond the higher degrees of change in responses to PTs compared with CxSs, we also found that PT response changes were more stimulus-specific than those of CxSs. These findings suggest that changes in responses to CxSs tended to be shaped more by global gain factors while changes in responses to PTs tended to reflect stimulus-tuning changes to a larger degree. If changes to

CxSs are correlated with changes to the tones that make up the CxSs, this finding may be influenced by the frequency overlap between CxSs, which is not the case for PTs. Although our finding that responses to CxSs do not significantly change as to their frequency-overlapping tones (Extended Data Fig. 2-1) argues against this possibility, the experiments described above could directly test it.

Beyond the acoustic differences between CxSs and PTs, a combination of evolution and previous experience may also have contributed to the enhanced stability of AC representations of CxSs compared with PTs. Future studies may test this hypothesis by comparing the degree of AC stability to sounds with similar spectrotemporal complexity but varying ethological relevance.

Our findings raise the question of whether enhanced stability of AC representations of CxSs are linked with the enhanced perceptual stability of these sounds. As the AC is important for sound perception in both humans (Kaga et al., 1997; Griffiths, 2012) and animal models (Ohl et al., 1999; Harrington et al., 2001; Rybalko et al., 2006; Frégnac and Bathellier, 2015; Kuchibhotla and Bathellier, 2018; Ceballo et al., 2019), it is tempting to speculate based on our findings that behavioral measures of perceptual stability, such as sound recognition across days, would be higher for CxSs compared with PTs. Testing this speculation may have important implications as PTs are not just widely used in auditory research but are also the standard in studies using classical conditioning and other learning paradigms.

References

Angeloni C, Geffen MN (2018) Contextual modulation of sound processing in the auditory cortex. Curr Opin Neurobiol 49:8–15.

Atencio CA, Sharpee TO, Schreiner CE (2008) Cooperative nonlinearities in auditory cortical neurons. Neuron 58:956–966.

Bandyopadhyay S, Shamma SA, Kanold PO (2010) Dichotomy of functional organization in the mouse auditory cortex. Nat Neurosci 13:361–368.

Barbour DL, Wang X (2003) Contrast tuning in auditory cortex. Science 299:1073–1075.

Bathellier B, Ushakova L, Rumpel S (2012) Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron 76:435–449.

Bergman AS (1990) Auditory scene analysis: the perceptual organization of sound. Cambridge, MA: MIT.

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

Blackwell JM, Taillefumier TO, Natan RG, Carruthers IM, Magnasco MO, Geffen MN (2016) Stable encoding of sounds over a broad range of statistical parameters in the auditory cortex. Eur J Neurosci 43:751–764.

Carruthers IM, Laplagne DA, Jaegle A, Briguglio JJ, Mwilambwe-Tshilobo L, Natan RG, Geffen MN (2015) Emergence of invariant representation of vocalizations in the auditory cortex. J Neurophysiol 114:2726–2740.

Ceballo S, Piwkowska Z, Bourg J, Daret A, Bathellier B (2019) Targeted cortical manipulation of auditory perception. Neuron 104:1168–1179.e5.

Chait M, de Cheveigné A, Poeppel D, Simon JZ (2010) Neural dynamics of attending and ignoring in human auditory cortex. Neuropsychologia 48:3262–3271.

Chapuis J, Wilson DA (2011) Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity. Nat Neurosci 15:155–161.

Chechik G, Nelken I (2012) Auditory abstraction from spectro-temporal features to coding auditory entities. Proc Natl Acad Sci U|S|A 109:18968–18973.

Dana H, Chen TW, Hu A, Shields BC, Guo C, Looger LL, Kim DS, Svoboda K (2014) Thy1-GCaMP6 transgenic mice for neuronal population imaging in vivo. PLoS One 9:e108697.

Deitch D, Rubin A, Ziv Y (2021) Representational drift in the mouse visual cortex. Curr Biol 31:4327–4339.e6.

Doupe AJ, Kuhl PK (1998) Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci 22:567–631.

Ehret G, Haack B (1982) Ultrasound recognition in house mice: key-stimulus configuration and recognition mechanism. J Comp Physiol 148:245–251.

Elhilali M, Fritz JB, Chi T-S, Shamma SA (2007) Auditory cortical receptive fields: stable entities with plastic abilities. J Neurosci 27:10372–10382.

Francis NA, Winkowski DE, Sheikhattar A, Armengol K, Babadi B, Kanold PO (2018) Small networks encode decision-making in primary auditory cortex. Neuron 97:885–897.e6.

Frégnac Y, Bathellier B (2015) Cortical correlates of low-level perception: from neural circuits to percepts. Neuron 88:110–126.

Griffiths TD (2012) The neural processing of complex sounds. In: The cognitive neuroscience of music (Peretz I, Zatorre RJ, eds), pp 168–180. Oxford, UK: Oxford UP.

Griffiths TD, Warren JD, Scott SK, Nelken I, King AJ (2004) Cortical processing of complex sound: a way forward? Trends Neurosci 27:181–185.

Guo W, Chambers AR, Darrow KN, Hancock KE, Shinn-Cunningham BG, Polley DB (2012) Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types. J Neurosci 32:9159–9172.

Gygi B, Kidd GR, Watson CS (2007) Similarity and categorization of environmental sounds. Percept Psychophys 69:839–855.

Harpaz M, Jankowski MM, Khouri L, Nelken I (2021) Emergence of abstract sound representations in the ascending auditory system. Prog Neurobiol 202:102049.

Harper NS, Schoppe O, Willmore BDB, Cui Z, Schnupp JWH, King AJ (2016) Network receptive field modeling reveals extensive integration and multi-feature selectivity in auditory cortical neurons. PLoS Comput Biol 12:e1005113.

Harrington IA, Heffner RS, Heffner HE (2001) An investigation of sensory deficits underlying the aphasia-like behavior of macaques with auditory cortex lesions. Neuroreport 12:1217–1221.

Heilbron M, Chait M (2018) Great expectations: is there evidence for predictive coding in auditory cortex? Neuroscience 389:54–73.

Kaga K, Shindo M, Tanaka Y (1997) Central auditory information processing in patients with bilateral auditory cortex lesions. Acta Otolaryngol 532:77–82.

Kanold PO, Nelken I, Polley DB (2014) Local versus global scales of organization in auditory cortex. Trends Neurosci 37:502–510.

Kato HK, Gillet SN, Isaacson JS (2015) Flexible sensory representations in auditory cortex driven by behavioral relevance. Neuron 88:1027–1039.

King AJ, Teki S, Willmore BDB (2018) Recent advances in understanding the auditory cortex. F1000Res 7:F1000.

Kuchibhotla K, Bathellier B (2018) Neural encoding of sensory and behavioral complexity in the auditory cortex. Curr Opin Neurobiol 52:65–71.

Lee J, Rothschild G (2021) Encoding of acquired sound-sequence salience by auditory cortical offset responses. Cell Rep 37:109927.

Leopold DA (2012) Primary visual cortex: awareness and blindsight. Annu Rev Neurosci 35:91–109.

Liu J, Kanold PO (2021) Diversity of receptive fields and sideband inhibition with complex thalamocortical and intracortical origin in L2/3 of mouse primary auditory cortex. J Neurosci 41:3142–3162.

Liu J, Whiteway MR, Sheikhattar A, Butts DA, Babadi B, Kanold PO (2019) Parallel processing of sound dynamics across mouse auditory cortex via spatially patterned thalamic inputs and distinct areal intracortical circuits. Cell Rep 27:872–885.e7.

Lütcke H, Margolis DJ, Helmchen F (2013) Steady or changing? Long-term monitoring of neuronal population activity. Trends Neurosci 36:375–384.

Maor I, Shalev A, Mizrahi A (2016) Distinct spatiotemporal response properties of excitatory versus inhibitory neurons in the mouse auditory cortex. Cereb Cortex 26:4242–4252.

Maor I, Shwartz-Ziv R, Feigin L, Elyada Y, Sompolinsky H, Mizrahi A (2020) Neural correlates of learning pure tones or natural sounds in the auditory cortex. Front Neural Circuits 13:82.

Meng X, Winkowski DE, Kao JPY, Kanold PO (2017) Sublaminar subdivision of mouse auditory cortex layer 2/3 based on functional translaminar connections. J Neurosci 37:10200–10214.

Merzenich MM, Kaas JH, Roth GL (1976) Auditory cortex in the grey squirrel: tonotopic organization and architectonic fields. J Comp Neurol 166:387–401.

Mizrahi A, Shalev A, Nelken I (2014) Single neuron and population coding of natural sounds in auditory cortex. Curr Opin Neurobiol 24:103–110.

Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. Curr Opin Neurobiol 14:474–480.

Nelken I (2008) Processing of complex sounds in the auditory system. Curr Opin Neurobiol 18:413–417.

Nelken I (2014) Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. Biol Cybern 108:655–663.

Nelken I, Bar-Yosef O (2008) Neurons and objects: the case of auditory cortex. Front Neurosci 2:107–113.

Nelken I, Rotman Y, Bar-Yosef O (1999) Responses of auditory-cortex neurons to structural features of natural sounds. Nature 397:154–157.

Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D (2003) Primary auditory cortex of cats: feature detection or something else? Biol Cybern 89:397–406.

Nelken I, Bizley J, Shamma SA, Shamma SA, Wang XQ, Wang XQ (2014) Auditory cortical processing in real-world listening: the auditory system going real. J Neurosci 34:15135–15138.

Ohl FW, Wetzel W, Wagner T, Rech A, Scheich H (1999) Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. Learn Mem 6:347–362.

Pachitariu M, Stringer C, Dipoppa M, Schröder S, Rossi LF, Dalgleish H, Carandini M, Harris K (2016) Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv. doi: 10.1101/061507.

Parras GG, Nieto-Diego J, Carbajal GV, Valdés-Baizabal C, Escera C, Malmierca MS (2017) Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. Nat Commun 8:2148.

Pérez-Ortega J, Alejandre-García T, Yuste R (2021) Long-term stability of cortical ensembles. Elife 10:e64449.

Peron SP, Freeman J, Iyer V, Guo C, Svoboda K (2015) A cellular resolution map of barrel cortex activity during tactile behavior. Neuron 86:783–799.

Ranson A (2017) Stability and plasticity of contextual modulation in the mouse visual cortex. Cell Rep 18:840–848.

Rauschecker JP (1998) Cortical processing of complex sounds. Curr Opin Neurobiol 8:516–521.

Rothschild G, Mizrahi A (2015) Global order and local disorder in brain maps. Annu Rev Neurosci 38:247–268.

Rothschild G, Nelken I, Mizrahi A (2010) Functional organization and population dynamics in the mouse primary auditory cortex. Nat Neurosci 13:353–360.

Rothschild G, Cohen L, Mizrahi A, Nelken I (2013) Elevated correlations in neuronal ensembles of mouse auditory cortex following parturition. J Neurosci 33:12851–12861.

Rule ME, O'Leary T, Harvey CD (2019) Causes and consequences of representational drift. Curr Opin Neurobiol 58:141–147.

Rybalko N, Šuta D, Nwabueze-Ogbo F, Syka J (2006) Effect of auditory cortex lesions on the discrimination of frequency-modulated tones in rats. Eur J Neurosci 23:1614–1622.

Sadagopan S, Wang X (2009) Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. J Neurosci 29:11192–11202.

Schoonover CE, Ohashi SN, Axel R, Fink AJP (2021) Representational drift in primary olfactory cortex. Nature 594:541–546.

Schreiner CE, Froemke RC, Atencio CA (2011) Spectral processing in auditory cortex. In: The auditory cortex (Winer JA, Schreiner CE, eds), pp 275–308. Boston: Springer.

Schwartz ZP, Buran BN, David SV (2020) Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. J Neurophysiol 123:191–208.

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

Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. Nat Neurosci 6:391–398.

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440–10453.

Wang X, Lu T, Snider RK, Liang L (2005) Sustained firing in auditory cortex evoked by preferred stimuli. Nature 435:341–346.