

Implicit Memory for Complex Sounds in Higher Auditory Cortex of the Ferret

Kai Lu,¹ Wanyi Liu,¹  Peng Zan,¹  Stephen V. David,²  Jonathan B. Fritz,¹ and  Shihab A. Shamma^{1,3}

¹Institute for Systems Research, University of Maryland, College Park, Maryland, 20740, ²Oregon Hearing Research Center, Oregon Health and Science University, Portland, Oregon, 97239, and ³Department of Cognitive Studies, Ecole Normale Supérieure, 75005 Paris, France

Responses of auditory cortical neurons encode sound features of incoming acoustic stimuli and also are shaped by stimulus context and history. Previous studies of mammalian auditory cortex have reported a variable time course for such contextual effects ranging from milliseconds to minutes. However, in secondary auditory forebrain areas of songbirds, long-term stimulus-specific neuronal habituation to acoustic stimuli can persist for much longer periods of time, ranging from hours to days. Such long-term habituation in the songbird is a form of long-term auditory memory that requires gene expression. Although such long-term habituation has been demonstrated in avian auditory forebrain, this phenomenon has not previously been described in the mammalian auditory system. Utilizing a similar version of the avian habituation paradigm, we explored whether such long-term effects of stimulus history also occur in auditory cortex of a mammalian auditory generalist, the ferret. Following repetitive presentation of novel complex sounds, we observed significant response habituation in secondary auditory cortex, but not in primary auditory cortex. This long-term habituation appeared to be independent for each novel stimulus and often lasted for at least 20 min. These effects could not be explained by simple neuronal fatigue in the auditory pathway, because time-reversed sounds induced undiminished responses similar to those elicited by completely novel sounds. A parallel set of pupillometric response measurements in the ferret revealed long-term habituation effects similar to observed long-term neural habituation, supporting the hypothesis that habituation to passively presented stimuli is correlated with implicit learning and long-term recognition of familiar sounds.

Key words: auditory cortex; habituation; long-term memory

Significance Statement

Long-term habituation in higher areas of songbird auditory forebrain is associated with gene expression and is correlated with recognition memory. Similar long-term auditory habituation in mammals has not been previously described. We studied such habituation in single neurons in the auditory cortex of awake ferrets that were passively listening to repeated presentations of various complex sounds. Responses exhibited long-lasting habituation (at least 20 min) in the secondary, but not primary auditory cortex. Habituation ceased when stimuli were played backward, despite having identical spectral content to the original sound. This long-term neural habituation correlated with similar habituation of ferret pupillary responses to repeated presentations of the same stimuli, suggesting that stimulus habituation is retained as a long-term behavioral memory.

Introduction

Neural activity in the mammalian primary auditory cortex encodes not only the acoustic features of sound but also stimulus

history and context (Ulanovsky et al., 2003, 2004; Nelken, 2014; Nieto-Diego and Malmierca, 2016; Parras et al., 2017). Most contextual effects in the mammalian auditory system typically exhibit a fairly brief time-span, from 100s of milliseconds to 10s of seconds (Brosch and Schreiner, 2000; Ulanovsky et al., 2003, 2004; Bartlett and Wang, 2005). By contrast, a form of stimulus-specific habituation with a much longer time-span of hours, even days, has been described in the secondary auditory forebrain [caudomedial neostriatum (NCM)] of songbirds (Chew et al., 1995, 1996; Phan et al., 2006), raising questions as to whether there are comparable forms of habituation in the mammalian auditory system. Although such stimulus long-term habituation was observed in NCM for conspecific songs, it was not found in

Received Aug. 16, 2018; revised Sept. 14, 2018; accepted Sept. 19, 2018.

Author contributions: S.A.S. edited the paper; K.L. designed research; K.L. and W.L. performed research; P.Z. and S.V.D. contributed unpublished reagents/analytic tools; K.L., J.B.F., and S.A.S. analyzed data; K.L., J.B.F., and S.A.S. wrote the paper.

This work was supported by Grants from the National Institutes of Health (R01 DC005779) and an Advanced ERC Grant (ADAM) to S.A.S.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Shihab A. Shamma, 2202 A.V. Williams Building, College Park, MD 20740. E-mail: sas@isr.umd.edu.

<https://doi.org/10.1523/JNEUROSCI.2118-18.2018>

Copyright © 2018 the authors 0270-6474/18/389955-12\$15.00/0

the primary auditory area of the songbird (Terleph et al., 2006). Avian long-term habituation has been shown to require RNA synthesis, and is associated with expression of the immediate early gene (*ZENK*), which is also necessary for long-term potentiation (Chew et al., 1995; Ribeiro et al., 2002; Phan et al., 2006). Such long-term habituation to conspecific songs is also known to correlate with behavioral responsiveness (Vicario, 2004), suggesting that it is a form of long-term auditory memory in songbirds that may be specialized for recalling the calls and songs of conspecifics.

Some reports have suggested that there may be a comparable long-term implicit auditory memory system in mammals. For example, monkeys implicitly learn acoustic presentations of simple artificial grammars (AGs) after an exposure of 20–30 min and show memory for the AGs for at least 30 min to 1 d afterward (Fitch and Hauser, 2004; Wilson et al., 2013). Pre-verbal infants can also process and recall different AGs and show implicit long-term statistical learning of words, auditory patterns, and musical sequences for over 2 weeks (Saffran et al., 1996, 1999; Jusczyk and Hohne, 1997). Adult humans show an extraordinarily robust and long-lasting implicit auditory memory that can develop rapidly even for random noise or time patterns (Agus et al., 2010; Kang et al., 2017). However, the existence of long-term habituation effects in the mammalian auditory system comparable to those observed in songbirds has remained unexplored at the single-neuronal level. The experiments described here seek to determine whether long-term stimulus habituation of neural responses also exists in ferret auditory cortex. We also wondered whether long-term stimulus habituation could be demonstrated in the ferret, and if so, whether such effects are only found in higher auditory cortical fields (as in songbirds), or are also present in the primary auditory cortex (A1). To answer these questions, we recorded responses in ferret A1, and also in two higher auditory cortical fields in the dorsal posterior ectosylvian gyrus (dPEG) of the ferret (Bizley et al., 2015) adjacent to A1. These two fields, posterior pseudosylvian field (PPF) and posterior suprasylvian field (PSF), comprise an area in dPEG that has been shown to exhibit enhanced response plasticity and selectivity during behavior (Atiani et al., 2014). We presented a range of complex sounds, including speech, animal vocalizations, and music. Stimuli were repeated in blocks, and then the habituated responses were tested for their strength and persistence (for up to 20 min). A parallel set of pupillometric response measurements were also conducted in the awake ferret to determine whether long-term behavioral habituation effects are similar to the neuronal ones, testing the hypothesis that habituation to passively presented stimuli may be correlated with long-term recognition of familiar sounds.

Materials and Methods

Subjects. Adult female ferrets, *Mustela putorius* ($n = 6$), were used for this study. Three ferrets were used for electrophysiological recordings from auditory cortex and three additional ferrets were used for pupillometry measurement. All ferrets used in this study had been previously trained on auditory streaming instrumental tasks (Lu et al., 2017) that were unrelated to this implicit habituation procedure. The ferrets were housed on a 12 h light/dark cycle. They were placed on a water-control protocol in which they received their daily liquid during experimental sessions on 2 d per week and obtained *ad libitum* water on the other 5 d per week. All experimental procedures were in accordance with National Institutes of Health policy on animal care and use and conformed to a protocol approved by the Institutional Animal Care and Use Committee of University of Maryland.

Surgeries. A head-post was surgically implanted to stabilize the head for neurophysiological and pupillary recordings. Animals were anesthe-

tized with isoflurane (1.5–2% in oxygen) and a customized stainless-steel head post was surgically implanted on the ferret skull under aseptic conditions. The skull covering the auditory cortex was exposed and covered by a thin layer of Heraeus Kulzer Charisma (1 mm thick) and surrounded by a thicker wall built with Charisma (3 mm thick). After recovery from surgery, the animals were habituated to head restraint in a customized holder, which fixed the head post stably in place and restrained the body in a plastic tube. Two days before electrophysiological recording, a small craniotomy (1–3 mm diameter) was made above auditory cortex for electrophysiological recordings. At the beginning and end of each recording session, the craniotomy was thoroughly rinsed with sterile saline. At the end of a recording session, the craniotomy and well was filled with topical antibiotics (in saline solution) that were rotated on a weekly basis (either Baytril or cefazolin). After drying with sterile eye spears, the well was filled with silicone (sterile vinyl polysiloxane impression material, Examix NDS), which provided a tight, protective seal for the well. The silicone plug could be removed easily at the next recording session. Necessary steps were taken to maintain sterility in the craniotomy region during all procedures. After recordings were completed in the original craniotomy, it was gradually enlarged by removing adjacent small adjacent bone sections of the skull over successive months of recording to eventually provide access for electrode recordings to the entire A1 and dPEG. The procedure for enlarging the craniotomy was identical to that followed for drilling the original craniotomy.

Electrophysiological recording. Electrophysiological recordings were made in a walk-in double-walled soundproof booth (IAC). The awake animal was immobilized in a comfortable tube and the implanted head post was clamped to fix the head in place relative to a stereotaxic frame. Recordings were conducted in primary A1 and in dPEG of both left and right hemispheres. In each recording, 4–8 tungsten microelectrodes (2–3 M Ω , FHC) were introduced through the craniotomies and controlled by independently moveable drives (Electrode Positioning System, Alpha-Omega). Raw neural activity traces were amplified, filtered, and digitally acquired by a data acquisition system (AlphaLab, Alpha-Omega). Multiunit neuronal responses were monitored online (including all spikes that rose above a threshold level of 3.5 SD of baseline noise). Single units were isolated online by searching for single-unit activity. Bandpass noise (0.2 s duration, 1 octave bandwidth) and pure tone (0.2 s duration) stimuli were presented to search for responsive sites. After recordings, single units were isolated again by off-line customized spike-sorting software, which was based on a PCA and template-matching algorithm (Meska-PCA, NSL). As auditory responses were found, the frequency tuning of neurons were determined by pure tone and temporally orthogonal ripple combinations (Depireux et al., 2001). Following characterization of receptive field properties and frequency tuning, stimulus playback experiments were performed.

Auditory field localization. The approximate location of A1 was initially determined by stereotaxic coordinates and then refined with neurophysiological mapping of the tonotopic map in A1 (Bizley et al., 2005). The posterior border of A1 (along the suprasylvian sulcus) was marked by the presence of visual responses from neighboring visual cortex. The location of dPEG was determined neurophysiologically by tonotopically mapping A1 and dPEG, which share a common low-frequency border (Fig. 1; Atiani et al., 2014; Bizley et al., 2005, 2015). The tonotopic gradient in A1 goes from high- to low-frequency along a rostrally tilted dorsoventral axis until the gradient reversed and best frequency increased again in the mirror frequency maps in the two areas that comprise the dPEG-PPF and PSF. In addition to the A1/dPEG tonotopic gradient reversal, transition to dPEG was also marked by somewhat longer latencies, greater sustained responses and weaker envelope phase-locking than in A1 (Atiani et al., 2014; Bizley et al., 2005). The ventral boundary of the dPEG was determined by abrupt transition from high-frequency tuning to broad low-frequency tuning (Atiani et al., 2014; Bizley et al., 2005, 2015).

Auditory stimuli and experimental design. Acoustic stimuli were comprised of 73 short samples of animal vocalizations, speech, vocal and instrumental music, and sampled at 40 kHz. Animal vocalizations and human speech were 2–4 s long (Fig. 2A). Instrumental music excerpts were 3–5 s long. All stimulus amplitudes were presented at 65 dB SPL

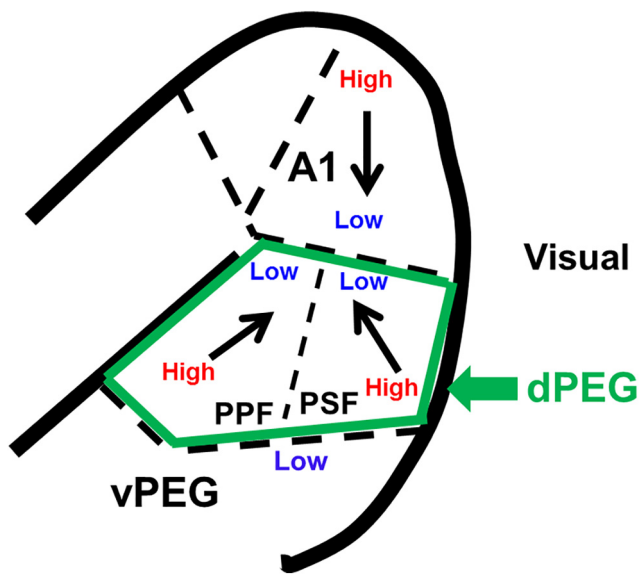


Figure 1. A. Schematic of location and tonotopic gradients in primary A1 and the two cortical fields in dPEG (PPF and PSF) in ferrets. Bold black trace indicates the border of the ectosylvian gyrus. A1 displays a clear tonotopic gradient, with best frequencies changing from high to low along the dorsoventral axis. The two cortical fields in dPEG share a common low-frequency border with A1 (slightly tilted horizontal dashed line). They have a reversed tuning-gradient from low to high frequencies (with the dorsoventral axis tilted anteroventral for PPF and posteroventral for PSF, arrows). The transition from dPEG to ventral PEG (vPEG) is marked by an abrupt transition to low-frequency tuning and longer response latencies (horizontal dashed line indicates border).

from a speaker placed 1 m in front of the animal in a large, walk-in double-walled soundproof booth (IAC).

The habituation paradigm used to habituate the animal to a specific complex sound sequence was very similar to the paradigm used in previous avian studies (Chew et al., 1995, 1996). It consisted of three phases, with three blocks in each phase (Fig. 2B). In Phase 1, three distinct stimuli, novel to the animal, were presented for 50 repetitions each, in three sequential blocks, with a 6 s interstimulus gap between stimuli (the same timing was used in the following 2 phases as well). Each block of 50 repetitions lasted ~10 min, and thus the entire Phase 1 lasted ~30 min. In Phase 2, familiarity to the same three original stimuli from Phase 1 was tested by presenting 25 repetitions in each block in the same order, so that each now-familiar stimulus was presented ~20 min after its last presentation in Phase 1. In Phase 3, the same three stimuli were reversed temporally to generate sounds with novel temporal features but minimally changed spectral characteristics. Reversed stimuli were presented in the same order as the three original stimuli from Phase 1, with 25 repetitions in each block.

It should be noted that we did not attempt to habituate cells with simple tonal stimuli because these stimuli were routinely used to search for units and characterize their tuning curves, so that pure tones were not novel for the animals. Furthermore, the time window available for retesting all neurons for persistence of memory effects was limited to the minimal time we could hold cells reliably, which was at least 1 h, hence the combined duration of the three Phases.

Data analysis. The amplitude of auditory responses was defined as the average spike-rate in the response window (from stimulus onset to 250 ms after stimulus offset). The baseline activity was quantified as the average spike-rate during the period of 1 s before stimulus onset. Any recording with significant change in baseline activity within or between habituation phases was excluded from further analysis (criterion within phases: Spearman correlation $p < 0.01$; between phases: Kolmogorov–Smirnov two-sample test, $p < 0.01$). Stability of single units throughout each recording was confirmed by stable spike waveforms. If there were changes in spike waveform during the recording, the neurons were not used for further analysis.

Quantification of habituation effect. To quantify habituation of auditory responses in Phase 1, the response rate for each single unit was quantified on a trial-by-trial basis (see Fig. 4C). Responses from trial 2 to 17 in each block (i.e., 1 stimulus type) in Phase 1 were averaged. These 16 trials (i.e., stimulus presentations) constituted the initial ~one-third of the stimulus presentations in the block, whose mean was defined as the starting-response for a given novel stimulus. Following paradigms used previously in the songbird literature (Phan et al., 2006; Phan and Vicario, 2010), the first stimulus was excluded from the analysis as it usually had a relatively strong and variable response that quickly adapted (in the Discussion, we will separately discuss and summarize responses to the first stimulus in the block). Similarly, response rates from the last 16 trials in each block were averaged and defined as the *end-response*. For each recorded neuron, three types of stimuli were tested. Both the starting- and end-responses were averaged across all these stimuli for a given neuron, so that the response amplitude change (i.e., the difference between starting- and end-response firing rates) was measured as a characteristic property of each neuron for each novel stimulus.

Dynamics of the habituation process were quantified by linear regression (following closely procedures used by Chew et al., 1995; Phan et al., 2006; Phan and Vicario, 2010) across all trials as follows: (1) The response rate of each trial in each block was divided by the average response amplitude at the beginning of each block. (2) A linear regression between trial number and normalized response amplitudes was performed for all trials (excluding the first trial) and the slope of the linear regression was calculated. (3) Regression slopes from the three tested stimuli (blocks) in each phase were averaged.

Regression slopes were compared across the three test phases and also compared between A1 and dPEG. For the cross-phase comparison, regression slopes for Phase 1 were only calculated from the first half of the trials, so that equal number of trials (25, because 25 repetitions were presented in Phases 2 and 3) were measured in all of the three phases (following procedure used by Phan et al., 2006; Velho et al., 2012). To compare regression slopes across different stimulus classes, all stimuli were assigned to one of three classes of auditory stimuli: (A) Animal vocalizations, (B) musical excerpts and vocal songs, and (C) human speech. Regression slopes within each stimulus class were pooled and compared.

Calculation of mutual information. We analyzed changes in mutual information (MI) carried by spike trains during habituation to novel stimuli in Phase 1. MI was calculated as the difference between the entropy associated with all responses of a neuron and the entropy of responses to a given stimulus, following standard procedures (Nelken and Chechik, 2007). The first 25 trials of the three novel stimuli in test Phase 1 were chosen for calculating MI for novel stimuli. MI was calculated from randomly selected subsets of data of three sizes [~50% (12), 80% (20), and 100% (25) of the first 25 trials]. Then a linear regression between MIs and the reciprocal of the sample size in each dataset was calculated. The intercept of the linear regression provided an estimate of the unbiased MI. Because the acoustic stimuli had different durations, only the first 800 ms was used for the MI calculations, with 80 ms as the bin size. This limited duration was used so as to use fewer trials for the MI calculation and capture its dynamics. Following the same procedure, MI for the familiar stimuli was calculated from the last 25 trials of the same stimuli in the block. In other words, for each neuron, we computed one MI value for the first 25 trials in Phase 1 and one for the last 25 trials in Phase 1. The difference between MI (novel) and MI (familiar) corresponds to the MI change (Δ MI). We averaged Δ MI over all neurons.

Pupillary measurements. We tested pupillary size changes during long-term stimulus habituation. Three naive ferrets were tested with the same paradigm, where animals were head-fixed, passively listening to sound presentations in three test phases. Pupillary changes were recorded at 3 Hz using a DCC1545M camera (Thorlabs; TML-HP 1_ Telecentric lens, Edmund Optics) with regular illumination. Recorded images were analyzed and pupillary size was measured offline. Baseline pupillary size was calculated by averaging measurements in 2 s windows before stimulus onset. Pupillary responses to stimuli were measured in 5 s windows after

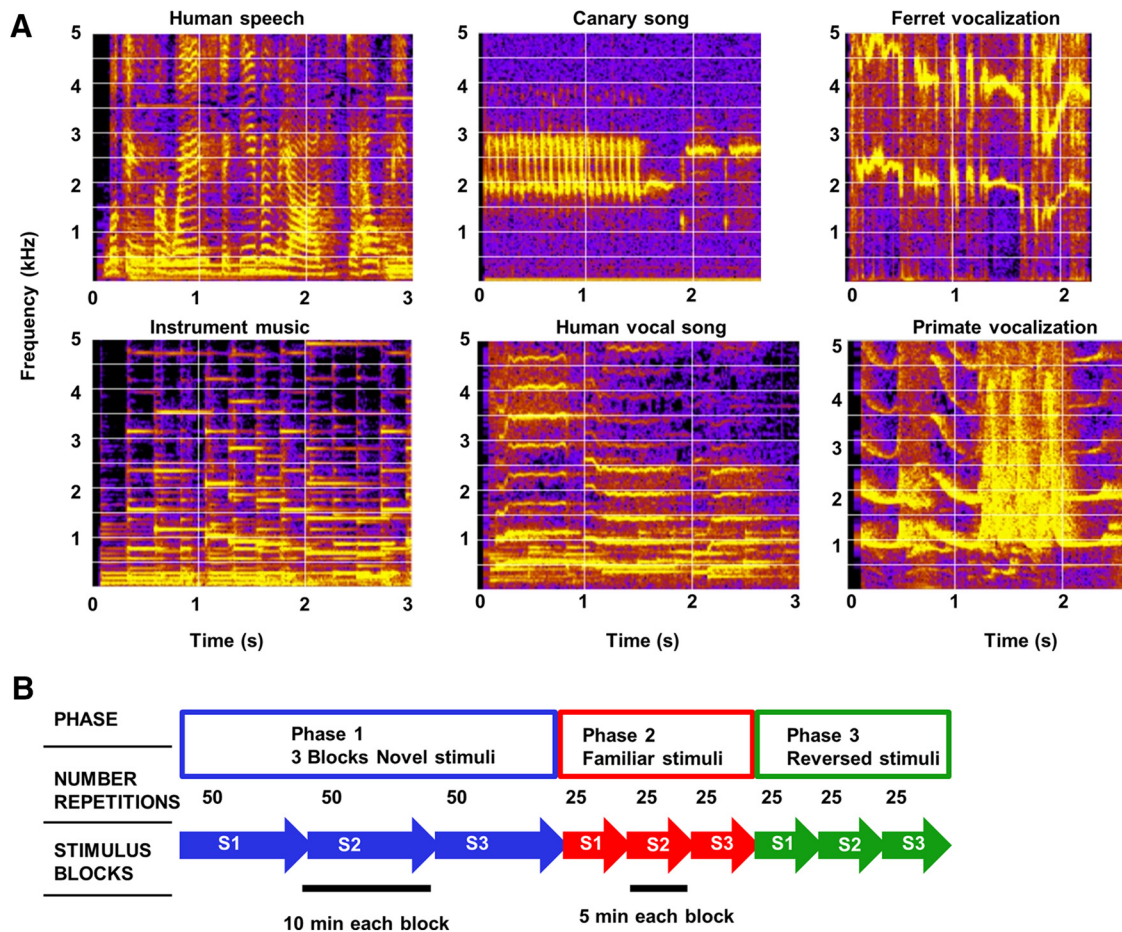


Figure 2. *A*, Spectrograms of six diverse broadband stimuli used in the habituation study. *B*, Schematic illustration of how stimuli were presented. Blocks of three Novel stimuli (blue arrows) were presented in Phase 1. In each block, there were 50 repetitions of one of the three novel stimuli. The now Familiar stimuli (red arrows) are then presented in blocks of 25 repetitions in Phase 2. The Reversed stimuli (green arrows) are then presented in blocks of 25 repetitions in Phase 3.

stimulus onset. Then pupillary size measured in the response window was adjusted by subtracting the baseline pupillary measurement in each trial. To compare the pupillary responses to novel stimuli and familiar stimuli, pupillary responses, as a function of time, were calculated in the same way as analysis of post-stimulus time histogram (PSTH) differences (except that more repetitions were used in the analysis because of the much lower sampling rate of pupillary recording compared with electrophysiology). Pupillary responses to the first 15 trials of novel sounds were pooled together and compared with pupillary responses to the last 15 trials of familiar sounds.

The latency of pupillary responses was determined in the same way as the analysis of temporal profile of neural habituation. Habituation curves of pupillary responses were estimated from trial-by-trial pupillary measurements as in the analysis of habituation curves of the neural data.

Statistical methods. Results were plotted as histograms or cumulative frequency distributions (which reveal details of multiple distributions) and also as conventional box-and-whisker plots. Because sample distributions in some tests did not satisfy criteria for parametric tests, non-parametric statistics were used. Changes in response rates over repetitive stimulation and slopes of linear regression from the neuron population were tested by the Wilcoxon matched pairs test (the nonparametric version of one-sample *t* test). Differences between test phases were similarly tested. Friedman ANOVA tests were used to test multiple matched samples. For non-matched samples, differences between the two brain regions (A1 and dPEG) were tested by the Kolmogorov–Smirnov two-sample test (the nonparametric version of two-sample *t* test). Median tests (similar to one-factor ANOVA) were applied to multiple non-matched samples.

Results

Basic response properties were measured in 83 single-units in dPEG (pooled data from PPF and PSF), and 81 single units in A1, in the auditory cortices of three ferrets. A1 neurons showed clear, narrowly tuned receptive fields (Fig. 3*A,B*, left column), compared with the more broadly-tuned receptive fields of dPEG neurons (Fig. 3*A,B*, right column), consistent with earlier reports (Atiani et al., 2014). Distribution of bandwidths for A1 and dPEG receptive field filters, as shown in Figure 3*C*, were significantly different (A1: median = 0.8, dPEG: median = 1.4; Kolmogorov–Smirnov two-sample test, $p = 0.023$). We note that we grouped the neurons in PPF and PSF together as one set in dPEG, because they showed similar properties in our experimental habituation paradigm.

Habituation to repeated acoustic stimuli

We measured the average responses to repeated presentations of novel stimuli in Phase 1, as illustrated for one dPEG neuron by the raster responses in Figure 4*A*. Both the temporal pattern of responses and spike waveforms remained consistent throughout the recording (Fig. 4*A,B*). Figure 4*C* showed spike-rate plots for an A1 neuron and a dPEG neuron. A typical pattern of response changes demonstrated in these plots is the habituation of responses of dPEG cells (but not of A1) as the number of repetitions increases. This was quantified by the difference between the av-

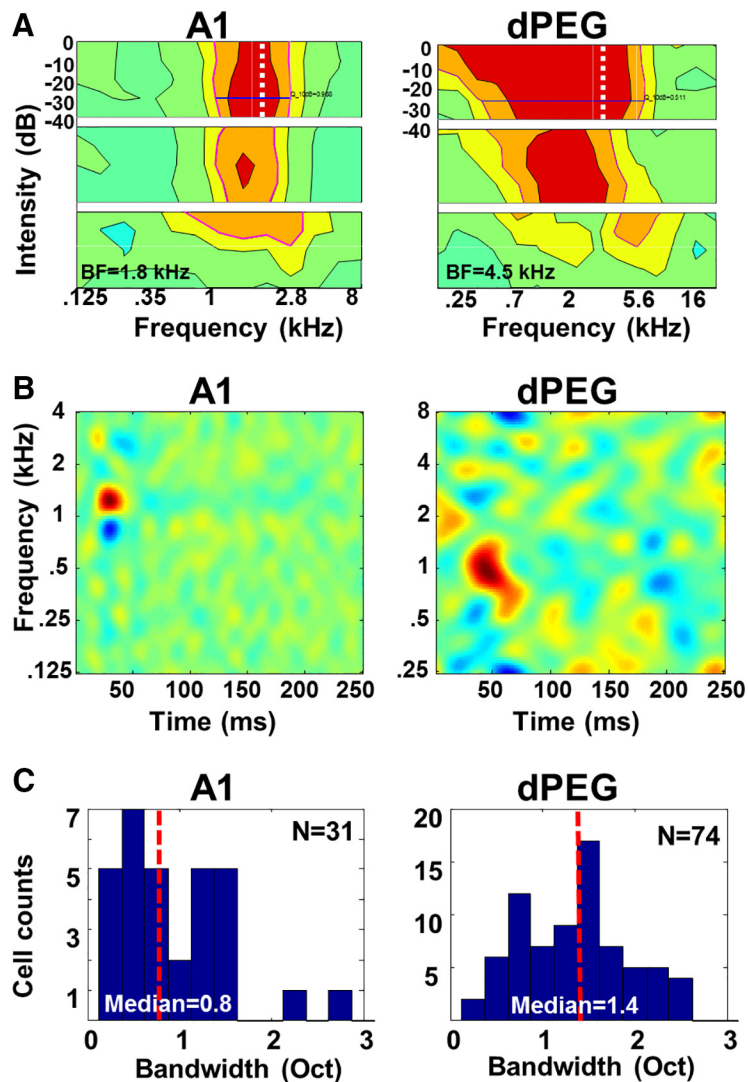


Figure 3. *A*, Examples of multi-intensity tuning curves of an A1 neuron (left) and a dPEG neuron (right). Tuning curves were displayed on the same color scale. Tuning curves shown in the three panels from top-to-bottom reflect, respectively, onset responses (0–100 ms after stimulus onset), sustained responses (100–200 ms after stimulus onset), and off-responses (0–50 ms after stimulus offset). *B*, Examples of spectro-temporal receptive fields (STRFs) of an A1 neuron (left) and a dPEG neuron (right). *C*, Histograms of tuning-widths of A1 (left) and dPEG (right) neurons. The red vertical dashed line indicates the median of tuning-width in each distribution in A1 and dPEG. Bandwidths of dPEG neurons are wider than in A1. The values are indicated at the bottom. A1 has fewer neurons ($N = 31$) and the mean for A1 was skewed by a few outliers.

erage response of the initial and last set of 16 trials, indicated by the green bars in the example units in Figure 4C. Overall, response-rates in dPEG decreased significantly (Wilcoxon test: $z = -4.77$, $p = 0.001$; Fig. 4D, right histogram). By comparison, they only slightly, and not significantly, decreased in A1 neurons (Wilcoxon test: $z = -1.68$, $p = 0.094$; Fig. 4D, left histogram), indicating that the degree of response habituation was quite different between the primary and secondary auditory cortical areas (Kolmogorov–Smirnov two-sample test, $p = 0.031$). We then quantified the trial-by-trial response change over repetitions using linear regression (Fig. 4C, red lines). Significant negative regression slopes were found only in dPEG (Wilcoxon test: $z = -4.23$, $p = 0.001$; Fig. 4E, right histogram). Slopes in A1 neurons, by contrast, were only slightly biased to the negative side (Wilcoxon test: $z = -0.76$, $p = 0.447$; Fig. 4E, left histogram). Differences in regression slopes between the two areas were significant (Kolmogorov–Smirnov two-sample test, $p = 0.041$) indicating

that far more of the dPEG neurons habituated to repetitive stimulation of novel sounds than A1 neurons.

Habituation was independent for each stimulus

Because both the frequency tuning of dPEG neurons and the overall average spectrum of the stimuli were both broad, dPEG neurons usually responded to all three types of stimuli (as shown in the example in right panel of Fig. 4C) in the first test Phase 1. If habituation were due to fatigue, presentation of stimuli in the first block would have interfered with responses to stimuli in the other two blocks in Phase 1. In particular, if habituation were due to fatigue, then stimuli in the second and third blocks in Phase 1 would have evoked weaker responses at the start of each block and should also have shown flatter regression slopes across each block. Instead, we found that habituation in dPEG neurons proceeded independently for each of the presented stimuli. As shown in Figure 4F, we did not find significant differences in initial response rate across the three blocks in Phase 1 (Friedman ANOVA: $\chi^2 = 0.521$, $p = 0.771$), so spike rate was not reduced across sequential blocks. There was also no significant difference in regression slopes across three sequential blocks (Fig. 4G; Friedman ANOVA: $\chi^2 = 3.479$, $p = 0.176$). Therefore, habituation to one stimulus did not interfere with either the spike rate or regression slopes of other incoming novel stimuli. Thus, habituation was stimulus independent for the diverse set of these complex stimuli.

Habituation was long lasting

Next, we compared habituation to the same stimulus in the three test Phases: as originally novel (Phase 1), as familiar (Phase 2) and as a reversed stimulus (Phase 3). Figure 5A illustrates the response habituation in two neurons in dPEG to the same stimulus in the three phases. In Phase 1 (blue dots), responses showed reduction with repetition. In the Phase 2 (red dots), the same stimuli were presented 20 min later, and responses remained habituated at the level measured at the end of the first phase. In addition, regression slopes of responses to these stimuli were much flatter than those for novel stimuli. This long-lasting habituation was clearly not due to fatigue of the stimulated neurons, because when the same stimuli were time-reversed and presented again (Phase 3, green dots), response rates dramatically increased initially, with regression slopes almost comparable with those of novel stimuli.

To quantify these effects at the population level, we compared regression slopes for novel, familiar, and time-reversed stimuli. Because familiar and reversed stimuli were presented over 25 repetitions, only the first 25 trials of the novel stimuli were used for comparison to ensure a matched sample (following the same

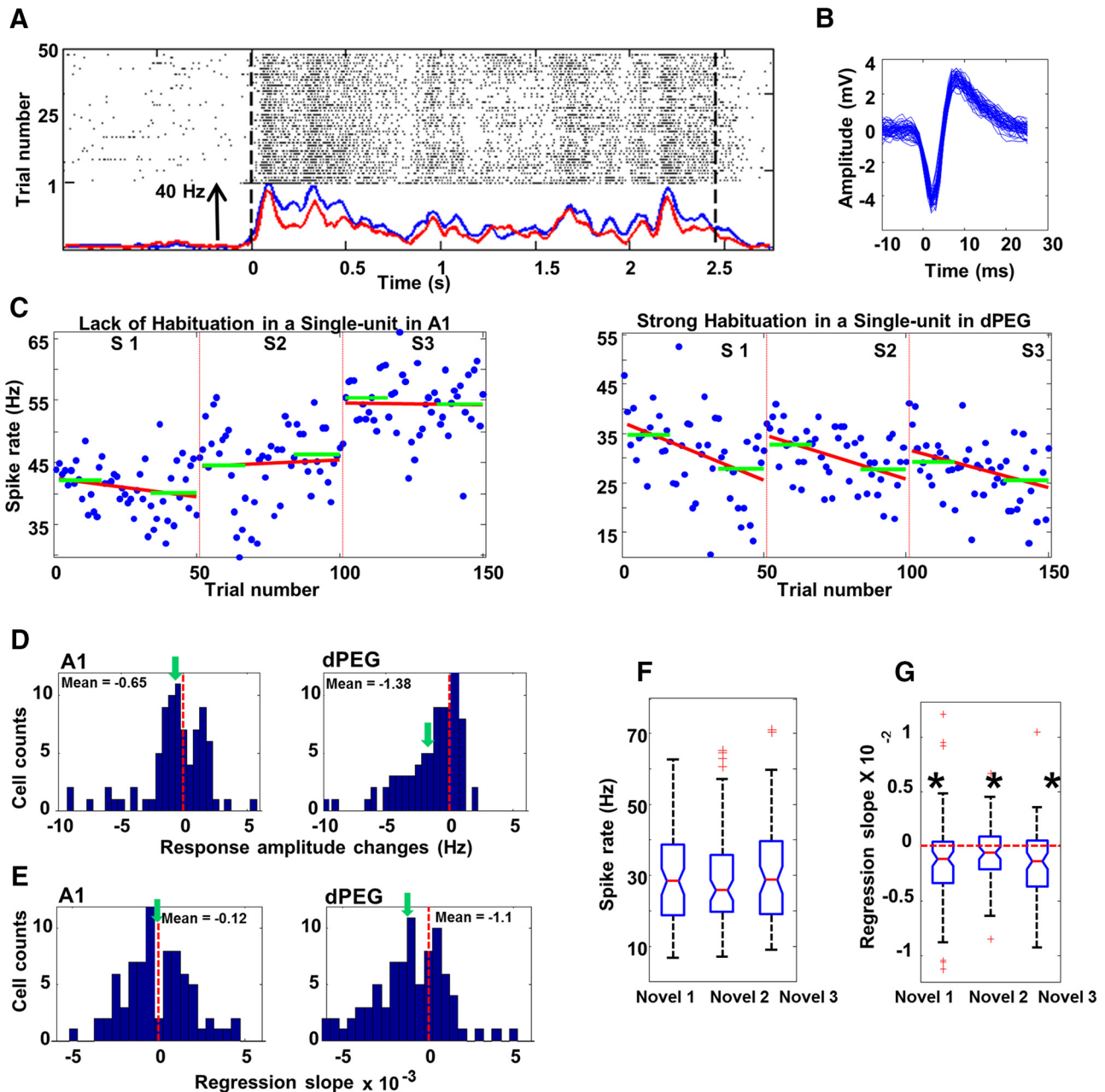


Figure 4. *A*, Examples of a dPEG response to novel stimuli. Raster shows spike times by trials. Two dashed lines indicate the onset and offset of the stimuli. Baseline activity was measured within the 1 s window before stimulus onset. Notice that the baseline is consistent in both PSTHs. Yet, spikes became sparser after 20–30 repetitions. In this example, PSTHs for the first eight trials (blue) in a block of identical stimuli is shows an increased response compared with the PSTH for the last eight trials in this block (red). *B*, Spike waveform for the neuron in 3A is consistent across the whole recording session for all phases. *C*, Examples of spike counts by trials for an A1 neuron and a dPEG neuron in response to three novel sounds in three sequential blocks. Red line shows the linear regression fit. Green horizontal lines show the average spike rate in the first and last 16 trials. *D*, Histograms of changes in response strength in A1 ($N = 81$) and dPEG ($N = 83$). Response change is calculated as the difference between the average spike count of the first 16 trials in the block and of the last 16 trials (*C*, green lines). The vertical dashed lines indicate zero. The green arrows indicated the mean of each distribution. *E*, Histograms of regression slopes in A1 ($N = 81$) and dPEG ($N = 83$). The vertical dashed lines indicate zero. The green arrows indicate the mean in each distribution. *F*, Box-and-whisker plot of spike counts in dPEG ($N = 73$) to three novel stimuli sequentially presented in Phase 1. There was no significant difference between blocks, indicating that habituation to the first stimulus did not affect responses to the second and third stimulus. *G*, Box-and-whisker plot of regression in dPEG ($N = 73$) slopes of three novel stimuli sequentially presented in Phase 1. The dashed red line indicated 0. Regression slopes in all three blocks were significantly lower than zero ($p < 0.001$ for all). There was no significant difference between blocks, indicating that each novel stimulus was habituated independently.

procedures previously used in songbirds: Phan et al., 2006; Velho et al., 2012). Figure 5, *B* and *C*, illustrates that the regression slopes to novel stimuli were significantly negative (Wilcoxon test: $z = -4.05, p = 0.001$; Fig. 5*B*, blue trace, *C*, left box). By contrast, regression slopes of familiar stimuli were evenly distributed around zero with no significant bias (Wilcoxon test: $z = -0.38, p = 0.702$; Fig. 5*B*, red trace, *C*, middle box). When stimuli were

time-reversed, regression slopes became significantly negative again (Wilcoxon test: $z = -3.66, p = 0.001$; Fig. 5*B*, green trace, *C*, right box).

To explore in more detail the differences between the habituation in the three test phases, we computed the grand average habituation curves from the entire population of dPEG cells for the various types of stimuli. As shown in Figure 5*D*, responses to

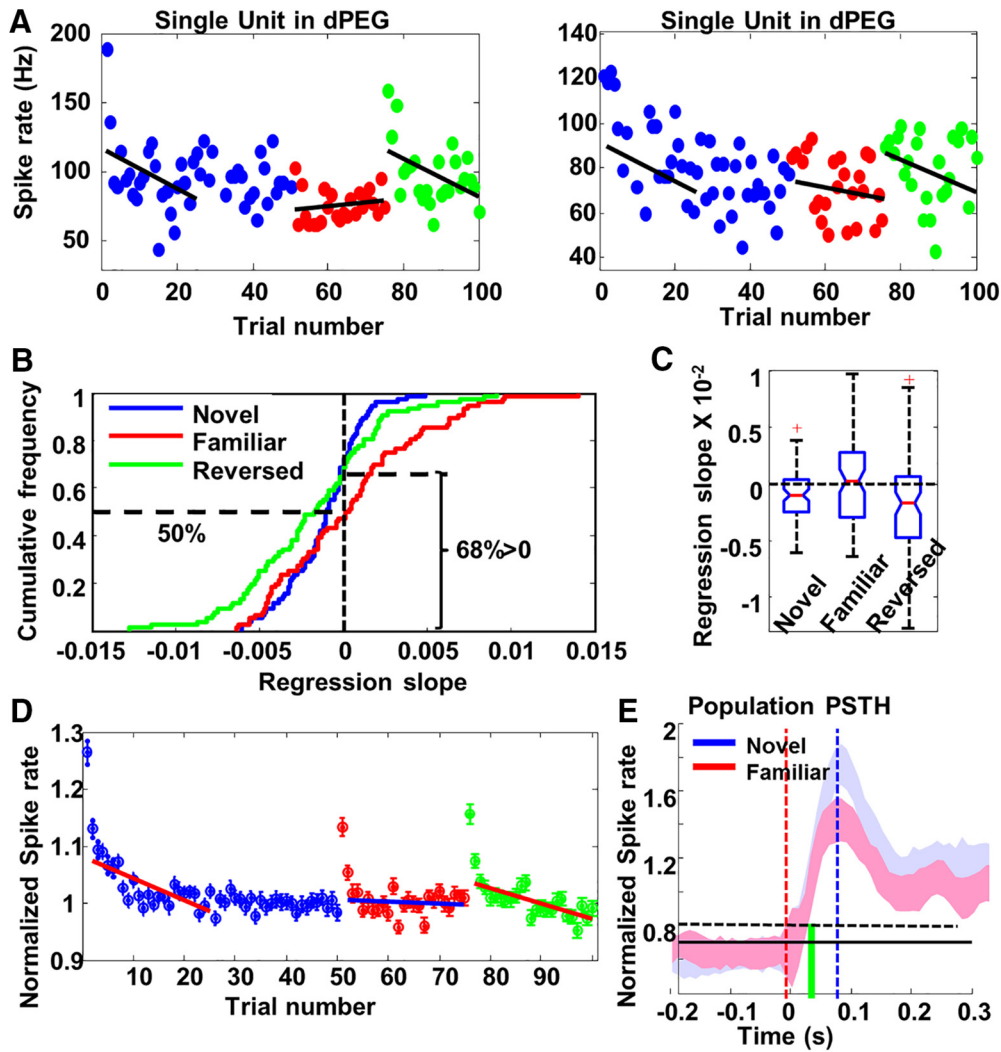


Figure 5. *A*, Examples of spike counts by trials in two dPEG neurons in response to a stimulus when it was novel (blue), familiar (red), or reversed (green). Note that the first presentation of the familiar stimulus in Phase 2 occurred 20 min after the last trial of the same stimulus when it was novel in Phase 1. *B*, Cumulative frequency distribution of regression slopes for novel stimuli (blue), familiar stimuli (red) and reversed stimuli (green) in dPEG. Horizontal dashed line on the right indicated that 68% regression slopes obtained from novel stimuli were lower than zero. Horizontal dashed line on the left indicated that only 50% (chance level) of the regression slopes obtained from familiar stimuli were lower than zero. *C*, Box-and-whisker plot of regression slopes in dPEG for novel (left), familiar (middle), and reversed (right) stimuli. *D*, Population habituation curve in PEG. Each circle indicates responses averaged from all neurons ($N = 83$). Error bars on top of circles show SE. The first red line (across the blue circles) shows the linear regression fit for the first 25 trials of novel stimuli. The blue line shows the linear regression fit for familiar stimuli. The second red line (across the green circles) shows the linear regression fit for reversed stimuli. The two red lines were clearly tilted. The blue line was flat. Notice that the very first trials in all three conditions were excluded from this analysis. *E*, The temporal profiles of auditory responses to novel stimuli in dPEG (red) and to familiar stimuli (blue), calculated from averaged PSTHs across all neurons and all stimuli. Red dashed line shows stimulus onset time. The Black dashed line is the threshold calculated from the mean in the baseline plus SE (in 10 ms bins). The latency of the responses is marked by green vertical line. The peak of the PSTHs is indicated by the blue vertical line.

novel stimuli (blue dots) decreased by $\sim 20\%$ in the first 25 trials that were used to calculate all regression slopes, reaching a plateau after ~ 25 – 30 trials. Responses to familiar stimuli (red dots) were approximately habituated to the same level as this plateau, except (as discussed later) for the first two trials, which showed an increased response. The regression slope for familiar stimuli was significantly flatter than the slope to novel stimuli. Finally, the regression slope for reversed stimuli (green dots) was similar to the slope for novel stimuli, suggesting that the long-lasting habituation to familiar stimuli was not due to fatigue, but likely reflected a form of long-term memory in dPEG (lasting longer than 20 min) analogous to that found in birds. The absence of long-lasting habituation in A1 and its presence in dPEG distinguished it from stimulus-specific adaptation (SSA), which is observed in many earlier auditory processing stages such as inferior colliculus and auditory thalamus (Antunes et al., 2010; Duque et al., 2012; Nelken, 2014).

However, the response to the very first trial was clearly much larger than the responses in the rest of trials of the block. We compared its amplitude with the mean of trials 2–25 (used in linear regression analysis). It was 23% larger than the rest of the trials for novel stimuli, 13% larger than the rest of trials for familiar stimuli and 15% larger than the rest of trials for reversed stimuli, which indicated that a possible SSA-like effect may occur at the beginning of each block.

Finally, we also examined the effects of the habituation on the temporal profile of the dPEG responses. The averaged PSTH of all dPEG responses to novel stimuli is shown in Figure 5*E* (blue trace), together with the same response to the familiar version of the stimuli (red trace). As noted earlier, the two responses appear very similar except for an overall 18% attenuation of the habituated responses. In A1, a similar analysis showed that there was no such attenuation.

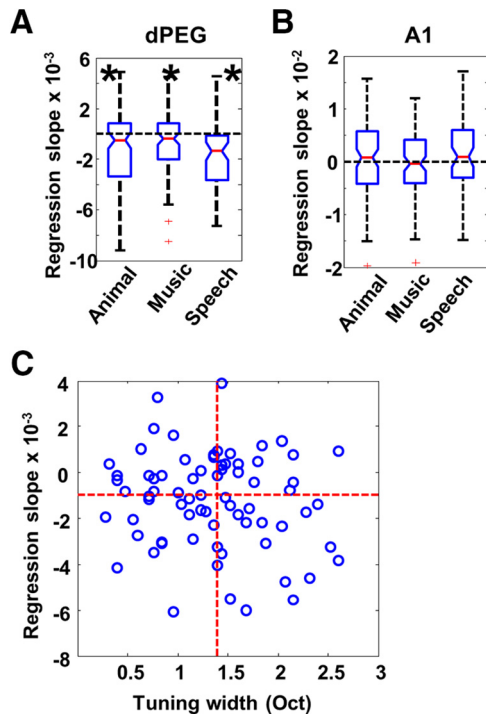


Figure 6. *A*, Box-and-whisker plots of regression slopes for animal vocalizations, music and speech, measured in dPEG. All are significantly lower than zero, and there are no significant differences among them. Animal: $p < 0.001$; Music: $p < 0.02$; Speech: $p < 0.001$. *B*, Box-and-whisker plot of regression slopes for animal vocalizations, music, and speech, measured in A1. None of them is significantly lower than zero. *C*, There was a marginal (not significant) negative relationship between regression slopes and tuning-width in dPEG neurons.

Habituation was similar for all stimulus types

As described in Materials and Methods, three different categories of novel stimuli (speech, music, animal vocalizations) were used in each recording. In the analysis discussed above, responses to these three different types of stimuli were averaged in each test phase. To compare habituation effects among these three stimulus categories, we analyzed data for each stimulus-type separately. However, we found no significant differences in regression slopes across the three classes, either in dPEG (Fig. 6*A*; Median test, $\chi^2 = 5.10$, $df = 2$, $p = 0.078$) or in A1 (Fig. 6*B*; Median test, $\chi^2 = 1.50$, $df = 2$, $p = 0.472$). Stimuli from all three types of sounds habituated equally in dPEG: animal vocalizations (Wilcoxon test: $z = -3.41$, $p = 0.001$); musical pieces and vocal songs (Wilcoxon test: $z = -2.31$, $p = 0.02$); speech (Wilcoxon test: $z = -4.13$, $p = 0.001$). Similarly, in A1, none of three types induced significant habituation: animal vocalization (Wilcoxon test: $z = -0.96$, $p = 0.338$); music pieces and vocal songs (Wilcoxon test: $z = -0.35$, $p = 0.730$); speech (Wilcoxon test: $z = -0.966$, $p = 0.334$). Therefore, habituation was independent of the type of sound presented.

Habituation was not correlated with the tuning of the neurons

We also tested to see whether habituation was related to neuronal frequency tuning by correlating the bandwidth of each recorded neuron in A1 and dPEG with its regression slope for novel stimuli (Phase 1). We did not find any significant correlation between the two properties in either cortical field: A1 ($\rho = -0.03$, $p = 0.858$) and dPEG ($\rho = -0.14$, $p = 0.222$). However, there is a slight negative trend for correlations calculated in dPEG neurons (Fig. 6*C*), indicating that neurons with wider tuning bandwidths may habituate slightly faster than narrowly tuned neurons.

Mutual information increased after habituation

So far, habituation of neural responses was quantified as the spike-rate change during stimulus presentations. We also examined the spike pattern changes during habituation. We hypothesized that as the spike rates decreased with habituation, response noise would also decrease such that the MI between spike-trains and stimuli would increase, as found in works with songbirds (Lu and Vicario, 2014). Thus, we calculated the MIs for the first half of trials of all three stimuli in test Phase 1 and compared them with the MIs for the second half of trials of the same stimuli in one block in Phase 1, when they were habituated. Consistent with the hypothesis, our analysis revealed that the MI did not significantly change in A1 (Fig. 7*A*, *C*, blue trace; Wilcoxon test: $z = -1.34$, $p = 0.181$), whereas it significantly increased in dPEG (Fig. 7*B*, *C*, red trace; Wilcoxon test: $z = -4.11$, $p = 0.001$).

Pupillary measurement was consistent with neural habituation results

Based on analogous findings in birds, the results thus far opened up the possibility that the long-lasting habituation of neural responses observed in dPEG might reflect a long-term auditory recognition memory for complex sounds that could be measured at a behavioral level. To confirm the presence of memory, we sought a behavioral measure that could demonstrate the animals' recognition of familiar stimuli compared with the novel ones. As pupillary size-changes in response to sound have been shown to be an indicator of arousal (Reimer et al., 2016) we therefore measured pupillary size-changes in three awake, quietly listening animals presented with the same sequence of stimuli used for the electrophysiology experiments. Animals' pupillary size increased significantly (13% increase relative to the baseline, Wilcoxon test: $z = -17.9$, $p = 0.001$) after presentation of novel sounds (Fig. 8*A*, blue trace) with a latency of 0.9 s. By comparison, pupillary responses to familiar sounds (same sounds presented ~ 20 min later in Phase 2) were much smaller and not significantly different from the baseline (Wilcoxon test: $z = -0.04$, $p = 0.968$; Fig. 8*A*, red trace), as a result of habituation. Therefore, habituation of pupillary responses was consistent with the pattern of long-term habituation of neural responses in dPEG. Trial-by-trial analysis of pupillary responses revealed a similar picture: pupillary responses to novel stimuli (blue trace) showed a clear reduction in the initial few trials; whereas pupillary size did not change significantly to familiar stimuli (red trace), except for the first trial. The habituation curve to the time-reversed stimuli (green trace) was similar to that of the novel stimuli, exhibiting recovered responses to those stimuli in the initial few trials.

However, one difference between pupillary and neural habituation was that habituation in pupillary responses to novel stimuli was much faster than neuronal habituation to the same stimuli. Many fewer trials were needed to induce pupillary habituation, which reached plateau after only the first five trials. By comparison, neural habituation to novel stimuli in dPEG did not plateau until 25–30 trials (Fig. 5*D*). This difference is highlighted by the difference between the fitted power-functions to the neuronal habituation (Fig. 8*C*, solid purple curve, replotted as *D*, dashed curve) and its pupillary habituation counterpart (Fig. 8*D*, solid blue curve). The dramatic difference between the two curves shows that the *global habituation* effect (as reflected by the pupillary modulation) occurred much faster, within only a few trials.

Discussion

Habituation experiments described in this report yielded several significant findings: (1) Long-term habituation occurred in re-

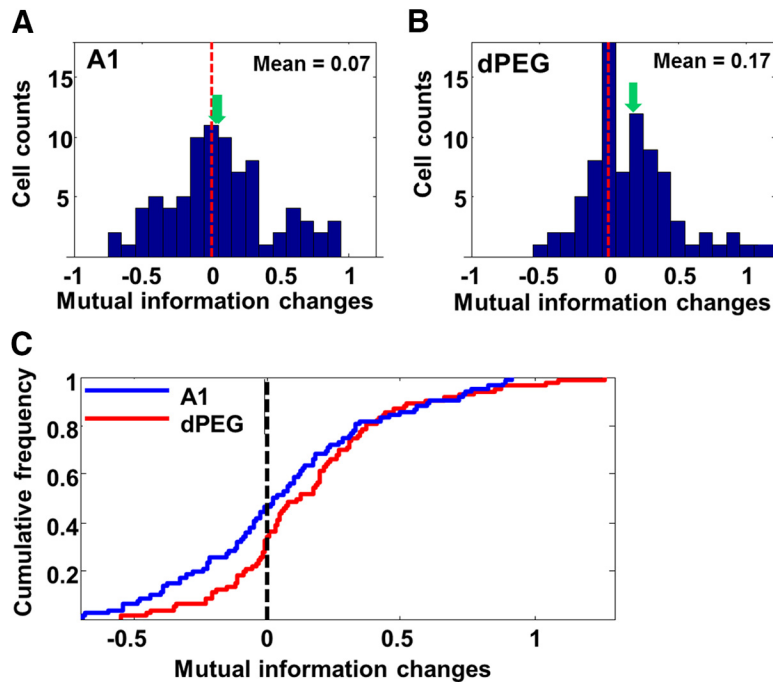


Figure 7. *A*, Histograms of mutual information changes (familiar – novel) during habituation in A1. There was no significant change in mutual information. The vertical dashed lines indicate zero, and the green arrows mark the mean of each distribution. *B*, Same plots as in *A*, except for all dPEG cells. *C*, Cumulative frequency distribution of mutual information changes in A1 (blue) and dPEG (red).

sponses of the secondary auditory cortical fields dPEG, but not in the primary auditory cortex A1. (2) The habituation to each novel stimulus in a given neuron showed no interference across stimuli. Habituation to each stimulus was independent of ongoing habituation to the other two novel stimuli in the current three-stimulus, three-block design, and hence the observed habituation was stimulus-specific for each set of three diverse stimuli. (3) Habituation effects lasted for at least 20 min. (4) The effects observed cannot be explained by simple fatigue in specific frequency channels, because time-reversed sounds (with comparable average spectral content) induced responses similar to novel sounds. (5) Pupillary changes revealed a rapid global habituation effect that correlated with the neural habituation effects, indicating links between arousal, neuronal activity, and potentially parallel behavioral correlates.

Stimulus-dependent attenuation of neuronal responses to repetitive presentation of visual and auditory stimuli has already been extensively studied and exploited as a powerful paradigm in fMRI, EEG experiments (for review, see Krekelberg et al., 2006; Grill-Spector et al., 2006; Summerfield and de Lange, 2014). There are at least two types of neuronal attenuation phenomena that have been previously studied: (1) Repetition suppression due to fatigue (for review, see Grill-Spector et al., 2006). (2) SSA because of contrast between oddball and standard stimuli (Ulanovsky et al., 2003, 2004). The long-term habituation that we describe here differs from both of these phenomena, as we shall describe below.

Differences between long-term habituation and repetition suppression due to fatigue

Repetition suppression has been used in fMRI and EEG studies, and is considered to be related to perceptual “after-effects”. For example, presenting gratings for 40 s and then switching to an orthogonal orientation induced large responses in V1 (Tootell et

al., 1998). It has been assumed that fatigue is the underlying cause of these “repetition suppression” effects, induced by repetitive activation of the same neuron population. However, as we described earlier, the habituation in our experiments cannot be explained by fatigue for three reasons: (1) In our experiments, both neuronal receptive fields and stimuli were often broad enough that neurons responded to all three stimuli in Phase 1. Thus, if habituation was the result of simple fatigue, then subsequently presented stimuli should have induced weaker responses and hence less habituation. But there was no such interaction between sequential stimulus blocks. (Fig. 4*F, G*). (2) Familiar stimuli were presented ~20 min later, yet the responses remained habituated (Fig. 5*C, D*). To our knowledge, such a long-term “fatigue” of responses from passive exposure to acoustic stimuli has not been previously reported at the single-unit level in the mammalian auditory system. (3) Time-reversed stimuli induced responses similar to novel stimuli in the same neurons that had already been habituated to the forward versions of the stimuli (Fig. 5*C, D*). However, there are still

many unanswered questions about the precision or degree of stimulus-generalization of the habituation process that we have observed in the ferret auditory cortex. Additional experiments, with more sequential blocks, with artificial stimuli that have varying degrees of stimulus similarity, will test what acoustic features, and at what spectral-temporal scale, may cause potential interference or generalization across stimuli during habituation. Such experiments will clarify the possible stimulus specificity of habituation and may also be beneficial for understanding the underlying mechanisms.

Differences between long-term habituation and stimulus-specific adaptation

SSA occurs when a small number of oddball stimuli are presented in the context of large numbers of standard stimuli. Oddball stimuli evoked larger responses than the adapted responses to the common standard stimuli and these oddball responses can be viewed as violations of learned expectations developed during repetition of the standard. SSA usually depends on the contrast between the oddball and the standard stimuli (e.g., frequency differences; Ulanovsky et al., 2003), and the underlying mechanisms may include release from suppression that is induced by the repetitive standard sounds within a short-time history (Yarden and Nelken, 2017). However, there are three factors distinguishing the long-term habituation we have observed from the SSA: (1) In our paradigm, a familiar sound was often presented immediately following a whole block of sounds that were of a different type, and hence contrasted strongly with them. If habituation followed the same pattern as SSA, then responses to familiar (i.e., previously habituated) sounds should be as strong as those to novel stimuli. Yet, the response to familiar sounds (apart from the first 1 or 2 presentations) appeared to be already habituated; i.e., in contrast, in SSA, the memory of habituation of a given stimulus in one block of repeated stimuli appears to be

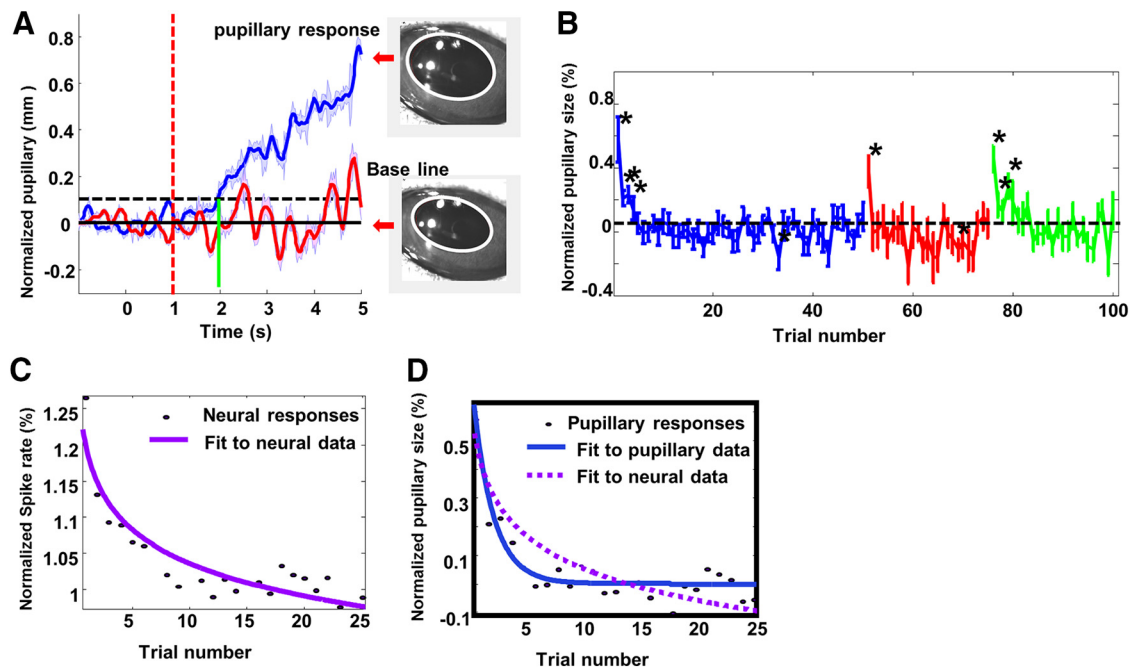


Figure 8. *A*, The temporal profile of pupillary responses in response to novel sounds (blue trace) and familiar sounds (red trace). Stimulus onset time (red dashed line), threshold (black dashed line), and latency (green vertical line) are all defined as in Figure 5*E*. An example of pupil image after presentation of a novel sound (top) and at baseline (before stimulus onset). Pupillary shape is outlined with a white trace. *B*, Population habituation curve of pupillary response to novel stimuli (blue), familiar stimuli (red), and reversed stimuli (green). Asterisks indicate the trials in which the averaged pupillary responses are significantly larger than baseline. Both novel stimuli and reversed stimuli induced significant pupillary responses in the initial few trials. Pupillary responses to familiar sounds are not significantly larger than the baseline, except for the first stimulus in a block. *C*, Best curve fit for the population habituation curve of dPEG neurons in response to novel stimuli (same as the curve consisting of blue dots in Fig. 5*D*). A power function is fit to the data (purple trace). *D*, Blue solid trace shows the best fit for the pupillary habituation curve in response to novel stimuli. The purple dashed line shows the best exponential fit for the population habituation curve of neural responses in *C*. Habituation in neural responses is much slower and continued even after habituation of pupillary responses plateaued.

wiped clean (*tabula rasa*) after a delay (15 s). Switching the role of the oddball and standard tones in SSA in successive blocks does not appear to affect the basic response pattern, except for a very small and nonspecific decline (Ulanovsky et al., 2004). Thus, in classic SSA, in part because of stimulus block design, the notion of long-term familiarized stimuli does not exist in the same way that we have shown for long-term habituation. (2) SSA typically acts over much shorter time windows lasting <2 s (Ulanovsky et al., 2003) although some influences can last much longer; up to a minute or more (Ulanovsky et al., 2004). The habituation we describe can last for at least 20 min, which is much longer than has been described in the SSA literature. (3) SSA was first found in A1 (Ulanovsky et al., 2003) and has also been described in higher auditory cortical areas (Nieto-Diego and Malmierca, 2016). But SSA has also been described in the midbrain (Malmierca et al., 2009; Ayala and Malmierca, 2013) and thalamus (Anderson et al., 2009; for review, see Malmierca et al., 2015). We found that long-term habituation was very weak in A1, becoming strong only in the secondary auditory cortical areas of dPEG in the ferret.

The habituation curve described in our results may reflect a combination of both local circuit and top-down effects. The response to the first trial for both novel and familiar stimuli (i.e., the same stimuli before and after habituation) elicited a significantly larger response, compared with the rest of the habituation curve (Fig. 5*D*). This may reflect a surprise or even the classic “oddball” SSA-like response, and an overall change in arousal due to the rapid switch between stimuli in successive test blocks. This is consistent with earlier findings in songbirds, which also showed that if a conspecific vocalization was presented after a few heterospecific vocalizations, a “surprise” effect could be induced,

independent of an ongoing habituation effect (Lu and Vicario, 2017).

As indicated above, long-term habituation to auditory stimuli has been described in the secondary auditory forebrain of songbirds but was also absent in the primary avian auditory forebrain, a counterpart of A1 (Terleph et al., 2006). In addition, long-lasting habituation has also been found in the optic tectum of owls (Reches and Gutfreund, 2008; Netser et al., 2011; Dutta and Gutfreund, 2014). In zebra finches, the long-term habituation was correlated with habituation in behavioral responses to sounds (Vicario, 2004) and imitation fidelity (Phan et al., 2006), reflecting long-term recognition memory. Long-term habituation was correlated with a decrease of immediate early gene expression in the same regions (Mello et al., 1992) and requires RNA and protein for expression (Chew et al., 1995) and is also dependent upon α -adrenergic transmission in NCM (Velho et al., 2012). Further behavioral, neurophysiological, and neuropharmacological experiments could test if similar gene and protein expression correlates are also present in the ferret dPEG. Such experiments may also clarify the mechanisms of habituation and distinguish it from other phenomenon (e.g., stimulus fatigue, SSA, etc.).

Comparison with studies of implicit (statistical) learning

Long-term habituation in birds is accompanied by extraction of recurrent statistical patterns (Lu and Vicario, 2014), resembling the implicit (statistical) learning that is found in animals and human behavior. Young, pre-verbal human infants, nonhuman primates and birds can process and recall different artificial grammars and exhibit implicit long-term statistical learning of

auditory patterns and sequences, or even random noise patterns (Saffran et al., 1996, 1999; Hauser et al., 2001; Fitch and Hauser, 2004; Newport et al., 2004; Agus et al., 2010; Abe and Watanabe, 2011; Wilson et al., 2013; Kang et al., 2017; Milne et al., 2018). The experiments we describe in this report may constitute the neuronal underpinnings of such long-term implicit learning phenomena, which have hitherto been unexplored at a cellular level in the mammalian auditory system.

Habituation in the visual system

The basic phenomenon of stimulus-dependent auditory habituation is likely related to various forms of plasticity observed in other modalities, such as the visual system. For example, a few days of exposure to sequences of visual patterns led to long-term potentiation in responses in V1 (Gavornik and Bear, 2014). Given our results, and the bird literature on habituation, both of which demonstrate the importance of higher auditory areas in long-term habituation, it is surprising and intriguing that such habituation can occur in the mouse primary visual cortex; although we speculate that V1 in the mouse may be more of an associative cortical area. In support of the importance of higher visual cortical areas in habituation, repetitive presentation of a visual object image led to habituation in neurons in higher visual cortical areas such as inferior temporal cortex in the macaque (Li et al., 1993; Miller and Desimone, 1994; Sobotka and Ringo, 1994; Ringo, 1996), a habituation that is long-lasting (5 min) even after 150 intervening stimuli (Li et al., 1993). However, these latter studies in the monkey are quite different from our paradigm in that they experimented with behaving (as opposed to passively viewing) monkeys performing delayed match-to-sample tasks, and hence the effects were dependent on the reward structure in the task. When a given visual object did not match the sample (irrelevant to the reward), the responses to repetition of this object image habituated. By contrast, if a given object was a match (led to reward), the responses increased with repetition. Finally, similar effects of long-lasting visual memories have been described in fMRI/EEG studies with attentive human subjects presented with repetitive or even single instance face or object stimuli (Henson et al., 2000; Jiang et al., 2000; van Turennout et al., 2000; Doniger et al., 2001; Schendan and Kutas, 2003).

Contribution of habituation to complementary processes of implicit acoustic scene analysis

We speculate that the combination of SSA and habituation to repetitive, behaviorally neutral stimuli may be an effective blend of implicit mechanisms for allocating attention to incoming stimuli in the environmental soundscape, each with their own time window. Although SSA detects violation of expectation in short time window, habituation avoids over-responsiveness to familiar sounds over a longer history. Thus, when a novel stimulus arrives, its biological relevance to the animal is initially unknown and hence it is beneficial to turn attention to the novel event with an enhanced oddball alerting response. However, if the same stimulus is repeated again and again without significant behavioral consequences, its immediate behavioral relevance is judged to be low and hence responses to this specific stimulus are habituated (filtered out), and stored in long-term memory as a signal that can be safely ignored, facilitating the processing of other potentially more relevant incoming signals. These implicit mechanisms are possibly complemented by processes during active engagement and attention where (as described earlier) association with a reward may potentiate the response.

References

- Abe K, Watanabe D (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat Neurosci* 14:1067–1074. [CrossRef Medline](#)
- Agus TR, Thorpe SJ, Pressnitzer D (2010) Rapid formation of robust auditory memories: insights from noise. *Neuron* 66:610–618. [CrossRef Medline](#)
- Anderson LA, Christianson GB, Linden JF (2009) Stimulus-specific adaptation occurs in the auditory thalamus. *J Neurosci* 29:7359–7363. [CrossRef Medline](#)
- Antunes FM, Nelken I, Covey E, Malmierca MS (2010) Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One* 5:e14071. [CrossRef Medline](#)
- Atiani S, David SV, Elgueta D, Locastro M, Radtke-Schuller S, Shamma SA, Fritz JB (2014) Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82:486–499. [CrossRef Medline](#)
- Ayala YA, Malmierca MS (2013) Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Front Neural Circuits* 6:89. [CrossRef Medline](#)
- Bartlett EL, Wang X (2005) Long-lasting modulation by stimulus context in primate auditory cortex. *J Neurophysiol* 94:83–104. [CrossRef Medline](#)
- Bizley JK, Bajo VM, Nodal FR, King AJ (2015) Cortico-cortical connectivity within ferret auditory cortex. *J Comp Neurol* 523:2187–2210. [CrossRef Medline](#)
- Bizley JK, Nodal FR, Nelken I, King AJ (2005) Functional organization of ferret auditory cortex. *Cereb Cortex* 15:1637–1653. [CrossRef Medline](#)
- Brosch M, Schreiner CE (2000) Sequence sensitivity of neurons in cat primary auditory cortex. *Cereb Cortex* 10:1155–1167. [CrossRef Medline](#)
- Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. *Proc Natl Acad Sci U S A* 92:3406–3410. [CrossRef Medline](#)
- Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the calls and songs of individual birds. *Proc Natl Acad Sci U S A* 93:1950–1955. [CrossRef Medline](#)
- Depireux DA, Simon JZ, Klein DJ, Shamma SA (2001) Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 85:1220–1234. [CrossRef Medline](#)
- Doniger GM, Foxe JJ, Schroeder CE, Murray MM, Higgins BA, Javitt DC (2001) Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *Neuroimage* 13:305–313. [CrossRef Medline](#)
- Duque D, Pérez-González D, Ayala YA, Palmer AR, Malmierca MS (2012) Topographic distribution, frequency, and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *J Neurosci* 32:17762–17774. [CrossRef Medline](#)
- Dutta A, Gutfreund Y (2014) Saliency mapping in the optic tectum and its relationship to habituation. *Front Integr Neurosci* 8:1. [CrossRef Medline](#)
- Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303:377–380. [CrossRef Medline](#)
- Gavornik JP, Bear MF (2014) Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci* 17:732–737. [CrossRef Medline](#)
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23. [CrossRef Medline](#)
- Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a nonhuman primate: statistical learning in cotton-top tamarins. *Cognition* 78:B53–B64. [CrossRef Medline](#)
- Henson R, Shallice T, Dolan R (2000) Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287:1269–1272. [CrossRef Medline](#)
- Jiang Y, Haxby JV, Martin A, Ungerleider LG, Parasuraman R (2000) Complementary neural mechanisms for tracking items in human working memory. *Science* 287:643–646. [CrossRef Medline](#)
- Jusczyk PW, Hohne EA (1997) Infants' memory for spoken words. *Science* 277:1984–1986. [Medline](#)
- Kang H, Agus TR, Pressnitzer D (2017) Auditory memory for random time patterns. *J Acoust Soc Am* 142:2219. [CrossRef Medline](#)
- Krekelberg B, Boynton GM, van Wezel RJ (2006) Adaptation: from single cells to BOLD signals. *Trends Neurosci* 29:250–256. [CrossRef Medline](#)
- Li L, Miller EK, Desimone R (1993) The representation of stimulus famil-

- ilarity in anterior inferior temporal cortex. *J Neurophysiol* 69:1918–1929. [CrossRef Medline](#)
- Lu K, Vicario DS (2014) Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. *Proc Natl Acad Sci U S A* 111:14553–14558. [CrossRef Medline](#)
- Lu K, Vicario DS (2017) Familiar but unexpected: effects of sound context statistics on auditory responses in the songbird forebrain. *J Neurosci* 37:12006–12017. [CrossRef Medline](#)
- Lu K, Xu Y, Yin P, Oxenham AJ, Fritz JB, Shamma SA (2017) Temporal coherence structure rapidly shapes neuronal interactions. *Nat Commun* 8:13900. [CrossRef Medline](#)
- Malmierca MS, Anderson LA, Antunes FM (2015) The cortical modulation of stimulus-specific adaptation in the auditory midbrain and thalamus: a potential neuronal correlate for predictive coding. *Front Syst Neurosci* 9:19. [CrossRef Medline](#)
- Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29:5483–5493. [CrossRef Medline](#)
- Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. *Proc Natl Acad Sci U S A* 89:6818–6822. [CrossRef Medline](#)
- Miller EK, Desimone R (1994) Parallel neuronal mechanisms for short-term memory. *Science* 28:263:520–522. [Medline](#)
- Milne AE, Petkov CI, Wilson B (2018) Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm. *Neuroscience*. 389:104–117. [CrossRef Medline](#)
- Nelken I (2014) Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol Cybern* 108:655–663. [CrossRef Medline](#)
- Nelken I, Chechik G (2007) Information theory in auditory research. *Hear Res* 229:94–105. [CrossRef Medline](#)
- Netser S, Zahar Y, Gutfreund Y (2011) Stimulus-specific adaptation: can it be a neural correlate of behavioral habituation? *J Neurosci* 31:17811–17820. [CrossRef Medline](#)
- Newport EL, Hauser MD, Spaepen G, Aslin RN (2004) Learning at a distance: II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn Psychol* 49:85–117. [CrossRef Medline](#)
- Nieto-Diego J, Malmierca MS (2016) Topographic distribution of stimulus-specific adaptation across auditory cortical fields in the anesthetized rat. *PLoS Biol* 14:e1002397. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Phan ML, Pytte CL, Vicario DS (2006) Early auditory experience generates long-lasting memories that may subserve vocal learning in songbirds. *Proc Natl Acad Sci U S A* 103:1088–1093. [CrossRef Medline](#)
- Phan ML, Vicario DS (2010) Hemispheric differences in processing of vocalizations depend on early experience. *Proc Natl Acad Sci U S A* 107:2301–2306. [CrossRef Medline](#)
- Reches A, Gutfreund Y (2008) Stimulus-specific adaptations in the gaze control system of the barn owl. *J Neurosci* 28:1523–1533. [CrossRef Medline](#)
- Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolia AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat Commun* 7:13289. [CrossRef Medline](#)
- Ribeiro S, Mello CV, Velho T, Gardner TJ, Jarvis ED, Pavlides C (2002) Induction of hippocampal long-term potentiation during waking leads to increased extrahippocampal *zif-268* expression during ensuing rapid-eye-movement sleep. *J Neurosci* 22:10914–10923. [CrossRef Medline](#)
- Ringo JL (1996) Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behav Brain Res* 76:191–197. [CrossRef Medline](#)
- Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928. [CrossRef Medline](#)
- Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70:27–52. [CrossRef Medline](#)
- Schendan HE, Kutas M (2003) Time course of processes and representations supporting visual object identification and memory. *J Cogn Neurosci* 15:111–135. [CrossRef Medline](#)
- Sobotka S, Ringo JL (1994) Stimulus specific adaptation in excited but not in inhibited cells in inferotemporal cortex of macaque. *Brain Res* 646:95–99. [CrossRef Medline](#)
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756. [CrossRef Medline](#)
- Terleph TA, Mello CV, Vicario DS (2006) Auditory topography and temporal response dynamics of canary caudal telencephalon. *J Neurobiol* 66:281–292. [CrossRef Medline](#)
- Tootell RB, Hadjikhani NK, Vanduffel W, Liu AK, Mendola JD, Sereno MI, Dale AM (1998) Functional analysis of primary visual cortex (V1) in humans. *Proc Natl Acad Sci U S A* 95:811–817. [CrossRef Medline](#)
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398. [CrossRef Medline](#)
- Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24:10440–10453. [CrossRef Medline](#)
- van Turennout M, Ellmore T, Martin A (2000) Long-lasting cortical plasticity in the object naming system. *Nat Neurosci* 3:1329–1334. [CrossRef Medline](#)
- Velho TA, Lu K, Ribeiro S, Pinaud R, Vicario D, Mello CV (2012) Noradrenergic control of gene expression and long-term neuronal adaptation evoked by learned vocalizations in songbirds. *PLoS One* 7:e36276. [CrossRef Medline](#)
- Vicario DS (2004) Using learned calls to study sensory-motor integration in songbirds. *Ann N Y Acad Sci* 1016:246–262. [CrossRef Medline](#)
- Wilson B, Slater H, Kikuchi Y, Milne AE, Marslen-Wilson WD, Smith K, Petkov CI (2013) Auditory artificial grammar learning in macaque and marmoset monkeys. *J Neurosci* 33:18825–18835. [CrossRef Medline](#)
- Yarden TS, Nelken I (2017) Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS Comput Biol* 13:e1005437. [CrossRef Medline](#)