

Familiar But Unexpected: Effects of Sound Context Statistics on Auditory Responses in the Songbird Forebrain

Kai Lu¹ and David S. Vicario²

¹Institute for Systems Research, University of Maryland, College Park, Maryland 20740, ²Psychology Department, Rutgers University, Piscataway, New Jersey 08854

Rapid discrimination of salient acoustic signals in the noisy natural environment may depend, not only on specific stimulus features, but also on previous experience that generates expectations about upcoming events. We studied the neural correlates of expectation in the songbird forebrain by using natural vocalizations as stimuli and manipulating the category and familiarity of context sounds. In our paradigm, we recorded bilaterally from auditory neurons in awake adult male zebra finches with multiple microelectrodes during repeated playback of a conspecific song, followed by further playback of this test song in different interleaved sequences with other conspecific or heterospecific songs. Significant enhancement in the auditory response to the test song was seen when its acoustic features differed from the statistical distribution of context song features, but not when it shared the same distribution. Enhancement was also seen when the time of occurrence of the test song was uncertain. These results show that auditory forebrain responses in awake animals in the passive hearing state are modulated dynamically by previous auditory experience and imply that the auditory system can identify the category of a sound based on the global features of the acoustic context. Furthermore, this probability-dependent enhancement in responses to surprising stimuli is independent of stimulus-specific adaptation, which tracks familiarity, suggesting that the two processes could coexist in auditory processing. These findings establish the songbird as a model system for studying these phenomena and contribute to our understanding of statistical learning and the origin of human ERP phenomena to unexpected stimuli.

Key words: adaptation; audition; context; expectation; multielectrode; surprise

Significance Statement

Traditional auditory neurophysiology has mapped acoustic features of sounds to the response properties of neurons; however, growing evidence suggests that neurons can also encode the probability of sounds. We recorded responses of songbird auditory neurons in a novel paradigm that presented a familiar test stimulus in a sequence with similar or dissimilar sounds. The responses encode, not only stimulus familiarity, but also the expectation for a class of sounds based on the recent statistics of varying sounds in the acoustic context. Our approach thus provides a model system that uses a controlled stimulus paradigm to understand the mechanisms by which top-down processes (expectation and memory) and bottom-up processes (based on stimulus features) interact in sensory coding.

Introduction

It is recognized that neurons in the auditory forebrain encode, not only the acoustic properties of sounds, but also the probability of those sounds and/or the transitions between sounds

(Ulanovsky et al., 2003, 2004; Gill et al., 2008; Beckers et al., 2010; Lu and Vicario, 2014). Ulanovsky (2003, 2004) suggested that the underlying mechanism encoding sound probability is stimulus-specific adaptation (repetition-induced suppression) to frequent, repeated sounds. However, Gill et al. (2008) showed that auditory responses reflect statistically unexpected events based on experience over longer time scales. It is difficult to separate these mechanisms experimentally because the most often repeated sound is usually the most expected one. This puzzle complicates the interpretation of single-unit studies in animals and is closely related to a long-standing question in auditory EEG research: whether the mismatch negativity (MMN) to an oddball sound is due to adaptation that reduces response to the more common sound, or to a violation of expectation that increases response to the oddball. A

Received Dec. 13, 2012; revised Aug. 30, 2017; accepted Sept. 29, 2017.

Author contributions: K.L. and D.S.V. designed research; K.L. performed research; K.L. and D.S.V. analyzed data; K.L. and D.S.V. wrote the paper.

We thank Charles R. Gallistel for insightful advice on statistics; Mimi L. Phan for critical reading of manuscript; and Tom Ziv, Manda Pierce, Kathleen Yoder, Lillian Yang, and Jianqiang Xiao for assistance with experiments.

The authors declare no competing financial interests.

Correspondence should be addressed to David S. Vicario, Psychology Department, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854. E-mail: vicario@rci.rutgers.edu.

DOI:10.1523/JNEUROSCI.5722-12.2017

Copyright © 2017 the authors 0270-6474/17/3712006-12\$15.00/0

recent MEG study (Todorovic and de Lange, 2012) showed that these two effects are separable in time, suggesting that two distinct mechanisms coexist. At the neural level, recent studies also showed that higher responses to oddball sounds found in auditory cortex of rats cannot be fully explained by repetition-induced suppression (Taaseh et al., 2011, Hershenhoren et al., 2014). The experiments described here study and potentially differentiate these effects at the neurophysiological level.

Even a familiar event may not be expected at a certain moment and the expectation more likely depends on the most recent context of events. Therefore, the brain appears to encode familiarity and expectation independently over different time scales. To investigate these processes, we implemented a novel experimental paradigm: a test sound was repeated during a first phase, presented interleaved with other “context” sounds in a second phase, and, in a final phase, repeated again. The effects of stimulus repetition were quantified by changes in response to the test between the first and final phases. We hypothesized that, when presented among context sounds, the test will be expected or unexpected depending on membership in the same acoustic category as context sounds. Therefore, responses to the test during the context phase may reflect both repetition and surprise effects. After removing the predictable effect of test repetition, any residual responses can be inferred to show the effect of surprise.

We tested our hypothesis by recording single-unit and multiunit activity in two areas of the zebra finch (ZF) auditory forebrain: the caudomedial nidopallium (NCM) and caudolateral mesopallium (CLM), both of which receive inputs from thalamo-recipient field L and thus may correspond to superficial layers of mammalian A1 or to a secondary auditory area (Wang et al., 2010, Theunissen and Shaevitz, 2006). In these areas, neural responses show long-lasting, stimulus-specific adaptation to repetition of specific songs in awake birds (Chew et al., 1995, Chew et al., 1996). These results depend on the large set of learned vocalizations in songbirds that provides a stimulus repertoire of distinct but related sounds. We now demonstrate that, when a natural vocalization is presented as a test stimulus among context sounds, such as songs of a different species, auditory responses are enhanced. Furthermore, the data show that surprise-induced enhancement and repetition-induced suppression do not interact, suggesting that different levels of familiarity and expectation may be encoded independently. This is consistent with, and could help to elucidate, the neural mechanisms of human MMN, which can be observed, not only for simple oddballs, but also for higher-order category violations (Näätänen et al., 2001).

Materials and Methods

Subjects. All animals used in our experiments were adult male ZFs ($n = 29$) bred in our aviary or obtained from the Rockefeller University Field Research Center. Animals were housed on a 12/12 h light/dark cycle in a general aviary, where they could see other birds and hear their vocalizations. Food and water were provided *ad libitum* and all procedures conformed to a protocol approved by the Institutional Animal Care and Use Committee of Rutgers University.

Surgery. In preparation for electrophysiological recording, each animal was anesthetized with isoflurane (2% in oxygen) and placed into a stereotaxic apparatus. Marcaine (0.04 cc, 0.25%) was injected under the scalp to provide local analgesia, the skin was incised, and a small craniotomy exposed the area of the bifurcation of the midsagittal sinus. Dental cement was used to attach a metal post to the skull rostral to the opening and to form a chamber around the recording area. The chamber was then sealed with silicone elastomer (Kwiksil; World Precision Instruments). To relieve postsurgical pain, Metacam (0.04 cc, 5 mg/ml) was adminis-

tered intramuscularly. Anesthesia was discontinued and the bird was allowed to recover under a heat lamp.

Electrophysiology. Two days after initial surgery (to allow for full recovery from anesthesia), electrophysiological recordings were made in a walk-in soundproof booth (IAC). The awake animal was immobilized in a comfortable tube and the implanted post was used to fix the head to a stereotaxic frame. Recordings were made at 16 sites, 4 each in the left and right NCM and 4 each in the left and right CLM (Fig. 1A), using glass insulated platinum/tungsten microelectrodes (2–3 M Ω impedance) independently advanced by a multielectrode microdrive (Ekhorn design; Thomas Recording). Electrode signals were amplified (19,000 \times) and band-pass filtered (0.5–5 kHz) and then acquired at 25 kHz using Spike 2 software (CED). White noise stimuli with the amplitude envelope of canary song were presented to search for responsive sites typical of the auditory forebrain. Once all electrodes were placed at responsive sites, stimulus playback experiments were performed. At the end of the recording, eight small electrolytic lesions (20 μ A for 15 s) were made to enable histological reconstruction of recording sites.

Histology. At the conclusion of the experiment, the animal was killed with an overdose of Nembutal, then perfused with saline and paraformaldehyde. Sagittal sections were cut from the fixed brains at 50 μ m on a Vibratome and then stained with cresyl violet. Lesion sites in NCM and in CLM were confirmed histologically based on cytoarchitectonic landmarks.

Auditory stimuli and experimental design. All sound stimuli consisted of natural ZF (conspecific) and canary (heterospecific) songs, which differ in their acoustic characteristics (Fig. 1B–D). Neurons in NCM are known to respond differently to these two types of songs (Chew et al., 1996). Stimuli lasted 0.77–1.21 s and were presented at 65 dB SPL (A scale). All experiments followed a similar protocol, which consisted of three phases: preadapting, context-modulated, and postcontext (Fig. 1E). In the preadapting phase, a test song (e.g., a novel ZF song) was repeated 20 times at a fixed interstimulus interval (ISI) of 7 s to establish initial adaptation. The ISI that we used is much longer than the maximum ISI for inducing forward suppression (several hundred milliseconds; Brosch and Schreiner, 1997) and stimulus-specific adaptation in rodents (up to 2 s; Ulanovsky, 2003, 2004). In the immediately following context-modulated phase, the test stimulus was again presented 19 times, but now in the context of other stimuli in random order (all at 7 s ISI). Finally, in the postcontext phase, the test song was again presented for 20 trials at a fixed 7 s ISI. Responses to the test song in the preadapting phase were used to compute an adaptation function (described below), the slope of which was used to estimate subsequent responses. Comparison of actual and estimated responses to the test song in the context-modulated phase was used to quantify the effects produced by context manipulations (detailed methods are described below).

In Experiment 1, three different context conditions were assessed in 11 birds. In the first condition, the canary context (Fig. 1E), the context stimulus set consisted of 19 repeats of a preadapted ZF test song and 20 repeats of each of seven novel canary songs, for a total of 159 trials presented in randomly shuffled order. Although the ISI was fixed at 7 s, the intervals between repeats of the one test song varied from 14 to 161 s. In the second condition, the silence context, all canary songs were replaced by silence, whereas the order and intervals were the same as in the first condition, effectively creating variable silences of 14–161 s between the onsets of test songs. In the third condition, the ZF context, the test song was played in the context of seven other novel ZF songs and the order of stimulus presentation was randomized, as in condition 1. Each subject was tested with all three conditions in random order and each condition used a different test song that was novel for the bird.

In Experiment 2, a new condition that reversed test and context song types, canary song in ZF context, was tested in six birds, together with the two conditions from Experiment 1: canary context and ZF context. All stimuli were novel songs as in Experiment 1. For canary song in ZF context, the test song was a novel canary song and ZF songs were used as context stimuli in the context-modulated phase. All other aspects of presentation and data analysis were the same as described above.

In Experiment 3, we tested whether the prior familiarity of the context songs played in the context-modulated phase influenced context effects

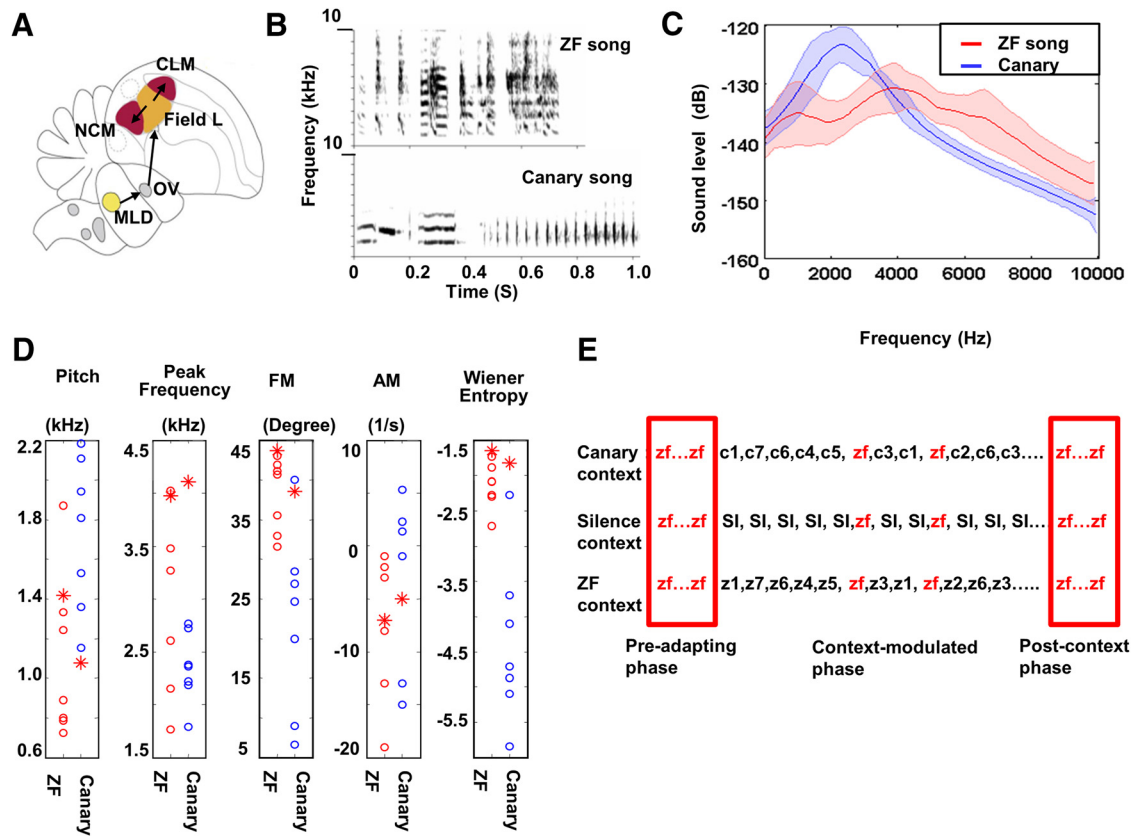


Figure 1. Avian auditory pathway, acoustic features of song stimuli, and experimental protocol. **A**, Ascending auditory pathways in songbirds are indicated by arrows. Auditory nuclei of avian hindbrain innervate MLD (homolog of the inferior colliculus). MLD innervates OV (homolog of the medial geniculate). OV projects to forebrain field L2 (orange; analog of lay IV of A1). Field L2 innervates L1, L3, NCM, and CLM (red; analogs of superficial layers of A1) (modified from Fig. 1 in Theunissen and Shavezit, 2006). **B**, Example spectrograms of a ZF song (top) and a canary song (bottom). Songs of the two species differ both in their spectra and temporal structure. **C**, Spectral envelopes of canary and ZF songs used as context stimuli in these experiments. **D**, Differences between canary and ZF songs used as context stimuli in these experiments quantified in five acoustic dimensions. Each vertical panel represents measurement in one dimension: pitch, peak frequency, frequency modulation, amplitude modulation, and Wiener entropy (a measure of the width and uniformity of the power spectrum). In each panel, the test song is shown as a red asterisk for comparison with ZF context songs (red circles) and canary context songs (blue circles). Values for ZF and canary songs overlap somewhat in each dimension but differ in central tendency and/or variance. The test song (red asterisk) in canary context is at the edge or outside of the distribution of canary songs in four of five dimensions. Specialized software was used to extract and quantify the different acoustic features (SAP2011; Tchernichovski et al., 2000). **E**, Schematic of the stimulus presentation sequence in Experiment 1. Each row represents one test condition. The stimulus order in the preadapting phase and the postcontext phase is the same for all conditions (red boxes, 20 repeats of the test stimulus), but differs in the context-modulated phase, as follows. Canary context: ZF test song in the context of 7 canary songs (c1 to c7) in shuffled order; silence context: ZF song in silent context in shuffled order; context stimuli were replaced by same duration of silence; ZF context: ZF song in the context of 7 ZF songs (z1 to z7) in shuffled order.

on responses to the test song. Seven birds were first tested with the ZF context condition (ZF context condition), in which the context songs were novel ZF songs exactly as in Experiment 1. Then, 50 repeats of the same 7 context songs were presented to the animal in shuffled order at 7 s ISI (350 stimuli total). The animals were again tested with the ZF context condition using a novel ZF test song and the now-familiar context songs in the context-modulated phase. Context effects on responses to the test song in the ZF context session before the familiarization training with the context songs were compared with context effects in the second ZF context session that used familiar context songs.

Data analysis: single units. Single units with spikes >3 SDs from the baseline were isolated from the electrode recordings offline using template-based digital clustering algorithms implemented in Spike2 software (CED). Single units were validated by analysis of the interspike interval (ISpI) histograms. To be accepted, a unit had to have a contamination rate (ISpIs <2 ms, corresponding to spike rates >500 Hz) <2%. The response amplitude of each unit was quantified as the spike rate in the response window (from stimulus onset to stimulus offset plus 100 ms) minus the spike rate in the 500 ms period preceding stimulus onset on each trial.

Data analysis: multiunit activity. Because the spikes of a single unit typically represent only ~10% of all multiunit spikes (that crossed a threshold) at each recording site, we not only report single-unit data, but

also multiunit data in parallel, to capture the activity of nonisolated neurons. For each channel, the root-mean-square (RMS) of the multiunit neural activity was calculated both over a baseline window (the 500 ms period before stimulus onset) and over a response window (from stimulus onset to stimulus offset plus 100 ms) on each trial. The RMS provides a method of rectifying the multiunit activity and computing its average power. Because our multiunit recordings typically were band-pass filtered (0.5–5 kHz), the RMS primarily measured action potentials (not LFPs or EEGs). Responses to song stimuli were quantified as the difference between the baseline RMS and response RMS measurements (Fig. 2A). A site was excluded if its response to the test song in the preadapting phase was not significantly different from the baseline. The baseline RMS was analyzed separately for comparison across the three phases of the experiment.

Effects of context modulation measured as delta-surprisals. The effect of different context manipulations on auditory responses was measured by quantifying how each response (single-unit spike rates or multiunit RMS) during the context-modulated phase deviated from the responses estimated from the responses in the preadapting phase and the postcontext phase. This was computed as the “surprisal,” a measure from information theory (Levy, 2008), according to the following procedure. First, the linear regression line for the responses in the preadapting phase was computed from the responses to the repeated “test” song during the

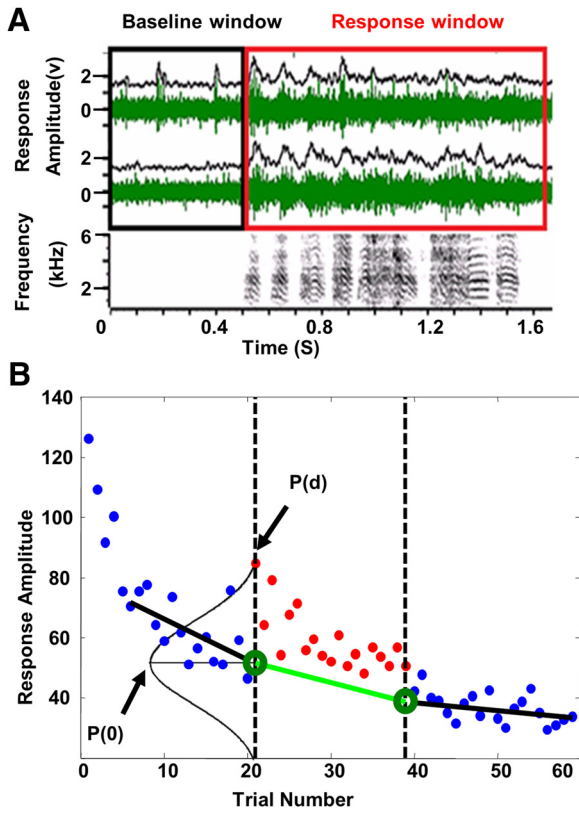


Figure 2. Response measurement and calculation of context effects. **A**, Example of electrophysiological recordings and multiunit RMS computation. For a single trial, the spectrogram of the song stimulus (bottom) and the multiunit activity recorded from two electrodes (green traces) is shown together with the moving RMS (window = 50 ms, only for display purposes) calculated for each recording (black traces). Single units were extracted from these recording. Responses to each song were defined as the difference between the spike counts (for single units) or RMS (for multiunits) during the response window (stimulus period plus 100 ms) minus the spike counts or RMS during the baseline window (500 ms preceding stimulus onset). **B**, Illustration of delta-surprisal computation. Each dot in the figure represents a response to the test song on one trial. Black dashed lines indicate the borders between the three phases. Solid black lines are regression lines obtained from responses (blue dots) in the preadapting and postcontext phases. Solid green line connecting the end points of these regressions (green circles) shows the linear estimate of the mean of the expected distribution of responses for each trial during the context-modulated phase if the context had no effect on the course of adaptation of the response to the test song. The extent to which the observed responses in the context-modulated phase (red dots) are unexpected is measured by their surprisal, which is $1/\log[P(d)]$, normalized by the surprisal of the expected response, which is $\log[1/P(0)]$. Note that the expectation of the signed surprisals is 0 (see text for further details).

linear portion of the adaptation function (Fig. 2B; trials 6–20, black line). This line was extrapolated to estimate the response on the first trial of the context-modulated phase (Fig. 2B, green circle at trial 21). Second, a second regression line was computed from responses to the test stimulus in the postcontext phase (trials 40–59, black line) and then extrapolated backwards to estimate the expected response on the last trial (Fig. 2B, green circle at trial 39) of the context-modulated phase. Third, the expected responses in the context-modulated phase were estimated by the line connecting the estimates for trials 21 and 39 (Fig. 2B, green line, called hereafter the interpolated regression). Fourth, the expected SD of the responses around the interpolated regression line was estimated by the SD of pooled residuals of the regressions of the pre and the post phases. Fifth, an observed response that falls on the interpolated regression line is the least surprising (most expected) response; the greater the deviation (d) of an observed response from this expectation, the more surprising it is. The degree to which it is surprising is a function of the probability of a deviation of magnitude d , namely, $\log[1/(P(d))]$, where $P(d)$ is the probability density of d in the assumed-to-be normal distri-

bution (Fig. 2B). Although the adaptation function during the context not be strictly linear, the interpolated linear regression line is a conservative estimate; if an exponential fit were used, then the observed deviations in the context block would be even greater. Sixth, the average magnitude of a surprisal is greater for distributions with large SDs than for narrower distributions, so the surprisals were normalized by subtracting the absolute value of the minimum surprisal, which is $\log[1/P(0)]$. Normalization makes the surprisal of an observed response that exactly conforms to an expectation equal to zero and it zeroes the expectation of the signed surprisals when observed responses are drawn from the expected distribution. Seventh, responses greater than expected are assigned positive surprisal, whereas responses less than expected are assigned negative surprisal. Eighth, therefore, our formula for the normalized signed delta-surprisal of an observed response is as follows:

$$\text{Delta surprisal} = \text{sign}(d) * \left(\log \frac{1}{P(d)} - \text{abs} \left(\log \frac{1}{P(0)} \right) \right)$$

where d is the deviation of the response from expectation.

Effects of stimulus repetition measured as an adaptation index. To test whether the modulations interact with stimulus-specific adaptation, we also computed an adaptation index for each condition at each recording site by dividing the response amplitude of the first test stimulus trial of the postcontext phase (trial 40) by the response on the last test stimulus trial of the preadapting phase (trial 20). This ratio provided an estimate of adaptation that occurred over the context phase (Fig. 2B). Notice that adaptation indices are different from the stimulus-specific adaptation indices used by Ulanovsky et al., 2003, in that adaptation indices reflect a reduction of neural responses to the same sound over repetitions, whereas stimulus-specific adaptation indices reflect differences in responses between oddball sounds and standard sounds. The same procedure and calculations were used for both single units and multiunit data.

Quantification of neurons' selectivity to the test song and context songs. To test whether the selectivity of neurons to the test sound and context sounds affected the enhancement effect, we also quantified D' for each multiunit site. D' measures the selectivity for one stimulus (A) over another stimulus (B) at each recording site and was calculated by the following formula (as described in Solis and Doupe, 1997):

$$D' = \frac{\text{Mean of Responses to A} - \text{Mean of Responses to B}}{\sqrt{\text{Variance of Responses to A} + \text{Variance of Responses to B}}}$$

A positive D' means that the neuron prefers stimulus A in its responses. To calculate D' , we first took the mean and variance of responses (multiunit RMS) to each song at each site (obtained from last 10 trials of the preadapting phase for the test song and first 10 trials of each context song in the context-modulated phase). D' for each test song with respect to the seven context songs was then calculated to produce seven D' values for the context-modulated phase, which were then averaged for each site.

Temporal profile of responses. We analyzed temporal characteristics of the responses seen with context manipulations by computing the difference between the averaged temporal waveform of responses to the test stimulus in the preadapting phase and in the context-modulated phase across all sites. We used the following procedure. First, we computed the moving average RMS (10 ms window) of the multiunit recording to produce a smoothed RMS waveform of the response to each stimulus at each site. Then, we computed averages of these waveforms both across the last six trials of the preadapting phase and across the first six trials of the context-modulated phase. Finally, averages of these waveforms were computed across all recording sites from all birds separately for each phase. These grand averages effectively eliminated response patterns due to characteristics of specific stimuli and/or specific recording sites. Therefore, the difference waveform between these two grand average waveforms shows the temporal profile of response enhancement caused by the surprise phenomenon.

Statistical methods. Data are graphed both as cumulative frequency distributions, which reveal the details of condition effects, and as

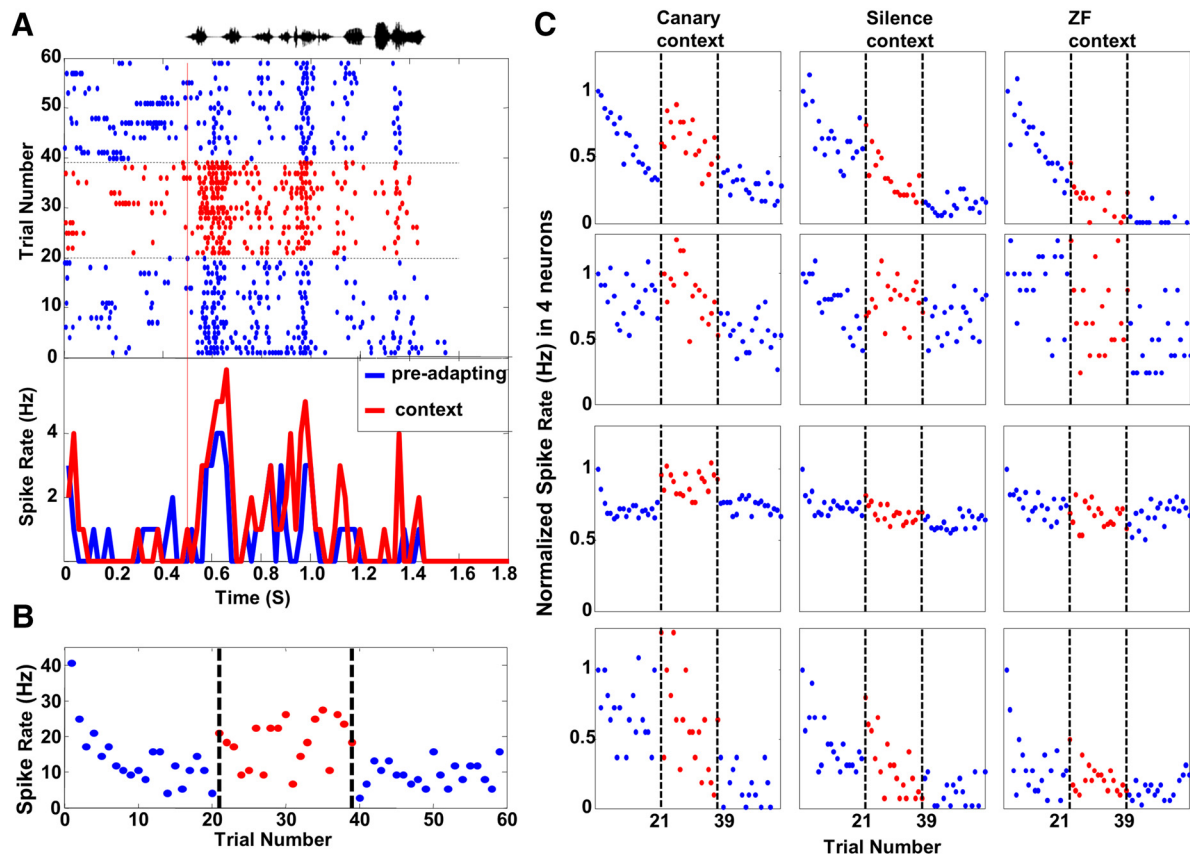


Figure 3. Single-unit responses to the test stimulus are increased in the context condition. **A**, Single-unit activity recorded in the canary context condition. The top trace (black) shows the amplitude waveform of the test song. The middle panel shows the raster plot of spike responses in the canary context condition for all three phases (blue dots for spikes in the preadapted and postcontext phase and red dots in the context-modulated phase), with trials ordered from bottom to top. The red vertical line indicates the stimulus onset. The horizontal dashed lines indicate the change between phases. Spiking increased during the context-modulated phase. The bottom panel compares the PSTH obtained from the preadapted phase (blue) and context-modulated phase (red). **B**, Plot of spike rates of the neuron shown in **A**. Spike rates are shown for the preadapting phase and the postcontext phase (blue dots) and for the context-modulated phase (red dots). Decreasing responses in the first phase demonstrate adaptation. Increased activity in the context-modulated phase demonstrates the enhancement effect (cf. Fig. 2B). **C**, Examples of spike rate plots for four other neurons (four rows) in three context conditions (three columns). Spike rates for each trial were normalized by the spike rate of the first trial in the same condition for convenience of visual comparison in this plot (conventions as in **B**). Responses in the canary and silence contexts clearly increased compared with responses in the preadapting phase, whereas smaller increases are seen in the ZF context. In the postcontext phase, responses resumed at the original adapted level or below.

conventional mean and SE plots. The distribution of samples in some conditions did not fully satisfy criteria for parametric tests. Therefore, appropriate nonparametric statistics were used throughout whenever possible. For Experiment 1, the delta-surprisals obtained for each recording site in each of the three conditions (across all three experiments) were treated as three repeated measures. The main effect was tested by the nonparametric Friedman's ANOVA, which does not require a normal data distribution. Differences between groups in which data were matched (e.g., different conditions recorded at the same electrode site) were tested by the Wilcoxon matched-pairs test. For group data in which samples were not explicitly matched, we used the Kolmogorov–Smirnov two-sample test. To quantify possible differences in the main effects between NCM and CLM, the interaction between regional difference and conditions was tested using repeated-measures ANOVA, in which region was treated as a factor and the delta-surprisals of the three conditions were repeated measures.

Results

Effects of context manipulations on auditory responses to a ZF song

Experiment 1 measured the effects of manipulating the acoustic and temporal context on responses to a preadapted test song. We obtained 68 isolated single units and 111 multiunit sites from brain regions NCM and CLM in 11 birds, each tested with the three different context conditions. Our measure of context effects

(delta-surprisal; see Materials and Methods) showed no significant differences between NCM and CLM in two-way repeated-measures ANOVAs (single units: $F_{(1,66)} = 0.29$, $p > 0.591$; multiunit: $F_{(1,109)} = 0.85$, $p > 0.357$) and no interaction between brain regions and context conditions (single units: $F_{(1,66)} = 1.62$, $p > 0.201$; multiunit: $F_{(1,109)} = 0.32$, $p > 0.728$). Therefore, data from NCM and CLM were combined for further analyses.

For single-unit data, an increased firing rate was seen in the context-modulated phase relative to the preadapting phase in raster plots and PSTHs (an example is shown in Fig. 3A) and in the plot of spike rates by trials (Fig. 3B,C). When the increased activity was quantified as delta-surprisals, there were significant differences between the 3 context conditions tested [Fig. 4A, Friedman's ANOVA, χ^2 ($n = 68$, $df = 2$) = 23.4, $p < 0.001$]. Most neurons in the canary context condition (72%, 49/68) and the silence condition (68% 46/68) showed positive delta-surprisals, indicating an increased firing rate during the context-modulated phase. In contrast, less than half (43%, 29/68) showed positive delta-surprisals in the ZF context. Further tests showed that delta-surprisals in the canary context were significantly larger than in the silence context (Wilcoxon, $z = 2.12$; $p < 0.034$), which in turn were significantly larger than ZF context (Wilcoxon, $z = 2.68$; $p < 0.007$).

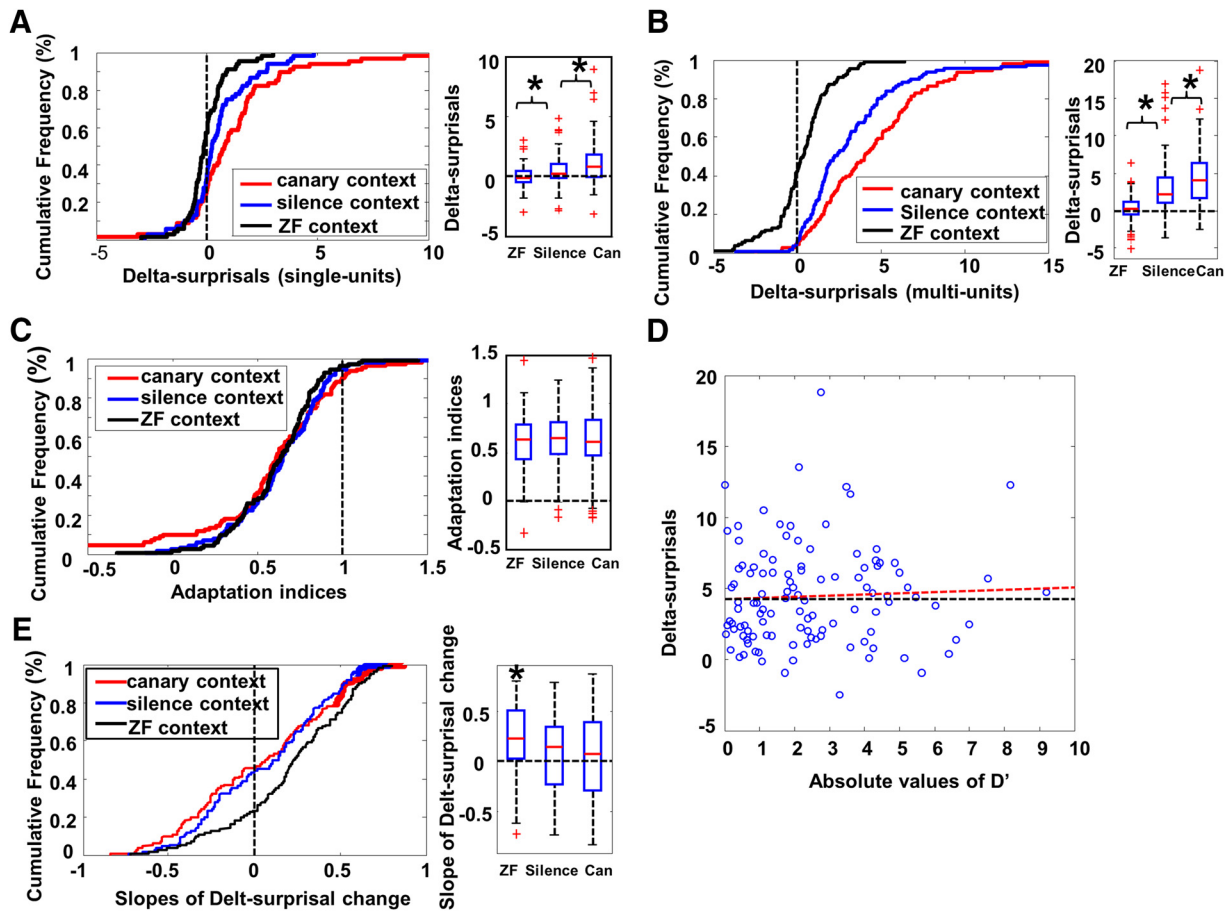


Figure 4. Response enhancement in the context-modulated phase measured as delta-surprisals. **A**, Cumulative frequency distributions of delta-surprisals obtained for single units in Experiment 1. The cumulative frequency for each x value as a percentage of the total is shown on the y -axis as a function of delta-surprisal values on the x -axis. Delta-surprisals for the canary (red), silence (blue), and ZF (black) context conditions are plotted. The box plot (right) shows the minimum (the bottom of the whisker), first quartile (the bottom of the box), median (the red line in the middle of the box), third quartile (the top of the box), and maximum (the top of the whisker) and possible outliers (red cross) of delta-surprisals in the ZF context condition (ZF, left box), the silence context condition (silence, middle box), and the canary context condition (Can, right box). **B**, Cumulative frequency distributions (left) and box plot (right) of delta-surprisals obtained for multiunits in the three conditions of Experiment 1. Multiunit data showed the same pattern of results as single-unit data. **C**, Cumulative frequency distributions (left) and box plot (right) of adaptation indices for the three conditions tested in Experiment 1. Adaptation indices were not different from each other in the three conditions. **D**, Scatter plot of absolute values of D' and delta-surprisals. The black dashed line is a flat reference line (slope zero). Red dashed line shows the best-fitting line from a linear regression, which shows no significant correlation between the two variables (Spearman's $r = 0.056$, $p > 0.557$). **E**, Cumulative frequency distributions (left) and box plot (right) of slopes of linear regression between trial number and trial-by-trial delta-surprisals in the three conditions of Experiment 1. In the canary and silence context conditions, there was no significant change in delta-surprisals over trials. In contrast, delta-surprisals increased significantly with trials in the ZF context condition.

When multiunit RMS were analyzed, even greater differences across the three conditions were observed [Fig. 4B, Friedman's ANOVA, χ^2 ($n = 111$, $df = 4$) = 129.9, $p < 0.001$]. More than 95% (106/111) of multiunit sites in the canary context and >94% (105/111) of multiunit sites in the silence context condition showed positive delta-surprisals, whereas only ~58% (64/111) of multiunit sites in the ZF context condition showed positive delta-surprisals. Delta-surprisals in the canary context were significantly larger than in the silence context condition (Wilcoxon $z = 3.13$; $p < 0.001$), which in turn were significantly larger than in the ZF context condition (Wilcoxon $z = 8.19$; $p < 0.001$), the same pattern as that seen for single units. It should be noted that the ISI for the test song during the context-modulated phase was variable and much longer (an average of 56 s) than in the pre-adapting phase (7 s). Although this longer effective ISI might contribute to the observed enhancement if it is considered a decay of adaptation, as in the silence condition with no intervening stimuli, this effect cannot explain the weaker enhancement in the ZF context or the higher enhancement in the canary condition. In all three cases, the ISI for the test song is equally long and variable,

but the enhancement differs in opposite directions for the two different types of context stimuli.

Baseline remains constant across three phases

Because multiunit responses to song stimuli were quantified as the difference between the baseline RMS and response RMS measurements, we also compared the baseline RMS both across the three context conditions and across three phases (preadapting phase, context-modulated phase, and postcontext phase) in a two-way repeated-measures ANOVA. We did not find significant changes in baseline activity across conditions ($F_{(2,660)} = 0.35$, $p > 0.706$) and across phases ($F_{(2,660)} = 0.45$, $p > 0.636$) or any interaction ($F_{(4,660)} = 1.71$, $p > 0.145$). Therefore, enhancement effects in the context-modulated phase were not due to changes in baseline activity.

Enhancement effects are independent from adaptation

Our experiment used the predicted trajectory of stimulus-specific adaptation as a baseline against which to calculate changes (as delta-surprisals showing enhancement) resulting from context manipu-

lations. Therefore, to fully interpret the results, it is essential to know whether the enhancement observed interacts with the adaptation process or is independent from it. For example, if they interact, then the enhancement effect might reduce or prevent the adaptation normally produced by presenting the same test song 19 times during the context-modulated phase (cf. precontext phase adaptation seen in Fig. 3C). If this is the case, then we predict that the context that produces the largest enhancement effect should also show the weakest adaptation. We measured adaptation (the drop in response amplitude) over the context-modulated phase as the adaptation index (see Materials and Methods). When we compared adaptation indices across the three conditions, the data showed no significant differences in either single-unit data [Friedman's ANOVA χ^2 ($n = 68$, $df = 2$) = 4.35, $p > 0.11$] or multiunit data [Fig. 4C, Friedman's ANOVA χ^2 ($n = 111$, $df = 2$) = 0.16, $p > 0.922$] and thus showed no relationship to the degree of enhancement across conditions. To examine this in more detail, we calculated the correlation coefficient (Spearman's rho) between the adaptation indices and the delta-surprisals for each site within each condition and each bird. Of the 33 correlations assessed (3 conditions \times 11 birds) 31 were not significant ($p > 0.05$). The two correlations that were significant (one negative and one positive) were not associated with any one bird or condition. Therefore, there was no systematic relationship between adaptation and the enhancement effect and we conclude that the two processes are independent from each other.

Enhancement effects are not associated with response selectivity between the test song and context songs

Previous work (Ulanovsky et al., 2003) showed that the effect of sound probability on auditory responses was positively correlated with the frequency separation between the standard tone and the oddball tone. If the enhancement effect were also affected by the spectral differences between the test song and the context songs, then we would expect that the larger the difference between responses to the test song (ZF song) and context songs (canary songs), the higher the enhancement effect that would be seen on a given neuron. To test this hypothesis, we calculated the selectivity of neurons to the test sound and context sounds, quantified as D' . The bias of neural responses to a tested ZF song or context canary songs was reflected by the absolute value of D' . We calculated the correlation between the absolute values of D' and the delta-surprisals from all multiunits for the canary context condition. Surprisingly, we did not find significant correlation between the absolute values of D' and the enhancement effect (Spearman's $r = 0.056$, $p > 0.557$, Fig. 5D). Therefore, enhancement effects cannot be fully explained by neurons' selective tuning toward ZF songs or canary songs.

Enhancement effects analyzed trial-by-trial

The enhancement effect described so far was quantified as delta-surprisals averaged across trials for each site. We further tested whether delta-surprisals increase or decrease with repetition of the test song in the context. For each multiunit site, we calculated the linear regression between trial number of the test song in the context and the delta-surprisal on each trial. Then, we pooled slopes of the regression from all multiunits and analyzed whether the slopes in any condition were significantly higher or lower than zero. As shown in Figure 5E, there were no significant increase or decrease in delta-surprisals with trials both in the canary context (Wilcoxon, $z = 0.94$; $p > 0.346$) and the silence context (Wilcoxon, $z = 0.22$; $p > 0.828$). In contrast, delta-surprisal in

the ZF context increased significantly with trials (Wilcoxon, $z = 5.74$; $p < 0.001$). This increase in delta-surprisals with trials in the ZF context may be due to the increasing familiarity of the context songs, which will be further discussed with the results of Experiment 3.

Temporal pattern of enhancement effects

In addition to quantifying the overall enhancement of responses by context manipulations, we examined the timing and waveform of the responses in the context-modulated phase to explore the possible mechanism of the enhancement. First, if the enhancement effect does not interact with adaptation (as shown above), then we would expect that the latency (measured from stimulus onset) of the enhanced component of the responses would be longer than for the auditory response itself because the surprisal effect may reflect top-down modulations. Second, in the silence context, only temporal uncertainty contributes to the enhancement effect (there are no intervening sounds, so no acoustic discrimination is needed), so we expected that the enhancement profile for this condition would have a shorter latency than in the canary context condition.

For each of the three conditions, the enhancement profile was computed as the difference between the averaged multiunit RMS waveforms between the context-modulated phase versus the preadapting phase across all multiunit sites ($n = 111$) (Fig. 5A, black trace). Using the mean of multiunit data effectively averages out temporal features associated with any specific stimulus or the tuning properties of individual neurons. The resulting "enhancement profile" shows the timing of enhancement due to modulation by the different acoustic contexts across all sites recorded. We found no differences between NCM and CLM in the enhancement profile for any condition, so data from the two brain areas were combined. First, for the canary context, we compared the latency of averaged multiunit RMS in the context-modulated phase (Fig. 5B, red trace) with the latency of the enhancement profile (Fig. 5B, black trace). The latency for the two waveforms was computed as the time from stimulus onset until the signal crossed a threshold, computed as the maximum value of the 99% confidence interval for each signal during the baseline window. For the canary context condition (Fig. 5C, red trace), the latency of the first increase in the enhancement profile was longer than the latency of the multiunit RMS waveform (27 vs 6 ms, respectively). The peak also occurred later (112 vs 85 ms, respectively), as shown in Figure 5D. In contrast, the ZF context condition showed no consistent change in the timing of the enhancement profile (Fig. 5C), which is consistent with the analysis based on delta-surprisals.

The enhancement profiles in the silence context condition differ from the enhancement profiles with in the canary context condition in two ways. First, the latency of enhancement in the silence context was 20 ms (shorter than in the canary context condition by 7 ms; Fig. 5D). To quantify the effect of this latency difference statistically, we compared response amplitudes in the window 10–20 ms after stimulus onset between the preadapting phase and the context-modulated phase for both the canary context and the silence context. We observed significant enhancement in this time window in the silence context (Wilcoxon $z = 2.8$; $p < 0.006$, Fig. 5F, left box), but not in the canary context (Wilcoxon $z = 1.5$; $p > 0.123$, Fig. 5F, right box). This observation supports the idea that early enhancement in the silence context condition is due to uncertainty about when the stimulus will occur, resulting in a very rapid detection of stimulus onset. In contrast, acoustic processing needed to detect a violation of the

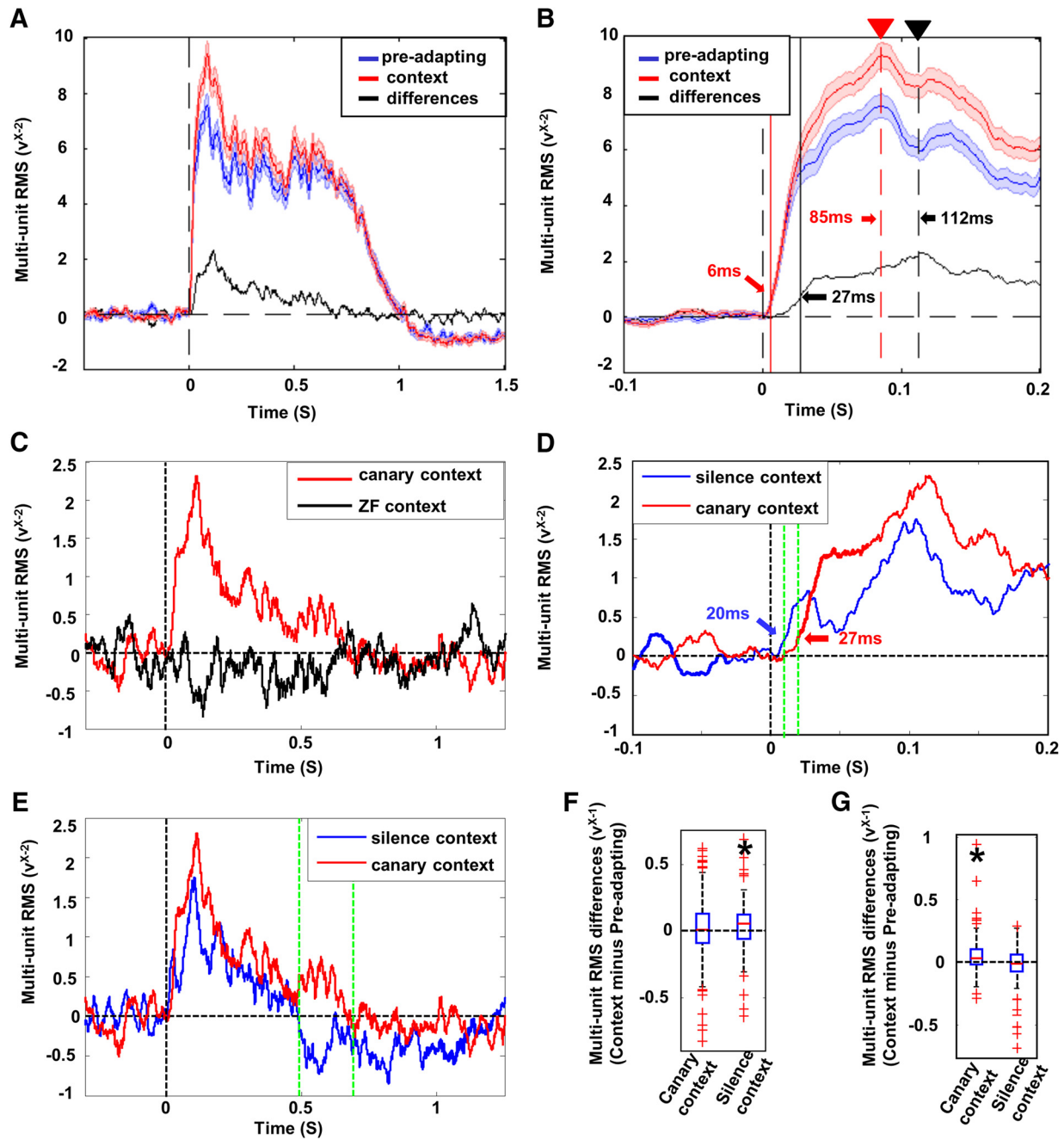


Figure 5. Response latency and timing of response enhancement. **A**, Enhancement profile (black trace) for the canary context is displayed as the difference between the averaged response waveforms for the test song in the preadapting phase (blue trace) and the context-modulated phase (red trace). Shaded areas along the traces show SEs. The vertical dashed line indicates the onset of the stimulus. **B**, Early phase of the averaged response waveforms and the enhancement profile for the canary context condition from **A** but at a finer time scale. The solid vertical lines indicate the latency of responses from stimulus onset (red) and the latency of the enhancement profile (black; 21 ms later than the red line). The latency was computed as the time from stimulus onset until the signal crossed a threshold, computed as the maximum value of the 99% confidence interval for each signal during the baseline window. Dashed red line with triangle indicates the peak of the response (85 ms after stimulus onset). The arrow and dashed black line with triangle indicate the peak of the enhancement profile (112 ms after stimulus onset). Both the latency and the peak of the enhancement profile occur later than the latency and the peak of the responses. **C**, Comparison between the enhancement profiles for the canary context condition (red trace) and for the ZF context condition (black trace). The ZF context condition shows no clear enhancement effect (note: gain is higher than in **A**). **D**, Comparison of the early phase of the enhancement profiles in the canary context condition (red trace) and in the silence context condition (blue trace) showing a shorter latency of the enhancement effect in the silence context condition. The green dashed lines indicate the window 10–20 ms after the stimulus onset, in which significant enhancement was observed in the silence context condition, but not in the canary context condition (also see **F**). **E**, Comparison between the enhancement profiles in the canary (red trace) and the silence context condition (blue trace) showing a shorter duration for the enhancement profile in the silence context condition. The green dashed lines indicate the window 490–690 ms after the stimulus onset, when the profile for the silence context condition decays to zero but the enhancement effect is still significant for the canary context condition (also see **F**). **F**, Box plot of multiunit RMS differences between the preadapting phase and the context-modulated phase in the window 10–20 ms after the stimulus onset (between dashed green lines in **D**) was compared between the canary context and the silence context. There was a significant increase in response amplitude in this time window in the silence context (left box; Wilcoxon $z = 2.8$; $p < 0.006$, $n = 98$), but not in the canary context (right box; Wilcoxon $z = 1.5$; $p > 0.123$, $n = 98$). **G**, Box plot of multiunit RMS differences between the preadapting phase and the context-modulated phase in the window 490–690 ms after the stimulus onset (between dashed green lines in **E**) was compared between the canary and silence contexts. There was a significant increase in response amplitude in this time window in the canary context (left box; Wilcoxon $z = 6.90$; $p < 0.001$, $n = 98$), but not in the silence context (right box; Wilcoxon $z = -1.24$; $p > 0.216$, $n = 98$).

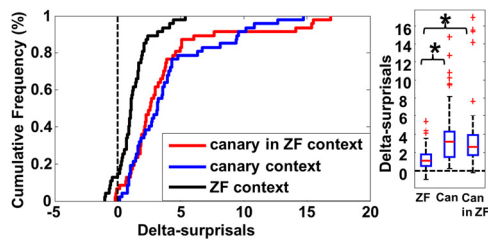


Figure 6. Delta-surprisals for the three conditions tested in Experiment 2. Cumulative frequency distributions and box plot show that delta-surprisals in the canary context condition (Can) and the canary in ZF context condition (Can in ZF) were not different from each other, but both were different from the ZF context condition (ZF).

acoustic context in the canary context condition requires more time. Second, the enhancement profile in the silence context had a shorter duration than the canary context profile, with a decay to zero 181 ms earlier than in the canary context (Fig. 5E). In the silence context condition, differences in responses between the context-modulated phase and the preadapting phase become undetectable in the window 490–690 ms after stimulus onset (Wilcoxon $z = -1.24$; $p > 0.216$, Fig. 5G), whereas, in the canary context condition, enhancement effect was still significant in the same window (Wilcoxon $z = 6.90$; $p < 0.001$). This is consistent with the idea that processing of acoustic features in the canary context continues during the evolving stimulus, whereas temporal uncertainty in the silence condition is largely detected at stimulus onset.

Canary test songs in the ZF context showed similar enhancement

To test whether enhancement is due to the fact that the context stimuli are canary songs and thus from a different species than the ZF subjects, Experiment 2 implemented the canary in the ZF context condition, in which the test song is a canary song and the context songs are novel ZF songs, as well as two conditions previously described: the canary context (ZF test song in canary contexts) and the ZF context (ZF test song in ZF context). In this experiment, 47 multiunit sites recorded in NCM and CLM in six birds were analyzed. Cumulative frequency distributions and the mean delta-surprisals for each condition are shown in Figure 6. There were significant differences in delta-surprisals across the three conditions [Friedman's ANOVA χ^2 ($n = 47$, $df = 2$) = 30.9, $p < 0.001$]. Both the canary in ZF context and the original canary context condition showed delta-surprisals significantly higher than those of the ZF context condition (Wilcoxon tests: canary in ZF context condition: $z = 4.31$; $p < 0.001$, canary context condition: $z = 5.09$; $p < 0.001$), but canary in ZF context and canary context conditions were not significantly different from each other (Wilcoxon tests: $z = -0.63$; $p > 0.526$). This result suggests that the strong enhancement effect seen in the canary context condition, described in Experiment 1 and 2, was due to the violation of ongoing expectations set up by context stimuli, rather than to a preexisting bias for conspecific songs.

Familiarity of context songs increases the enhancement effect

In Experiment 3, 67 multiunit sites recorded in NCM and CLM in seven birds were analyzed. This experiment tested the effect of the familiarity of the context songs on context-induced enhancement in the ZF context. First, the ZF context condition was tested with novel context songs. Then, these same context songs were repeated 50 times each to make them very familiar and the ZF context condition was tested again now with a novel test song and

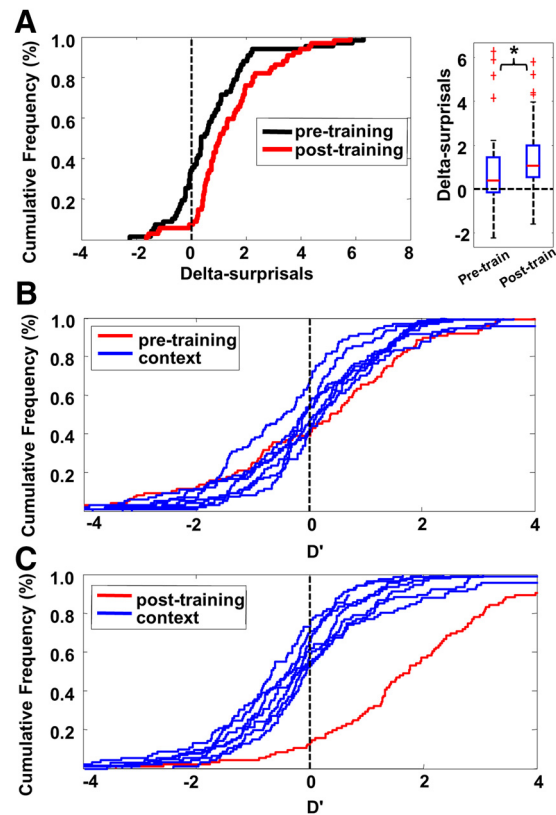


Figure 7. Effects of preexposure to context stimuli tested in Experiment 3. **A**, Cumulative frequency distributions and box plot of delta-surprisals for the ZF context condition tested before and after extensive exposure to context songs in Experiment 3. Delta-surprisals increased when context songs were familiar due to prior exposure. **B**, Cumulative frequency distributions of D' obtained from the context songs (blue curves) and the test song (red curve) before exposure training are not distinguishable. **C**, Cumulative frequency distributions of D' obtained from the context songs and the test song after exposure training are significantly different.

the familiar context songs. Cumulative frequency distributions and the mean delta-surprisals for the two ZF context conditions are shown in Figure 7A. The context-modulated delta-surprisals obtained in the ZF context condition with familiar context songs were significantly higher than delta-surprisals obtained when the songs were novel (Wilcoxon $z = 3.22$; $p = 0.001$).

The results of Experiments 1 and 2 above showed strong enhancements both for a ZF test song presented in a canary context and for a canary test song in a ZF context. These enhancements appeared to reflect a categorical contrast in stimulus statistics between the test song and context songs from a different species. However, there was no such contrast in Experiment 3 (all stimuli were ZF songs), which showed that the familiarity of context songs increased enhancement significantly. This suggests that familiarity itself, which reduces response strength through stimulus-specific adaptation, might also function as a contrast dimension. To assess this, we calculated the response selectivity as the D' between songs heard in the context-modulated phase both before and after exposure to the same context songs. D' for each test song with respect to the seven context songs was calculated in the same manner as in Experiment 1. In addition, D' values for each context song with respect to the other context songs were calculated. Therefore, we averaged eight D' values (one for the test song and seven for the context songs) for each recording site in one session. These means reflected whether this song elicited responses stronger or weaker than other songs on average. We

found that, before the training, the D' of the test song only differed from D' values of 3 out of 7 context songs (Fig. 7B, tested by Wilcoxon test: $p < 0.05$). In contrast, after exposure to context songs, the D' of the test song differed from D' values of all 7 context songs (Wilcoxon, $p < 0.001$ for all; Fig. 7C). Therefore, prior exposure to context songs increased the contrast in responses between the test song and context songs and this may have made the test song perceptually different from the context songs, producing greater enhancement for the test song in the context phase, as seen in Figure 7A.

Discussion

Our results show that neural responses in songbird forebrain areas can be strongly modulated by the ongoing acoustic context in which a given sound occurs. The degree of modulation depends on the acoustic and temporal characteristics of the context and is independent of ongoing stimulus-specific adaptation. Response enhancement is greatest either when a sound violates acoustic expectations that reflect the category of the recent sound context (canary vs ZF songs) or under conditions of temporal uncertainty (random timing in silence context) and these two factors can interact. Furthermore, prior familiarity of specific context stimuli can create a contrast that induces enhancement, even when the test song shares stimulus statistics with the context.

Enhancement in canary versus ZF contexts reflects expectations for context stimuli

In Experiment 1, canary and ZF contexts showed the largest difference in enhancement. Auditory neurons responded to the test song differently depending on recent exposure to context songs from the same versus different category as the test song. Little enhancement occurred in the ZF context despite the unique features of each context song, which are sufficient to differentiate these conspecific songs during stimulus-specific adaptation (Chew et al., 1995). Apparently, exposure to the acoustic features shared by the context songs produces an expectation that the next song will be from the same category as the context. Test songs that violate that expectation in the canary context are surprising and elicit enhanced responses. This is more complex than a simple oddball effect because it reflects a violation of the expected stimulus category, not simply of the expected stimulus.

Longer latency of the enhancement profile suggests a top-down influence

Analysis of the temporal profile of enhancement showed that the latency of enhancement is longer than for auditory responses (27 vs 6 ms, respectively) and the enhancement peak comes later (112 vs 85 ms, respectively). This implies that the recognition of unexpected events (e.g., acoustic and temporal context violations) requires processing time, consistent with the observation that the peak of MMN follows that of N1, a main ERP component (Näätänen et al., 2005). We hypothesize that the recognition of unexpected stimuli may require a top-down process that takes time and may originate from anatomical areas (cf. Bar, 2003, 2004; Turk-Browne et al., 2009) not assessed in these experiments.

Enhancement in the silence context condition may reflect temporal uncertainty

Significant response enhancement also occurred for the silence context, when only the test song was presented in the context phase but with random intervals. This is unlikely to be simply an

effect of longer ISIs for two reasons. First, enhancement measurements were controlled for adaptation. Delta-surprisals were calculated from the differences between observed responses and estimates of the adaptation trajectory from samples at the end of the first and beginning of the last phase. If long intervals led to recovery from adaptation, then early responses in the last phase would also increase and so the estimated trajectory would also be higher. Second, there was no systematic relationship between delta-surprisals and adaptation, as measured by adaptation indices; therefore, response enhancement did not interact with adaptation. Because there was no recovery from adaptation during the context-modulated phase, enhancement in the silence context is likely due to the temporal uncertainty of test song onset. In addition, the enhancement profile in the silence context had a shorter latency and duration than in the canary context, implying that temporal surprise occurs earlier than discriminating test songs in the canary context. This may be because, in the silence condition, only test song onset needs to be detected, not its acoustic features. Further studies will test whether any subpopulation of neurons is more sensitive to temporal versus acoustic feature surprise.

Stimulus repetition effects are independent from expectation effects

The neural mechanisms that represent the probability of external world events have long been a focus of both human EEG studies and extracellular recordings in animal models (Naatanen, 1995; Ulanovsky et al., 2003, 2004). Studies in the auditory system typically use an oddball paradigm that compares neural responses to a sound when it is an oddball (occurring infrequently and unpredictably) versus when it is common. The response difference could reflect either repetition-induced suppression to the sound when common or surprise-induced enhancement to the rare oddball. Therefore, the mechanisms of MMN and related phenomena at the neuronal level have been debated vigorously (Nelken and Ulanovsky, 2007; May and Tiitinen, 2010; Fishman, 2014). In human MEG studies that carefully controlled both repetition and expectancy, Todorovic and de Lange (2012) and Symonds et al. (2017) showed that both mechanisms could coexist. However, recent extracellular recording studies showed mixed results: although some failed to find surprise-induced enhancement in auditory cortex (Farley et al., 2010; Fishman and Steinschneider, 2012), others suggested that higher responses to oddball sounds in cortical neurons cannot be fully explained by repetition-induced suppression (Taaseh et al., 2011; Hershenhoren et al., 2014; Rubin et al., 2016).

In our paradigm, we separated the effects of repetition and expectancy into three phases, allowing us to determine, in neuronal activity, whether and how surprise effects interact with repetition-induced suppression (aka stimulus-specific adaptation). Our results clearly show that surprise effects for an unexpected stimulus are independent of and do not interact with repetition-induced suppression for the very same stimulus. We also found that enhancement induced by violation of prediction peaks much later than the response itself (112 vs 85 ms after stimulus onset, respectively). Our results not only confirmed recent work that successfully revealed effects of surprise-induced enhancement (Taaseh et al., 2011; Hershenhoren et al., 2014; Rubin et al., 2016), but also demonstrate a way to measure surprise-induced enhancement more explicitly so that it can be studied independently from repetition-induced suppression.

Repetition-induced suppression and surprise-induced enhancement may be two independent neural mechanisms that

represent statistical properties of the sensory environment at different levels. Repetition-induced suppression for passively heard sounds seems to reflect a memory process that encodes the long-term familiarity that underlies recognition of a given sound (Chew et al., 1996, Phan et al., 2006), whereas surprise-induced enhancement independently encodes the expectancy for that sound in the ongoing context. This dual-coding scheme allows expectations to update dynamically while maintaining memory of previous experience. Our recordings provide evidence for both processes and thus demonstrate the value of our novel paradigm for studying the neural mechanism of probability coding, oddball effects, and potentially MMN.

Enhancement is also induced by differential familiarity between context and test stimuli

We showed that hearing a sound at an unexpected time or with unexpected features enhances responses. In addition, we found much greater enhancement in the ZF context when context stimuli were familiar than when they were novel (Fig. 7). Apparently, the familiarity produced by prior repetition of context stimuli (eliciting adaptation) produced an expectation (that the next stimulus would be equally familiar) that was violated when the test song was heard. This suggests that adaptation can change perceptual properties of stimuli and may subserve a form of implicit memory, which in turn may contribute to schema-based auditory scene analysis that increases the chance of detecting novel sounds in a familiar acoustic environment (Corbetta and Shulman, 2002; Lu and Vicario, 2011; Pérez-González and Malmierca, 2014).

Implications

Traditional auditory neurophysiology has mapped acoustic stimulus properties to the response properties of neurons, such as in tonotopic maps. However, growing evidence suggests that auditory responses can also encode the probability of sounds and/or sound transitions (Ulanovsky et al., 2003; Gill et al., 2008; Beckers and Gahr, 2010; Lu and Vicario, 2014). Our current results suggest that neurons can do more than just predict the probability of one sound based on its repetition: they can also represent an expectation for a class of sounds (canary vs ZF songs) based on the statistical similarity (and/or relative familiarity) of varying sounds in the context over at least several seconds. If the incoming sound violates the prediction, then responses are enhanced (after a short processing delay), which may serve to redirect attention to the novel target. As a result, rapid identification of a sound from a new category is achieved. Such a process may contribute importantly to auditory perception in the noisy natural acoustic environment. Our results are consistent with recent studies showing that the auditory cortex of human and animals is sensitive to the statistical context at large time scales (Herrmann et al., 2015; Yaron et al., 2012; Rubin et al., 2016). Moreover, we show that surprise-induced enhancement in auditory responses does not interact with repetition-induced suppression: test stimuli with adapted responses maintain adaptation and even adapt further despite eliciting larger responses on trials in contexts that render the stimulus surprising. The independence of these two processes enables the brain to represent stimulus familiarity through suppressed responses while concurrently modulating those responses to bias attention based on violation of prediction. Our work provides a model that uses responses to a fixed set of stimuli to better understand the mechanisms by which top-down processes (expectation and memory) and bottom-up processes (based on stimulus features) interact in sensory coding.

References

- Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15:600–609. [CrossRef Medline](#)
- Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617–629. [CrossRef Medline](#)
- Beckers GJ, Gahr M (2010) Neural processing of short-term recurrence in songbird vocal communication. *PLoS One* 5:e11129. [CrossRef Medline](#)
- Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol* 77:923–943. [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](#)
- Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215. [Medline](#)
- Farley BJ, Quirk MC, Doherty JJ, Christian EP (2010) Stimulus-specific adaptation in auditory cortex is an NMDA-independent process distinct from the sensory novelty encoded by the mismatch negativity. *J Neurosci* 30:16475–16484. [CrossRef Medline](#)
- Fishman YI (2014) The mechanisms and meaning of the mismatch negativity. *Brain Topogr* 27:500–526. [CrossRef Medline](#)
- Fishman YI, Steinschneider M (2012) Searching for the mismatch negativity in primary auditory cortex of the awake monkey: deviance detection or stimulus specific adaptation? *J Neurosci* 32:15747–15758. [CrossRef Medline](#)
- Gill P, Woolley SM, Fremouw T, Theunissen FE (2008) What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. *J Neurophysiol* 99:2809–2820. [CrossRef Medline](#)
- Herrmann B, Henry MJ, Fromboluti EK, McAuley JD, Obleser J (2015) Statistical context shapes stimulus-specific adaptation in human auditory cortex. *J Neurophysiol* 113:2582–2591. [CrossRef Medline](#)
- Hershenhoren I, Taaseh N, Antunes FM, Nelken I (2014) Intracellular correlates of stimulus-specific adaptation. *J Neurosci* 34:3303–3319. [CrossRef Medline](#)
- Levy R (2008) Expectation-based syntactic comprehension. *Cognition* 106:1126–1177. [CrossRef Medline](#)
- Lu K, Vicario DS (2011) Toward a neurobiology of auditory object perception: what can we learn from the songbird forebrain. *Current Zoology* 57:671–683. [CrossRef](#)
- 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](#)
- May PJ, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47:66–122. [CrossRef Medline](#)
- Näätänen R (1995) The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear* 16:6–18. [Medline](#)
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) “Primitive intelligence” in the auditory cortex. *Trends Neurosci* 24:283–288. [CrossRef Medline](#)
- Näätänen R, Jacobsen T, Winkler I (2005) Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology* 42:25–32. [CrossRef Medline](#)
- Nelken I, Ulanovsky N (2007) Mismatch negativity and stimulus-specific adaptation in animal models. *J Psychophysiol* 21:214–223. [CrossRef](#)
- Pérez-González D, Malmierca MS (2014) Adaptation in the auditory system: an overview. *Front Integr Neurosci* 8:19. [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](#)
- Rubin J, Ulanovsky N, Nelken I, Tishby N (2016) The representation of prediction error in auditory cortex. *PLoS Comput Biol* 12(8).
- Solis MM, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci* 17:6447–6462. [Medline](#)
- Symonds RM, Lee WW, Kohn A, Schwartz O, Witkowski S, Sussman ES (2017) Distinguishing neural adaptation and predictive coding hypoth-

- eses in auditory change detection. *Brain Topogr* 30:136–148. [CrossRef](#) [Medline](#)
- Taaseh N, Yaron A, Nelken I (2011) Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One* 6:e23369. [CrossRef](#) [Medline](#)
- Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. *Anim Behav* 59:1167–1176. [CrossRef](#) [Medline](#)
- Theunissen FE, Shaevitz SS (2006) Auditory processing of vocal sounds in birds. *Curr Opin Neurobiol* 16:400–407. [CrossRef](#) [Medline](#)
- Todorovic A, de Lange FP (2012) Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci* 32:13389–13395. [CrossRef](#) [Medline](#)
- Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK (2009) Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* 21:1934–1945. [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](#)
- Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in avian brain. *Proc Natl Acad Sci U S A* 107:12676–12681. [CrossRef](#) [Medline](#)
- Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76:603–615. [CrossRef](#) [Medline](#)