

# Acoustic Context Modulates Natural Sound Discrimination in Auditory Cortex through Frequency-Specific Adaptation

 Luciana López-Jury,  Francisco García-Rosales,  Eugenia González-Palomares, Manfred Kössl, and  Julio C. Hechavarría

Institute for Cell Biology and Neuroscience, Goethe University, 60438 Frankfurt am Main, Germany

Sound discrimination is essential in many species for communicating and foraging. Bats, for example, use sounds for echolocation and communication. In the bat auditory cortex there are neurons that process both sound categories, but how these neurons respond to acoustic transitions, that is, echolocation streams followed by a communication sound, remains unknown. Here, we show that the acoustic context, a leading sound sequence followed by a target sound, changes neuronal discriminability of echolocation versus communication calls in the cortex of awake bats of both sexes. Nonselective neurons that fire equally well to both echolocation and communication calls in the absence of context become category selective when leading context is present. On the contrary, neurons that prefer communication sounds in the absence of context turn into nonselective ones when context is added. The presence of context leads to an overall response suppression, but the strength of this suppression is stimulus specific. Suppression is strongest when context and target sounds belong to the same category, e.g., echolocation followed by echolocation. A neuron model of stimulus-specific adaptation replicated our results *in silico*. The model predicts selectivity to communication and echolocation sounds in the inputs arriving to the auditory cortex, as well as two forms of adaptation, presynaptic frequency-specific adaptation acting in cortical inputs and stimulus-unspecific postsynaptic adaptation. In addition, the model predicted that context effects can last up to 1.5 s after context offset and that synaptic inputs tuned to low-frequency sounds (communication signals) have the shortest decay constant of presynaptic adaptation.

**Key words:** auditory cortex; bats; integrate-and-fire; natural sounds; neuroethology

## Significance Statement

We studied cortical responses to isolated calls and call mixtures in awake bats and show that (1) two neuronal populations coexist in the bat cortex, including neurons that discriminate social from echolocation sounds well and neurons that are equally driven by these two ethologically different sound types; (2) acoustic context (i.e., other natural sounds preceding the target sound) affects natural sound selectivity in a manner that could not be predicted based on responses to isolated sounds; and (3) a computational model similar to those used for explaining stimulus-specific adaptation in rodents can account for the responses observed in the bat cortex to natural sounds. This model depends on segregated feedforward inputs, synaptic depression, and postsynaptic neuronal adaptation.

## Introduction

Studying how vocalizations are processed in the brain of listeners is key to understanding the neurobiology of acoustic communication. Bats are highly vocal social animals. They live in complex auditory environments filled with mixtures of their own calls and vocalizations emitted by conspecifics. Bat vocalizations can

be classified into two categories, namely, echolocation and communication. Echolocation is primarily used for navigation (Neuweiler, 1989; Thomas et al., 2005), and communication calls convey a variety of messages to conspecifics (Gillam and Fenton, 2016; Chaverri et al., 2018).

Previous studies in the mustached bat described that neurons in specialized areas of the auditory cortex (AC) not only exhibit combination sensitivity for echolocation pulses (O'Neill and Suga, 1979; Fitzpatrick et al., 1993) but also to communication calls (Ohlemiller et al., 1996; Kanwal, 1999). Although several studies clearly indicated that within the AC both echolocation and communication signals can be processed by the same neurons (Esser et al., 1997; Medvedev and Kanwal, 2004, 2008; Kanwal, 2012), how these multifunctional neurons process combinations of these two signals is still unknown. The latter constitutes a relevant question for neuroethology because acoustic

Received Apr. 23, 2021; revised Oct. 21, 2021; accepted Oct. 26, 2021.

Author contributions: L.L.-J., F.G.-R., E.G.-P., M.K., and J.C.H. designed research; L.L.-J. performed research; L.L.-J. analyzed data; and L.L.-J. wrote the paper.

This work was supported by Deutsche Forschungsgemeinschaft Project No. 428645493 (J.C.H.). We thank Michael Pasek for assistance during the writing of this manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Luciana Lopez-Jury at lopezjury@bio.uni-frankfurt.de or Julio C. Hechavarría at hechavarría@bio.uni-frankfurt.de.

<https://doi.org/10.1523/JNEUROSCI.0873-21.2021>

Copyright © 2021 the authors

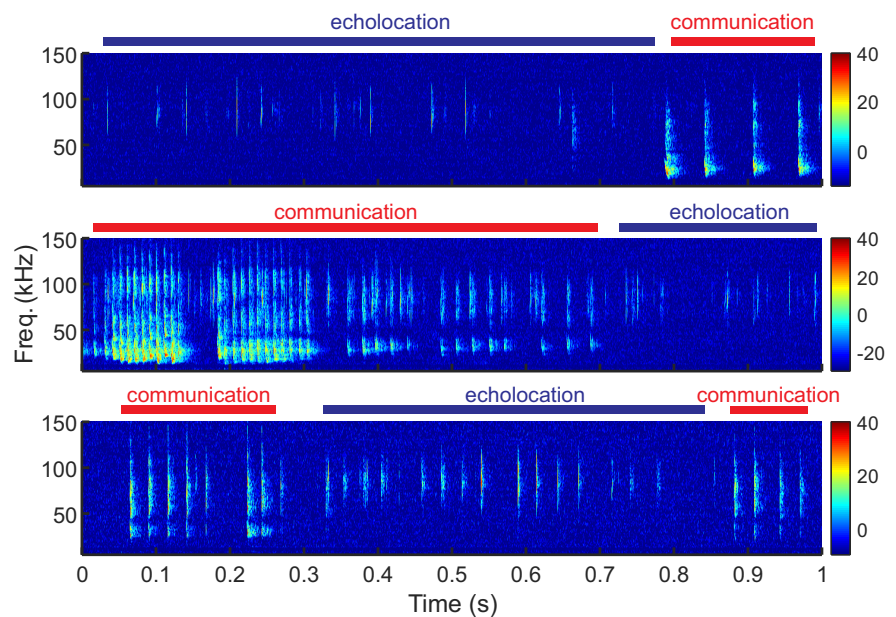
transitions, that is, communication sounds following echolocation sequences or vice versa, are common in the natural environments of bats (Fig. 1).

The main aim of the present study is to characterize how AC neurons respond to the acoustic transitions mentioned above. The study was performed in the bat species *Carollia perspicillata*. The echolocation and communication calls of *C. perspicillata* differ markedly in frequency composition, with energy peaking between 12 and 40 kHz in communication sounds and above 60 kHz in echolocation signals (Porter, 2010; Knörnschild et al., 2014; Hechavarría et al., 2016). Neurons in nontotopic areas (high-frequency fields) of the AC of *C. perspicillata* exhibit multi-peaked frequency tuning curves peaking at frequencies close to 20 and 60 kHz, which closely match the peak frequency of communication and echolocation sounds, respectively (Hagemann et al., 2010, 2011).

We tested how neurons in nontotopic areas of the AC respond to combinations of natural echolocation and communication sounds. As an example of communication calls, we chose distress vocalizations, a type of communication sound emitted under duress, which carries strong communicative power and has been described extensively (Knörnschild et al., 2014; Hechavarría et al., 2016, 2020; González-Palomares et al., 2021). Our goal was to test how AC neurons with multi-peaked frequency tuning, which could potentially respond to both echolocation and communication signals, process sound mixtures in which a leading sequence of natural sounds (termed context sequence) is followed by a probe sound. Context and probe could match (e.g., echolocation following echolocation) or mismatch each other (e.g., communication following echolocation).

Single-neuron recordings revealed that leading echolocation and communication sequences modulate cortical responses to lagging sounds. Context-triggered suppression was stimulus specific, with suppression being strongest when both context and probe sounds belonged to the same acoustic category (matched treatment). Context effects suppressed responses to probe sounds even 400 ms after context offset. An integrate-and-fire neuron model, similar to synaptic models designed to explain stimulus-specific adaptation to artificial sounds in rodents (Taaseh et al., 2011; Hershshoren et al., 2014), could replicate our empirical results. The model predicted high selectivity to communication and echolocation sounds in the inputs arriving in AC neurons of bats, as well as two forms of adaptation, presynaptic frequency-specific adaptation acting in cortical inputs and stimulus-unspecific postsynaptic adaptation. The model also predicted the temporal course of synaptic depression, which differed between the two types of acoustic contexts tested, with faster adaptation in response to communication than to echolocation sounds.

Our findings have two main implications. (1) Context-driven suppression can change the abilities of neurons to discriminate between ethologically relevant natural sounds in awake, passive listening animals, and (2) models of stimulus-specific adaptation developed for explaining responses to artificial stimuli in classic laboratory animals (i.e., rodents) also can explain responses to natural ethologically relevant sounds in bats.



**Figure 1.** Sound transitions in natural environments of bats. Three 1 s snapshots of audio recordings in a group of bats (6 animals in a keeping box). Spectrograms show vocalization type transitions between the categories of echolocation and communication. The color bar indicates sound amplitude (arbitrary units).

## Materials and Methods

### Animal preparation

Six bats (two female, *Carollia perspicillata*) were used in this study. They were taken from a breeding colony in the Institute for Cell Biology and Neuroscience at the Goethe University in Frankfurt am Main, Germany. All experiments were conducted in accordance with the Declaration of Helsinki and local regulations in the state of Hessen (Experimental permit FU1126, Regierungspräsidium Darmstadt).

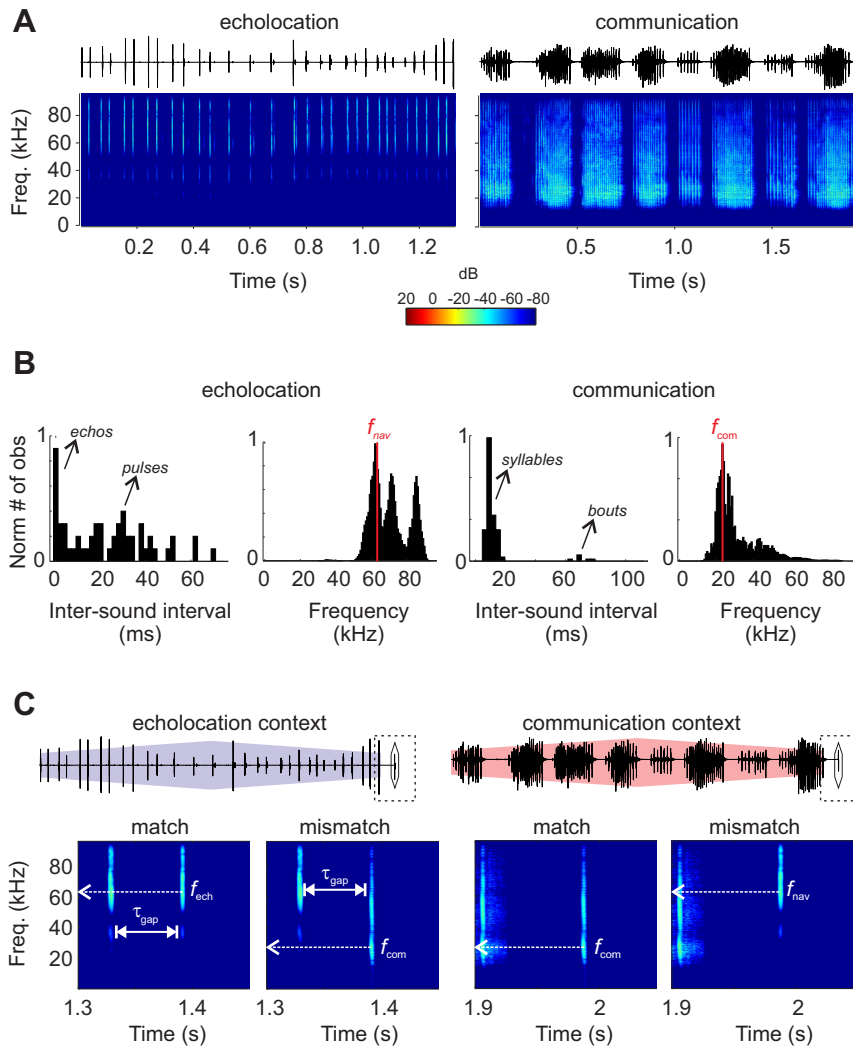
The bats were anesthetized with a mixture of ketamine (100 mg/ml Ketavet; Pfizer) and xylazine hydrochloride (23.32 mg/ml Rompun; Bayer). Under deep anesthesia, the dorsal surface of the skull was exposed. The underlying muscles were retracted with an incision along the midline. A custom-made metal rod was glued to the skull using dental cement to fix the head during electrophysiological recordings. After the surgery, the animals recovered for 2 d before participating in the experiments.

On the first day of recordings, a craniotomy was performed using a scalpel blade on the left side of the cortex in the position corresponding to the auditory region. Particularly, the caudoventral region of the auditory cortex was exposed, spanning primary and secondary auditory cortices (AI and AII, respectively), the dorsoposterior field and high-frequency (HF) fields. The location of these areas was made by following patterns of blood vessels and the position of the pseudocentral sulcus (Esser and Eiermann, 1999; Hagemann et al., 2010).

### Neuronal recordings

In all six bats, recordings were performed over a maximum of 14 d. Experiments were conducted in awake animals. Bats were head fixed and positioned in a custom-made holder over a warming pad set to a temperature of 27°C. Local anesthesia (Ropivacaine 1%, AstraZeneca) was administered topically over the skull before each session. Each recording session lasted a maximum of 4 h.

All experiments were performed in an electrically isolated and soundproof chamber. For neural recordings, carbon-tip electrodes (impedance ranged from 0.4 to 1.2 M $\Omega$ ) were attached to an electrode holder connecting the electrode with a preamplifier to a DAGAN four-channel amplifier (Dagan EX4-400 Quad Differential Amplifier, gain = 50, filter low cut = 0.03 Hz, high cut = 3 kHz). Analog-to-digital conversion was achieved using a sound card (RME ADI-2 Pro, sample rate = 192 kHz). Electrodes were driven into the cortex with the aid of a Piezo



**Figure 2.** The stimulation paradigm. **A**, Oscillogram and spectrogram of natural calls of *C. perspicillata* associated with echolocation and communication behaviors; a sequence of echolocation pulses (left) and a distress call (right). **B**, Distribution of intersound intervals and frequencies of the sequences above. Arrows indicate independent components within the sequences. Red lines indicate the peak of the distribution ( $f_{ech} = 66$  kHz and  $f_{com} = 23$  kHz). **C**, Top, Oscillograms of the stimulation protocol. Natural sequences preceding a short probe (highlighted by a diamond). Bottom, spectrograms of the last 150 ms of stimulus (dashed rectangle) point out the last element of each sequence and the subsequent probe for the contextual match and mismatch conditions. Time between the offset of the sequence and the onset of the probe is indicated as  $\tau_{gap}$  in white, and the frequency at peak energy of the probes is indicated by dashed arrows.

manipulator (PM 10/1, Science Products). Single-unit auditory responses were located at depths of  $308 \pm 79 \mu\text{m}$ , mean  $\pm$  SD, using a broadband search stimulus (downward frequency modulated communication sound of the same bat species) that triggered activity in both low- and high-frequency-tuned neurons. A similar paradigm has been used in previous studies to locate neuronal responses (Martin et al., 2017).

#### Acoustic stimulation

We used natural sounds to trigger neural activity during the recordings. The natural sounds were obtained from the same species in previous studies from our lab (Beetz et al., 2016; Hechavarría et al., 2016). Acoustic signals were generated with an RME ADL2 Pro Sound card and amplified by a custom-made amplifier. Sounds were then produced by a calibrated speaker (NeoCD1.0 Ribbon Tweeter; Fountek Electronics), which was placed 15 cm in front of the bat right ear of the bat. The calibration curve of the speaker was calculated with a microphone (model #4135, Brüel & Kjær).

Once an auditory neuron was located, we determined the isolevel frequency tuning of the unit, with 20 ms pure tones (0.5 ms rise/fall time) presented randomly in the range of frequencies from 10 to 90 kHz

(5 kHz steps, 20 trials) at a fixed sound pressure level of 80 dB SPL. This was done only in a subset of the neurons recorded (55 of a total of 74). Our previous studies indicate that 80 dB SPL represents a good compromise as it is strong enough to drive activity in most AC neurons and allows to differentiate between single-peaked and double-peaked tuning curves typical of the AC of this bat (López-Jury et al., 2020). The repetition rate for stimulus presentation was 2 Hz.

We studied context-dependent auditory responses using two types of acoustic contexts, sequences of echolocation pulses and sequences of communication (distress calls). We refer to these two sequences of natural vocalization as context because they preceded the presentation of probe sounds that were either a single echolocation pulse or a single distress syllable. The properties of the contexts used for stimulation are shown in Figure 2A,B. In a nutshell, the echolocation sequence was composed of high-frequency vocalizations and echoes of the vocalizations (carrier frequencies  $> 50$  kHz) repeated at intervals of  $\sim 30$  ms (Fig. 2A,B, left). The echolocation sequence was recorded from a bat swung in a pendulum following procedures described previously (Beetz et al., 2016). The distress sequence, on the other hand, was composed of individual syllables with peak frequencies of  $\sim 23$  kHz (Fig. 2A,B, right). Distress syllables occurred in groups (so-called bouts) repeated at intervals of  $\sim 60$  ms, and within the bouts, syllables occurred at rates of  $\sim 15$  ms (Hechavarría et al., 2016). We chose to study these two acoustic contexts because they rely on fundamentally different acoustic parameters and are linked to distinct behaviors, that is, echolocation and calling under duress (distress).

Probe sounds (sounds that followed the context) were single echolocation and distress syllable, each obtained from the respective context sequences (Fig. 2C). The sound level of both probes was fixed to 73 dB SPL. These levels resemble a bat vocalizing 1.5 m away from the listener. Distress calls at similar intensities have been recorded in our previous work with a microphone 1.5 m away from the bats (median of 70 dB SPL; Hechavarría et al., 2016). On the other hand, the emitted call intensity of echolocation pulses in this bat species has been reported as source levels (referenced to 10 cm away from the mouth of the bat) averaging 99 dB SPL (Brinklov et al., 2011). Considering spherical spreading (6 dB/doubling of distance) and atmospheric attenuation at peak frequency of 90 kHz ( $\sim 2.2$  attenuation/m; Lawrence and Simmons, 1982), 73 dB SPL would correspond to a bat echolocating 1.6 m away.

We tested two temporal separations (gaps) between the context offset and probe onset. A previous study in the same region of the AC of *C. perspicillata* showed neuronal responses to a masker-probe stimulation paradigm using systematically spaced delays. The neurons started showing responses to the probes only 50 ms after the masker offset; at shorter delays it was difficult to disentangle responses to maskers (i.e., the last elements of the context sequence) from those evoked by the probe sounds (Martin et al., 2017). We decided to use gaps above 50 ms in the same order of magnitude (60 ms) or much longer (416 ms) to measure the context effect in our paradigm. Especially in bats, short delays could

lead to response facilitation that although interesting was not in the main scope of this study and has been characterized before in the context of echolocation, that is, delay tuning (Wenstrup and Portfors, 2011; Suga, 2015), and communication (Esser et al., 1997). Therefore, a total of eight context stimuli (2 contexts  $\times$  2 probes  $\times$  2 gaps) were randomly presented and repeated 20 times to awake bats during electrophysiological recordings—echolocation context followed by echolocation probe, echolocation context followed by distress probe, distress context followed by echolocation probe, and distress context followed by distress probe. In addition, we presented each probe after 3.5 s of silence (no context) and repeated it 20 times.

#### Data analysis

**Spike clustering.** All the recording analyses, including spike sorting, were conducted using custom-written MATLAB scripts (R2018b, MathWorks). The raw signal was filtered between 300 Hz and 3 kHz using a bandpass Butterworth filter (third order). To extract spikes from the filtered signal, we detected negative peaks that were at least 3 SDs above the recording baseline; times spanning 1 ms before the peak and 2 ms after were considered as one spike. The spike waveforms were sorted using an automatic clustering algorithm, KlustaKwik, which uses results from principal component analysis to create spike clusters (Harris et al., 2000). For each recording, we considered only the spike cluster with the highest number of spikes.

**Neuronal classification.** A total of 74 units were considered responsive to at least one of the probe sounds tested. A unit was considered responsive if the number of spikes fired in response to the sound in question was above the 95% confidence level calculated for spontaneous firing for the same unit (calculated along 200 ms before the start of each trial). Evoked firing had to surpass this threshold for at least 8 ms after probe onset for a unit to be considered responsive. To test for natural sound preferences in each unit (i.e., whether units preferred echolocation vs distress sounds or vice versa), we used the responses to the probe sounds when they were preceded by silence (no context). The spike counts during 50 ms after each probe onset were compared using a nonparametric effect size metric, Cliff's delta. Cliff's delta quantifies effect size by calculating how often values in one distribution are larger than values in a second distribution. Distribution here refers to the number of spikes fired in presentations of echolocation and communication probe trials. The statistic gives values from  $-1$  to  $1$ , with identical groups rendering values of zero. Following previous studies (Romano et al., 2006), if the effect size between the two distributions was negligible or small [ $\text{abs}(\text{Cliff's delta}) \leq 0.3$ ], the unit was classified as equally responsive to both probes. On the other hand, if  $\text{abs}(\text{Cliff's delta}) > 0.3$ , the unit was classified either as preference to echolocation or preference to communication. The preference was assigned to the probe that evoked the highest number of spikes across trials. If a unit was responsive to only one of the probes, it was classified as responsive to only echolocation or only communication. We checked the classification criterion by using nonparametric Wilcoxon rank-sum tests. Rank-sum test failed to reject the null hypothesis when comparing spike counts of the responses to both probes in all the units classified as equally responsive based on the Cliff's delta metric. In contrast, in all 16 units considered as preference to communication, the null hypothesis was rejected, that is, they showed significant differences in spike counts across trials ( $p < 0.05$ , Wilcoxon rank-sum test).

The isolevel tuning curves were obtained from the average of spike counts across trials  $x_i$  in response to pure tones at  $N$  frequencies, rescaled by minimum/maximum normalization as follows:

$$x_{i,norm} = \frac{x_i - \min(x_1, \dots, x_N)}{\max(x_1, \dots, x_N) - \min(x_1, \dots, x_N)}$$

Units were classified as low-frequency tuned if the normalized response was higher than or equal to 0.6 at any frequency lower than 50 kHz and lower than 0.6 for all frequencies higher than or equal to 50 kHz. High-frequency tuned neurons were those in which  $x_{i,norm}$  was above 0.6 for any frequency higher than or equal to 50 kHz and lower than 0.6 for all of the frequencies lower than 50 kHz. Multi-peaked

neurons were those in which  $x_{i,norm}$  exceeded 0.6 in both frequency bands ( $<50$  and  $\geq 50$  kHz). A similar classification has been used in previous studies (López-Jury et al., 2020).

**Quantifying the effects of leading acoustic context.** To test for effects of leading acoustic context on probe-triggered responses, we tested (using a nonparametric Wilcoxon rank-sum test) whether the response to the probe, in terms of number of spikes across trials occurring after the context, was significantly different from the responses to the same sound presented without any context. To quantify the magnitude of the effect, we used the following equation:

$$\text{context effect} = \frac{s(\text{probe, context}) - s(\text{probe, silence})}{s(\text{probe, context}) + s(\text{probe, silence})},$$

where  $s(\text{probe, context})$  corresponds to the number of spikes during 50 ms from the onset of the corresponding probe following the corresponding context and  $s(\text{probe, silence})$  following silence. Effects of the same context on different probe-triggered responses were statistically compared per unit using paired statistics (Wilcoxon signed-rank test). In addition, we compared the calculated effect size (Cliff's delta test) between probes after silence with the same measurement after context, also using paired statistics (Wilcoxon signed-rank test). According to our convention, negative Cliff's delta values indicate higher responses to the communication probe than to the echolocation probe. Positive Cliff's delta values indicate the opposite.

**Call-evoked responses in equally responsive neurons.** To determine the latency and duration of single-unit responses to the probes in the no-context condition, we convolved the peristimulus time histograms (PSTHs) with a Gaussian kernel with a bin size of 1 ms and a bandwidth of 5 ms to obtain a spike density function. The latency of the response corresponded to the time at which the spike density function exceeded 50% of the peak response. The response ended when the density function decreased the 50% threshold. Response duration was calculated as the difference between the end and the latency of the response. Together with the spiking data, local field potentials (LFPs) were obtained from the same electrode with a low bandpass filter (0.1–300 Hz). Call-evoked LFPs were averaged and  $z$  normalized across trials. To estimate the energy of LFPs, we calculated the absolute value of the Hilbert transform of the signal within a time window of 0–150 ms. To quantify aspects of the temporal structure of the call-evoked LFPs, we calculated the times at negative and positive peaks, and at which the potential became positive between the mentioned peaks. Paired statistics were used to compare such measurements between echolocation and communication probes per unit (Wilcoxon signed-rank test).

#### Modeling possible synaptic origins of contextual neural response modulation

We modeled a broadly tuned cortical neuron that reproduces the behavior of equally responsive neurons observed in our experiments, using an integrate-and-fire neuron model. Our model (described in detail below) receives input from two narrow frequency channels and includes synaptic adaptation (i.e., activity-dependent depression) in the cortical input synapses. Frequency-specific adaptation in multiple synaptic channels has been proposed to mediate stimulus-specific adaptation in auditory cortex, including during speech processing (Taaseh et al., 2011; May and Tiitinen, 2013). All simulations were performed using the Python package Brian2, version 2.3 (Stimberg et al., 2019).

In the model proposed, the membrane potential ( $V_m$ ) of the neuron evolves according to the following differential equation:

$$\frac{dV_m}{dt} = \frac{g_L(E_L - V_m) + g_e(E_e - V_m)}{C_m} + \sigma \sqrt{2/\tau_\sigma} x_i,$$

where  $g_L$  is the leak conductance,  $E_L$  is the resting potential,  $C_m$  is the membrane capacitance,  $E_e$  is the excitatory reversal potential, and  $g_e$  is the synaptic conductance. The second term of the equation corresponds to the stochastic current added to the neuron model, and  $x_i$  is a

Gaussian random variable specified by an Ornstein–Uhlenbeck process, with variance  $\sigma$  and correlation time  $\tau_\sigma$ .

The neuron fires when  $V_m$  reaches the membrane potential threshold  $\omega_{th}$ , and the variable  $V_m$  is reset to the fixed value  $V_r$ . To implement an adaptive threshold to the model that decreases the probability of firing with postsynaptic activity. The firing threshold  $\omega_{th}$  starts at  $V_{th}$ , and after the occurrence of a spike, it is increased by the amount  $\Delta_{th}$  and returns back to  $V_{th}$  with a time constant  $\tau_{th}$ , that is, as follows:

$$\frac{d\omega_{th}}{dt} = \frac{(V_{th} - \omega_{th})}{\tau_{th}}.$$

The model includes conductance-based excitatory synapses; every synapse creates a fluctuation of conductance  $g_e$  that changes in time with time constant  $\tau_e$  as follows:

$$\frac{dg_e}{dt} = -\frac{g_e}{\tau_e}.$$

When an action potential arrives at a synapse, the excitatory conductance  $g_e$  increases in the postsynaptic neuron, according to the following:

$$g_e \leftarrow g_e + w_e X_s(t),$$

where  $w_e$  is the maximum synaptic weight and is modulated by a time-dependent adaptation term,  $X_s(t)$ . The effective synaptic weight between the inputs and the cortical neuron is the product  $w_e X_s(t)$  and depends on the presynaptic activity. The model assumes that  $X_s$  starts at one and a presynaptic spike arrival at synapse induces a decrement of  $X_s$  value by the amount  $\Delta_s$ , which recovers in time by the following:

$$\frac{dX_s}{dt} = \Omega_s(1 - X_s).$$

The neuron model has two excitatory synapses that receive inputs, which differ by preference of the inputs for natural sound stimulation. The input of the synapse  $j$  corresponds to a spike train that follows an inhomogeneous Poisson process with rate  $\lambda_j$  as follows:

$$\lambda_j(t) = s_{sound}(t)(k_{j,sound}v) + v_{spont},$$

where  $s_{sound}(t)$  corresponds to the envelope in time of the respective natural sound,  $k_{j,sound}$  is proportional to the response of input  $j$  to the respective sound,  $v$  is the average firing rate of the inputs, and  $v_{spont}$  is the spontaneous firing rate of the inputs.

#### Model parameters and data fitting

All the parameters used in the model are indicated in Table 1. The parameters associated with intrinsic properties of the neuron model were set to qualitatively mimic the extracellular activity of cortical neurons. The parameters that determine the dynamic of the adaptive threshold ( $\Delta_{th}$  and  $\tau_{th}$ ) were set to produce a reduction of 40–50% in the response to the probes after any context, as observed empirically.

In our model the reversal potential  $E_e$  of excitatory synapses was set to a standard value used for modeling AMPA synapses (Clopath et al., 2010). The time constant  $\tau_e$  was adjusted to obtain spiking responses to the probes qualitatively similar in duration with the observed experimental data.

The maximum synaptic weight  $w_e$  of each synapse was set to fit experimental data. We ran 100 simulations, systematically changing the maximum synaptic weight of each synapse independently from 1 to 20 nS, in steps of 2 nS. For each simulation, we compared statistically (using Wilcoxon rank-sum test) 50 neuron models with the 45 units were classified as equally responsive, regarding Cliff's delta values between probe responses. In addition, the chosen synaptic weights were constrained to

**Table 1. Parameters used in the neuron model**

Parameters	Values
<i>Neuron model</i>	
$C_m$ , membrane capacitance	100 pF
$E_L$ , leak reversal potential	−55 mV
$g_L$ , leak conductance	5 nS
$V_{th}$ , firing threshold potential	−50 mV
$V_r$ , reset potential	−55 mV
$\sigma$ , sigma noise	2 mV
$\tau_{\sigma}$ , time constant noise	10 ms
<i>Neuronal adaptation</i>	
$\Delta_{th}$ , adaptive threshold increment	0.25 mV
$\tau_{th}$ , adaptive threshold time constant	550 ms
<i>Excitatory synapses</i>	
$E_e$ , excitatory reversal potential	0 mV
$\tau_e$ , excitatory time constant	10 ms
<i>Input</i>	
$w_e$ , synaptic weight	8 nS
$\nu$ , average input rate	2/ms
$\nu_{spont}$ , spontaneous input rate	1/s
$k_{i,com}$ , response input $i$ to com context	0.1
$k_{j,com}$ , response input $j$ to com context	0.0165
$k_{i,ech}$ , response input $i$ to ech context	0
$k_{j,ech}$ , response input $j$ to ech context	1.5
$k_{i,com}$ , response input $i$ to com probe	0.7
$k_{j,com}$ , response input $j$ to com probe	0.1
$k_{i,ech}$ , response input $i$ to ech probe	0
$k_{j,ech}$ , response input $j$ to ech probe	1.5
<i>Synaptic adaptation</i>	
$\Omega_i, s$ , adaptation recovery rate input $i$	1.6/s
$\Omega_j, s$ , adaptation recovery rate input $j$	1.0/s
$\Delta_i, s$ , synaptic decrement input $i$	0.045
$\Delta_j, s$ , synaptic decrement input $j$	0.040

fit approximately the number of spikes evoked by each probe observed empirically.

We assumed that the spiking of the inputs was tuned to the spectral content of natural sounds; high-frequency tuned ( $\geq 45$  kHz) and low-frequency tuned ( $< 45$  kHz). Considering the spectral distribution of the vocalizations used as context (Fig. 2B), high-frequency-tuned input was set to be responsive to echolocation as well as to communication signals, and low-frequency-tuned input to only communication. The input selectivity to natural sounds is given by the parameter  $k_{j,sound}$  and varies between context and probe. We tested several ratios of selectivity to communication context between the inputs ( $k_{low,com}/k_{high,com}$ ), ranging from equally responsive to low-frequency input exclusively responsive. The ratios were tested for several synaptic decrement values as the behavior of the model is highly dependent on this parameter. Ratios close to one were unable to reproduce our data after communication context. We chose a ratio of  $\sim 6.5$  because it is similar to the proportion of low-frequency components relative to high frequencies in our sequence of communication. The absolute values of selectivity were set to obtain a similar number of spikes in response to each context observed empirically. The average rate of the inputs was calibrated to generate spike trains with one or two spikes during the probe stimulation ( $\sim 1.5$  ms duration).

Our model includes short-term synaptic adaptation (i.e., activity-dependent depression) that depends essentially on two parameters, the synaptic decrement  $\Delta_s$  and the adaptation recovery rate  $\Omega_s$ . The interplay between these two parameters was systematically tested by changing them on each synapse. Results in terms of the discriminability index between the two probes after 60 and 416 ms of context offset were analyzed. The chosen values do not show significant differences with experimental data after both contexts and both gaps. In addition to the discriminability index, we calculated the context effect from the same simulations. We used such measurements to constrain the range of fitting parameters that satisfy experimental data.

Finally, we modified our model to reproduce experimental data obtained in units classified as preference to communication. We

**Table 2. Parameters that were modified from the fitted neuron model**

Models varying degree of input selectivity	
<i>None</i>	
$k_{i,comv}$ , response input $i$ to com (context)	0.1
$k_{j,comv}$ , response input $j$ to com (context)	0.1
$k_{i,echv}$ , response input $i$ to ech (context)	0.8
$k_{j,echv}$ , response input $j$ to ech (context)	0.8
$k_{i,comv}$ , response input $i$ to com (probe)	0.4
$k_{j,comv}$ , response input $j$ to com (probe)	0.4
$k_{i,echv}$ , response input $i$ to ech (probe)	0.8
$k_{j,echv}$ , response input $j$ to ech (probe)	0.8
<i>High</i>	
$k_{i,comv}$ , response input $i$ to com (context)	0.4
$k_{j,comv}$ , response input $j$ to com (context)	0
$k_{i,echv}$ , response input $i$ to ech (context)	0
$k_{j,echv}$ , response input $j$ to ech (context)	2
$k_{i,comv}$ , response input $i$ to com (probe)	0.8
$k_{j,comv}$ , response input $j$ to com (probe)	0
$k_{i,echv}$ , response input $i$ to ech (probe)	0
$k_{j,echv}$ , response input $j$ to ech (probe)	1.5
<i>Low</i>	
$k_{i,comv}$ , response input $i$ to com (context)	0.1
$k_{j,comv}$ , response input $j$ to com (context)	0.0165
$k_{i,echv}$ , response input $i$ to ech (context)	0
$k_{j,echv}$ , response input $j$ to ech (context)	1.5
$k_{i,comv}$ , response input $i$ to com (probe)	0.7
$k_{j,comv}$ , response input $j$ to com (probe)	0.1
$k_{i,echv}$ , response input $i$ to ech (probe)	0
$k_{j,echv}$ , response input $j$ to ech (probe)	1.5
<b>Models varying adaptation</b>	
<i>None</i>	
$\Delta_{thr}$ , adaptive threshold increment	0 mV
$\Delta_i$ , $s$ , synaptic decrement input $i$	0.0
$\Delta_j$ , $s$ , synaptic decrement input $j$	0.0
<i>Post</i>	
$\Delta_{thr}$ , adaptive threshold increment	0.25 mV
$\Delta_i$ , $s$ , synaptic decrement input $i$	0.0
$\Delta_j$ , $s$ , synaptic decrement input $j$	0.0
<i>Pre</i>	
$\Delta_{thr}$ , adaptive threshold increment	0 mV
$\Delta_i$ , $s$ , synaptic decrement input $i$	0.045
$\Delta_j$ , $s$ , synaptic decrement input $j$	0.040
<i>Post + Pre</i>	
$\Delta_{thr}$ , adaptive threshold increment	0.25 mV
$\Delta_i$ , $s$ , synaptic decrement input $i$	0.045
$\Delta_j$ , $s$ , synaptic decrement input $j$	0.040
<b>Model of a preference for communication neuron</b>	
$k_{i,comv}$ , response input $i$ to com (context)	0.15
$k_{i,comv}$ , response input $i$ to com (probe)	1.2
$w_{i,er}$ , synaptic weight of $i$	15 nS
$w_{j,er}$ , synaptic weight of $j$	9 nS

Post, postsynaptic; Pre, presynaptic.

increased the input selectivity for communication sounds (Table 2) and ran several simulations changing the maximum synaptic weight  $w_e$  of each synapse. A higher synaptic weight in the low-frequency synapse and lower weight in the high-frequency synapse turned our neuron model into a preference to communication unit with similar firing rates observed in the experiments. Without changing any other parameter, we observed that the model reproduces the behavior of these neurons after context.

## Results

We investigated how auditory cortex neurons of the bat *C. perspicillata* process two types of natural vocalizations, echolocation pulses and distress calls (a type of communication call, hereafter referred to as communication calls). The goal was to assess how

leading context of behaviorally relevant sounds—echolocation and communication vocalizations—affect cortical responses to lagging sounds (called probes throughout the text; Fig. 2). Probe sounds could either match the previous sequence (e.g., echolocation occurring after echolocation; Fig. 2C) or mismatch it (e.g., communication following echolocation). In *C. perspicillata*, the main energy of echolocation and communication calls occurs at different frequencies,  $\sim 66$  kHz (Fig. 2B, second histogram) and  $\sim 23$  kHz (Fig. 2B, fourth histogram), respectively. In addition to frequency composition, the sequences used in this study also differ in temporal structure (Fig. 2B; first and third histograms). Probes were presented at two lags after context offset, 60 and 416 ms. These two delay values were chosen because they allow a clear separation between responses to the last elements of the context sequence and responses to lagging probe sounds. A systematic mapping of forward suppression covering multiple delays between leading and lagging sounds was presented previously, albeit in anesthetized animals and using only pairs of communication calls (Martin et al., 2017). Studying the effects of multiple delays between context and probe sounds is done here by means of modeling (see below, Timescale of synaptic depression predicted by the model).

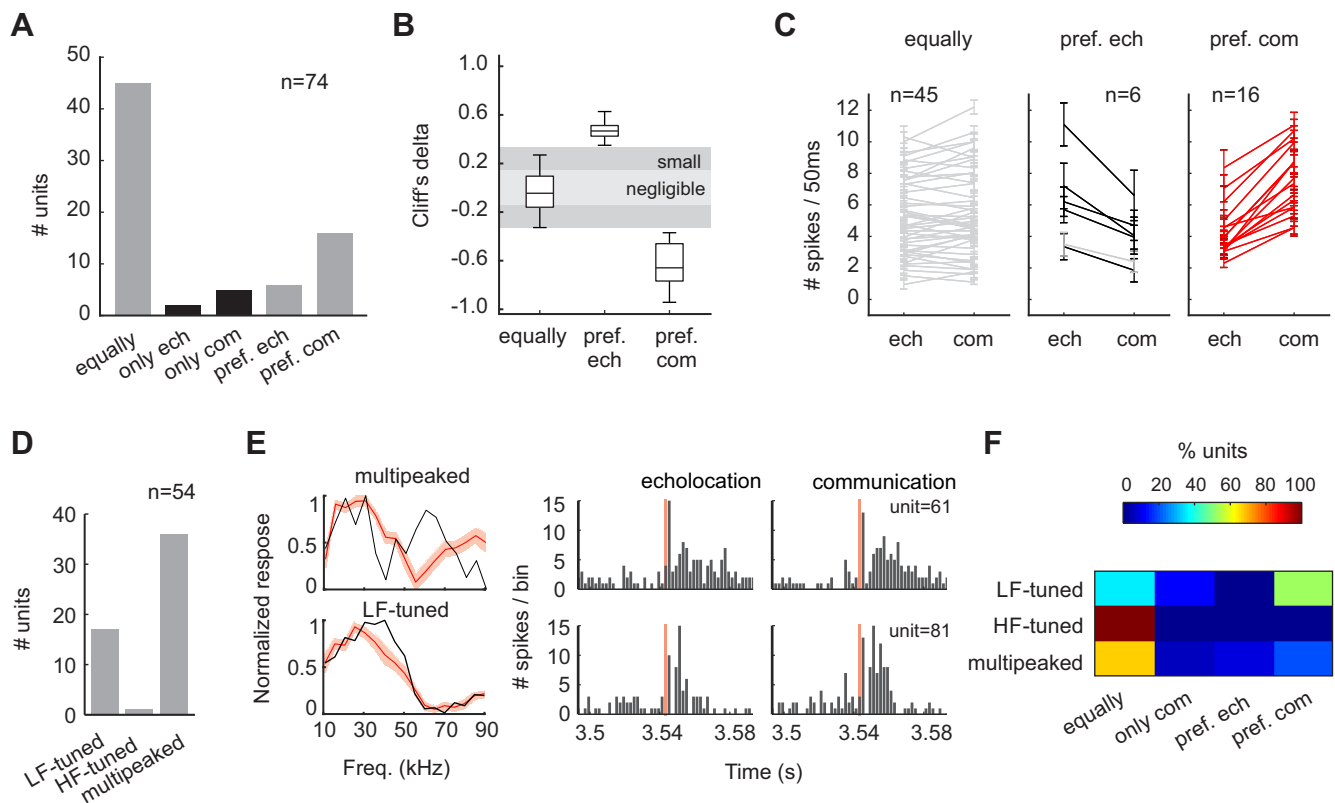
### Cortical responses to echolocation and communication syllables after silence

The extracellular activity of 74 units was recorded in the AC of awake bats. Neurons were classified into five groups regarding their responses to the probes presented in isolation, that is, preceded by 3.5 s of silence (see above, Materials and Methods). The majority of units (91%, 67 of 74) responded to both sounds (Fig. 3A, gray bars), and the remaining 9% (7/74 units) responded exclusively to echolocation or communication (Fig. 3A, black bars). Note that we selectively targeted high-frequency regions of the auditory cortex that respond to both low and high frequencies (Hagemann et al., 2010, 2011) and thus this result was expected.

In each unit, the number of spikes evoked by the two probes, that is, single echolocation pulse and communication syllable, were statistically compared using the Cliff's delta; metric. Units with negligible and small effect size [ $\text{abs}(\text{Cliff's delta}) \leq 0.3$ ] were defined as equally responsive to both sounds (45/74 units). Otherwise, units with a larger effect size were classified as preference for communication (16/74 units) or preference for echolocation (6/74 units) by comparing mean spike counts (Fig. 3B). Figure 3C shows the spike count of all units split into the three groups mentioned.

Next, we tested whether the responsiveness of the units could be explained based on the overlap between the frequency spectrum of call components and the isolevel frequency tuning curves measured at 80 dB SPL. We chose this SPL value because it provides a readout that approximates the maximum broadness of the frequency-receptive fields. The level of natural vocalizations used as probe sounds was also fixed to SPL (root mean square) values of 73 dB SPL, resembling a situation in which a conspecific is vocalizing  $\sim 1.5$  m away from the listener (see above, Materials and Methods).

Isolevel frequency tuning curves were calculated using pure tones covering frequencies from 10 to 90 kHz (5 kHz steps) for a subset of neurons ( $n = 54$  of 74 units). The classification of the units according to the shape of their isolevel tuning curve is shown in Figure 3D (see above, Materials and Methods). Sixty-seven percent of the units (36/54 units) presented multip peaked frequency tuning curves (Fig. 3E, top, example unit). This feature has been associated with nontopographically arranged areas in the AC (high-frequency



**Figure 3.** Neuronal classification of the responses to natural sounds without the acoustic context. **A**, Number of units classified according to responses to echolocation and communication probes. Categories were defined considering the number of spikes evoked during 50 ms from the sound onset [equally:  $abs(Cliff's\ delta) \leq 0.3$ ; pref.:  $abs(Cliff's\ delta) > 0.3$ ]. Gray bars correspond to units responsive to both sounds. Black bars correspond to units that are only responsive to one sound. **B**, Cliff's delta calculated for neuronal classification for units responsive to both sounds. Size effect is indicated by the shaded areas. **C**, Average number of spikes across trials in response to both probes per classification. Gray lines indicate nonsignificant differences between spike counts ( $p > 0.05$ , Wilcoxon rank-sum test). Red and black lines correspond to units that showed significant differences ( $p < 0.05$ , Wilcoxon rank-sum test). Error bars indicate SEs across trials. **D**, Number of units according to isolevel frequency tuning shape. Low- and high-tuned correspond to one-peak tuning curves, peaking at  $< 50$  and  $> 50$  kHz, respectively. **E**, Two examples of units classified as multi-peaked (black top) and low-frequency tuned (black bottom). The red isolevel tuning curves indicate the average across units, and the shaded area indicates SE. At the right, the corresponding PSTH in response to each probe. **F**, Percentage of units classified by response of units to natural sounds per isolevel tuning curve shape. Abs, absolute value; pref., preference.

fields and dorsoposterior area) in *C. perspicillata* (Hagemann et al., 2010; López-Jury et al., 2020). The remaining units studied (18/54) exhibited single-peaked tuning, associated with neurons in tonotopic areas (primary and secondary AC and anterior auditory field). Of those, 17 were low-frequency tuned (best frequency  $< 50$  kHz; Fig. 3E, example unit, bottom) and only one had best frequency  $\geq 50$  kHz. The mean isolevel tuning curves of all multi-peaked and low-frequency-tuned units is shown in red in Figure 3E.

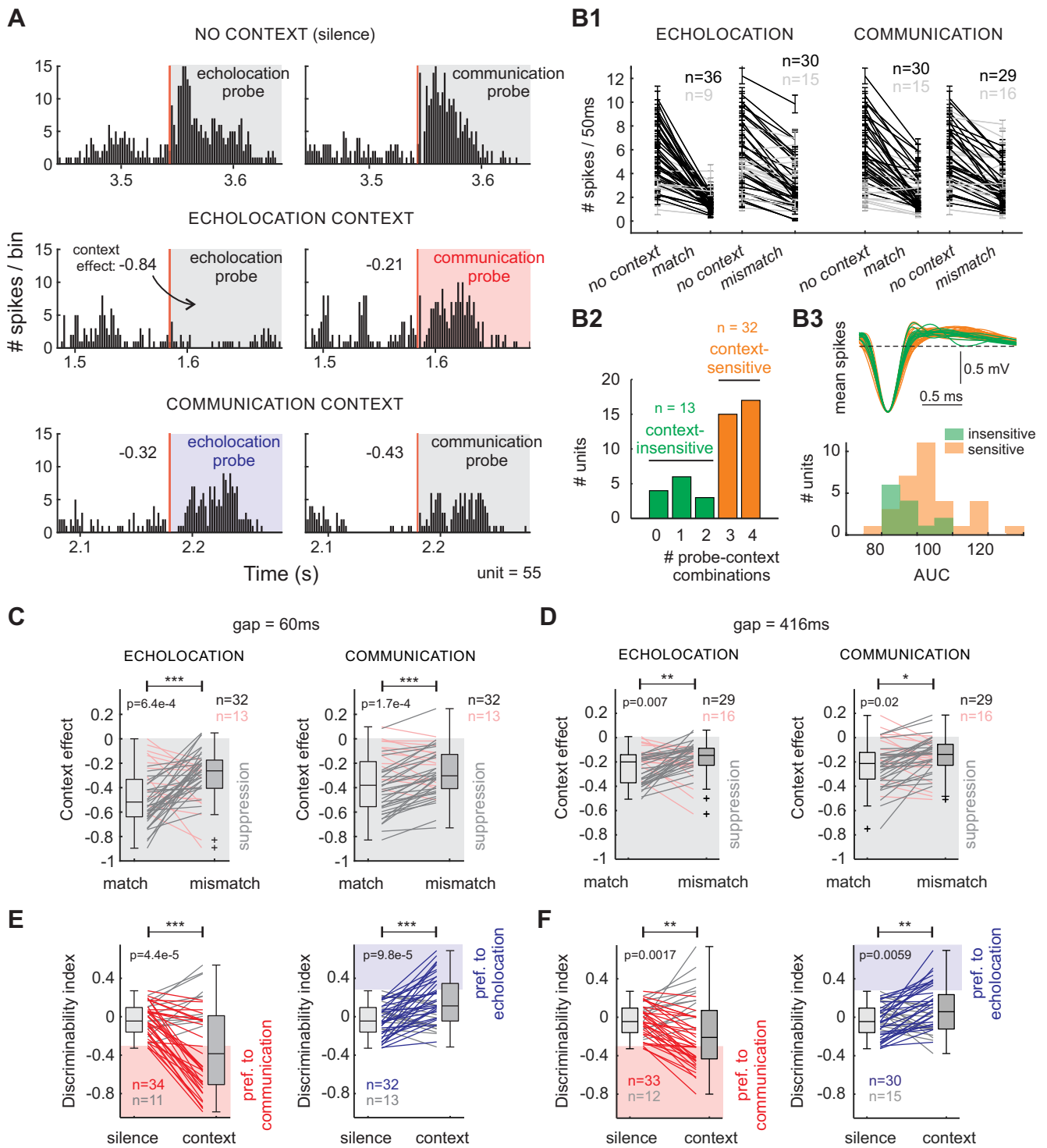
As expected, the majority of neurons with multi-peaked tuning curves (67%) belonged to the group equally responsive (Fig. 3E, top). In addition, 53% of the neurons with a low-frequency isolevel tuning curve presented preferential or exclusive responses for communication sounds (Fig. 3E, bottom). These results indicate that although simple, isolevel tuning curves partially predict neuronal responsiveness to natural sounds presented without context, that is, preceded by silence. The percentages of responsiveness to natural sounds for each tuning curve group are shown in Figure 3F.

### Natural acoustic context suppresses cortical responses to lagging probe sounds

The effect of acoustic context, that is, echolocation or communication sound sequences preceding the presentation of probe sounds, was first quantified for the 45 units that responded equally well to both probes when presented after silence (Fig. 3C, left). Figure 4A shows the PSTHs of an example equally

responsive unit in response to echolocation and communication probe sounds preceded either by silence (top), by an echolocation context (middle), or by a communication context (bottom) 60 ms after the offset of the context. In this unit, the response to both probes was reduced after the context sequences. In fact, the response to the echolocation probe was completely abolished after the echolocation context (Fig. 4A, middle, left). Overall, the presence of a leading context decreased responsiveness in most cortical neurons regardless of the type of probe sound that followed the context sequences ( $N$  values between 36 and 29 of 45 units,  $p < 0.05$ , Wilcoxon rank-sum test). Note that all black lines (significant changes) point downward in Figure 4B1. However, a subset of neurons did not present significant variations in the response to the probe after context ( $p > 0.05$ , Wilcoxon rank-sum test), the respective spike counts are plotted with gray lines in Figure 4B1 ( $N$  values between 9 and 16).

Thirty-two units showed significant reduction of responses in at least three context-probe combinations and were considered context-sensitive neurons. The rest of the neurons ( $n = 13$ ) were classified as context insensitive (Fig. 4B2). To study if these two classes of neurons differed from each other on their intrinsic properties, we compared best frequency, spontaneous firing rate, and spike waveform. Only the shape of the spikes was significantly different between the two groups ( $p = 0.01$ , Wilcoxon rank-sum test, Fig. 4B3). Context-sensitive neurons showed higher values of area under the spike waveform than context-



**Figure 4.** Stimulus-specific suppression on neuronal responses by leading acoustic context. **A**, Responses to echolocation (ech) and communication (com) probes of a unit classified as equally responsive after 3.5 s of silence (top), after a sequence of echolocation (middle), and after a sequence of communication (bottom). **B1**, Average number of spikes across trials in response to a probe after silence (no context) and after context (match or mismatch). Match corresponds to ech-ech or com-com (context-probe) and mismatch to ech-com or com-ech. The respective context is indicated above. Black lines correspond to significant decrement ( $p < 0.05$ , Wilcoxon signed-rank test) and gray corresponds to not significant. Error bar indicates SE across trials. **B2**, Number of units that presented significant suppression in 0–4 context-probe combinations. **B3**, Spike waveforms (average per isolated single unit) of context-insensitive neurons (green) and context-sensitive neurons (orange). Bottom, Distributions of the corresponding areas under the waveform. **C**, Context effect per probe of equally responsive units 60 ms after the offset of the context. Lines join values from the same neuron. Gray lines correspond to units with a higher suppression on matching sounds. **D**, Same as in **C** but 416 ms after the offset of the context. **E**, Discriminability index of equally responsive units comparing probe responses after silence and 60 ms after the offset of the context. **F**, Same as in **E** but 416 ms after the offset of the context. Significance level is indicated above **C–F**;  $p$  values were obtained using the paired Wilcoxon signed-rank test.



insensitive ones, suggesting that context-sensitive neurons could be classified as putative broad-spiking pyramidal cells, whereas context-insensitive neurons correspond to putative narrow-spiking interneurons (Tsunada et al., 2012).

### Context increases cortical discrimination of natural sounds

Although the example in Figure 4A shows reduction of probe-evoked responses by leading context, the magnitude of the suppression was different when comparing between probes. The matching probes in both context categories were more suppressed than the mismatching probes relative to the responses after silence (see context effect suppression values at the top of each PSTH). The context-dependent effect illustrated in Figure 4A was not unique to this example neuron. In fact, stronger reduction of responses to matching probes were common among the neurons studied (Fig. 4C), regardless of the type of context that preceded the probes ( $p = 6.4 \times 10^{-4}$  for echolocation and  $p = 1.7 \times 10^{-4}$  for communication, Wilcoxon signed-rank test). Relative to the probes preceded by silence, there was a median of 48% of suppression for matching sounds versus 30% for mismatching sounds after echolocation context, and 37 and 28% for communication context, respectively. As expected, the amount of suppression driven by the context sequence decreased when the interval between context offset and probe was set to 416 ms (compare Fig. 4D, 416 ms delay, with data obtained with 60 ms delay in Fig. 4C). A decrease in context effect was apparent in all context-probe combinations [ $p = 6.8 \times 10^{-8}$  for echolocation (ech)-ech,  $p = 2.0 \times 10^{-5}$  for ech-communication (com),  $p = 0.0013$  for com-com, and  $p = 3.9 \times 10^{-5}$  for com-ech; Wilcoxon signed-rank test comparing data obtained with 60 ms delay vs data obtained with 416 ms delay]. Interestingly, the context suppression was still stimulus specific in the 416 ms gap treatment ( $p = 0.007$  for echolocation and  $p = 0.02$  for communication, Wilcoxon signed-rank test) indicating long-lasting effects of natural context on upcoming sounds.

So far, we have shown that nonselective neurons for natural sounds of different behavioral categories, echolocation and communication, exhibit suppression by leading acoustic context that is stimulus specific. We reasoned that such a context-specific suppression could actually render cortical neurons more selective to natural sounds than what could be predicted by presenting single sounds preceded by silence. To quantify whether this was the case, we compared the Cliff's delta values (used here as a discriminability index) obtained when comparing spike counts between the isolated probes versus the value obtained when there was leading context (Fig. 4E). After both contexts, the discriminability between the probes significantly increased (values closer to  $-1$  or  $1$ ) compared with the discriminability exhibited in silence ( $p = 4.4 \times 10^{-5}$  after echolocation and  $p = 9.8 \times 10^{-5}$  after communication, Wilcoxon signed-rank test). The results showed that under the echolocation context, Cliff's delta values became more negative from a median of  $-0.045$  (in silence) to  $-0.38$ , which is indicative of a higher number of spikes in response to the communication probe. On the other hand, the presence of the communication context shifted Cliff's delta values closer to  $1$  (from a median of  $-0.045$  to  $0.11$ ), which corresponds to a higher number of spikes in response to the echolocation probe. Such an increment in the discriminability index after context was observed as well in the 416 ms treatment (Fig. 4F;  $p = 0.0017$  after echolocation and  $p = 0.0059$  after communication, Wilcoxon signed-rank test), suggesting again that context effects are long lasting. Overall, these results show that responses to communication or echolocation syllables presented alone are more difficult to

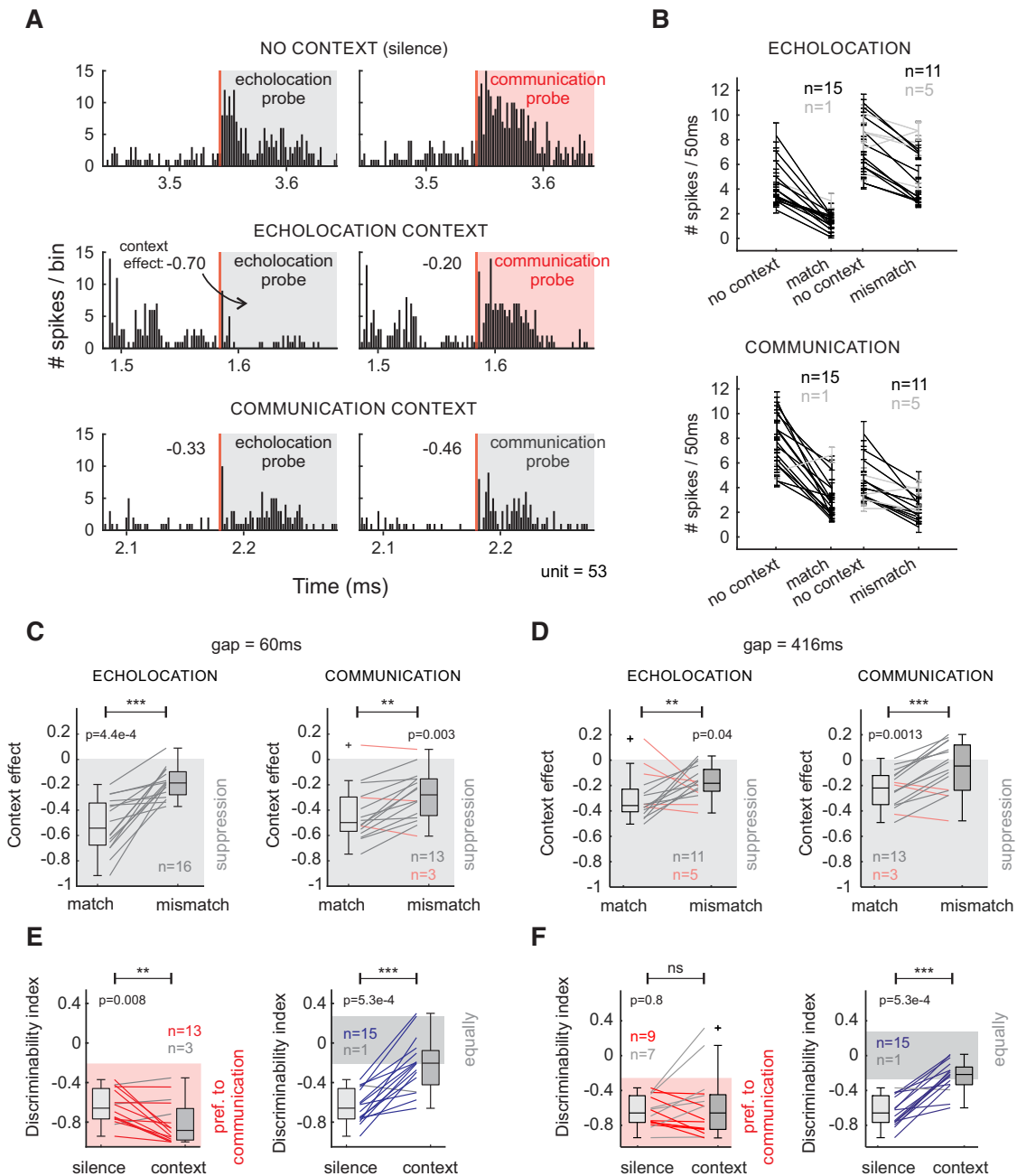
discern than when the syllables are presented after a behaviorally relevant acoustic sequence. In other words, acoustic contexts decreased the ambiguity of the response of nonselective cortical neurons by increasing the sensitivity to temporal transitions between these two sound categories.

### Effects of acoustic context on neurons with preference for communication sounds

We also examined the effects of context sequences on the response of neurons classified as preference for communication sounds. These neurons formed the second largest neuronal group observed when presenting the sounds without context (22% of the neurons studied; Fig. 3C, right). A typical example neuron that preferred communication probes can be observed in Figure 5A (top). Note that although this neuron has a stronger response to the communication probe, it still responded well to the echolocation probe presented without any acoustic context. In this example, as in the previous neuronal group (equally responsive neurons; Fig. 4), the leading context suppressed the response to both probes, and the suppression was stronger on matching sounds (Fig. 5A, second and third row). The responses to the probes after context were analyzed for all the preference for communication units ( $n = 16$ ) and significant suppression after context was found in the majority of the neurons (Fig. 5B). Moreover, stimulus-specific suppression was found also at the population level (Fig. 5C), following the same trend exhibited by the example unit in Figure 5, as well as in the equally responsive neurons group (Fig. 4C). Stimulus-specific suppression was also observed in the 416 ms treatment ( $p = 0.04$  after echolocation and  $p = 0.0013$  after communication, Wilcoxon signed-rank test; Fig. 4D). Comparing the amount of suppression (context effect) in the 60 ms and 416 ms conditions resulted in significant differences in three of the four comparisons conducted ( $p = 0.002$  for ech-ech,  $p = 0.6$  for ech-com,  $p = 7.8 \times 10^{-4}$  for com-com, and  $p = 4.4 \times 10^{-4}$  com-ech, Wilcoxon signed-rank test), indicating again that suppression decreases with lengthening the delay between context offset and probe occurrence.

The clearest difference between selective and nonselective neurons occurred when the probe sounds were preceded by the communication context. To illustrate, in the example in Figure 5A (bottom), when preceded by a communication context, the responses to communication and echolocation probes became similar, although they were notably different when the sounds were presented without context (Fig. 5A, top). This result stems from the unbalanced suppression on probe-evoked responses ( $-0.46$  on matching vs  $-0.33$  on mismatching) together with the intrinsic neuronal preference for communication sounds, which brought spike outputs in response to each probe to the same level when they were presented after the communication sequence.

Equally strong responses to the probes after the communication context were also present at the population level. To quantify the difference between probe-evoked responses, we calculated the discriminability index after silence and after context (Fig. 5E). In agreement with the example neuron shown in Figure 5A, at the population level when the context was communication, Cliff's delta values went from very negative values (preference for communication when preceded by silence) to higher values ( $p = 5.3 \times 10^{-4}$ , Wilcoxon signed-rank test), closer to zero, that is, similar probe-evoked responses (Fig. 5E, right). This effect on probe discriminability was also observed when the delay was longer (416 ms; Fig. 5F, right). In addition, when the context was the



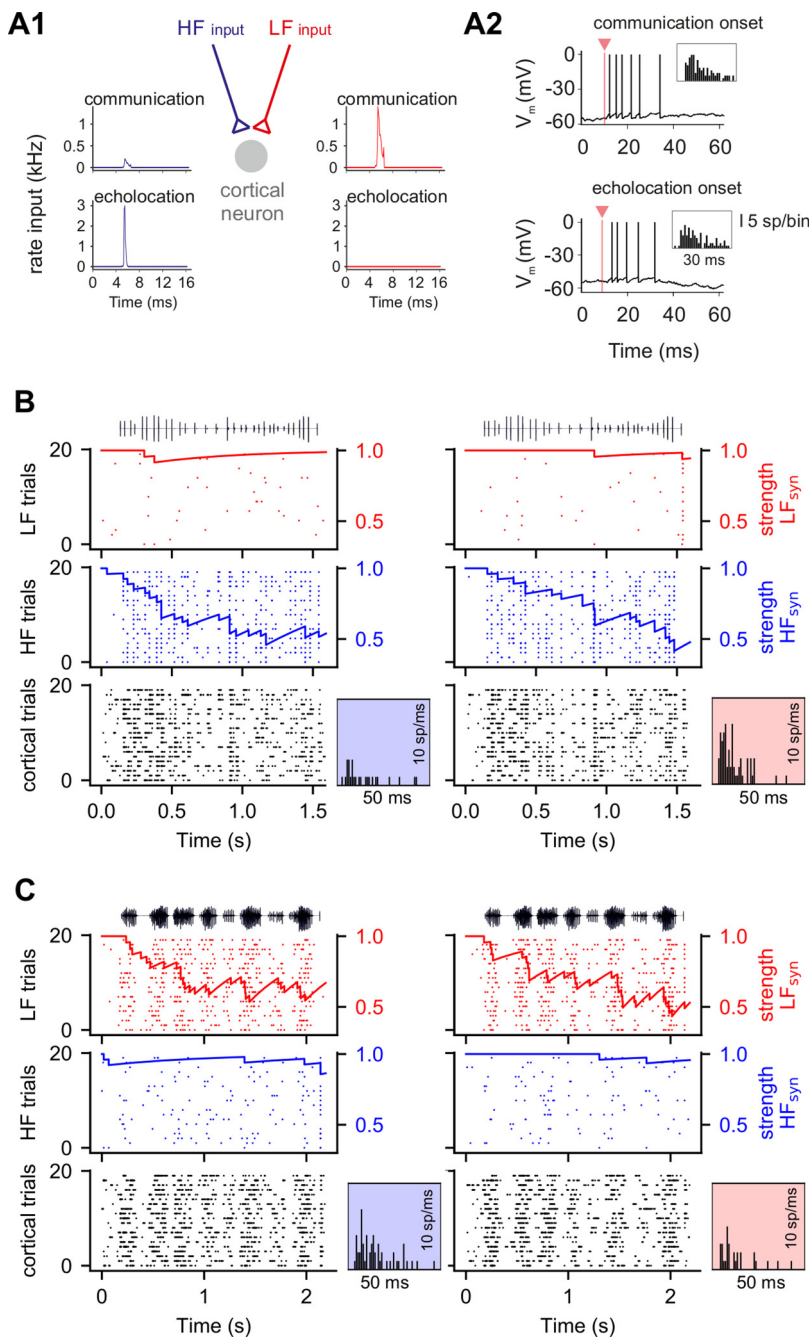
**Figure 5.** Context modulation in selective neurons decreased neuronal discriminability. **A**, Responses to echolocation and communication probes of a unit classified as preference for communication after 3.5 s of silence (top), after a sequence of echolocation (middle), and after a sequence of communication (bottom). **B**, Average number of spikes across trials in response to a probe after silence (no context) and after context (match or mismatch). Black lines correspond to significant decrement ( $p < 0.05$ , Wilcoxon signed-rank test), and gray indicates not significant. Error bar indicates SE across trials. **C**, Context effect per probe of preference for communication units 60 ms after the offset of the context. Lines join values from the same neuron. **D**, Same as in **C** but 416 ms after the offset of the context. **E**, Discriminability index of preference for communication units comparing probe responses after silence and 60 ms after the offset of the context. **F**, Same as in **E** but 416 ms after the offset of the context. Significance level is indicated above **C–F**;  $p$  values were obtained using the paired Wilcoxon signed-rank test.

echolocation sequence, Cliff's  $\delta$  values became even more negative in the 60 ms gap treatment ( $p = 0.008$ , Wilcoxon signed-rank test), indicating better discrimination and higher responsivity to mismatching (communication) sounds (Fig. 5E, left). However, this effect was not visible after a 416 ms delay (Fig. 5F, left). To summarize, our results indicate that good neuronal echolocation versus communication classifiers, according to spike counts measured with single sounds presented in silence, became in fact poor category discriminators when tested in a context of communication sounds.

**A computational model with frequency-specific adaptation explains context-dependent suppression observed *in vivo***

Although single echolocation pulse and communication syllable evoked similar responses in most cortical neurons, after the acoustic context the responses shifted toward the syllable that corresponded to the mismatching sound. To identify which mechanisms can account for these context-dependent effects, we modeled an equally responsive cortical neuron using a leaky integrate-and-fire neuron model (Abbott, 1999).

First, we modeled a single cortical neuron, set to be equally responsive to both sounds categories when presented without



**Figure 6.** Computational modeling of stimulus-specific suppression by leading context on cortical neurons. **A1**, Diagram of the components of the model. The rate (in kHz) of the spiking of the respective input in response to echolocation and communication probes. **A2**, Membrane potential of the neuron model in response to both probes. Insets, Simulated PSTH across 20 trials for the corresponding sounds. **B**, Example of 20 simulations of neuronal responses to an echolocation sequence followed by an echolocation probe (left) and by a communication probe (right). Raster plots indicate the spiking of the inputs (top and middle) and of the cortical neuron model (bottom). On top of the rasters, the time evolution of the synaptic strength associated to the spiking of the first trial of each input. Insets, The PSTH obtained from the neuron model during 50 ms after the onset of the probe for each stimulation. The parameters used in the model are indicated in Table 1. **C**, Same as in **B** but in response to a communication sequence followed by an echolocation probe (left) and a communication probe (right).

context and with two synaptic inputs whose firing are highly selective for either communication or echolocation signals (Fig. 6A1). Each synapse receives spike trains whose firing rate is proportional to the degree of selectivity that each input was assumed to have for each sound category. The temporal pattern of the spike trains was assumed to follow the envelope of the respective natural sound used in each simulation. We also assumed that the degree of

the input selectivity is correlated with the spectral components of the natural sounds used in our experiments. Thus, we distinguished two input classes, high-frequency tuned ( $>45$  kHz) and low-frequency tuned ( $<45$  kHz). The rate of the inputs in response to each probe are shown in Figure 6A1. Note that the firing rate of both inputs increases in response to the communication syllable, albeit to a higher degree in the low-frequency-tuned input. The latter is a consequence of the spectral broadness of distress calls that carry energy in both low- and high-frequency portions of the spectrum, as opposed to echolocation signals that in this bat species are limited to frequencies above 50 kHz (Fig. 2, natural sounds).

In the neuron model, each presynaptic spike induces an increment of excitatory conductance. The amplitude of this conductance change was adjusted so that the neuron responded equally well to both probe sounds. Similar spiking responses to both probe sounds in one trial simulation are illustrated in Figure 6A2. We ran 20 trial simulations per probe stimulation, and the resultant PSTHs are shown in the respective insets in Figure 6A2.

Simulated responses to combinations of context followed by probe are shown in Figure 6, *B* and *C*. The spiking of the low-frequency-tuned input (red), high-frequency-tuned input (blue), and that of the cortical neuron (black) are shown as raster plots (top, middle, and bottom, respectively). Simulations of echolocation context followed by matching and mismatching probes are shown in Figure 6*B*; the spiking of the low-frequency input (top) across 20 trials corresponds only to spontaneous activity of the input, and the spiking of the high-frequency channel (middle) tightly follows the envelope of the echolocation sequence. The spiking of the cortical neuron (bottom) also exhibited a response to the echolocation sequence because it integrates both low- and high-frequency inputs. In contrast, simulations of communication context evoked spiking responses mostly in the low-frequency input (Fig. 6*C*) and in the cortical neuron.

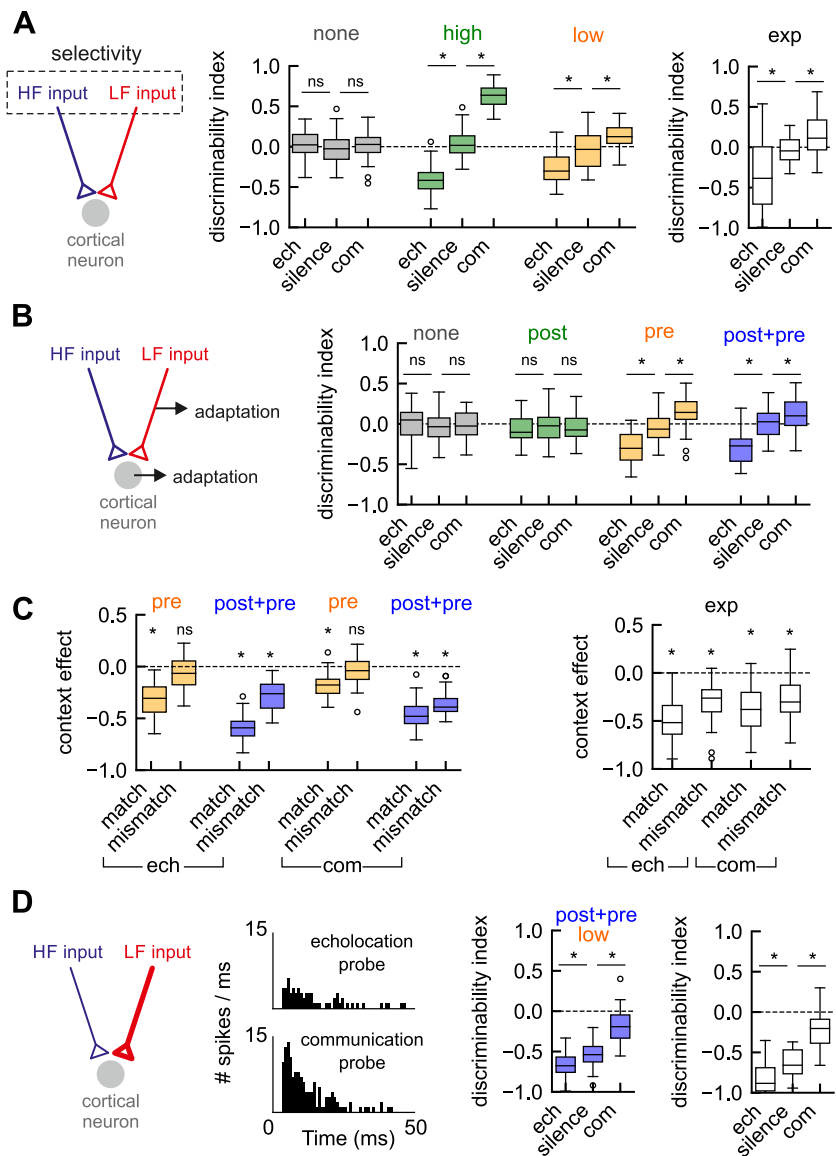
To replicate *in vivo* results *in silico*, we implemented an activity-dependent adaptation to the model. A likely mechanism of adaptation is synaptic depression, which has been proposed to underlie context-dependent processing (Ulanovsky et al., 2004; Wehr and Zador, 2005). The temporal course of the synaptic strength associated with the spiking of the first simulation trial of each input is indicated by solid lines at the top of the corresponding raster plots (Fig. 6*B,C*). Because synaptic

depression is activity dependent, the more activity a synapse receives from an input, the more adapted it is (lower values of synaptic strength). Therefore, during the echolocation context, the high-frequency-tuned input synapse is more adapted than the low-frequency-tuned input synapse (Fig. 6B), and it is the opposite during the communication context (Fig. 6C). The synaptic strength is proportional to the amplitude of the excitatory postsynaptic potential that each presynaptic spike produces in the cortical neuron; low values indicate less probability of spiking, high values indicate the opposite. The recovery time of the adaptation implemented allowed to affect output responses to forthcoming probe sounds 60 ms after the offset of the context sequence (and 416 ms, see below). Consequently, the spiking associated to the matching probes generated less spikes in the cortical neuron than the spiking in response to the mismatching probes (Fig. 6B,C, PSTHs, right of each raster plot).

**Requirements of the model to reproduce data**

Our model suggests that stimulus-specific suppression arises from (1) segregated inputs selective for each sound category and (2) synaptic adaptation. To illustrate that these conditions are necessary to reproduce our empirical results, we ran different simulations modifying synaptic properties and quantified the same indexes obtained *in vivo* with the metrics context effect and discriminability index.

First, we varied the degree of input selectivity and compared the discriminability index obtained after silence versus context (Fig. 7A). The simulations showed that when the inputs have no preference for any of the signals (both inputs respond to both), the discriminability index did not change significantly after context ( $p > 0.05$ , Wilcoxon rank-sum test; Fig. 7A, none, gray box plots). These results proved that input selectivity was necessary to obtain stimulus-specific context effects. On the other hand, if the inputs were selective exclusively to one sound category (either echolocation or communication), the discriminability index was significant and drastically different after context compared with silence ( $p < 0.001$ , Wilcoxon rank-sum test; Fig. 7A, high, green box plots), following the trend found empirically. However, only simulations with the echolocation context showed no statistical differences with our data ( $p = 0.3$  echolocation;  $p = 6.0e-11$  communication, Wilcoxon rank-sum test). A more biologically plausible model would assume inputs responding to the spectral features of sounds and not to sound categories, probably a middle point between no selectivity and high selectivity. Indeed, the latter model rendered significant changes after



**Figure 7.** Requirements of the model to reproduce empirical results. **A**, Discriminability index after silence, after echolocation, and after communication context using models with different degrees of selectivity in the inputs. None, Model in which inputs were not selective for any type of signals, echolocation, or communication, although responded to both of them; High, inputs exhibited exclusive selectivity for one signal; Low, input selectivity correlates with the spectral components of the natural sounds. Right, Box plots show the discriminability index obtained from our empirical data. **B**, Discriminability indexes (as in **A**) using models with different forms of adaptation. None, Model without any type of adaptation; Post, model with neuronal adaptation; Pre, model with synaptic adaptation; Post + Pre, Model with neuronal and synaptic adaptation. **C**, Context effect per probe obtained from Pre and Post + Pre models. Right, Box plots show experimental data. **D**, Simulated PSTHs in response to both probes and discriminability index obtained from a model in which the synaptic weights were unbalanced, as illustrated in the diagram). The neuron model assumes low and Post + Pre conditions. Statistical comparison between discriminability indexes were obtained using nonpaired Wilcoxon rank-sum test for the simulations, and the paired Wilcoxon signed-rank for real data; \**p* values below 0.05. The medians of the context effect that differ from zero are indicated by an asterisk (one-sample Wilcoxon signed-rank test). The parameters used on each model are indicated in Table 2. Post, postsynaptic; Pre, presynaptic.

context ( $p < 0.001$ , Wilcoxon rank-sum test; Fig. 7A, low, orange box plots) but a less dramatic effect than in the previous model with high selectivity. That is, when cortical inputs showed an intermediate level of selectivity, simulations with both contexts showed no differences with the empirical data ( $p = 0.1$  echolocation;  $p = 0.7$  communication, Wilcoxon rank-sum test). Note that in this case, the input tuned to high frequencies would respond at least to some extent to the high-frequency components found in distress signals.

Second, we tested how different forms of adaptation influence the discriminability index (Fig. 7B) and the context effect index (Fig. 7C) in the simulations. As expected, a model that lacks any type of adaptation showed that the discriminability index was zero in all conditions tested (Fig. 7B, none, gray box plots). Similarly, a model that includes neuronal adaptation, dependent on postsynaptic (post) spiking activity, did not show significant differences between silence and context (Fig. 7B, post, green box plots). This indicates that adaptation at the cortical neuron level allows context suppression but not stimulus specificity. Alternatively, if the adaptation depends on presynaptic (pre) activity, the context effect becomes stimulus specific, and the discriminability index increased after context (Fig. 7B, pre, orange box plots). The same effect was observed when the model includes adaptation both at the level of the cortical neuron and at the presynaptic level (Fig. 7B, post + pre, blue box plots).

It is worth mentioning that, although in both models (post and post + pre) the discriminability index increased after context, the model that included both types of adaptations exhibited significant suppression of the mismatching probes ( $p < 0.05$ , Wilcoxon one-sample signed-rank test; Fig. 7C, left), which agrees with our experimental data (Fig. 7C, right). In other words, synaptic adaptation is necessary to create stimulus-specific suppression but not enough to explain a reduction of the responses to mismatching probes.

To test whether our computational model could also explain the decrease of the discriminability index observed after the communication context in preference for communication neurons, we modified the maximum synaptic weight ( $w_e$ ) of both synapses to obtain a neuron model that possesses selectivity for communication sounds (Fig. 7D). After increasing the  $w_e$  of the low-frequency-tuned synapse and decreasing the  $w_e$  of the high-frequency-tuned synapse, our neuron model presented stronger responses to communication stimulation than to echolocation after silence (Fig. 7D, PSTHs). Without changing any other parameters in the model, we calculated the discriminability index between probes after silence and after context. The results are shown in the Figure 7D (left, box plots). Comparable to our data (Fig. 7D, right, box plots), the neuron model increased its discriminability for the probes after the echolocation context (more negative Cliff's delta values in comparison with silence) but decreased the discriminability after the communication context (Cliff's delta values close to zero). The results suggest that although these two classes of neurons (equally responsive and preference for communication) possess different degrees of selectivity for single isolated natural sounds, the effects of a leading context on their responses to forthcoming sounds can be explained by a common canonical mechanism.

### Time course of context-driven suppression and neuronal discriminability changes

Our empirical data indicate that context-dependent effects were also present on the probe responses occurring 416 ms after context offset (Figs. 4D,F, 5D,F). To reproduce this result, we fitted the parameters that describe the time course of the adaptation implemented in the model to replicate discriminability and context effect indexes obtained after the time intervals tested (i.e., 60 and 416 ms).

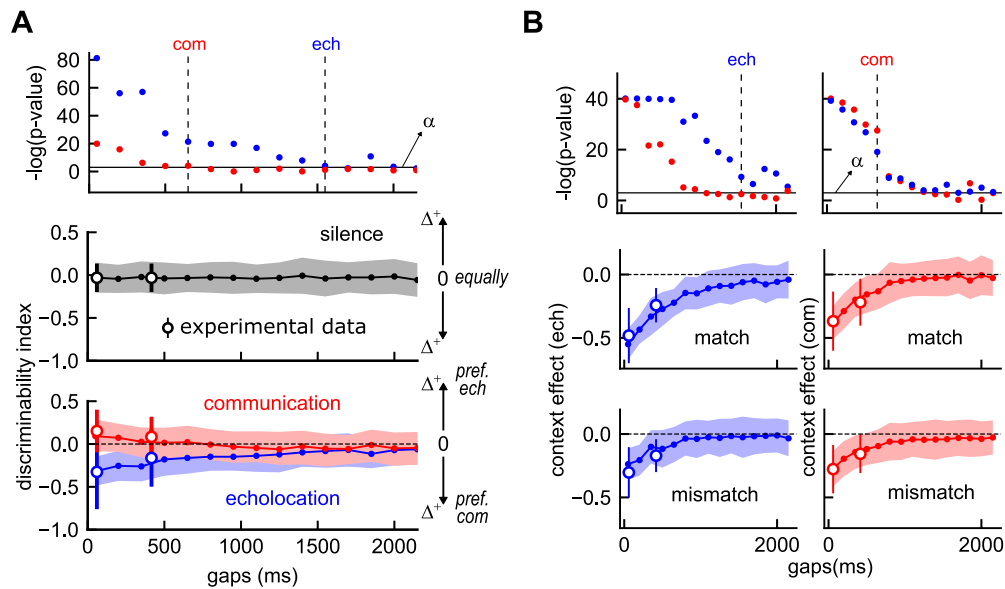
We reasoned that the fitted model could be used to systematically test the effect of multiple gap values. Several simulations were run changing the temporal separation (gap) between context offset and probe occurrence. Figure 8A shows the temporal course of discriminability index after silence (middle) and after the two types of context tested (echolocation in blue and

communication in red) obtained from the simulations. The average across 100 simulated units tightly matched the empirical average (open circles), and the shaded area indicates the SD across the simulations, which also replicated relatively well the variability observed *in vivo* (vertical solid bars indicate SD). The discriminability index after silence remains constant (top graph) at values  $\sim 0$ , and it was compared statistically with values obtained after each context (bottom graph). The respective  $p$  values are plotted at the top. The model predicted that the increment in discriminability, which can be represented by values closer to 1 or  $-1$ , disappears  $\sim 650$  ms after communication and  $\sim 1550$  ms after echolocation context (last gaps at which  $p < 0.05$  Wilcoxon rank-sum test between respective context and silence). In addition to the discriminability index, we examined the suppression strength as a function of gap length. Figure 8B shows the time course of the recovery of the context effect obtained from the same simulations. The model showed that at gaps when there was no effect in terms of discriminability index (1550 ms for echolocation and 650 ms for communication context), suppression was significant (Wilcoxon one-sample rank-sum test; Fig. 8A, top). However, the strength of the suppression was not enough to create significant differences in the spike count between probe responses.

### Time scale of synaptic depression predicted by the model

To visualize the time course of adaptation during and after context resulting from the fitting procedure of model parameters (see above, Materials and Methods), we simulated the presentation of the context sequences followed by 5 s of silence and plotted the average spiking threshold (postsynaptic adaptation) and the average synaptic strength (presynaptic adaptation) across 20 trials (Fig. 9A). The spiking threshold increased during both contexts and recovered following the same time constant ( $\tau_{th} = 550$  ms; Fig. 9A, top). The synaptic strengths of the stimulated input channels were found to decrease during context and recovered following different time constants ( $\tau_{HF} = 1$  s and  $\tau_{LF} = 0.6$  s; Fig. 9A, bottom). More specifically, the synaptic depression associated to the dominant input activity recovered more slowly after the echolocation than communication context (Fig. 9A, compare blue trace, left, with red trace, right), which correlates with a longer lasting context effect after echolocation compared with communication (Fig. 9A). In other words, our model predicts that high-frequency-tuned inputs adapt and recover slower than low-frequency inputs.

To test this prediction in our data, we compared the low-frequency components of the neural recording in response to the probes presented after silence (LFPs). It has been speculated that LFPs could reflect differences in synaptic input properties (Haider et al., 2016; Arroyo et al., 2018) and thus were used here as a readout of network synaptic activity to test the predictions of our model. The average evoked potentials across equally responsive units that presented a significant context effect ( $n = 32$ ; Fig. 4B2) are shown in Figure 9B. Although there were no significant differences in terms of the energy of the evoked potential during 150 ms after the onset of the probe ( $p = 0.06$ , Wilcoxon signed-rank test), we did find differences in the temporal course of the LFPs (Fig. 9C,D). LFPs in response to the echolocation probes recovered slower than in response to communication. To quantify this, we calculated the time at which the LFP first crossed zero and the time at which the positive peak occurs. Both measures gave significant differences between the probes, namely, higher values for echolocation ( $p = 0.005$  and  $p = 0.04$ , respectively,



**Figure 8.** Model predicts differences in the time course of the context effects. **A**, Top, The  $p$  values obtained from Wilcoxon rank-sum tests comparing the discriminability indices (middle and bottom) between the conditions of silence and echolocation (blue circles) and silence versus communication (red circles) across time intervals between context and probe (gaps). The solid line indicates the significance level ( $\alpha = 0.05$ ). Dashed lines correspond to the longest gap that exhibited significant  $p$  value for each condition (echolocation and communication contexts). Middle and bottom, Filled circles indicate the average discriminability index across 100 simulated neurons (shaded area indicates SE) at different gaps. Black, Probes preceded by silence; red, preceded by communication; and in blue, preceded by echolocation context. Note that discriminability increases opposite to zero; negative values correlate with stronger responses to communication (pref. com) and positive values with stronger responses to echolocation probes (pref. ech). Open circles correspond to the empirical values obtained 60 and 416 ms after the end of the context. **B**, Top, The  $p$  values obtained from one-sample Wilcoxon rank-sum tests comparing the context effect indices (below) against a null distribution. Blue circles correspond to  $p$  values calculated with echolocation probe and red circles, with communication probe. Solid and dashed lines indicate the same as in **A**. Middle, bottom, Average context effect obtained from same simulations as in **A**. Blue and red correspond to echolocation and communication context effects, respectively.

Wilcoxon signed-rank test). To test whether such a difference could be because of differences in the spiking, we calculated the duration and the latency of the spiking response, and we found no significant differences ( $p = 0.8$  for latency and  $p = 0.7$  for duration, Wilcoxon signed-rank test; Fig. 9E,F), suggesting that the observed differences in the temporal pattern of the LFPs are because of different properties of the synaptic inputs.

## Discussion

It is known that in the bat AC there are multifunctional neurons specialized to process sounds linked to echolocation and communication. How these neurons respond to transitions between these two sounds remains unknown. This is especially relevant for communication in complex auditory environments where calls occur within the context of other natural sounds. This study demonstrates that the presence of leading acoustic context increases neuronal discriminability in neurons that are nonselective when tested in a silent condition (i.e., no context) and has the opposite effect on selective neurons. A neuron model explains our results by input selectivity to sounds based on their frequency composition, as well as both presynaptic and postsynaptic adaptation. Furthermore, our model predicts the temporal course of the context dependence and proposes that the time constant of adaptation is shortest for low-frequency sounds, that is, communication signals.

### Cortical neurons process both echolocation and communication signals

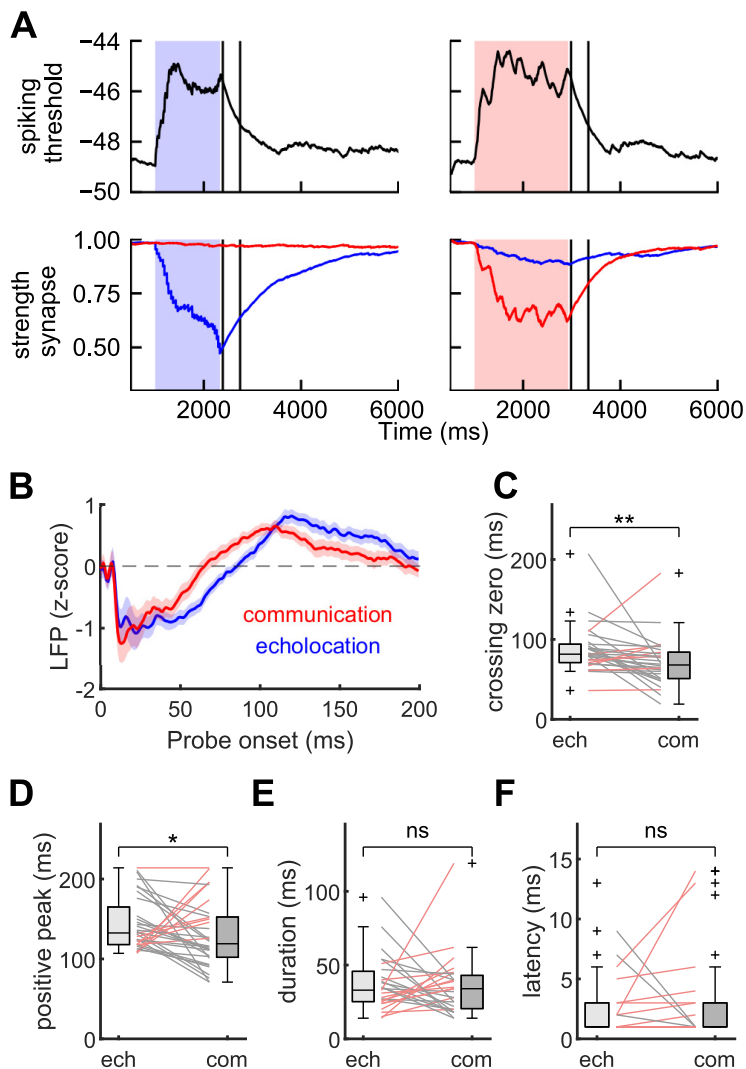
Previous studies have shown that neurons in the AC of mustached bats that are highly specialized for processing echolocation pulses also present specializations for communication calls

processing (Ohlemiller et al., 1996; Esser et al., 1997; Kanwal, 1999, 2012). Here, we provide strong evidence that this is also the case for a frugivorous bat, the *C. perspicillata*. The majority of the neurons tested in this study (91%) presented responses to conspecific echolocation pulses and to a particular type of communication call, distress calls.

In the mustached bat, it has been well studied how these multifunctional neurons use combination sensitivity as a basic mechanism for call selectivity (Kanwal, 1999, 2006). Comparable physiological responses have been described in *C. perspicillata* (Kossl et al., 2014). Cortical delay tuning in the form of FM-FM combination sensitivity neurons is robust in dorsal HF areas of the AC (Hagemann et al., 2010, 2011), where we conducted our neural recordings. Therefore, the responses described here could be specific for these neurons, and their underlying neural mechanisms could be similar to those well studied in the mustached bat.

A large proportion of the neurons studied here exhibited multiplexed frequency tuning curves, which is consistent with the fact that we recorded mostly from HF fields of the brain of *C. perspicillata* (Hagemann et al., 2010; 2011). In agreement with several studies, in a subset of neurons, probe-evoked responses were unpredicted based on their response profile to pure tones (Suga et al., 1978; O'Neill and Suga, 1979; Machens et al., 2004; Sadagopan and Wang, 2009; Laudanski et al., 2012; Feng and Wang, 2017).

In our data, the majority of the selective multifunctional neurons fired more to communication than to echolocation signals. Cortical processing of social calls is known to be lateralized to the left side (Kanwal, 2012; Washington and Kanwal, 2012). All recordings included in this article were, in fact, conducted in the left hemisphere. Accordingly, one could speculate that recordings in the right hemisphere could show a different picture with a higher proportion of



**Figure 9.** Model predicts shortest time scale in synaptic depression associated with communication signals. **A**, Top, Time course of neuronal adaptation (measure as mean spiking threshold) along 6 s of simulation during and after the echolocation context (blue shaded area) and in the communication (red area). Bottom, Time course of synaptic adaptation (measured as strength of synapse) from same simulations above. Blue lines indicate adaptation in the high-frequency input, red lines indicate adaptation in the low-frequency input. **B**, Average evoked potential in response to probes after silence across neurons that were equally responsive and context sensitive. **C–F**, Paired statistical comparisons of temporal properties of the call-evoked LFPs (**C**, **D**) and spiking responses (**E**, **F**). Significance level is indicated at the top of **C–F**; *p* values were obtained from the Wilcoxon signed-rank test.

neurons preferring high-frequency echolocation sounds, but this remains to be tested.

### Acoustic context modulates natural sound discrimination in auditory cortex neurons

The data presented in this article advances our thinking on how context modulates natural sound selectivity by showing that the presence of leading context yields nonselective multifunctional neurons into selective ones and has the opposite effect on selective neurons. Moreover, such modulation appears to be related to neuronal type (i.e., putative pyramidal vs interneuron; Fig. 4B). These findings could have bearings on our interpretation on how the auditory cortex operates in natural conditions. Note that here selectivity refers only to differences in spike count obtained when presenting echolocation and communication sounds. Future studies are needed to test whether similar results are obtained when using different types of communication calls.

We propose that nonselective neurons are important for coding violations to regularities in the sound stream, that is, transitions from echolocation to communication (or vice versa). Meanwhile, neurons that appear to be selective to specific natural sound categories in the absence of context would be worse detectors of acoustic transitions but probably key for sound discrimination in silent environments. Accordingly, it has been claimed that the primary function of the AC is not to map parameters or features, such as frequency, but to integrate sounds on different time scales (Ulanovsky et al., 2003, 2004; Kanwal, 2012; Washington and Kanwal, 2012). Our findings are probably the outcome of sound processing at several levels both below and above the AC. Cortical areas upstream of the AC, such as the frontal cortex, are involved in the production of both vocal and behavioral responses (Eiermann and Esser, 2000; Kanwal et al., 2000; Garcia-Rosales et al., 2020; Weineck et al., 2020) and could potentially modulate how novel sounds are processed in the AC.

### Mechanisms underlying context-specific response modulation

We show that in context-sensitive neurons, the presence of a leading context always reduced responses to forthcoming sounds independently of the neuronal tuning. We propose that there is a common mechanism that underlies such a context effect across all cortical neurons of nontopographic areas. Interestingly, a significant reduction of the probe response was observed both 60 and 416 ms after the context end. Whole-cell recordings have shown that such a long-lasting context dependence (>100 ms) in the AC depends on synaptic depression, whereas at shorter intervals the effects are attributed to postsynaptic (GABAergic) inhibition (Wehr and Zador, 2005; Asari and Zador, 2009).

We found that the magnitude of the suppression driven by context sequences on responses to probe sounds was stimulus specific; there was stronger suppression on responses to probes that matched the context in comparison with those that mismatched. Similar results were found in studies using artificial sounds in the framework of stimulus specific adaptation (SSA) and predictive coding (Calford and Semple, 1995; Auksztulewicz and Friston, 2016; Carbajal and Malmierca, 2018). It has been proposed that SSA underlies enhanced responses to unexpected events occurring as deviant stimuli against a background of repetitive signals (Ayala and Malmierca, 2012). The auditory stimulation paradigm used by us shares similarities with SSA classic paradigms, especially if one considers the context sequences as a form of standard and the mismatched probes as deviant. We propose that highly specialized neurons in the bat AC follow

general neuronal principles that explain auditory deviance detection. It should be noted that combination sensitivity present in these neurons involve nonlinear interactions at intersound delays covering time scales between 1 and 30 ms (Hagemann et al., 2010). This time scale is outside the range of delays studied here (60 and 416 ms in the *in vivo* experiments, longer than to 2 s *in silico*).

Our model predicts that context modulation in neuronal responses to natural sounds results from two mechanisms, adaptation dependent on the postsynaptic activity and adaptation dependent on the presynaptic activity. The model assumes that the processing of natural sounds is segregated in different frequency channels that converge into a cortical neuron that exhibits context sensitivity. Thus, hearing the context unbalances the input-dependent adaptation and allows context-specific effects. However, the suppression of mismatching sounds cannot be explained by frequency-specific adaptation alone. We predict that postsynaptic-dependent adaptation, which attenuates every input arriving in the cortical neurons, is needed to explain the totality of our results.

We should point out that in our model, synaptic depression is assumed to occur in excitatory synapses in cortical cells. However, inhibitory synapses could also reproduce our predictions. Indeed, the role of GABAergic neurotransmission has been demonstrated in similar context-dependent situations (Perez-Gonzalez et al., 2012). It remains a challenge for future work to discover the biological correlate(s) of the synapses. Whether the synaptic depression depends on the activity of thalamic or cortical neurons remains unsolved. With our data we also cannot establish whether the effects observed originate at the cortical level or whether these are simply inherited from subcortical neurons. What we do propose is that regardless of the input origin, the selectivity to natural sounds in the presynaptic sites is tuned to the spectral components of natural sounds more than to sound categories (either to echolocation or communication signals). Our results showed that preference but not exclusivity to natural sounds is sufficient to reproduce and better fit the observed context-specific modulatory effects. The latter is in accordance with several reports in lower areas of the auditory pathway and in cortical areas, where the neurons exhibit high selectivity to natural sounds (Klug et al., 2002; Mayko et al., 2012; Salles et al., 2020).

Our neuron model also predicts that the increment of neuronal discriminability with acoustic context lasts longer for echolocation than for communication context, 1.6 s and 650 ms, respectively. Similar time scales have been described in the AC as important for processing natural auditory scenes that require integration over seconds (Ulanovsky et al., 2004; Asari and Zador, 2009). In addition, the model suggests different time scales of synaptic depression in the low- and high-frequency inputs of cortical neurons, slowest in high-frequency inputs. This last prediction was corroborated by studying LFP responses evoked by the communication and echolocation sound. Predictions arising from this study could be used as guiding hypotheses in future work using direct measurements of membrane potentials in cortical cells.

## References

- Abbott LF (1999) Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Res Bull* 50:303–304.
- Arroyo S, Bennett C, Hestrin S (2018) Correlation of synaptic inputs in the visual cortex of awake, behaving mice. *Neuron* 99:1289–1301. e1282.
- Asari H, Zador AM (2009) Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. *J Neurophysiol* 102:2638–2656.
- Auksztulewicz R, Friston K (2016) Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80:125–140.
- Ayala YA, Malmierca MS (2012) Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Front Neural Circuits* 6:89.
- Beetz MJ, Hechavarría JC, Kossel M (2016) Cortical neurons of bats respond best to echoes from nearest targets when listening to natural biosonar multi-echo streams. *Sci Rep* 6:35991.
- Brinklov S, Jakobsen L, Ratcliffe JM, Kalko EK, Surlykke A (2011) Echolocation call intensity and directionality in flying short-tailed fruit bats, *Carollia perspicillata* (Phyllostomidae). *J Acoust Soc Am* 129:427–435.
- Calford MB, Semple MN (1995) Monaural inhibition in cat auditory cortex. *J Neurophysiol* 73:1876–1891.
- Carbajal GV, Malmierca MS (2018) The neuronal basis of predictive coding along the auditory pathway: from the subcortical roots to cortical deviance detection. *Trends Hear* 22:2331216518784822.
- Chaverri G, Ancillotto L, Russo D (2018) Social communication in bats. *Biol Rev Camb Philos Soc* 93:1938–1954.
- Clopath C, Busing L, Vasilaki E, Gerstner W (2010) Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat Neurosci* 13:344–352.
- Eiermann A, Esser KH (2000) Auditory responses from the frontal cortex in the short-tailed fruit bat *Carollia perspicillata*. *Neuroreport* 11:421–425.
- Esser KH, Eiermann A (1999) Tonotopic organization and parcellation of auditory cortex in the FM-bat *Carollia perspicillata*. *Eur J Neurosci* 11:3669–3682.
- Esser KH, Condon CJ, Suga N, Kanwal JS (1997) Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proc Natl Acad Sci U S A* 94:14019–14024.
- Feng L, Wang X (2017) Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc Natl Acad Sci U S A* 114:E840–E848.
- Fitzpatrick DC, Kanwal JS, Butman JA, Suga N (1993) Combination-sensitive neurons in the primary auditory cortex of the mustached bat. *J Neurosci* 13:931–940.
- García-Rosales F, López-Jury L, González-Palomares E, Cabral-Calderin Y, Hechavarría JC (2020) Fronto-temporal coupling dynamics during spontaneous activity and auditory processing in the bat *Carollia perspicillata*. *Front Syst Neurosci* 14:14.
- Gillam E, Fenton MB (2016) Roles of acoustic social communication in the lives of bats. In: *Bat bioacoustics* (Fenton MB, Grinnell AD, Popper AN, Fay RR, eds), pp. 117–139. New York: Springer.
- González-Palomares E, López-Jury L, Wetekam J, Kiai A, García-Rosales F, Hechavarría JC (2021) Male *Carollia perspicillata* bats call more than females in a distressful context. *R Soc Open Sci* 8:202336.
- Hagemann C, Esser KH, Kossel M (2010) Chronotopically organized target-distance map in the auditory cortex of the short-tailed fruit bat. *J Neurophysiol* 103:322–333.
- Hagemann C, Vater M, Kossel M (2011) Comparison of properties of cortical echo delay-tuning in the short-tailed fruit bat and the mustached bat. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 197:605–613.
- Haider B, Schulz DP, Haussler M, Carandini M (2016) Millisecond coupling of local field potentials to synaptic currents in the awake visual cortex. *Neuron* 90:35–42.
- Harris KD, Henze DA, Csicsvari J, Hirase H, Buzsáki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 84:401–414.
- Hechavarría JC, Beetz MJ, Macias S, Kossel M (2016) Distress vocalization sequences broadcasted by bats carry redundant information. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 202:503–515.
- Hechavarría JC, Jerome Beetz M, García-Rosales F, Kössel M (2020) Bats distress vocalizations carry fast amplitude modulations that could represent an acoustic correlate of roughness. *Sci Rep* 10:7332.
- Hershenhoren I, Taaseh N, Antunes FM, Nelken I (2014) Intracellular correlates of stimulus-specific adaptation. *J Neurosci* 34:3303–3319.
- Kanwal JS (1999) Processing species-specific calls by combination-sensitive neurons in an echolocating bat. In: *The design of animal communication* (Hauser MD, Konishi M, eds), pp. 135–157. Cambridge, MA: MIT.
- Kanwal JS, Gordon M, Peng JP, Heinz-Esser K (2000) Auditory responses from the frontal cortex in the mustached bat, *Pteronotus parnellii*. *Neuroreport* 11:367–372.



- Kanwal JS (2006) A distributed cortical representation of social communication calls. In: Behaviour and Neurodynamics for Auditory Communication (Kanwal JS and Ehret G, eds), pp 156–188. Cambridge University Press.
- Kanwal JS (2012) Right-left asymmetry in the cortical processing of sounds for social communication vs. navigation in mustached bats. *Eur J Neurosci* 35:257–270.
- Klug A, Bauer EE, Hanson JT, Hurley L, Meitzen J, Pollak GD (2002) Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition. *J Neurophysiol* 88:1941–1954.
- Knörnschild M, Feifel M, Kalko EK (2014) Male courtship displays and vocal communication in the polygynous bat *Carollia perspicillata*. *Behav* 151:781–798.
- Kossl M, Hechavarría JC, Voss C, Macias S, Mora EC, Vater M (2014) Neural maps for target range in the auditory cortex of echolocating bats. *Curr Opin Neurobiol* 24:68–75.
- Laudanski J, Edeline JM, Huetz C (2012) Differences between spectro-temporal receptive fields derived from artificial and natural stimuli in the auditory cortex. *PLoS One* 7:e50539.
- Lawrence BD, Simmons JA (1982) Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. *J Acoust Soc Am* 71:585–590.
- López-Jury L, Mannel A, García-Rosales F, Hechavarría JC (2020) Modified synaptic dynamics predict neural activity patterns in an auditory field within the frontal cortex. *Eur J Neurosci* 51:1011–1025.
- Machens CK, Wehr MS, Zador AM (2004) Linearity of cortical receptive fields measured with natural sounds. *J Neurosci* 24:1089–1100.
- Martin LM, García-Rosales F, Beetz MJ, Hechavarría JC (2017) Processing of temporally patterned sounds in the auditory cortex of Seba's short-tailed bat, *Carollia perspicillata*. *Eur J Neurosci* 46:2365–2379.
- May PJ, Tiitinen H (2013) Temporal binding of sound emerges out of anatomical structure and synaptic dynamics of auditory cortex. *Front Comput Neurosci* 7:152.
- Mayko ZM, Roberts PD, Portfors CV (2012) Inhibition shapes selectivity to vocalizations in the inferior colliculus of awake mice. *Front Neural Circuits* 6:73.
- Medvedev AV, Kanwal JS (2004) Local field potentials and spiking activity in the primary auditory cortex in response to social calls. *J Neurophysiol* 92:52–65.
- Medvedev AV, Kanwal JS (2008) Communication call-evoked gamma-band activity in the auditory cortex of awake bats is modified by complex acoustic features. *Brain Res* 1188:76–86.
- Neuweiler G (1989) Foraging ecology and audition in echolocating bats. *Trends Ecol Evol* 4:160–166.
- Ohlemiller KK, Kanwal JS, Suga N (1996) Facilitative responses to species-specific calls in cortical FM-FM neurons of the mustached bat. *Neuroreport* 7:1749–1755.
- O'Neill WE, Suga N (1979) Target range-sensitive neurons in the auditory cortex of the mustache bat. *Science* 203:69–73.
- Perez-Gonzalez D, Hernandez O, Covey E, Malmierca MS (2012) GABA(A)-mediated inhibition modulates stimulus-specific adaptation in the inferior colliculus. *PLoS One* 7:e34297.
- Porter FL (2010) Social behavior in the leaf-nosed bat, *Carollia perspicillata*: II. social communication. *Zeitschrift für Tierpsychologie* 50:1–8.
- Romano J, Kromrey JD, Coraggio J, Skowronek J (2006) Appropriate statistics for ordinal level data: Should we really be using *t*-test and Cohen's *d* for evaluating group differences on the NSSE and other surveys. Paper presented at the annual meeting of the Florida Association of Institutional Research, Cocoa Beach, FL, February.
- Sadagopan S, Wang X (2009) Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. *J Neurosci* 29:11192–11202.
- Salles A, Park S, Sundar H, Macias S, Elhilali M, Moss CF (2020) Neural response selectivity to natural sounds in the bat midbrain. *Neuroscience* 434:200–211.
- Stimberg M, Brette R, Goodman DF (2019) Brian 2, an intuitive and efficient neural simulator. *Elife* 8:e47314.
- Suga N (2015) Neural processing of auditory signals in the time domain: delay-tuned coincidence detectors in the mustached bat. *Hear Res* 324:19–36.
- Suga N, O'Neill WE, Manabe T (1978) Cortical neurons sensitive to combinations of information-bearing elements of biosonar signals in the mustache bat. *Science* 200:778–781.
- Taaseh N, Yaron A, Nelken I (2011) Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One* 6:e23369.
- Thomas JA, Moss CF, Vater M, Moore PW (2005) Echolocation in bat and dolphins. *J Acoust Soc Am* 118:2755–2756.
- Tsunada J, Lee JH, Cohen YE (2012) Differential representation of auditory categories between cell classes in primate auditory cortex. *J Physiol* 590:3129–3139.
- Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6:391–398.
- Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24:10440–10453.
- Washington SD, Kanwal JS (2012) Sex-dependent hemispheric asymmetries for processing frequency-modulated sounds in the primary auditory cortex of the mustached bat. *J Neurophysiol* 108:1548–1566.
- Wehr M, Zador AM (2005) Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron* 47:437–445.
- Weineck K, García-Rosales F, Hechavarría JC (2020) Neural oscillations in the fronto-striatal network predict vocal output in bats. *PLoS Biol* 18:e3000658.
- Wenstrup JJ, Portfors CV (2011) Neural processing of target distance by echolocating bats: functional roles of the auditory midbrain. *Neurosci Biobehav Rev* 35:2073–2083.