



---

**Research Articles: Systems/Circuits**

**Familiar but Unexpected: Effects of Sound Context Statistics on Auditory Responses in the Songbird Forebrain.**

Kai Lu and David S. Vicario

Kai Lu, Institute for Systems Research, University of Maryland, College Park, MD, 20740 USA  
David S. Vicario, Psychology Department, Rutgers University, Piscataway, NJ, 08854 USA

DOI: 10.1523/JNEUROSCI.5722-12.2017

Received: 13 December 2012

Revised: 30 August 2017

Accepted: 29 September 2017

Published: 8 November 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.

**Conflict of Interest:** The authors declare no competing financial interests.

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

Corresponding author: David S. Vicario, 152 Frelinghuysen Road, Piscataway, NJ 08854, Ph (848)-445-2907, [vicario@rci.rutgers.edu](mailto:vicario@rci.rutgers.edu)

**Cite as:** J. Neurosci ; 10.1523/JNEUROSCI.5722-12.2017

**Alerts:** Sign up at [www.jneurosci.org/cgi/alerts](http://www.jneurosci.org/cgi/alerts) to receive customized email alerts when the fully formatted version of this article is published.

# 1 Journal of Neuroscience Manuscript Submission Title Page

## 2 Title:

## 3 Familiar but Unexpected: Effects of Sound Context Statistics on

### 4 Auditory Responses in the Songbird Forebrain.

5 Kai Lu and David S. Vicariot

6

## 7 Abbreviated title:

## 8 Effects of Sound Context Statistics

## **9 Authors, Affiliations, and author addresses:**

10 Kai Lu, Institute for Systems Research, University of Maryland, College Park, MD,  
11 20740 USA

12 David S. Vicario, Psychology Department, Rutgers University, Piscataway, NJ,  
13 08854 USA

## 14 Corresponding author

15            David S. Vicario  
16            152 Frelinghuysen Road  
17            Piscataway, NJ 08854  
18            Ph (848)-445-2907  
19            [vicario@rci.rutgers.edu](mailto:vicario@rci.rutgers.edu)

20 Number of pages: 38(total) 23(main text)

## 21 Number of figures: 7

## 22 Number of words for

23 Abstract: 248  
24 Introduction: 584  
25 Discussion: 1498  
26

<sup>27</sup> Conflict of Interest: No conflict of interest.

28

29    **Acknowledgements**

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

33

34    **Abstract**

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

54

55     **Significance**

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

65

66

67

68

69

70

71

72

73

74

75

76 **Introduction**

77 It is recognized that neurons in the auditory forebrain encode not only the acoustic  
78 properties of sounds but also the probability of those sounds and/or the transitions between  
79 sounds (Ulanovsky et al., 2003, 2004; Gill et al., 2008; Beckers et al., 2010; Lu and Vicario  
80 2014). Ulanovsky, (2003, 2004) suggested that the underlying mechanism encoding sound  
81 probability is stimulus-specific adaptation (repetition-induced suppression) to frequent, repeated  
82 sounds. However, Gill et al. (2008) showed that auditory responses reflect statistically  
83 unexpected events based on experience over longer time scales. It is difficult to separate these  
84 mechanisms experimentally because the most often repeated sound is usually the most  
85 expected one. This puzzle complicates the interpretation of single-unit studies in animals and is  
86 closely related to a long-standing question in auditory EEG research: whether the mismatch-  
87 negativity (MMN) to an oddball sound is due to adaptation that reduces response to the more  
88 common sound, or to a violation of expectation that increases response to the oddball. A recent  
89 MEG study (Todorovic and de Lange, 2012) showed that these two effects are separable in time,  
90 suggesting that two distinct mechanisms co-exist. At the neural level, recent studies also  
91 showed that higher responses to oddball sounds found in auditory cortex of rats cannot be fully  
92 explained by repetition-induced suppression (Taaseh et al., 2011, Hershenhoren et al., 2014).  
93 The experiments in this paper will study and potentially differentiate these effects at the  
94 neurophysiological level.

95 Even a familiar event may not be expected at a certain moment, and the expectation  
96 more likely depends on the most recent context of events. Thus, the brain appears to encode  
97 familiarity and expectation independently over different time scales. To investigate these  
98 processes, we implemented a novel experimental paradigm: a test sound was repeated during a  
99 first phase, then presented interleaved with other “context” sounds in a second phase, and, in a  
100 final phase, repeated again. The effects of stimulus repetition were quantified by changes in

101 response to the test between the first and final phases. We hypothesized that, when presented  
102 among context sounds, the test will be expected or unexpected depending on membership in  
103 the same acoustic category as context sounds. Thus, responses to the test during the context  
104 phase may reflect both repetition and surprise effects. After removing the predictable effect of  
105 test repetition, any residual responses can be inferred to show the effect of surprise.

106 We tested our hypothesis by recording single-unit and multi-unit activity in two areas of  
107 the zebra finch auditory forebrain: caudomedial nidopallium (NCM) and caudolateral  
108 mesopallium (CLM); both receive inputs from thalamo-recipient Field L, and thus may  
109 correspond to superficial layers of mammalian A1 or to a secondary auditory area (Wang et al.,  
110 2010, Theunissen and Shaevitz, 2006). In these areas, neural responses show long-lasting  
111 stimulus-specific adaptation to repetition of specific songs in awake birds (Chew et al., 1995,  
112 Chew et al., 1996). These results depend on the large set of learned vocalizations in songbirds  
113 which provides a stimulus repertoire of distinct but related sounds. We now demonstrate that,  
114 when a natural vocalization is presented as a test stimulus among context sounds, e.g. songs of  
115 a different species, auditory responses are enhanced. Furthermore, the data show that surprise-  
116 induced enhancement and repetition-induced suppression do not interact, suggesting that  
117 different levels of familiarity and expectation may be encoded independently. This is consistent  
118 with, and could help to elucidate, the neural mechanisms of human MMN, which can be  
119 observed not only for simple oddballs but also for higher order category violations (Näätänen et  
120 al., 2001).

121

122 **Materials and Methods**123 *Subjects*

124 All animals used in our experiments were adult male zebra finches (n=29) bred in our aviary  
125 or obtained from the Rockefeller University Field Research Center. Animals were housed on a  
126 12:12h light-dark cycle in a general aviary, where they could see other birds and hear their  
127 vocalizations. Food and water were provided *ad libitum* and the all procedures conformed to a  
128 protocol approved by the Institutional Animal Care and Use Committee of Rutgers University.

129 *Surgery*

130 In preparation for electrophysiological recording, each animal was anesthetized with  
131 isoflurane (2% in oxygen) and placed into a stereotaxic apparatus. Marcaine (0.04 cc, 0.25%)  
132 was injected under the scalp to provide local analgesia, the skin was incised, and a small  
133 craniotomy exposed the area of the bifurcation of the mid-sagittal sinus. Dental cement was  
134 used to attach a metal post to the skull rostral to the opening and to form a chamber around the  
135 recording area. The chamber was then sealed with silicone elastomer (Kwiksil, World Precision  
136 Instruments, USA). To relieve post-surgical pain, Metacam (0.04 cc, 5 mg/mL) was  
137 administered i.m. Anesthesia was discontinued and the bird was allowed to recover under a  
138 heat lamp.

139 *Electrophysiology*

140 Two days after initial surgery (to allow for full recovery from anesthesia),  
141 electrophysiological recordings were made in a walk-in soundproof booth (IAC, New York). The  
142 awake animal was immobilized in a comfortable tube and the implanted post was used to fix the  
143 head to a stereotaxic frame. Recordings were made at 16 sites, 4 each in the left and right NCM,  
144 4 each in the left and right CLM (**Figure 1A**), using glass insulated platinum/tungsten  
145 microelectrodes (2-3 M $\Omega$ ms impedance) independently advanced by a multielectrode

146 microdrive (Ekhorn design, Thomas Recording , Germany). Electrode signals were amplified  
147 (x19,000) and filtered (bandpass 0.5-5 kHz), then acquired at 25kHz using Spike 2 software  
148 (CED, Cambridge, England). White noise stimuli with the amplitude envelope of canary song  
149 were presented to search for responsive sites typical of the auditory forebrain. Once all  
150 electrodes were placed at responsive sites, stimulus playback experiments were performed. At  
151 the end the recording, eight small electrolytic lesions (20 uA for 15 s) were made to enable  
152 histological reconstruction of recording sites.

153 *Histology*

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

158 *Auditory Stimuli and Experimental design*

159 All sound stimuli consisted of natural zebra finch (conspecific) and canary (heterospecific)  
160 songs which differ in their acoustic characteristics (**Figure 1B, 1C and 1D**). Neurons in NCM  
161 are known to respond differently to these two types of songs (Chew et al., 1996). Stimuli lasted  
162 0.77-1.21 s and were presented at 65 dB SPL (A scale). All experiments followed a similar  
163 protocol, which consisted of 3 phases: Pre-adapting, Context-modulated, and Post-context  
164 (**Figure 1E**). In the Pre-adapting phase, a test song (e.g. a novel zebra finch song) was  
165 repeated 20 times at a fixed inter-stimulus-interval (ISI) of 7 s to establish initial adaptation.  
166 Notice that the ISI we used is much longer than the maximum ISI for inducing forward  
167 suppression (several hundred milliseconds; Brosch and Schreiner, 1997) and stimulus-specific  
168 adaptation in rodentsy (up to 2 seconds; Ulanovsky (2003, 2004, ). In the immediately following  
169 Context-modulated phase, the test stimulus was again presented 19 times, but now in the

170 context of other stimuli in random order (all at 7 s ISI). Finally, in the Post-context phase, the  
171 test song was again presented for 20 trials at a fixed 7 s ISI. Responses to the test song in the  
172 Pre-adapting phase were used to compute an adaptation function (described below) whose  
173 slope was used to estimate subsequent responses. Comparison of actual and estimated  
174 responses to the test song in the Context-modulated phase was used to quantify the effects  
175 produced by context manipulations (detailed methods below).

176 In Experiment 1, three different context conditions were assessed in 11 birds. In the first  
177 condition, **Canary context (Figure 1E)**, the Context stimulus set consisted of 19 repeats of a  
178 pre-adapted zebra finch test song and 20 repeats of each of 7 novel canary songs, for a total of  
179 159 trials presented in randomly shuffled order. Although the ISI was fixed at 7 s, the intervals  
180 between repeats of the one test song varied from 14 -161 s. In the second condition, **Silence**  
181 **context**, all canary songs were replaced by silence, while the order and intervals were the same  
182 as in the first condition, effectively creating variable silences of 14 – 161 s between the onsets  
183 of test songs. In the third condition, **Zebra finch context**, the test song was played in the  
184 context of 7 other novel zebra finch songs, and the order of stimulus presentation was  
185 randomized, as in condition 1. Each subject was tested with all 3 conditions in random order,  
186 and each condition used a different test song that was novel for the bird.

187 In Experiment 2, a new condition that reversed test and context song types, **Canary song**  
188 **in Zebra finch context**, was tested in 6 birds, together with the two conditions from Experiment  
189 1: Canary context and Zebra finch context. All stimuli were novel songs as in Experiment 1. For  
190 **Canary song in Zebra finch context**, the test song was a novel canary song and zebra finch  
191 songs were used as context stimuli in the Context-modulated phase. All other aspects of  
192 presentation and data analysis were the same as described above.

193 In Experiment 3, we tested whether the prior familiarity of the context songs played in the  
194 Context-modulated phase influenced context effects on responses to the test song. Seven birds  
195 were first tested with the Zebra finch context condition (ZF context condition), in which the  
196 context songs were novel zebra finch songs, exactly as in Experiment 1. Then 50 repeats of the  
197 same 7 context songs were presented to the animal in shuffled order at 7 s ISI (350 stimuli total).  
198 Then the animals were again tested with the ZF context condition, using a novel zebra finch test  
199 song and the now-familiar context songs in the Context-modulated phase. Context effects on  
200 responses to the test song in the Zebra finch context session prior to the familiarization training  
201 with the context songs were compared to context effects in the second Zebra finch context  
202 session that used familiar context songs.

203 *Data analysis: single-units*

204 Single-units with spikes greater than 3 standard deviations from the baseline were isolated  
205 from the electrode recordings off-line, using template-based digital clustering algorithms  
206 implemented in Spike2 software (CED, Cambridge, UK). Single-units were validated by analysis  
207 of the inter-spike interval (ISpi) histograms. To be accepted, a unit had to have a contamination  
208 rate (ISpIs less than 2ms, corresponding to spike rates > 500 Hz) lower than 2%. The response  
209 amplitude of each unit was quantified as the spike rate in the response window (from stimulus  
210 onset to stimulus offset plus 100 ms) – minus the spike rate in the 500 ms period preceding  
211 stimulus onset, on each trial.

212 *Data analysis: multi-unit activity*

213 Since the spikes of a single-unit typically represent only ~10% of all multi-unit spikes (that  
214 crossed a threshold) at each recording site, we not only report single-unit data but also multi-  
215 unit data in parallel, in order to capture the activity of non-isolated neurons. For each channel,  
216 the root-mean-square (RMS) of the multi-unit neural activity was calculated both over a baseline

217 window (the 500 ms period prior to stimulus onset) and over a response window (from stimulus  
218 onset to stimulus offset plus 100 ms) on each trial. The RMS provides a method of rectifying the  
219 multi-unit activity and computing its average power. Because our multi-unit recordings typically  
220 were bandpass filtered (0.5-5 kHz), the RMS primarily measured action potentials (not LFP's or  
221 EEG). Responses to song stimuli were quantified as the difference between the baseline RMS  
222 and response RMS measurements (**Figure 2A**). A site was excluded if its response to the test  
223 song in the Pre-adapting phase was not significantly different from the baseline. The baseline  
224 RMS was separately analyzed for comparison across the three phases of the experiment.

225 *Effects of context modulation measured as Delta-surprises*

226 The effect of different context manipulations on auditory responses was measured by  
227 quantifying how each response (single-unit spike rates or multi-unit RMS) during the Context-  
228 modulated phase deviated from the responses estimated from the responses in the Pre-  
229 adapting phase and the Post-context phase. This was computed as the “surprise”, a measure  
230 from information theory (Levy, 2008), according to the following procedure: 1) The linear  
231 regression line for the responses in the Pre-adapting phase was computed from the responses  
232 to the repeated “test” song during the linear portion of the adaptation function (**Figure 2B**; trials  
233 6-20, black line). This line was extrapolated to estimate the response on the first trial of the  
234 Context-modulated phase (**Figure 2B**, green circle at trial 21). 2) A second regression line was  
235 computed from responses to the test stimulus in the Post-context phase (trials 40-59, black line),  
236 and then extrapolated backwards to estimate the expected response on the last trial (**Figure 2B**,  
237 green circle at trial 39) of the Context-modulated phase. 3) The expected responses in the  
238 Context-modulated phase were estimated by the line connecting the estimates for trials 21 and  
239 39 (**Figure 2B**, green line, called hereafter the interpolated regression). 4) The expected  
240 standard deviation of the responses around the interpolated regression line was estimated by  
241 the standard deviation of pooled residuals of the regressions of the Pre- and the Post- phases; 5)

242 An observed response that falls on the interpolated regression line is the least surprising (most  
243 expected) response; the greater the deviation (d) of an observed response from this expectation,  
244 the more surprising it is. The degree to which it is surprising is a function of the probability of a  
245 deviation of magnitude d, namely,  $\log(1/P(d))$ , where P(d) is the probability density of d in the  
246 assumed-to-be normal distribution (see **Figure 2B**). While the adaptation function during the  
247 context not be strictly linear, the interpolated linear regression line is a conservative estimate; if  
248 an exponential fit were used, the observed deviations in the context block would be even  
249 greater. 6) The average magnitude of a surprisal is greater for distributions with large standard  
250 deviations than for narrower distributions, so the surprisals were normalized by subtracting the  
251 absolute value of the minimum surprisal, which is  $\log(1/P(0))$ . Normalization makes the surprisal  
252 of an observed response that exactly conforms to an expectation equal to zero, and it zeroes  
253 the expectation of the signed surprisals when observed responses are drawn from the expected  
254 distribution. 7) Responses greater than expected are assigned positive surprisal, while  
255 responses less than expected are assigned negative surprisal. 8) Thus, our formula for the  
256 normalized signed Delta-surprisal of an observed response is :

257 
$$\text{Delta surprisal} = \text{sign}(d) * (\log \frac{1}{P(d)} - \text{abs}(\log \frac{1}{P(0)}))$$

258 where d = the deviation of the response from expectation

259 *Effects of stimulus repetition measured as an Adaptation index*

260 In order to test whether the modulations interact with stimulus-specific adaptation, we also  
261 computed an Adaptation Index for each condition at each recording site by dividing the  
262 response amplitude of the first test stimulus trial of the Post-context phase (trial 40) by the  
263 response on the last test stimulus trial of the Pre-adapting phase (trial 20). This ratio provided  
264 an estimate of adaptation that occurred over the context phase (see **Figure 2B**). Notice that

265 adaptation indices are different from the stimulus-specific adaptation indices (SI) used by  
266 Ulanovsky et al. 2003, in that adaptation indices reflects reduction of neural responses to the  
267 same sound over repetitions, while SI reflects differences in responses between oddball sounds  
268 and standard sounds. The same procedure and calculations were used for both single-units and  
269 multi-unit data.

270 *Quantification of neurons' selectivity to the test song and context songs*

271 In order to test whether the selectivity of neurons to the test sound and context sounds  
272 affected the enhancement effect, we also quantified D' for each multi-unit site. D' measures the  
273 selectivity for one stimulus (A) over another stimulus (B) at each recording site, and was  
274 calculated by the following formula (as described in Solis and Doupe, 1997):

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

275  
276 A positive D' means that the neuron prefers stimulus A in its responses. To calculate D',  
277 we first took the mean and variance of responses (multi-unit RMS) to each song at each site  
278 (obtained from last ten trials of the Pre-adapting phase for the test song and first ten trials of  
279 each context song in the Context-modulated phase). Then D' for each test song with respect to  
280 the 7 context songs was calculated to produce 7 D's for the context-modulated phase, and then  
281 these were averaged for each site.

282 *Temporal Profile of Responses*

283 We analyzed temporal characteristics of the responses seen with context manipulations by  
284 computing the difference between the averaged temporal waveform of responses to the test  
285 stimulus in the Pre-adapting phase and in the Context-modulated phase across all sites. We  
286 used the following procedure: 1) We first computed the moving average RMS (10 ms window) of

287 the multi-unit recording to produce a smoothed RMS waveform of the response to each stimulus  
288 at each site; 2) Averages of these waveforms were then computed both across the last 6 trials  
289 of the Pre-adapting phase and across the first 6 trials of the Context-modulated phase. 3) Then  
290 averages of these waveforms were computed across all recording sites from all birds separately  
291 for each phase. These grand averages effectively eliminated response patterns due to  
292 characteristics of specific stimuli and/or specific recording sites. Thus, the difference waveform  
293 between these two grand average waveforms shows the temporal profile of response  
294 enhancement caused by the surprise phenomenon.

295 *Statistical methods*

296 Data are graphed both as cumulative frequency distributions, which reveal the details of  
297 condition effects, and as conventional mean and standard error plots. The distribution of  
298 samples in some conditions did not fully satisfy criteria for parametric tests. Therefore,  
299 appropriate non-parametric statistics were used throughout whenever possible (Siegel, 1956).  
300 For Experiment 1, the Delta-surprisals obtained for each recording site in each of the three  
301 conditions (across all three experiments) were treated as three repeated measures. The main  
302 effect was tested by the non-parametric Friedman ANOVA, which does not require a normal  
303 data distribution. Differences between groups where data were matched (e.g. different  
304 conditions recorded at the same electrode site) were tested by the Wilcoxon matched pairs test.  
305 For group data where samples were not explicitly matched, we used the Kolmogorov-Smirnov  
306 two-sample test. In order to quantify possible differences in the main effects between NCM and  
307 CLM, the interaction between regional difference and conditions was tested using a Repeated-  
308 Measures ANOVA, in which region was treated as a factor and the Delta-surprisals of the three  
309 conditions were repeated measures.

310

311 **Results**312 Effects of Context Manipulations on Auditory Responses to a Zebra Finch Song

313 Experiment 1 measured the effects of manipulating the acoustic and temporal context on  
314 responses to a pre-adapted test song. We obtained 68 isolated single-units and 111 multi-unit  
315 sites from brain regions NCM and CLM in 11 birds, each tested with the 3 different context  
316 conditions. Our measure of context effects (Delta-surprisal, see Methods) showed no significant  
317 differences between NCM and CLM in two-way repeated-measures ANOVAs (single-units:  $F(1, 66) = 0.29, p > 0.591$ ; multi-unit:  $F(1, 109) = 0.85, p > 0.357$ ) and no interaction between brain  
318 regions and context conditions (single-units:  $F(1, 66) = 1.62, p > 0.201$ ; multi-unit:  $F(1, 109) =$   
319  $0.32, p > 0.728$ ). Thus data from NCM and CLM were combined for further analyses.

321 For single-unit data, an increased firing rate was seen in the Context-modulated phase  
322 relative to the Pre-adapting phase in raster plots and PSTHs (example in **Figure 3A**) and in the  
323 plot of spike rates by trials (**Figure 3B and 3C**). When the increased activity was quantified as  
324 Delta-surprisals, there were significant differences between the 3 context conditions tested  
325 (**Figure 4A**, Friedman ANOVA, Chi Sq. ( $N = 68, df = 2$ ) = 23.4,  $p < 0.001$ ). Most neurons in the  
326 Canary context condition (72%, 49/68) and the Silence condition (68% 46/68) showed positive  
327 Delta-surprisals, indicating an increased firing rate during the context-modulated phase. In  
328 contrast, less than half (43%, 29/68) showed positive Delta-surprisals in the Zebra finch context.  
329 Further tests showed that Delta-surprisals in the Canary context were significantly larger than in  
330 the Silence context (Wilcoxon,  $z = 2.12; p < 0.034$ ), which in turn were significantly larger than  
331 Zebra finch context (Wilcoxon,  $z = 2.68; p < 0.007$ ).

332 When multi-unit RMS were analyzed, even greater differences across the three  
333 conditions were observed (**Figure 4B**, Friedman ANOVA, Chi Sq. ( $N = 111, df = 4$ ) = 129.9,  $p$   
334  $< 0.001$ ). More than 95% (106/111) of multi-unit sites in the Canary context and more than 94%

335 (105/111) of multi-unit sites in the Silence context condition showed positive Delta-surprisals,  
336 while only about 58% (64/111) of multi-unit sites in the ZF context condition showed positive  
337 Delta-surprisals. Delta-surprisals in the Canary context were significantly larger than in the  
338 Silence context condition (Wilcoxon  $z = 3.13$ ;  $p < 0.001$ ), which in turn were significantly larger  
339 than in the ZF context condition (Wilcoxon  $z = 8.19$ ;  $p < 0.001$ ), the same pattern as that seen  
340 for single-units. It should be noted that the ISI for the test song during the context-modulated  
341 phase was variable and much longer (an average of 56 s) than in the pre-adapting phase (7 s).  
342 While this longer effective ISI might contribute to the observed enhancement if it is considered a  
343 decay of adaptation, as in the Silence condition with no intervening stimuli, this effect cannot  
344 explain the weaker enhancement in the Zebra finch context, or the higher enhancement in the  
345 Canary condition. In all three cases, the ISI for the test song is equally long and variable, but the  
346 enhancement differs in opposite directions for the two different types of context stimuli.

347 The baseline remains constant across three phases.

348 Because multi-unit responses to song stimuli were quantified as the difference between  
349 the baseline RMS and response RMS measurements, we also compared the baseline RMS  
350 both across the three context conditions and across three phases (pre-adapting phase, the  
351 context-modulated phase and the post-context phase) in a two-way repeated-measures ANOVA.  
352 We did not find significant changes in baseline activity across conditions ( $F(2, 660) = 0.35$ ,  $p >$   
353  $0.706$ ) and across phases ( $F(2, 660) = 0.45$ ,  $p > 0.636$ ) or any interaction ( $F(4, 660) = 1.71$ ,  $p >$   
354  $0.145$ ). Therefore, enhancement effects in the context-modulated phase were not due to  
355 changes in baseline activity.

356 Enhancement effects are independent from adaptation

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

359 context manipulations. Thus, to fully interpret the results, it is essential to know whether the  
360 enhancement observed interacts with the adaptation process or is independent from it. For  
361 example, if they interact, the enhancement effect might reduce or prevent the adaptation  
362 normally produced by presenting the same test song 19 times during the Context-modulated  
363 phase (cf. pre-Context phase adaptation seen in Figure 3C). If this is the case, we predict that  
364 the context that produces the largest enhancement effect should also show the weakest  
365 adaptation. We measured adaptation (the drop in response amplitude) over the Context-  
366 modulated phase as the Adaptation Index (see Methods). When we compared Adaptation  
367 Indices across the three conditions the data showed no significant differences in either single-  
368 unit data (Friedman ANOVA Chi Sq. ( $N = 68$ ,  $df = 2$ ) = 4.35,  $p > 0.11$ ) or multi-unit data (**Figure**  
369 **4C**, Friedman ANOVA Chi Sq. ( $N = 111$ ,  $df = 2$ ) = 0.16,  $p > 0.922$ ) and thus showed no  
370 relationship to the degree of enhancement across conditions. To examine this in more detail,  
371 we calculated the correlation coefficient (Spearman rho) between the adaptation indices and the  
372 Delta-surprisals for each site within each condition and each bird. Of the 33 correlations  
373 assessed (3 conditions X 11 birds) 31 were not significant ( $p > 0.05$ ). The 2 correlations that  
374 were significant (1 negative and 1 positive) were not associated with any one bird or condition.  
375 Therefore, there was no systematic relationship between adaption and the enhancement effect  
376 and we conclude that the two processes are independent from each other.

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

379 Previous work (Ulanovsky et al., 2003) showed that the effect of sound probability on  
380 auditory responses was positively correlated with the frequency separation between the  
381 standard tone and the oddball tone. If the enhancement effect was also affected by the spectral  
382 differences between the test song and the context songs, we would expect that the larger the  
383 difference between responses to the test song (ZF song) and context songs (Canary songs), the

384 higher the enhancement effect that would be seen on a given neuron. To test this hypothesis,  
385 we calculated the selectivity of neurons to the test sound and context sounds, quantified as D'.  
386 The the bias of neural responses to a tested ZF song or context canary songs was reflected by  
387 the absolute value of D'. We calculated the correlation between the absolute values of D' and  
388 the Delta-surprisals from all multi-units for the Canary context condition. Surprisingly, we did not  
389 find significant correlation between the absolute values of D' and the enhancement effect  
390 (Spearman:  $r = 0.056$ ,  $p > 0.557$ , Figure 5D). Therefore, enhancement effects cannot be fully  
391 explained by neurons' selective tuning toward ZF songs or canary songs.

392 Enhancement effects analyzed trial-by-trial.

393 The enhancement effect described so far was quantified as Delta-surprisals averaged  
394 across trials for each site. We further tested if Delta-surprisals increase or decrease with  
395 repetition of the test song in the context. For each multi-unit site, we calculated the linear  
396 regression between trial number of the test song in the context and the Delta-surprisal on each  
397 trial. Then we pooled slopes of the regression from all multi-units and analyzed whether the  
398 slopes in any condition were significantly higher or lower than zero. As shown in Figure 5E,  
399 there were no significant increase or decrease in Delta-surprisals with trials both in the Canary  
400 context (Wilcoxon,  $z = 0.94$ ;  $p > 0.346$ ) and the Silence context (Wilcoxon,  $z = 0.22$ ;  $p > 0.828$ ).  
401 In contrast, Delta-surprisal in the ZF context significantly increased with trials (Wilcoxon,  $z =$   
402  $5.74$ ;  $p < 0.001$ ). This increase in Delta-surprisals with trials in the ZF context may be due to  
403 the increasing familiarity of the context songs, which will be further discussed with results of  
404 Experiment 3.

405 Temporal pattern of enhancement effects

406 In addition to quantifying the overall enhancement of responses by context manipulations,  
407 we examined the timing and waveform of the responses in the Context-modulated phase in

408 order to explore the possible mechanism of the enhancement. First, if the enhancement effect  
409 does not interact with adaptation (as shown above), we expect that the latency (measured from  
410 stimulus onset) of the enhanced component of the responses will be longer than for the auditory  
411 response itself because the surprisal effect may reflect top-down modulations. Second, in the  
412 Silence context, only temporal uncertainty contributes to the enhancement effect (there are no  
413 intervening sounds, so no acoustic discrimination is needed) so we expected the enhancement  
414 profile for this condition would have a shorter latency than in the Canary context condition.

415 For each of the 3 conditions, the enhancement profile was computed as the difference  
416 between the averaged multi-unit RMS waveforms between the Context-modulated phase versus  
417 the Pre-adapting phase across all multi-unit sites ( $N = 111$ ) (**Figure 5A, black trace**). Using the  
418 mean of multi-unit data effectively averages out temporal features associated with any specific  
419 stimulus or the tuning properties of individual neurons. The resulting “enhancement profile”  
420 shows the timing of enhancement due to modulation by the different acoustic contexts across all  
421 sites recorded. We found no differences between NCM and CLM in the enhancement profile for  
422 any condition, so data from the two brain areas were combined. First, for the canary context, we  
423 compared the latency of averaged multi-unit RMS in the context-modulated phase (red trace in  
424 **Figure 5B**) to the latency of the enhancement profile (black trace in **Figure 5B**). The latency for  
425 the two waveforms was computed as the time from stimulus onset until the signal crossed a  
426 threshold, computed as the maximum value of the 99% confidence interval for each signal  
427 during the baseline window. For the Canary context condition (**Figure 5C, red trace**), the  
428 latency of the first increase in the enhancement profile was longer than the latency of the multi-  
429 unit RMS waveform (27 ms vs. 6 ms). The peak also occurred later (112 ms vs. 85 ms), as  
430 shown in **Figure 5D**. In contrast, the ZF context condition showed no consistent change in the  
431 timing of the enhancement profile (**Figure 5C**), consistent with the analysis based on Delta-  
432 surprisals.

433 The enhancement profiles in the Silence context condition differ from the enhancement  
434 profiles with in the Canary context condition in two ways. First, the latency of enhancement in  
435 the Silence context was 20ms (shorter than in the Canary context condition by 7 ms), (**Figure**  
436 **5D**). To quantify the effect of this latency difference statistically, we compared response  
437 amplitudes in the window 10- 20 ms after stimulus onset between the Pre-adapting phase and  
438 the Context-modulated phase for both the Canary context and the Silence context. We  
439 observed significant enhancement in this time window in the Silence context (Wilcoxon  $z = 2.8$ ;  
440  $p < 0.006$ , **Figure 5F, left box**), but not in the Canary context (Wilcoxon  $z = 1.5$ ;  $p > 0.123$ ,  
441 **Figure 5F, right box**). This observation supports the idea that early enhancement in the Silence  
442 context condition is due to uncertainty about when the stimulus will occur, resulting in a very  
443 rapid detection of stimulus onset. In contrast, acoustic processing needed to detect a violation of  
444 the acoustic context in the Canary context condition requires more time. Second, the  
445 enhancement profile in the Silence context had a shorter duration than the Canary context  
446 profile, with a decay to zero 181 ms earlier than in the Canary context (**Figure 5E**). In the  
447 Silence context condition, differences in responses between the Context-modulated phase and  
448 the Pre-adapting phase become undetectable in the window 490- to 690 ms after stimulus onset  
449 (Wilcoxon  $z = -1.24$ ;  $p > 0.216$ , **Figure 5G**), while, in the Canary context condition,  
450 enhancement effect was still significant in the same window (Wilcoxon  $z = 6.90$ ;  $p < 0.001$ ). This  
451 is consistent with the idea that processing of acoustic features in the canary context continues  
452 during the evolving stimulus, while temporal uncertainty in the silence condition is largely  
453 detected at stimulus onset.

454 Canary test songs in the zebra finch context showed similar enhancement

455 In order to test whether enhancement is due to the fact that the context stimuli are Canary  
456 songs and thus from a different species than the zebra finch subjects, Experiment 2  
457 implemented the Canary in ZF context condition in which the test song is a canary song and the

458 context songs are novel Zebra finch songs, as well as two conditions previously described: the  
459 Canary context (Zebra finch test song in canary contexts) and ZF context conditions (zebra  
460 finch test song in zebra finch context). In this experiment, 47 multi-unit sites recorded in NCM  
461 and CLM in 6 birds were analyzed. Cumulative frequency distributions and the mean Delta-  
462 surprisals for each condition are shown in **Figure 6**. There were significant differences in Delta-  
463 surprisals across the three conditions (Friedman ANOVA Chi Sq. ( $N = 47$ ,  $df = 2$ ) = 30.9,  $p <$   
464 0.001). Both the Canary in ZF context and the original Canary context condition showed Delta-  
465 surprisals significantly higher than those of the ZF context condition (Wilcoxon tests: Canary in  
466 ZF context condition:  $z = 4.31$ ;  $p < 0.001$ . Canary context condition:  $z = 5.09$ ;  $p < 0.001$ ), but  
467 Canary in ZF context and Canary context conditions were not significantly different from each  
468 other (Wilcoxon tests:  $z = -0.63$ ;  $p > 0.526$ ). This result suggests that the strong enhancement  
469 effect seen in the Canary context condition, described in Experiment 1 and 2, was due to the  
470 violation of ongoing expectations set up by context stimuli, rather than to a pre-existing bias for  
471 conspecific songs.

472 Familiarity of context songs increases the enhancement effect

473 In Experiment 3, 67 multi-unit sites recorded in NCM and CLM in 7 birds were analyzed.  
474 This experiment tested the effect of the familiarity of the context songs on context-induced  
475 enhancement in the Zebra finch context. First, the ZF context condition was tested with novel  
476 context songs. Then these same context songs were repeated 50 times each to make them  
477 very familiar and the ZF context condition was tested again now with a novel test song and the  
478 familiar context songs. Cumulative frequency distributions and the mean Delta-surprisals for the  
479 two ZF context conditions are shown in **Figure 7A**. The Context-modulated Delta-surprisals  
480 obtained in the ZF context condition with familiar context songs were significantly higher than  
481 Delta-surprisals obtained when the songs were novel (Wilcoxon  $z = 3.22$ ;  $p = 0.001$ ).

482 The results of Experiments 1 and 2 above showed strong enhancements both for a zebra  
483 finch test song presented in a canary context and for a canary test song in a zebra finch context.  
484 These enhancements appeared to reflect a categorical contrast in stimulus statistics between  
485 the test song and context songs from a different species. However, there was no such contrast  
486 in Experiment 3 (all stimuli were zebra finch songs), which showed that the familiarity of context  
487 songs significantly increased enhancement. This suggests that familiarity itself, which reduces  
488 response strength through stimulus-specific adaptation, might also function as a contrast  
489 dimension. To assess this, we calculated the response selectivity as the D' between songs  
490 heard in the context-modulated phase both before and after exposure to the same context  
491 songs. D' for each test song with respect to the 7 context songs was calculated in the same  
492 manner as in Experiment 1. In addition, D's for each context song with respect to the other  
493 context songs were calculated. Thus, we averaged 8 D's (1 for the test song and 7 for the  
494 context songs) for each recording site in one session. These means reflected whether this song  
495 elicited responses stronger or weaker than other songs on average. We found that, before the  
496 training, the D' of the test song only differed from D's of 3 out of 7 context songs (**Figure 7B**,  
497 tested by Wilcoxon test:  $p < 0.05$ ). In contrast, after exposure to context songs, the D' of the test  
498 song differed from D's of all 7 context songs (Wilcoxon,  $p < 0.001$  for all, **Figure 7C**). Therefore,  
499 prior exposure to context songs increased the contrast in responses between the test song and  
500 context songs, and this may have made the test song perceptually different from the context  
501 songs, producing greater enhancement for the test song in the context phase, as seen in **Figure**  
502 **7A.**

503

504 **Discussion**

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

514 *Enhancement in Canary vs. Zebra finch contexts reflects expectations for context stimuli*

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

525 *The longer latency of the enhancement profile suggests a top-down influence.*

526 Analysis of the temporal profile of enhancement showed that the latency of enhancement is  
527 longer than for auditory responses (27 ms vs. 6 ms) and the enhancement peak comes later

528 (112 ms vs. 85 ms). This implies that the recognition of unexpected events (e.g. acoustic and  
529 temporal context violations) requires processing time, consistent with the observation that the  
530 peak of MMN follows that of N1, a main ERP component (Näätänen et al., 2005). We  
531 hypothesize that the recognition of unexpected stimuli may require a top-down process that  
532 takes time, and may originate from anatomical areas (cf. Bar, 2003; 2004; Turk-Brown et al.,  
533 2009) not assessed in these experiments.

534 *Enhancement in the Silence context condition may reflect temporal uncertainty*

535 Significant response enhancement also occurred for the Silence context, when only the test  
536 song was presented in the Context phase but with random intervals. This is unlikely to be simply  
537 an effect of longer inter-stimulus intervals for two reasons. First, enhancement measurements  
538 were controlled for adaptation. Delta-surprisals were calculated from the differences between  
539 observed responses and estimates of the adaptation trajectory from samples at the end of the  
540 first and beginning of the last phase. If long intervals led to recovery from adaptation, then early  
541 responses in the last phase would also increase and so the estimated trajectory would also be  
542 higher. Second, there was no systematic relationship between Delta-surprisals and adaptation,  
543 as measured by Adaptation Indices; thus, response enhancement did not interact with  
544 adaptation. Since there was no recovery from adaptation during the Context-modulated phase,  
545 enhancement in the Silence context is likely due to the temporal uncertainty of test song onset.  
546 In addition, the enhancement profile in the Silence context had a shorter latency and duration  
547 than in the Canary context, implying that temporal surprise occurs earlier than discriminating  
548 test songs in the Canary context. This may be because, in the Silence condition, only test song  
549 onset needs to be detected, not its acoustic features. Further studies will test if any  
550 subpopulation of neurons is more sensitive to temporal versus acoustic feature surprise.

551 *Stimulus repetition effects are independent from expectation effects*

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

567        In our paradigm, we separated the effects of repetition and expectancy into three phases,  
568    allowing us to determine, in neuronal activity, whether and how surprise effects interact with  
569    repetition-induced suppression (aka stimulus specific adaptation). Our results clearly show that  
570    surprise effects for an unexpected stimulus are independent of and do not interact with  
571    repetition-induced suppression for the very same stimulus. We also found that enhancement  
572    induced by violation of prediction peaks much later than the response itself (112 vs. 85ms after  
573    stimulus onset). Our results not only confirmed recent work that successfully revealed effects of  
574    surprise-induced enhancement (Taaseh et al., 2011, Hershenhoren et al., 2014; Rubin, et al.,  
575    2016), but also demonstrate a way to measure surprise-induced enhancement more explicitly,  
576    so that it can be studied independently from repetition-induced suppression.

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

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

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

598 *Implications*

599        Traditional auditory neurophysiology has mapped acoustic stimulus properties to the  
600 response properties of neurons, e.g. in tonotopic maps. However, growing evidence suggests

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

621

622

623

624

625 **References**

626

627 Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object  
628 recognition. *J Cogn Neurosci* 4:600-9.

629 Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617-29.

630 Beckers GJ, Gahr M (2010) Neural processing of short-term recurrence in songbird vocal  
631 communication. *PLoS One* 17;5(6):e11129.

632 Brosch M, Schreiner CE (1997) Time course of forward masking tuning curves in cat primary  
633 auditory cortex. *J Neurophysiol.* 77(2):923-43.

634 Chew SJ, Mello C, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses  
635 to a repeated conspecific song are long-lasting and require two periods of protein synthesis in  
636 the songbird forebrain. *Proc Natl Acad Sci USA* 92:3406-10.

637 Chew SJ, Vicario DS, Nottebohm F (1996) A large-capacity memory system that recognizes the  
638 calls and songs of individual birds. *Proc Natl Acad Sci USA* 93:1950-5.

639 Corbetta M, Shulman GL.(2002) Control of goal-directed and stimulus-driven attention in the  
640 brain. *Nat Rev Neurosci.*3(3):201-15.

641 Farley BJ, Quirk MC, Doherty JJ, Christian EP (2010) Stimulus-specific adaptation in auditory  
642 cortex is an NMDA-independent process distinct from the sensory novelty encoded by the  
643 mismatch negativity. *J Neurosci.* 30(49):16475-84.

644 Fishman YI, Steinschneider M (2012) Searching for the mismatch negativity in primary auditory  
645 cortex of the awake monkey: deviance detection or stimulus specific adaptation? *J Neurosci.*  
646 32(45):15747-58.

- 647 Fishman YI. (2014) The mechanisms and meaning of the mismatch negativity. *Brain Topogr.*  
648 27(4):500-26.
- 649 Gill P, Woolley SM, Fremouw T, Theunissen FE (2008) What's that sound? Auditory area CLM  
650 encodes stimulus surprise, not intensity or intensity changes. *J Neurophysiol* 99:2809-20.
- 651 Herrmann B, Henry MJ, Fromboluti EK, McAuley JD, Obleser J (2015) Statistical context  
652 shapes stimulus-specific adaptation in human auditory cortex. *J Neurophysiol.* 113(7):2582-91.
- 653 Hershenhoren I, Taaseh N, Antunes FM, Nelken I (2014) Intracellular correlates of stimulus-  
654 specific adaptation. *J Neurosci.* 34(9):3303-19.
- 655 Levy R (2008) Expectation-based syntactic comprehension. *Cognition* 106:1126–1177.
- 656 Lu K, Vicario DS (2011) Toward a neurobiology of auditory object perception: what can we learn  
657 from the songbird forebrain. *Current Zoology* 57: 671–683.
- 658 Lu K, Vicario DS (2014) Statistical learning of recurring sound patterns encodes auditory objects  
659 in songbird forebrain. *Proc Natl Acad Sci U S A.* 111(40):14553-8.
- 660 Näätänen R (1995) The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear*  
661 *Hear* 16:6-18.
- 662 Näätänen R, Jacobsen T, Winkler I (2005) Memory-based or afferent processes in mismatch  
663 negativity (MMN): a review of the evidence. *Psychophysiology.* 42:25-32.
- 664 Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence"  
665 in the auditory cortex. *Trends Neurosci* 24:283-8.
- 666 Nelken I, Ulanovsky N (2007) Mismatch negativity and stimulus-specific adaptation in animal  
667 models. *J Psychophysiol* 21:214 –223.

- 668 May PJ, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection,  
669 explained. *Psychophysiology*. 47(1):66-122.
- 670 Pérez-González D, Malmierca MS (2014) Adaptation in the auditory system: an overview. *Front  
671 Integr Neurosci*. 8:19.
- 672 Phan ML, Pytte CL, Vicario DS. (2006) Early auditory experience generates long-lasting  
673 memories that may subserve vocal learning in songbirds. *Proc Natl Acad Sci U S A*.  
674 103(4):1088-93.
- 675 Rubin J, Ulanovsky N, Nelken I, Tishby N (2016) The Representation of Prediction Error in  
676 Auditory Cortex. *PLoS Comput Biol*. 12(8)
- 677 Solis MM, Doupe AJ (1997) Anterior forebrain neurons develop selectivity by an intermediate  
678 stage of birdsong learning. *J Neurosci* 17:6447-62.
- 679 Symonds RM, Lee WW, Kohn A, Schwartz O, Witkowski S, Sussman ES. (2017) Distinguishing  
680 Neural Adaptation and Predictive Coding Hypotheses in Auditory Change Detection. *Brain  
681 Topogr*. 30(1):136-148.
- 682 Taaseh N, Yaron A, Nelken I (2011) Stimulus-specific adaptation and deviance detection in the  
683 rat auditory cortex. *PLoS One*.6(8):e23369.
- 684 Tchernichovski O, Nottebohm F, Ho CE, Bijan P, Mitra PP (2000). A procedure for an  
685 automated measurement of song similarity. *Animal Behaviour* 59:1167-1176.
- 686 Theunissen FE, Shaevitz SS (2006) Auditory processing of vocal sounds in birds. *Curr Opin  
687 Neurobiol*. 16:400-7.
- 688 Todorovic A, de Lange FP (2012) Repetition suppression and expectation suppression are  
689 dissociable in time in early auditory evoked fields. *J Neurosci*. 32(39):13389-95.

- 690 Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK (2009) Neural evidence of statistical  
691 learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* 10:1934-  
692 45.
- 693 Ulanovsky N, Las L, Nelken I (2003). Processing of low-probability sounds by cortical neurons.  
694 *Nat. Neurosci* 6:391-8.
- 695 Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory  
696 cortex neurons. *J Neurosci*. 24:10440-53.
- 697 Wang Y, Brzozowska-Prechtl A, Karten HJ (2010) Laminar and columnar auditory cortex in  
698 avian brain. *Proc Natl Acad Sci U S A*. 107(28):12676-81.
- 699 Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex statistical regularities in rat  
700 auditory cortex. *Neuron*. 76(3):603-15.
- 701
- 702
- 703
- 704
- 705
- 706
- 707
- 708
- 709

710 **Figure Legends**

711

712 **Figure 1. Avian auditory pathway, acoustic features of song stimuli and experimental**  
713 **protocol.**

714 **A:** Ascending auditory pathways in songbirds are indicated by arrows. Auditory nuclei of avian  
715 hindbrain innervate MLd (homolog of the inferior colliculus). MLd innervates OV (homolog of the  
716 medial geniculate). OV projects to forebrain field L2 (orange; analog of lay IV of A1). Field L2  
717 innervates L1, L3, NCM and CLM (red; analogs of superficial layers of A1). (modified from  
718 Figure1, Theunissen and Shaevitz, 2006).

719 **B:** Example spectrograms of a zebra finch song (upper panel) and a canary song (lower panel).  
720 Songs of the two species differ both in their spectra and temporal structure.

721 **C:** Spectral envelopes of canary and zebra finch songs used as context stimuli in these  
722 experiments.

723 **D:** Differences between canary and zebra finch songs used as context stimuli in these  
724 experiments, quantified in five acoustic dimensions. Each vertical panel represents  
725 measurement in one dimension: pitch, peak frequency, frequency modulation, amplitude  
726 modulation and Wiener entropy (a measure of the width and uniformity of the power spectrum).  
727 In each panel, the test song is shown as a red asterisk, for comparison with zebra finch context  
728 songs (red circles) and canary context songs (blue circles). Values for zebra finch and canary  
729 songs overlap somewhat in each dimension but differ in central tendency and/or variance. The  
730 test song (red asterisk) in canary context is at the edge or outside of the distribution of canary  
731 songs in 4 out 5 dimensions. Specialized software was used to extract and quantify the different  
732 acoustic features (SAP2011;Tchernichovski et al., 2000)

733 **E:** Schematic of the stimulus presentation sequence in Experiment 1. Each row represents one  
734 test condition. The stimulus order in the Pre-adapting phase and the Post-context phase is the  
735 same for all conditions (red boxes, 20 repeats of the test stimulus), but differs in the Context-  
736 modulated phase, as follows. **Canary context:** zebra finch test song in the context of 7 canary  
737 songs (c1 to c7) in shuffled order; **Silence context:** zebra finch song in silent context in shuffled  
738 order; context stimuli were replaced by same duration of silence; **ZF context:** zebra finch song  
739 in the context of 7 zebra finch songs (z1 to z7) in shuffled order.

740

741 **Figure 2. Response measurement and calculation of context effects.**

742 **A:** Example of electrophysiological recordings and multi-unit RMS computation. For a single trial,  
743 the spectrogram of the song stimulus (bottom) and the multi-unit activity recorded from two  
744 electrodes (green traces) is shown together with the moving RMS (window = 50 ms, only for  
745 display purposes) calculated for each recording (black traces). Single units were extracted from  
746 these recording. Responses to each song were defined as the difference between the spike  
747 counts (for single-units) or RMS (for multi-units) during the response window (stimulus period  
748 plus 100ms) minus the spike counts or RMS during the baseline window (500 ms preceding  
749 stimulus onset).

750 **B:** Illustration of Delta-surprisals computation. Each dot in the figure represents a response to  
751 the test song on one trial. Black dashed lines indicate the borders between the three phases.  
752 Solid black lines are regression lines obtained from responses (blue dots) in the pre-adapting  
753 and post-context phases. The solid green line connecting the end points of these regressions  
754 (green circles) shows the linear estimate of the mean of the expected distribution of responses  
755 for each trial during the Context-Modulated phase if the context had no effect on the course of  
756 adaptation of the response to the test song. The extent to which the observed responses in the

757 Context-modulated phase (red dots) are unexpected is measured by their surprisal, which is  
758  $1/\log(P(d))$ , normalized by the surprisal of the expected response, which is  $\log(1/P(0))$ . Note  
759 that the expectation of the signed surprisals is 0 (see text for further details).

760

761

762 **Figure 3. Single-unit responses to the test stimulus are increased in the context  
763 condition.**

764 **A:** Single-unit activity recorded in the Canary context condition. The top trace (black) shows the  
765 amplitude waveform of the test song. The middle panel shows the raster plot of spike responses  
766 in the Canary context condition for all three phases (blue dots for spikes in the pre-adapted and  
767 post-context phase, and red dots in the context-modulated phase), with trials ordered from  
768 bottom to top. The red vertical line indicates the stimulus onset. The horizontal dashed lines  
769 indicate the change between phases. Spiking increased during the context-modulated phase.  
770 The bottom panel compares the PSTH obtained from the Pre-adapted phase (blue) and  
771 Context-modulated phase (red).

772 **B:** Plot of spike rates of the neuron shown in A. Spike rates are shown for the Pre-adapting  
773 phase and the Post-context phase (blue dots), and for the Context-modulated phase (red dots).  
774 Decreasing responses in the first phase demonstrate adaptation. Increased activity in the  
775 Context-modulated phase demonstrates the enhancement effect (cf Figure 2B).

776 **C:** Examples of spike rate plots for 4 other neurons (4 rows) in 3 context conditions (3 columns).  
777 Spike rates for each trial were normalized by the spike rate of the first trial in the same condition  
778 for convenience of visual comparison in this plot (conventions as in B). Responses In the  
779 Canary and Silence contexts clearly increased compared to responses in the pre-adapting

780 phrase, while smaller increases are seen in the ZF context. In the post-context phase,  
781 responses resumed at the original adapted level or below. .

782 **Figure 4. Response enhancement in the Context-modulated phase measured as Delta-**  
783 **surprisals .**

784 **A:** Cumulative frequency distributions of Delta-surprisals obtained for single-units in Experiment  
785 1. The cumulative frequency for each x value as a percent of the total is shown on the Y axis as  
786 a function of Delta-surprisal values on the X axis. Delta-surprisals for Canary (red),  
787 silence( blue) and ZF (black) context conditions are plotted. The box plot (right) shows the  
788 minimum (the bottom of the whisker), first quartile (the bottom of the box), median (the red line  
789 in the middle of the box), third quartile (the top of the box), and maximum (the top of the whisker)  
790 and possible outliers (red cross).of Delta-surprisals in the ZF context condition (ZF, left box),  
791 Silence context condition (Silence, middle box) and Canary context condition (Can, right box)..

792 **B:** Cumulative frequency distributions (left) and box plot (right) of Delta-surprisals obtained for  
793 multi-units in the three conditions of Experiment 1. Multi-unit data showed the same pattern of  
794 results as single-unit data.

795 **C:** Cumulative frequency distributions (left) and box plot (right) of Adaptation Indices for the  
796 three conditions tested in Experiment 1. Adaptation indices were not different from each other in  
797 the three conditions.

798 **D:** Scatter plot of absolute values of D' and Delta-surprisals. The black dashed line is a flat  
799 reference line (slope zero). Red dashed line shows the best fitting line from a linear regression,  
800 which shows no significant correlation between the two variables (Spearman:  $r = 0.056, p >$   
801 0.557).

802 **E:** Cumulative frequency distributions (left) and box plot (right) of slopes of linear regression  
803 between trial number and trial-by-trial Delta-surprisals in the three conditions of Experiment 1. In  
804 the Canary context condition and Silence context condition, there was no significant change in  
805 Delta-surprisals over trials. In contrast, Delta-surprisals significantly increased with trials in the  
806 ZF context condition.

807 **Figure 5. Response latency and timing of response enhancement.**

808 **A:** The enhancement profile (black trace) for the Canary context is displayed as the difference  
809 between the averaged response waveforms for the test song in the Pre-adapting phase (blue  
810 trace) and the Context-modulated phase (red trace). Shaded areas along the traces show  
811 standard errors. The vertical dashed line indicates the onset of the stimulus.

812 **B:** The early phase of the averaged response waveforms and the enhancement profile for the  
813 Canary context condition, from Figure 5A, at a finer time scale. The solid vertical lines indicate  
814 the latency of responses from stimulus onset (red) and the latency of the enhancement profile  
815 (black; 21ms later than the red line). The latency was computed as the time from stimulus onset  
816 until the signal crossed a threshold, computed as the maximum value of the 99% confidence  
817 interval for each signal during the baseline window. Dashed red line with triangle indicates the  
818 peak of the response (85 ms after stimulus onset). The arrow and dashed black line with  
819 triangle indicate the peak of the enhancement profile (112 ms after stimulus onset). Both the  
820 latency and the peak of the enhancement profile occur later than the latency and the peak of the  
821 responses.

822 **C:** Comparison between the enhancement profiles for the Canary context condition (red trace)  
823 and for the Zebra finch (ZF) context condition (black trace). The ZF context condition shows no  
824 clear enhancement effect (note: gain is higher than in **A**).

825 **D:** Comparison of the early phase of the enhancement profiles in the Canary context condition  
826 (red trace) and in the Silence context condition (blue trace) showed a shorter latency of the  
827 enhancement effect in the Silence context condition. The green dashed lines indicate the  
828 window 10-20 ms after the stimulus onset, in which significant enhancement was observed in  
829 the Silence context condition, but not in the Canary context condition (also see **F**).  
  
830 **E:** Comparison between the enhancement profiles in the Canary (red trace) and the Silence  
831 context condition (blue trace) showed a shorter duration for the enhancement profile in the  
832 Silence context condition. The green dashed lines indicate the window 490-690 ms after the  
833 stimulus onset when the profile for the Silence context condition decays to zero, but the  
834 enhancement effect is still significant for the Canary context condition (also see **G**).  
  
835 **F:** Box plot of multi-unit RMS differences between the Pre-adapting phase and the Context-  
836 modulated phase in the window 10-20 ms after the stimulus onset (between dashed green lines  
837 in **D**) was compared between the Canary context and the Silence context. There was a  
838 significant increase in response amplitude in this time window in the Silence context (left box;  
839 Wilcoxon  $z = 2.8$ ;  $p < 0.006$ ,  $n = 98$ ), but not in the Canary context (right box; Wilcoxon  $z = 1.5$ ;  
840  $p > 0.123$ ,  $n = 98$ ).  
  
841 **G:** Box plot of multi-unit RMS differences between the Pre-adapting phase and the Context-  
842 modulated phase in the window 490-690 ms after the stimulus onset (between dashed green  
843 lines in **E**) was compared between the Canary and Silence contexts. There was a significant  
844 increase in response amplitude in this time window in the Canary context (left box; Wilcoxon  $z =$   
845  $6.90$ ;  $p < 0.001$ ,  $n = 98$ ), but not in the Silence context (right box; Wilcoxon  $z = -1.24$ ;  $p > 0.216$ ,  
846  $n = 98$ ).  
  
847

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

852

853 **Figure 7. Effects of pre-exposure to context stimuli tested in Experiment 3.**

854 **A:** Cumulative frequency distributions and box plot of Delta-surprisals for the ZF context  
855 condition tested before and after extensive exposure to context songs in Experiment 3. Delta-  
856 surprisals increased when context songs were familiar due to prior exposure.

857 **B:** Cumulative frequency distributions of **D'** obtained from the context songs (blue curves) and  
858 the test song (red curve) before exposure training are not distinguishable.

859 **C:** Cumulative frequency distributions of **D'** obtained from the context songs and the test song  
860 after exposure training are significantly different.

861

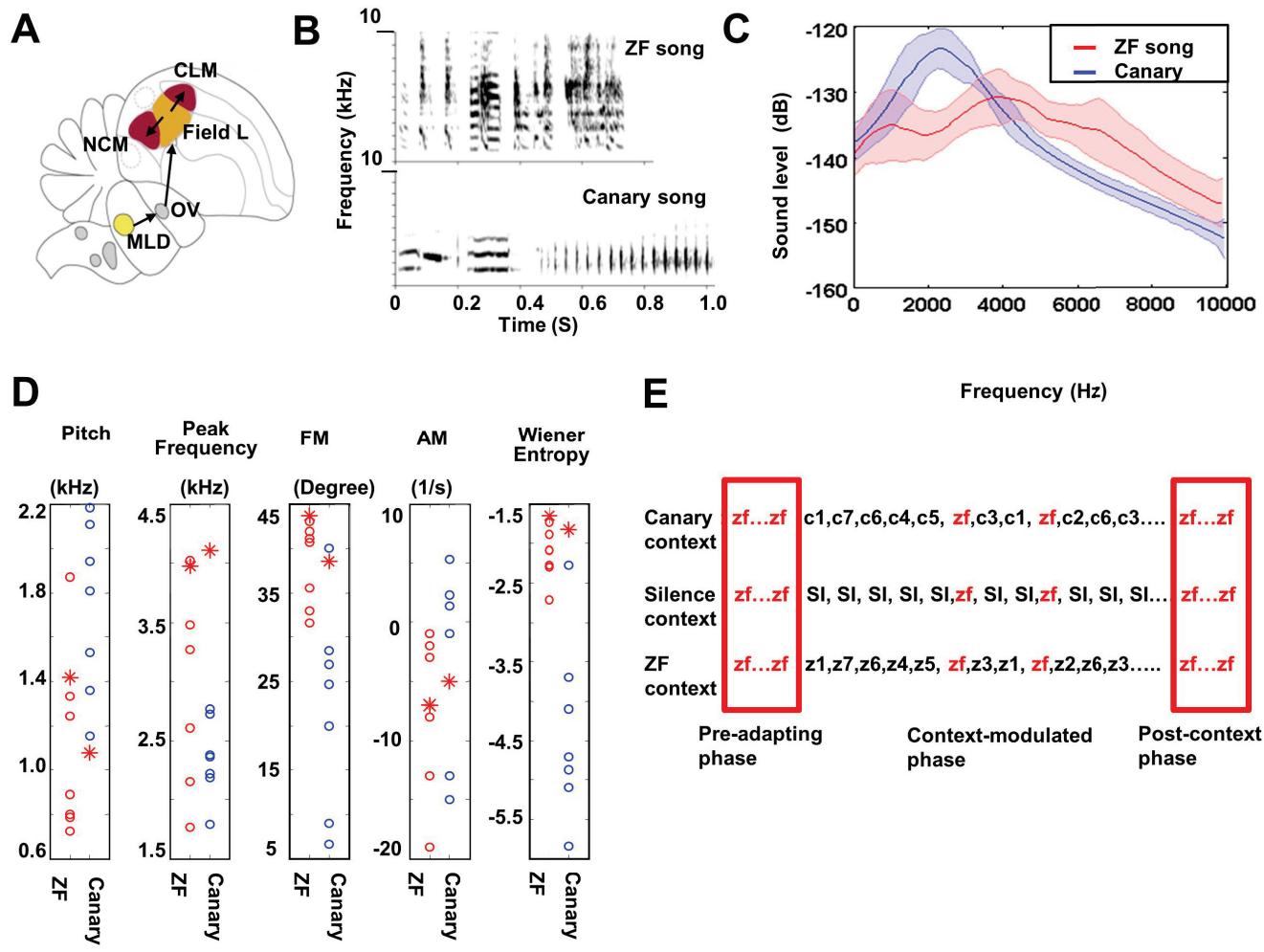


Figure 1  
2 Column

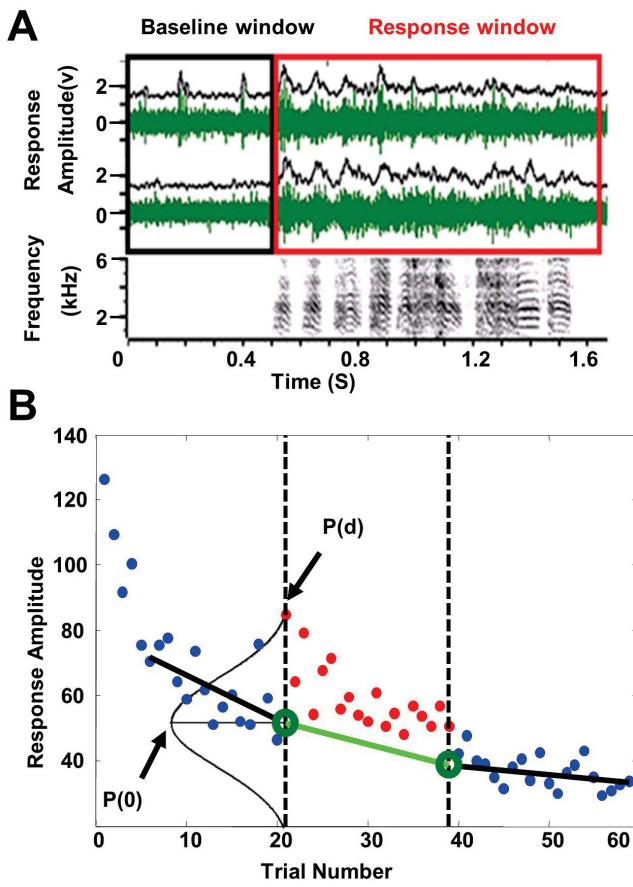


Figure 2  
1 Column

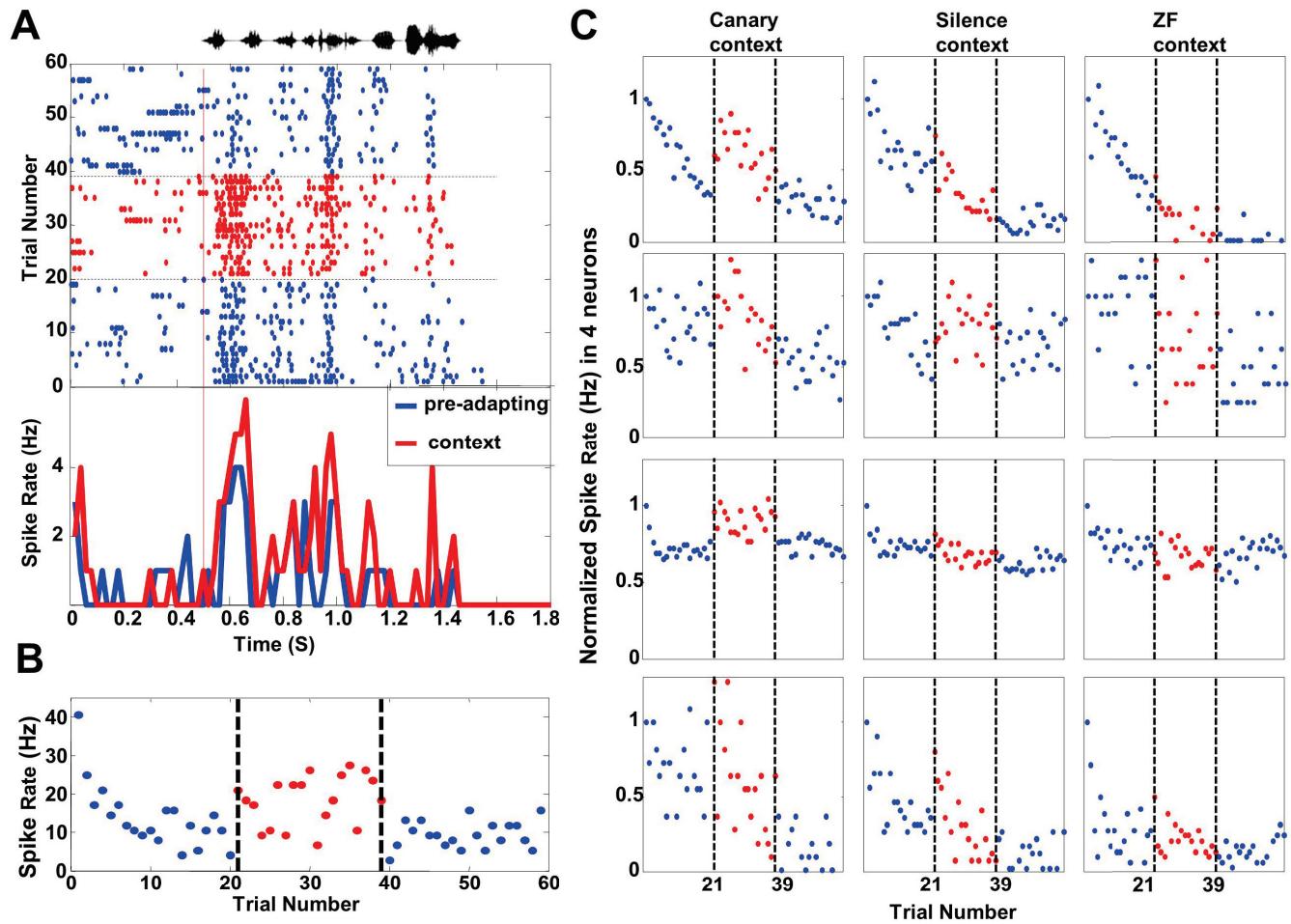


Figure 3  
2 Column

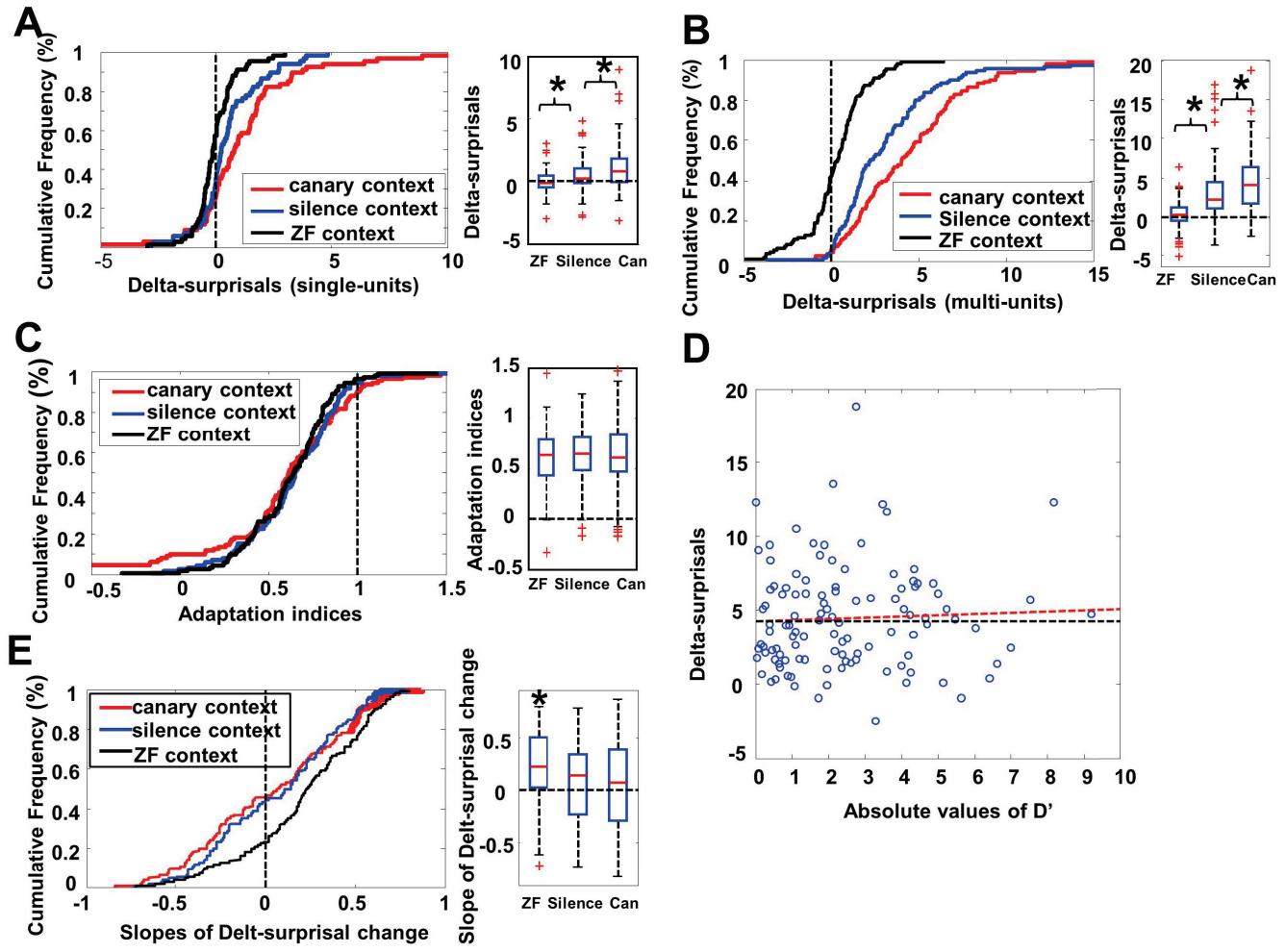


Figure 4  
2 Column

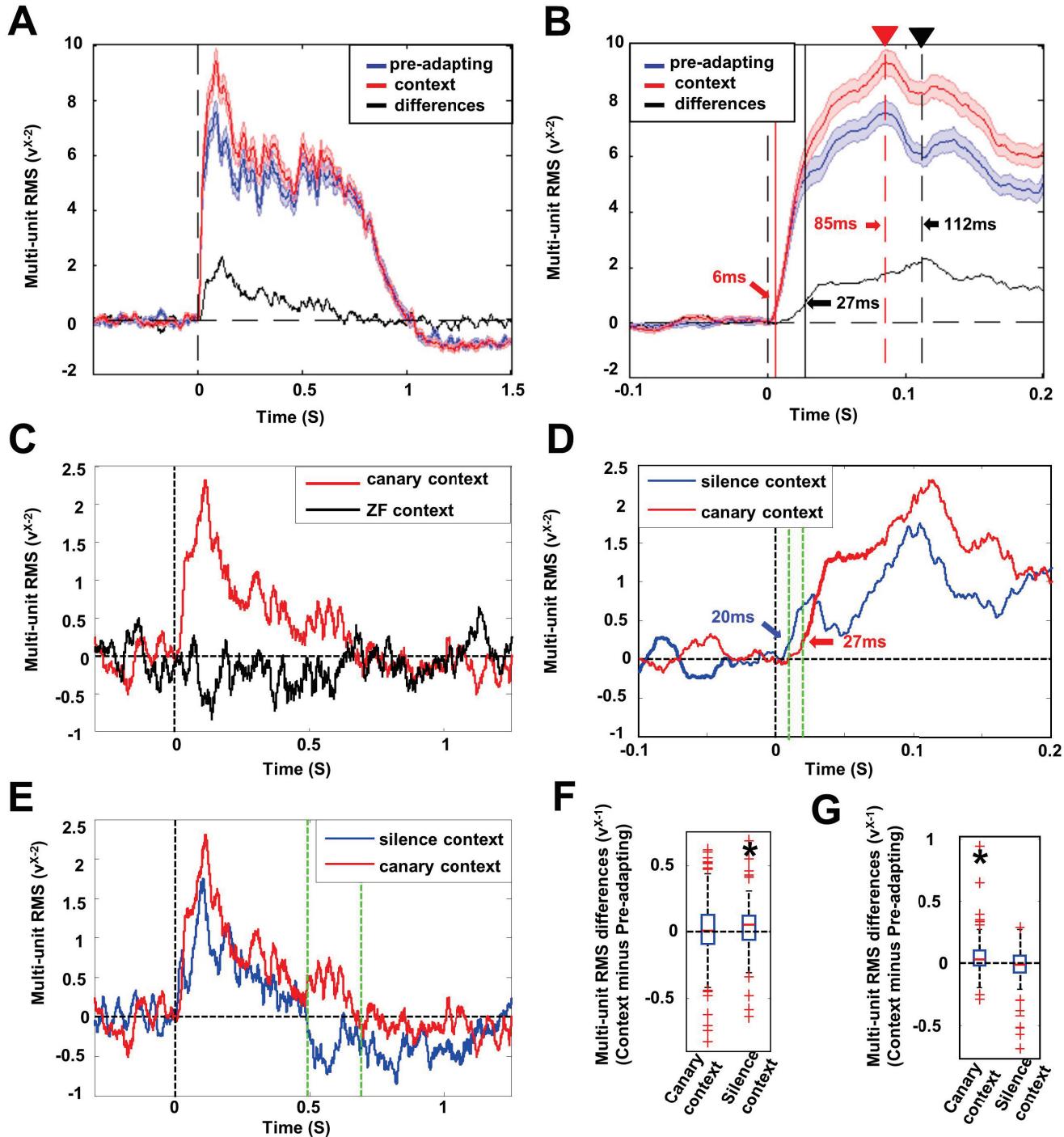


Figure 5  
2 Column

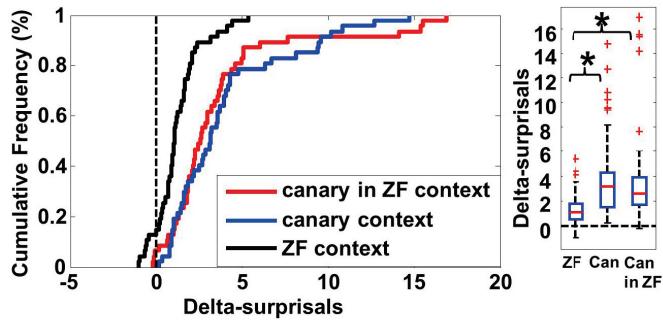


Figure 6  
1 Column

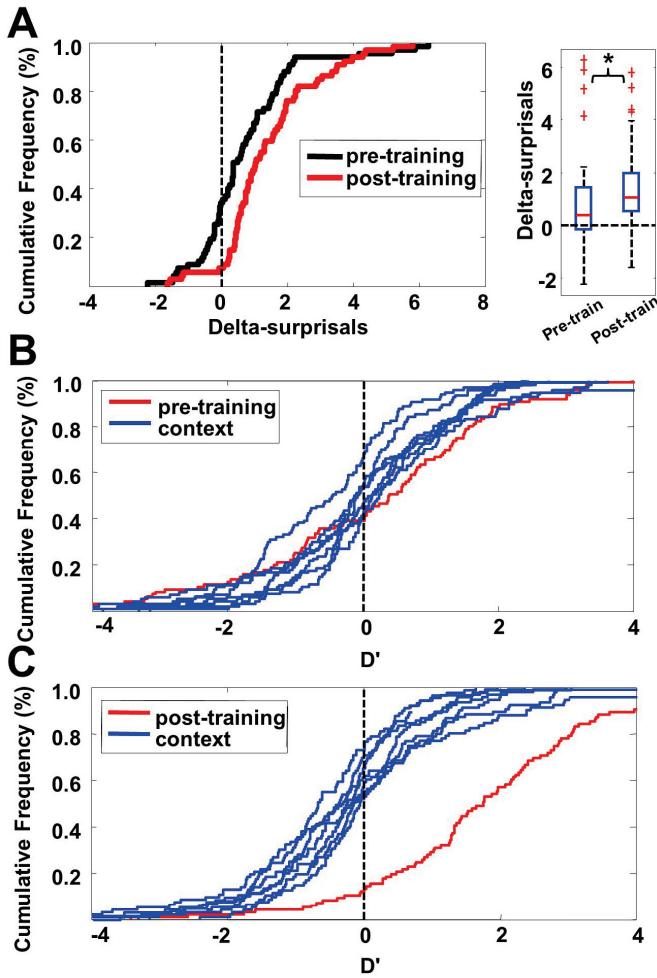


Figure 7  
1 Column