

# Adaptive Efficient Coding of Correlated Acoustic Properties

Kai Lu,<sup>1</sup> Wanyi Liu,<sup>1</sup>  Kelsey Dutta,<sup>1</sup>  Peng Zan,<sup>1</sup>  Jonathan B. Fritz,<sup>1</sup> and  Shihab A. Shamma<sup>1,2,3</sup>

<sup>1</sup>Institute for Systems Research, <sup>2</sup>Electrical and Computer Engineering Department, University of Maryland, College Park, Maryland 20740, and <sup>3</sup>Department of Cognitive Studies, Ecole Normale Supérieure, 75005 Paris, France

Natural sounds such as vocalizations often have covarying acoustic attributes, resulting in redundancy in neural coding. The efficient coding hypothesis proposes that sensory systems are able to detect such covariation and adapt to reduce redundancy, leading to more efficient neural coding. Recent psychoacoustic studies have shown the auditory system can rapidly adapt to efficiently encode two covarying dimensions as a single dimension, following passive exposure to sounds in which temporal and spectral attributes covaried in a correlated fashion. However, these studies observed a cost to this adaptation, which was a loss of sensitivity to the orthogonal dimension. Here we explore the neural basis of this psychophysical phenomenon by recording single-unit responses from the primary auditory cortex in awake ferrets exposed passively to stimuli with two correlated attributes, similar in stimulus design to the psychoacoustic experiments in humans. We found: (1) the signal-to-noise ratio of spike-rate coding of cortical responses driven by sounds with correlated attributes remained unchanged along the exposure dimension, but was reduced along the orthogonal dimension; (2) performance of a decoder trained with spike data to discriminate stimuli along the orthogonal dimension was equally reduced; (3) correlations between neurons tuned to the two covarying attributes decreased after exposure; and (4) these exposure effects still occurred if sounds were correlated along two acoustic dimensions, but varied randomly along a third dimension. These neurophysiological results are consistent with the efficient coding hypothesis and may help deepen our understanding of how the auditory system encodes and represents acoustic regularities and covariance.

**Key words:** auditory cortex; efficient coding; ferret; single-unit

## Significance Statement

The efficient coding (EC) hypothesis (Attneave, 1954; Barlow, 1961) proposes that the neural code in sensory systems efficiently encodes natural stimuli by minimizing the number of spikes to transmit a sensory signal. Results of recent psychoacoustic studies in humans are consistent with the EC hypothesis in that, following passive exposure to stimuli with correlated attributes, the auditory system rapidly adapts so as to more efficiently encode the two covarying dimensions as a single dimension. In the current neurophysiological experiments, using a similar stimulus design and the experimental paradigm to the psychoacoustic studies of Stilp et al. (2010) and Stilp and Kluender (2011, 2012, 2016), we recorded responses from single neurons in the auditory cortex of the awake ferret, showing adaptive efficient neural coding of two correlated acoustic attributes.

## Introduction

Perception of natural sounds and images often relies on correlated cues (Kluender et al., 2013). For example, the perception of slopes of objects can be cued both by binocular disparity and texture gradient (Hillis et al., 2002). In speech, many acoustic features covary to give rise to the percept of different phonemes. This is critical because a single acoustic feature often does not

provide reliable information to distinguish a phoneme because they are susceptible to changes due to preceding and/or following articulations in various ways (Kluender and Lotto, 1999; Kluender and Kiefe, 2006; Kluender and Alexander, 2007). By contrast, the covariance among multiple features can reliably capture the differences among speech sounds (Sussman et al., 1998). Therefore, understanding how the nervous system adapts to encode the covarying features of a stimulus is a key question in the study of perception and communication.

Attneave (1954) and Barlow (1961) proposed the efficient coding (EC) hypothesis, where they conjectured that spikes in sensory systems form an efficient code to represent natural stimuli and that sensory processing is optimized for natural stimuli. Consistent with EC, there is evidence that neural responses in the auditory and visual system are indeed apparently optimized to encode natural sounds and images (Olshausen and Field, 1997;

Received Jan. 16, 2019; revised Aug. 26, 2019; accepted Aug. 30, 2019.

Author contributions: K.L. designed research; W.L. and K.D. performed research; K.L., W.L., P.Z., and S.A.S. analyzed data; K.L., J.B.F., and S.A.S. wrote the paper.

This work was supported by Grants from the National Institutes of Health (R01 DC005779) and an Advanced ERC Grant (NEUME) to S.A.S. We thank Dr. Keith Kluender for the insightful and detailed comments.

The authors declare no competing financial interests.

Correspondence should be addressed to Shihab A. Shamma at [sas@isr.umd.edu](mailto:sas@isr.umd.edu).

<https://doi.org/10.1523/JNEUROSCI.0141-19.2019>

Copyright © 2019 the authors

Vinje and Gallant, 2000; Lewicki, 2002; Smith and Lewicki, 2006). And because many such stimuli have correlated attributes, it has been proposed that sensory systems are able to actively recalibrate coding to enhance coding efficiency (Barlow and Földiák, 1989) and some recent experiments in the visual system support this dynamic version of EC (Coen-Cagli et al., 2015).

To understand how EC contributes to the encoding of covariant features, recent psychoacoustic studies (Stilp et al., 2010; Stilp and Kluender, 2011, 2012, 2016) tested subjects who either received passive exposure to, or provided continuous discrimination judgments for sets of complex sounds with covarying spectrotemporal attributes (spectral-shapes and amplitude attack-decay ratios). Following passive exposure, and also over the course of active discrimination, subjects' acuity in discriminating sounds along the correlated dimensions remained intact, but discrimination of pairs on the orthogonal dimension was significantly impaired when those pairs were proximal to the principle vector of covariance. These findings suggested that experience with correlated attributes (spectral shape and attack-decay) induced the auditory system to collapse the covarying dimensions into a single dimension at the expense of lost sensitivity to the orthogonal dimension. In the present study, we tested whether neural responses at a single-unit level in an animal model replicate this psychoacoustic phenomenon observed in humans. We did so by measuring neural responses in primary auditory cortex (A1) of awake ferrets, using the same passive stimulus exposure paradigm as used in the original study of Stilp et al. (2010).

The current study comprises three key experiments. In Experiment 1, we measured baseline auditory cortical responses to sounds with a correlation between two acoustic attributes, spectral and temporal, which were the peak frequency of the spectral envelope (SP) and the amplitude modulation (AM) rate. We found that cortical responses became adapted to sounds along the dimension of the correlated attributes. Consistent with the results of psychoacoustic studies, the signal-to-noise ratios (SNRs) along this dimension remained intact, whereas those along the dimension orthogonal to it decreased. Also consistent with human behavioral studies (Stilp et al., 2010), the performance of a decoder trained with spike data to discriminate stimuli along the orthogonal dimension was reduced after exposure. Finally, correlations between neurons tuned to the same two attributes decreased after exposure. A control experiment (Experiment 2) tested whether passive exposure to stimuli varying along only a single dimension, i.e., holding other parameters constant, induced effects similar to those for the correlation dimension in Experiment 1. The goal of Experiment 2 was to ascertain whether the effects observed in Experiment 1 could have arisen due to simple stimulus adaptation, or whether they required covariance between the two attributes. A final experiment (Experiment 3) tested whether the exposure effects of the two covarying attributes observed in Experiment 1 persisted in the presence of a third acoustic attribute that varied randomly along a separate third additional acoustic dimension (e.g., the fundamental frequency).

## Materials and Methods

### Subjects

Experiments used adult female ferrets ( $n = 4$ ) housed with a 12 h light/dark cycle. Two ferrets (F-1 and F-2) were used in Experiment 1. Two ferrets (F-2 and F-3) were used in Experiment 2. Two ferrets (F-3 and F-4) were used in Experiment 3. Ferrets used in this study were previously trained on unrelated auditory tasks (Lu et al., 2017). Neurophysiological recording sessions (4–8 h in duration) occurred on 2

nonconsecutive days per week. All procedures were in accord with National Institutes of Health policy on experimental animal care and use and conformed to a protocol approved by the Institutional Animal Care and Use Committee of the University of Maryland.

### Surgeries

To stabilize the head for electrophysiological recording, a headpost was implanted in a surgery that occurred at least 1 month before the initiation of recordings. Animals were anesthetized with isoflurane (1–2% in oxygen), and a customized stainless steel headpost was surgically implanted on the skull under aseptic conditions. The skull over the auditory cortex was exposed and covered with a thin layer of Heraeus Kulzer Charisma (1 mm) surrounded by a thicker wall built with UV-curable Charisma (3 mm thick). After recovery from surgery, animals were gradually habituated to restraint in a customized head-fixed holder. After successful habituation, 1–2 d before electrophysiological recording, a small craniotomy (1–2 mm diameter) was made above the primary auditory cortex. At the beginning and end of each recording session, the craniotomy was thoroughly rinsed with sterile saline. At the end of a recording session, the craniotomy and the well were filled with topical antibiotics that were rotated on a weekly basis (Baytril and cefazolin). The area containing the hole was then filled with sterile vinyl polysiloxane impression material (Examix NDS, GC America Inc.) that maintained a tight seal and kept the brain protected between experiments. After recordings were completed in the original craniotomy, adjacent 0.5 mm bone sections were carefully removed over successive months of neurophysiological recording so that eventually the enlarged craniotomy (~4 mm diameter) encompassed the entire primary auditory cortex.

### Electrophysiological recording

Electrophysiological recordings were made in a double-walled sound-proof room (IAC). The awake animal was placed in a horizontal, Lexan tube, and the implanted headpost was used to stabilize and fix the head in position, relative to a stereotaxic frame. Recordings were conducted in the A1 of both left and right hemispheres over a period of 3–6 months. For each recording, 4–8 tungsten microelectrodes (2–3 MΩ; FHC) were introduced through the craniotomies and controlled by independently moveable drives (Electrode Positioning System, Alpha-Omega). Raw neural activity traces were amplified, filtered, and digitally acquired by a data acquisition system (AlphaLab, Alpha-Omega). Multiunit neuronal activity (including all spikes that rose above a threshold level of 3.5 SD of baseline noise) was monitored online. In addition, single units were identified online using Alpha-Omega spike waveform profiling to isolate and to monitor single neuron responses. Bandpass noise (0.2 s, 1 octave) and pure tone (0.2 s duration) stimuli were presented to search for responsive sites. Because of evidence that neurons in supragranular layers (II–III) of A1 show greater plasticity than neurons in deeper layers (Francis et al., 2018), most of the recording depths in this study were within 100–400 μm of the cortical surface, and electrodes were advanced only to the most superficial position where single-unit responses to bandpass noise and tones were found. Once clear, stable auditory responses were obtained, the experimental stimuli were presented. All stimulus amplitudes were presented at 65 dB sound pressure level from a speaker placed 1 m in front of the animal. After recordings were completed, single units were isolated again by off-line customized spike-sorting software, involved three steps: (1) we extracted three principal components from the spike waveform from each recording channel. (2) We used the  $k$  means clustering approach to split multiunit spikes into single-unit clusters and then created the spike templates from the center of each single unit cluster. (3) A template-matching algorithm was used to assigned spikes into each spike template (Meska-PCA, NSL). Single units were confirmed by interspike interval histogram (no more than 2% spike in the 1 ms bin) and consistency of spike waveforms.

### Auditory stimuli and experimental design

In each of the three sets of experiments that were performed, animals were exposed to a generated set of acoustic stimuli, that differed in fundamental ways in each experiment. In Experiment 1, two acoustic attributes were correlated (peak frequency of the SP and AM). In this experiment, we explored the effects of this correlation on the coding of

the stimuli in A1. In contrast, Experiment 2 was a control experiment, in which only one acoustic attribute was varied to test whether simple adaptation in A1 could explain the results obtained in Experiment 1. Experiment 3 was also a control experiment, in which we presented stimuli with the same correlated acoustic attributes used in Experiment 1, but now we explored the effects (or lack thereof) of adding a new, third independently varying (uncorrelated) attribute to the acoustic stimuli, on the effects observed in Experiment 1. Thus, in Experiment 3, all stimuli were characterized by three acoustic attributes, two of which were correlated, and one of which was not. We tested whether the presence of an uncorrelated acoustic attribute would change the effects of exposure to stimuli with two correlated acoustic attributes, as observed in Experiment 1.

### Experiment 1

**Stimuli and optimal stimulus design.** In each recording session, we generated a new “optimal” stimulus matrix with two manipulated attributes: AM rate and peak frequency of the SP. All stimuli were harmonic complexes with an initially flat spectrum modulated/filtered by the two attributes. Exposure stimuli were generated with orthogonal correlations between attributes. AM and SP were either positively correlated (increasing rate with increasing peak frequency; Fig. 1A, represented by red dots in C) or negatively-correlated (increasing rate with decreasing peak frequency; Fig. 1C, blue dots). “Test” stimuli were generated with all combinations of the two attributes, so that they were uniformly distributed in the 2-dimensional space of the two parameters (Fig. 1C, black dots).

To design an optimal stimulus matrix for a given recording site, it was first necessary to evaluate the frequency response profile of the site. Thus, at the beginning of each recording session, several neurons were isolated; best frequencies (BFs) measured; and, median BFs estimated. Once frequency tuning was measured, tone-complexes ranging across  $\pm 3$  octaves around the median BF (Fig. 1B) were created. Nineteen frequencies were defined as SP peaks with equal (1/3 octave) steps across this range, and 19 triangular SP filter shapes were designed (Fig. 1A, 3 examples are shown in the second column).

Nineteen AM rates were set at equal log steps from 5 to 120 Hz (Fig. 1A, first column). In each recording session, the fundamental frequency ( $f_0$ ) of the harmonic complex was randomly selected from a range of 200–500 Hz following the criterion that  $f_0$  must be 0.1 octave lower than the lowest peak frequency of spectral functions (Fig. 1B, bottom, green dashed line). Roving  $f_0$  across sessions minimized the effects of frequency-specific influences (e.g., peak harmonic relative to the peak of spectral envelope) and the possibility of retaining a memory of previous stimulus sets, because each stimulus set was session-unique. Stimuli were 500 ms duration with 5 ms cosine onset and offset ramps, and sampled at 40 kHz. Test stimuli included 25 stimuli selected from the combination of the steps 2, 6, 10, 14, and 18 from both two dimensions forming a  $5 \times 5$  matrix uniformly sampled from the  $19 \times 19$  exposure matrix. Finally, because all exposure stimuli were located along diagonals of the stimulus matrix, additional eight test stimuli were selected along these two diagonals of the stimulus matrix at steps 4, 8, 12, and 16.

**Passive stimulus exposure and testing procedure.** As mentioned in the previous section, in Experiment 1, two ferrets were first tested with the full stimulus matrix shown in Figure 1C to measure neurons’ tuning properties before stimulus exposure. This was followed by exposure to 80 repetitions of 19 stimuli along one diagonal (either red dots or blue dots). Next, a first post-exposure test included all 33 stimuli in the full matrix (as in the Pre-exposure test) to measure the effects of passive exposure to the 19 stimuli with correlated attributes. Because of uncertainty regarding the duration of persistence of exposure effects, we repeated the Exposure/Post-exposure sequence. In summary, the complete sequence protocol for stimulus presentations is shown in Figure 1D: (1) pre-exposure test stimuli, including the 33 testing stimuli ( $(5 \times 5) + 8$ ) were presented 20 times each with 1 s silence (ISI) between sounds, all over a period of 16.5 min. (2) First passive exposure session: 19 exposure stimuli were presented 80 times each with 0.25 s ISI over a period of 19 min. (3) Post-exposure test using the same 33 stimuli of the Pre-exposure test with 10 repetitions over a period of 8.25 min. (4) A second exposure session (over a period of 9.5 min) was conducted immediately afterward, using

the same 19 stimuli (as in Step 2) repeated 40 times to reassess the effects of the first exposure combined with the second exposure. (5) A final (second) post-exposure test, using the same 33 stimuli of the Pre-exposure test, repeated 10 times over a period of 8.25 min. Data collected from the two Post-exposure tests (Steps 3,5) were pooled together for comparison with the Pre-exposure test (1). In both the Pre-exposure and Post-exposure tests and two exposure sessions, all stimuli were presented in a randomly-shuffled order.

### Experiment 2

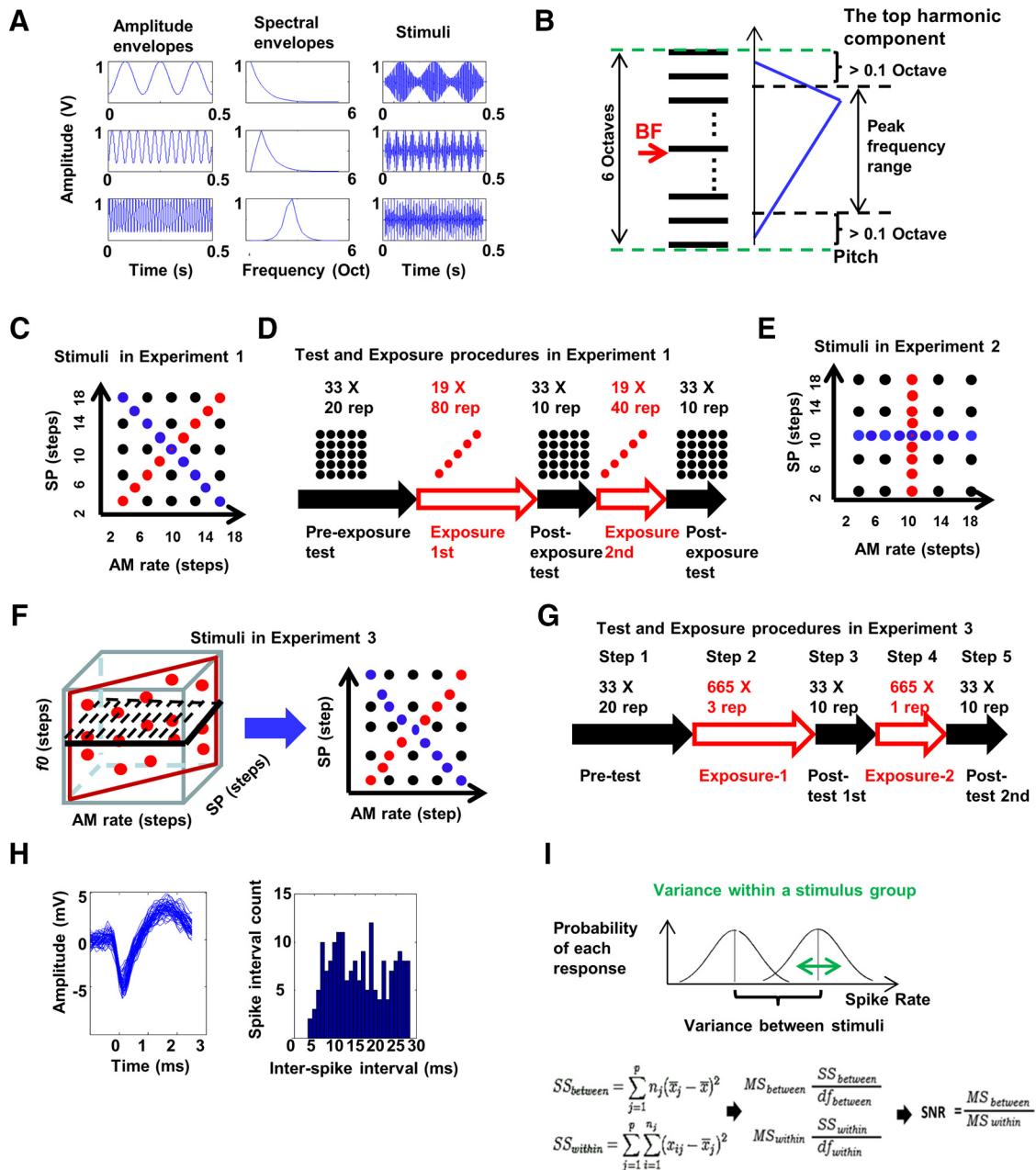
A possible confound that we considered is that the effects observed in Experiment 1 could simply have arisen from adaptation to repetitive sounds. As a control, two more ferrets were tested (Experiment 2) with 19 sounds during the Passive exposure phase, which in this experiment, varied along only a single dimension (either AM or SP, balanced across recordings). The parameter of the other dimension (the orthogonal dimension) was held constant in a given recording session (Fig. 1E), but was different across sessions. Testing stimuli in Experiment 2 included a  $5 \times 5$  stimulus matrix as in Experiment 1. However, in contrast with the stimuli used in Experiment 1, in Experiment 2 the additional eight testing stimuli were sampled from a single row or column of the stimuli (rather than along the diagonals) at the step 4, 8, 12, and 16 to generate a total of 33 test stimuli. The basic sequence, test procedures, and ISI were the same as in Experiment 1, with the key difference being that in Experiment 2, exposure stimuli were now along the vertical or horizontal direction, rather than along diagonal direction (Experiment 1). By comparing the results of Experiments 1 and 2, we could assess the specific effects of the two correlated attributes.

### Experiment 3

In Experiment 3, modeled on the earlier psychoacoustic study of Stilp and Kluender (2011), we tested, in recordings from two ferrets, whether the exposure effects caused by covariation in stimuli with two acoustic dimensions (as observed in Experiment 1) would be sustained in the presence of substantial variation in new stimuli with a third acoustic dimension that introduced widely varying physical acoustic properties. Here, we introduced this variation by varying  $f_0$  from trial to trial. As in Experiment 1, 19 AM and SP combinations were selected, in which the two dimensions were either positively or negatively correlated. In addition, 35  $f_0$  parameters were generated in 2% increments between the two nearest steps. The fundamental frequency range was restricted by the criteria that the lowest  $f_0$  had to be 1.5 octaves higher than the maximum AM rate, and the highest  $f_0$  had to be 0.1 octave lower than the lowest spectral peak selected for the SP dimension. A 2-dimensional matrix of 665 ( $19 \times 35$ ) stimuli with 19 AM/SP combinations as one dimension and 35  $f_0$ s as the second dimension was created as the exposure stimulus set (Fig. 1F, red vertical rectangle). Note that this matrix effectively created acoustic stimuli with three distinct attributes (i.e., AM rate, SP, and frequency). For testing, we randomly selected one  $f_0$  of the 35  $f_0$  values in each recording. Then, testing stimuli were created based on the selected  $f_0$  in the exact same way as in Experiment 1 (Fig. 1F, black horizontal rectangular). Thus, the *exposure* stimuli in Experiment 3 existed in a 2-dimensional space that was orthogonal to the space of *testing* stimuli. Pre-exposure and Post-exposure test procedures were the same as in Experiment 1. Note that in Experiment 1, all 19 AM/SP combinations in passive exposure were presented 120 times. To ensure approximately equal covariance exposure, all 665 exposure stimuli in the first passive exposure session of Experiment 3 were presented three times (Fig. 1G). Then, as in the procedure in Experiment 1, the 665 stimuli were presented one more time in a second passive exposure session (140 repeats of 19 AM/SP combinations in total, but with different  $f_0$ s).

### Data analysis of single-unit responses (spike rate)

As indicated in the section of electrophysiological recordings, we isolated single units off-line from our recordings, and all the data analysis was conducted on isolated single neurons from A1 (Fig. 1H). For each trial, we measured spike rate in a 650 ms response window (a window that began 50 ms after stimulus onset, and continued throughout the rest of the 500 ms stimulus until 200 ms after stimulus offset) and the spike rate in a baseline window (200 ms before stimulus onset). The response am-



**Figure 1.** **A**, Schematic illustration of how stimuli were generated. Amplitude envelopes with different AM modulation rates were generated (left column). Spectral envelopes with different peak frequencies were generated (middle column). Sounds were generated by combining the two acoustic attributes (right column). **B**, Illustration of how stimuli parameters were chosen. Black horizontal lines indicate harmonic components consisting of the basic structure of a stimulus. The center of the harmonic tone was chosen around the BF of recorded neurons (red arrow). The blue triangle represented one of the 19 spectral envelopes used to generate stimuli. The limit of the peak frequencies was indicated by the two horizontal black dashed lines. The range of the harmonic structure was 6 octaves. The limits of the lowest harmonic component ( $f_0$ ) and highest component were indicated by the green dashed line, at least 0.1 octave away from the limits of the peak frequencies. **C**, Testing stimulus space in Experiment 1. Notice that, in one-half of the recordings, 19 exposure stimuli had positively correlated attributes and thus were lined up with those red dots. In this case, the orthogonal dimension was indicated by blue dots. In the other half of the recordings, the opposite was true: the exposure dimension was lined up with blue dots and the orthogonal dimension was lined up with red dots. These two types of exposure were shuffled across recordings in a balanced design. **D**, Schematic illustration of exposure and testing procedure in Experiment 1. Black arrows indicate the testing session; red arrows indicate the exposure sessions. The matrix of black dots on top of the black arrow represents 33 test stimuli. Red dots on top of red arrows represent 19 exposure stimuli. **E**, Testing stimulus space in Experiment 2. In one-half of the recordings, sounds with constant AM and varying SP (lined up with red dots) were exposure sounds, whereas those with constant SP and varying AM (lined up with blue dots) were on the orthogonal dimension. In the other half of recordings, the opposite was true. Two types of exposure were also shuffled in a balanced design. **F**, Testing stimuli space in Experiment 3. The red parallelogram indicates exposure sounds, which had perfect positive or negative correlation between AM and SP, and varied  $f_0$ . The black parallelogram with dashed lines indicates testing sounds, in which only one  $f_0$  was randomly chosen for testing in each recording. The structure of the testing space is the same as in Experiment 1, indicated on the right by the blue arrow. **G**, Exposure and testing procedure in Experiment 3. Black arrows indicate the testing sessions; red arrows indicate the exposure sessions. The total stimuli number in exposure was approximately the same as that in Experiment 1 (2280 in Experiment 1; 2660 in Experiment 3). **H**, An example of an isolated single-unit. Spike waveforms (left) for the neuron were consistent across the whole recording session. Histograms of interspike intervals (right) were used to verify isolation quality. **I**, Illustration of how SNRs of spike rate coding for the exposed dimension and the orthogonal dimension were calculated. The bell-shaped curves illustrate the possible spike rate variation within the group, the sum of which across all groups yielded the “Sum of Squares (SS) within”. The distance between the centers of the bell curves indicated the variance between the groups, explained by stimulus differences along the given dimension. The sum of all such variance yielded the “SS between”. Then SNR was calculated from SS within and SS between as shown in the equation on the right. MS, Mean of Squares; df, Degree of freedom.

plitude was defined as the spike rate in the response window minus the baseline spike rate, and was averaged over 20 repetitions. Any trials with response amplitude less than or  $>5$  standard deviation (SD) from mean amplitude were excluded from the analysis. Response amplitudes in the Pre-exposure test and Post-exposure test were calculated separately. Trials from two Post-exposure tests (3, 5) were combined for analysis. Baseline spike rates were compared across all three test sessions. Units with significant changes between baseline measures were excluded from further analysis.

**Quantification of SNR in spike-rate coding.** We used three neuronal response measures and explored how each of these measures of neural activity was affected by stimulus exposure in each of the three experiments. The first measure is the SNR of spike rate coding, defined as the ratio of the response variance to stimuli along one acoustic dimension, to the overall variance within each stimulus group (Fig. 1*I*). The calculation procedure resembled a two-way analysis of variance (ANOVA) (Privitera, 2017), in which stimulus levels along the exposure (correlated) dimension and the orthogonal (uncorrelated) dimension were treated as two independent variables, and the response amplitude of each stimulus was treated as the dependent variable. The analysis of SNR was performed for each neuron.

The same SNR analysis was conducted for data from both the Pre-exposure test and the Post-exposure tests, and differences in SNRs (change in SNRs) before and after exposure were measured. SNR changes due to exposure were analyzed separately for exposure dimension and for orthogonal dimension. In addition, SNRs themselves were compared between the Pre-exposure test and the Post-exposure test. Because the distribution of SNRs did not fully satisfy criteria for parametric tests, the Wilcoxon test (a nonparametric version of paired *t* test) was performed to compare SNRs before and after exposure.

**Discrimination performance by the neural decoder.** To compare results of neural data and previous psychoacoustic experiments in humans (Stilp et al., 2010), we trained a decoder to discriminate stimuli along the exposed dimension and the orthogonal dimension, based on a maximum likelihood estimation (MLE) method (Geman and Hwang, 1982). This approach has two advantages over other approaches: (1) it allows a decoder to learn to discriminate stimuli from very limited trials of training ( $n = 10$ ); (2) it does not require the decoder to be trained for multiple epochs, whereas other methods usually require hundreds of epochs of training. The decoder was trained independently for each neuron for Pre-tests before and Post-tests after exposure. In each test, two stimuli (Stimulus 1 and Stimulus 2) on adjacent positions along the exposed dimension or the orthogonal dimension were chosen (see Fig. 4*A*). Their identities were decoded based on the spike rate for each neuron, measured in the same way as for the SNR calculation. First, responses to two adjacent stimuli were separated evenly into training and testing sets, 10 trials of training and 10 trials of testing for a 20 trial recording of responses to each stimulus. Second, a Bayesian approach was performed to solve the two-class classification problem in the testing set based on the statistics of the training set. Specifically, a Bayesian approach solves the maximum a posteriori probability problem (Geman and Hwang, 1982) as follows:

$$\begin{aligned} \max p(\theta|x) &= \max_{\theta} \frac{p(\theta, x)}{p(x)} \\ &= \max_{\theta} \left\{ \frac{p(x|\theta=0)p(\theta=0)}{p(x)}, \frac{p(x|\theta=1)p(\theta=1)}{p(x)} \right\}, \quad (1) \end{aligned}$$

where  $\theta \in \{0,1\}$  is the stimulus identity (Stimulus 1 as 0 and Stimulus 2 as 1), and  $x$  is the spike number. This approach finds the  $\theta$  that maximize a posteriori probability  $p(\theta|x)$ . Because the probability was drawn from training data, which contain the same number of trials for Stimulus 1 and Stimulus 2,  $p(\theta=0) = p(\theta=1) = \frac{1}{2}$ . Therefore, Equation 1 becomes:

$$\max p(\theta|x) = \max_{\theta \in \{0,1\}} p(x|\theta), \quad (2)$$

which is also known as the MLE equation. In practice, a probability histogram was extracted for  $p(x|\theta)$  from the training set for each stimulus, and the center of each histogram bin was recorded. A classification decoder was then computed by assigning the class number with a larger conditional probability to each bin. When a testing sample was fed to the decoder, it was assigned the stimulus number of the bin it fell into. Accuracy of discrimination was calculated based on the performance of the decoder.

Accuracy of discrimination of all stimulus pairs along the exposure dimension and orthogonal dimension were averaged separately for each neuron. Analyses were performed in the same way for each neuron before and after exposure. Wilcoxon tests were used to compare the accuracy of discrimination along each dimension obtained before and after exposure.

**Correlation of tuning in simultaneously recorded neurons.** Finally, the third response measure used in this analysis was the correlation coefficient between the tuning functions to the AM and SP parameters. We measured the correlation coefficient for either the same neuron or for neuron pairs recorded simultaneously [neuron pairs were distinct neurons recorded simultaneously either from adjacent electrodes or separate single units (based on waveform and tuning) recorded from the same electrode]. For analysis of correlation in the same neuron, the tuning functions to AM and SP were calculated based on the averaged response at each AM/SP level. The correlation between the two functions was then calculated (with Spearman correlation) for each neuron, and for each Pre-exposure and Post-exposure test. Next, we computed the differences in correlation coefficients obtained from all neurons before and after exposure. Finally, correlation coefficients from recordings with different stimulus exposures, e.g., positively versus negatively correlated, were computed and compared with each other. Because the population of correlation coefficients was not normally distributed, the Mann–Whitney *U* test (a nonparametric version of two-sample *t* test), which provides a more conservative evaluation than a traditional *t* test, was used for the positive/negative comparison. To analyze correlations between AM and SP from different simultaneously recorded neurons, we first paired neurons recorded simultaneously in each session, one for AM and one for SP. Then, the correlation between AM and SP functions was calculated for the neuron pairs and averaged across all pairs. The differences between correlation coefficients before and after exposure were compared between recordings with exposure to positively and negatively correlated stimuli.

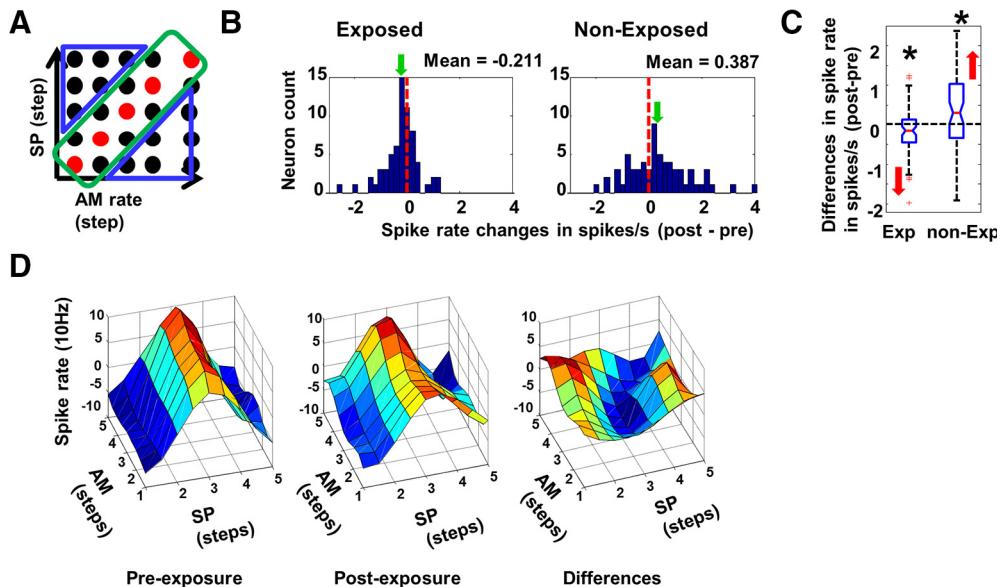
## Results

### Experiment 1: effects of exposure to stimuli with two correlated acoustic attributes

In the first experiment, we examined the effects of exposure to stimuli with two correlated acoustic attributes on single-unit neuronal responses in A1. We first describe the results of exposure on (1) changes in the tuning properties of the neurons with respect to the AM and SP parameters of the exposure stimulus set, then examine (2) changes in coding quality, i.e., SNR, and (3) changes in the discrimination accuracy of the neuronal responses using a MLE decoder, and finally (4) describe the changes in inter-neuronal correlations between simultaneously recorded neurons after stimulus exposure. These multiple measured changes reveal the effects of stimulus exposure on neural coding of correlated features, in a manner consistent with the EC hypothesis and previous psychoacoustic studies (Stilp et al., 2010; Stilp and Kluender, 2012, 2016).

### Effect of stimulus exposure on neuronal tuning properties

As described in detail in Materials and Methods, we first exposed ferrets to a set of covarying acoustic stimuli optimized based on the neurons' BF and then recorded post-exposure changes in responses from single neurons ( $n = 65$ ) in the A1 of two ferrets. Our first question was whether there might be a nonspecific effect following exposure on the overall responsiveness (firing rate) of



**Figure 2.** *A*, Schematic illustration of stimuli used in the tuning analysis in Experiment 1. Stimuli within the green diagonal rectangle indicated exposure stimuli (5 stimuli directly on the diagonal, and 8 nearby). Spike rates of these 13 composite stimuli were averaged and compared between tests before and after exposure. Stimuli within the two blue triangles were non-exposure stimuli (far from the exposure stimuli), from which the spike rates were averaged for comparison between tests before and after exposure. *B*, Histograms of spike rate changes after exposure in Experiment 1 (left, exposure stimuli; right, non-exposure stimuli). The red dashed line indicates zero change. The distribution of spike rate changes for exposure stimuli was significantly biased to the negative side, whereas the distribution for non-exposure stimuli was biased to the positive side. The green arrow indicates the mean (also indicated on the top of the histogram) for each distribution. *C*, Boxplot of spike rate changes after exposure in Experiment 1 (left, exposure and nearby stimuli; right, non-exposed stimuli). The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. Red crosses indicate possible outliers. The star on the top indicates that the distribution was significantly different from zero. The red arrow on the side indicates the direction of significant changes. Exp ( $p = 0.001$ ); non-Exp ( $p = 0.016$ ). Spike rates significantly decreased for the exposure stimuli and significantly increased for non-exposure stimuli. *D*, Surface plot of the 2-dimensional tuning map in Experiment 1: before exposure (left), and after exposure (middle). The unit of AM and SP was the step used to create the stimulus matrix. The differences between the two maps were taken (post – pre) and showed on the right. The dark blue area in the plot on the right indicates reduced responses along the diagonal along the exposed dimension. The changes in the tuning map were consistent with quantification shown in *B*.

A1 neurons. To answer this question, we compared the response amplitude averaged across all stimuli before and after exposure. We did not observe any significant changes in overall response amplitude after exposure (Wilcoxon test:  $z = -0.8332$ ,  $p = 0.404$ ) and hence concluded that there were no nonspecific changes in neuronal responsiveness (firing rate) following exposure.

Second, we tested the key question of whether there was a selective adaptation effect in A1 neuronal responses that was specifically related to the exposed stimuli, by separately examining responses to two groups of stimuli, e.g., the acoustic stimuli to which the ferrets had been exposed compared with the stimuli to which they had not been exposed. Specifically, when we contrasted responses to “exposure” stimuli (Fig. 2A, 13 combinations within the diagonal green rectangle) versus the responses to non-exposure stimuli (Fig. 2A, 12 combinations within blue triangles), we found a clear effect of selective adaptation: responses to the exposure stimuli (both positively or negatively correlated) significantly decreased in the post-exposure tests (Wilcoxon test:  $z = -2.6$ ,  $p = 0.001$ ; Fig. 2B, left-histogram, C, left-box), whereas responses to non-exposure stimuli significantly increased (Wilcoxon test:  $z = -2.4$ ,  $p = 0.016$ ; Fig. 2B, right-histogram, C, right-box).

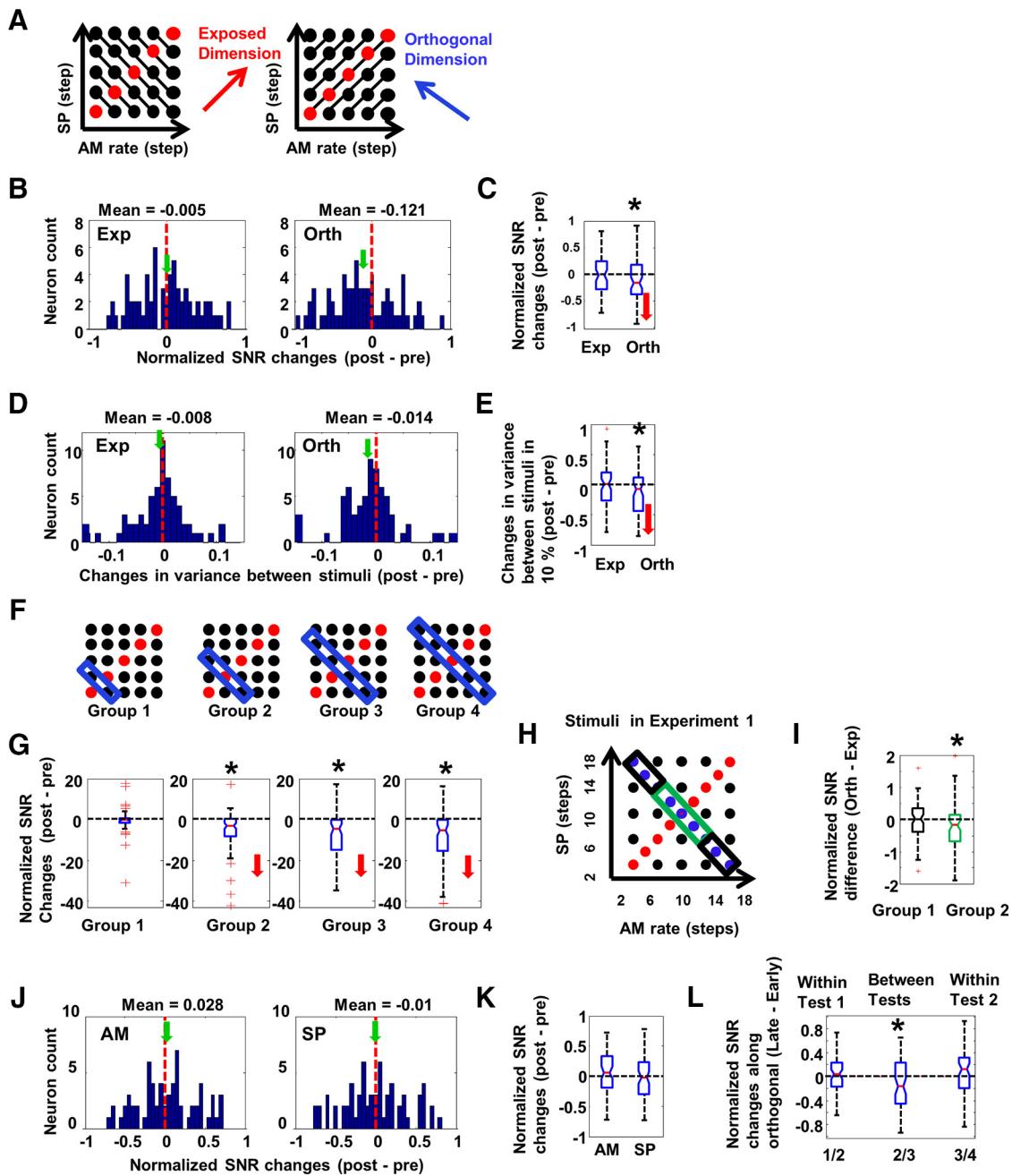
This result is consistent with previous findings by Dragoi et al. (2000), where adaptation was also shown to cause lateral shifts in neuronal tuning functions, thus significantly modifying them, as illustrated by the surface plot of response amplitudes to each test stimulus in Figure 2D. Note that the tuning map from recordings with negative correlated exposure was flipped so that the exposure dimension from two types of exposure (positive and negative) could be aligned and the results from the two types of

exposure could be pooled. Thus, our results showed that the exposure specifically modified the tuning properties of neurons, by suppressing responses to exposed stimuli and increasing responses to non-exposed stimuli.

#### Effect of exposure on the coding quality (SNR)

We also analyzed the effect of stimulus exposure on the SNR of each neuron. Early psychoacoustic studies (Stilp et al., 2010) had shown that, after passive exposure to sounds with correlated properties, the auditory system captured the covariance of the two acoustic attributes and treated them as a single perceptual dimension while simultaneously losing discriminability along the orthogonal dimension. Therefore, we hypothesized that this effect would be manifested at a neuronal level, and would cause a reduction in the SNR of spike rate coding along the orthogonal dimension after exposure to the correlated stimuli.

To test this hypothesis, we calculated the SNR for each neuron, with stimulus levels along the exposure dimension and the orthogonal dimension as two independent variables (Fig. 3A). Data from exposure to positively- and negatively-correlated properties were combined in this analysis. The results confirmed the prediction that the SNR on the exposure dimension remained unchanged (Wilcoxon test:  $z = -1.1$ ,  $p = 0.268$ ), whereas it significantly decreased along the orthogonal dimension (Wilcoxon test:  $z = -2.0$ ,  $p = 0.042$ ). Figure 3B illustrates this result with a histogram of the normalized SNR changes (divided by the sum of the pre- and post-SNRs). On the exposure dimension (left histogram), SNR changes were symmetrically distributed around zero, whereas along the orthogonal dimension they were significantly biased to the negative side (right histogram). This pattern is also demonstrated in Figure 3C, boxplots. Furthermore, as



**Figure 3.** **A**, Schematic illustration of the calculation of SNR along the exposed (left) and the orthogonal dimension (right). The  $5 \times 5$  matrix of dots represents the 33 stimuli tested. Each line connected with dots indicates stimuli to be treated as one single group (one level on the exposure or orthogonal dimension). The variance between stimuli was compared between lines (groups). The red arrow indicates the direction of the exposure dimension; blue arrow indicates the direction of the orthogonal dimension. **B**, Histograms of SNR changes in Experiment 1. SNRs for the exposure dimension were symmetrically distributed around zero (left, red dashed line indicates 0). SNRs for the orthogonal dimension were significantly reduced (right). The green arrow indicates the mean (also indicated on the top of histogram) for each distribution. The SNR change ( $\text{post} - \text{pre}$ ) was taken for each neuron and the difference was normalized by dividing by the mean of the two values. Normalization was only applied for display, while statistical analysis was done without normalization. **C**, Boxplot of SNR changes in Experiment 1. SNRs for the orthogonal dimension were significantly reduced (right). The asterisk on the top indicated that the distribution was significantly different from zero ( $*p = 0.042$ ). The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. The red arrow on the side indicates the direction of significant changes. **D**, Histogram of changes in variance (in spike rates) between stimuli in Experiment 1. Only the histogram on the right (for orthogonal stimuli) was significantly biased to the negative side, resembling the plot for SNR changes (above). **E**, Boxplot of changes in variance (in spike rates) between stimuli in Experiment 1. Variance between stimuli for orthogonal dimension was significantly reduced (right). The asterisk on the top indicated that the distribution was significantly different from zero ( $*p = 0.035$ ).

**F**, Schematic illustration of the subgroup selection, indicated by the blue box, for the tests of SNR changes in Experiment 1. The  $5 \times 5$  matrix of dots represents the 33 stimuli tested. The stimulus group on the left was located at the corner of the test space, whereas the stimulus group on the right was located on the diagonal line of the testing space.

**G**, Boxplot showed the effect of exposure on SNR changes on the orthogonal dimension in each subgroup illustrated on the top.  $*p = 0.001$  for all.

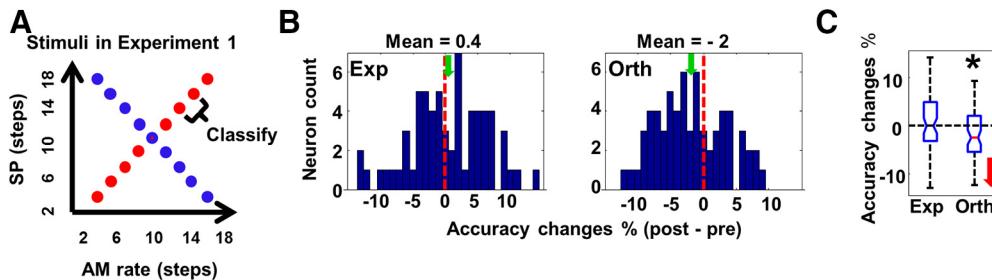
**H**, Schematic illustration of the subgroup selection along the orthogonal dimension, indicated by the black boxes for stimuli at the ends of the (orthogonal) diagonal and the green box for stimuli close to the center of the (orthogonal) diagonal.

**I**, Boxplot illustrates that only SNRs from stimuli close to the center of the diagonal exhibit significant differences.  $*p = 0.005$ .

**J**, Histograms of normalized SNR changes for AM (left) and SP (right) in Experiment 1. There was no significant change for either dimension. SNR changes were all symmetrically distributed around zero in both plots. The mean for each distribution is indicated both on the top of the plot and by the green arrow.

**K**, Boxplot of normalized SNR changes for AM (left) and SP (right) in Experiment 1. There was no significant change for either dimension.

**L**, Boxplot shows the dynamics of SNR changes along the orthogonal dimension. The Pre-exposure test and the Post-exposure were split into four test blocks. The comparison between Block 1 and Block 2 (left) showed consistency of SNRs obtained within the Pre-exposure test. The comparison between Block 2 and Block 3 (Middle) showed significant SNRs differences between the two test sessions due to the first exposure session. Blocks 3 and 4 were separated by the second exposure session (shorter exposure). Comparison between the two blocks (right) showed a nonsignificant trend of reduction of the effect of exposure.  $*p = 0.035$ .



**Figure 4.** **A**, Schematic illustration of the stimuli discriminated by a decoder trained with spike data in Experiment 1. Red dots indicate exposure stimuli; blue dots indicate orthogonal stimuli. **B**, Histograms of accuracy changes by the decoder trained with spike data in Experiment 1. The decoder was trained to discriminate stimuli along the exposed dimension or those along the orthogonal dimension. The red dashed line indicates zero. The green arrow indicates the mean. The accuracy of discrimination for stimuli along the orthogonal dimension reduced in 65% of neurons (right), whereas the distribution of accuracy changes for the exposed dimension was almost symmetrically centered at zero (left). **C**, The boxplot of accuracy changes in Experiment 1. Accuracy significantly reduced for orthogonal stimuli, but not for exposed stimuli. The star on the top indicates that the distribution was significantly different from zero. The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. The red arrow on the side indicates the direction of significant changes.  $*p = 0.008$ .

predicted, there were no significant SNR changes for the interaction between the exposure and orthogonal dimensions (Wilcoxon test:  $z = -0.1$ ,  $p = 0.909$ ). In summary, the pattern of neuronal adaptation is consistent with the findings of the psychoacoustic studies (Stilp et al., 2010).

Although these results clearly demonstrated SNR changes after stimulus exposure, they nevertheless raised a question about the origin of these SNR changes. Because SNR is a ratio of two variances (within/between stimuli), it was not immediately clear what exactly caused the SNR changes, i.e., which of the two variances dominated the overall change in SNR. By conducting a further analysis in which we examined the two variances separately, we found that the SNR changes in the orthogonal dimension were due to reduced variance between responses to stimuli along the diagonal (Fig. 3D, E; Wilcoxon test:  $z = -2.1$ ,  $p = 0.035$ ), whereas variance within each stimulus remained the same after exposure (Wilcoxon test:  $z = -1.2$ ,  $p = 0.243$ ).

Therefore, the SNR reduction was primarily due to a reduced signal representation on the orthogonal axis (i.e., diminished response differences between stimuli), rather than because of changes in noise level (i.e., the variance of responses within each stimulus). The size of this SNR decrease depended upon the number of stimuli distributed along the orthogonal dimension. Because stimuli located in the middle of the diagonal could be influenced by more stimuli along the orthogonal dimension than those located near the corners (Fig. 3F), we conjectured that stronger effects would be seen in the middle of the diagonal. To test this conjecture, we calculated and compared SNRs before and after exposure for four separate groups of stimuli, that were defined by their distance to the diagonal. We also averaged the SNR changes from stimuli that were symmetrically placed on either side of the diagonal. As predicted (Fig. 3G) the effects of exposure along the orthogonal dimension peaked at the diagonal (Wilcoxon test:  $z = -4.6$ ,  $p = 0.001$ ), becoming weaker toward the corners (Wilcoxon test:  $z = -1.8$ ,  $p = 0.078$ ).

In human psychoacoustic studies (Stilp and Kluender, 2016), the effect of exposure on discrimination also changed along the orthogonal dimension compared with discrimination along the exposed dimension. Therefore, we further calculated SNRs for two separate groups of stimuli along the orthogonal dimension that located at different distances from the center of the diagonal (Fig. 3H). SNR from stimuli that were symmetrically placed on either side of the center were averaged together. SNR from stimuli in the two groups were compared with the averaged SNRs from the exposure dimension, as in the comparison by Stilp and

Kluender (2016). Similar to earlier psychoacoustic results, the effects of exposure along the orthogonal dimension peaked for stimuli close to the center (Fig. 3I, green box; Wilcoxon test:  $z = -2.8$ ,  $p = 0.005$ ), becoming weaker toward the two ends of the diagonal (Fig. 3I, black boxes; Wilcoxon test:  $z = -0.15$ ,  $p = 0.878$ ), although we did not find a reversal at the end of the orthogonal diagonal as by Stilp and Kluender (2016).

In further analysis, we also measured the SNR changes separately for stimuli along the pure AM and SP single dimensions. The results revealed no significant changes along either axis (Fig. 3J, K; Wilcoxon test:  $z = -0.11$ ,  $p = 0.914$  for AM, and Wilcoxon test:  $z = -1.19$ ,  $p = 0.235$  for SP), with no significant changes in interactions between them (Wilcoxon test:  $z = -0.2$ ,  $p = 0.847$ ). This was also consistent with earlier psychophysical results (Stilp and Kluender, 2012, 2016).

Finally, we examined the dynamics of SNR changes by splitting the two test sessions into four blocks. SNRs were calculated in the first 10 trials and last 10 trials in the Pre-exposure test (Test 1) and Post-exposure test (Test 2). We did not find significant changes in SNR along the orthogonal dimension within Test 1 (Blocks 1 and 2; Fig. 3L, left boxplot; Wilcoxon test:  $z = -0.8$ ,  $p = 0.427$ ). Consistent with results in Figure 3C, SNRs along the orthogonal dimension decreased significantly between the two tests (Block 2 vs Block 3; Fig. 3L, middle boxplot; Wilcoxon test:  $z = -2.1$ ,  $p = 0.035$ ). Finally, we compared SNRs along the orthogonal dimension within Test 2 (Block 3 vs Block 4), which was separated by the second exposure that was only half the length of the first exposure. We found a nonsignificant trend of a reduced SNR in the last test block (Fig. 3L, right, boxplot; Wilcoxon test:  $z = 1.7$ ,  $p = 0.078$ ).

#### Performance of the decoder decreased for stimuli along the orthogonal dimension

We next examined whether there was any change in discriminability by the decoder trained with spike data to discriminate stimuli along the exposed dimension or those along the orthogonal dimension (see Materials and Methods). This decoder analysis allowed us to compare the effects of exposure in the neural data with the effects in previous experiments in human subjects (Stilp et al., 2010). We expected that the performance of the decoder would decrease for stimuli along the orthogonal dimension after exposure due to reduced SNR in spike rate coding. The decoder was trained to discriminate the adjacent stimuli along each dimension (Fig. 4A), mimicking the task of human subjects (Stilp et al., 2010).

We found that, consistent with the measurements of SNR described, the discriminability along the orthogonal dimension was significantly reduced. As shown in the histogram of Figure 4*B* (right), 65% of neurons showed decreased accuracy along the orthogonal axis (Wilcoxon test:  $z = -2.65, p = 0.008$ ). By contrast, as shown in the histogram of Figure 4*B* (left), only 46% of neurons exhibited decreased accuracy along the exposed dimension (Wilcoxon test:  $z = -0.70, p = 0.487$ ). Thus, these results from the decoder analysis are also consistent with the results of the psychoacoustic experiments with human subjects (Stilp et al., 2010).

#### Decorrelated tuning functions between neurons

Continuing our investigation of the neural effects of stimulus exposure in Experiment 1, we also examined the effects of stimulus exposure at the A1 neuronal population level. We reasoned that if two neurons were tuned to two different sound attributes (e.g., AM or SP), then if these two sound attributes became correlated, one neuron's responses would become predictable from the responses of the other, making one neuron's responses redundant. Efficient coding theory predicts that as stimulus parameters become correlated, tuning functions of different neurons should become less alike (or decorrelated) so as to reduce response correlations (and predictability) and increase coding efficiency (Barlow and Földiák, 1989). To test this conjecture, we calculated tuning functions along the AM and SP dimensions separately for each unit. We then paired simultaneously recorded neurons, and calculated the correlation coefficient between the AM tuning function in one neuron and the SP tuning in the other (Fig. 5*A*), both for the Pre-exposure and Post-exposure tests, and finally computed the difference in the correlations before and after exposure.

Because correlations estimated from different neuron pairs in each recording were not independent, we averaged changes in correlation coefficients from all simultaneously recorded neuron pairs for statistical analysis. We hypothesized that exposure to positively correlated attributes would lead to a decreased correlation coefficient, whereas exposure to negatively correlated attributes would lead to an increased correlation coefficient, because of decorrelation of initially negatively correlated tuning functions. The results of this analysis are plotted in a cumulative frequency distribution (Fig. 5*B*). After exposure to positively correlated attributes, the majority of recordings (8/11) exhibited reduced correlations between AM and SP (Fig. 5*B*, blue trace), whereas after exposure to negatively correlated attributes, 7/10 recordings exhibited positive correlation changes (Fig. 5*B*, red trace), indicating a significant contrast between the two groups (Mann–Whitney *U* test,  $p = 0.018$ ). Figure 5*C* summarizes the results in a boxplot. Note that the opposite direction of changes after exposure to negatively-correlated attributes and positively-correlated attributes actually reflects the same basic effect: tuning functions along the AM and SP dimensions in different neurons became decorrelated, or more dissimilar, after exposure. As a control, we also computed the correlations between the AM-tuning function and SP-tuning functions within the same neuron for positively- and negatively-correlated exposures (Fig. 5*D*). The results of this control analysis showed no significant difference in the correlations between the two types of exposure (Fig. 5*E, F*; Mann–Whitney *U* test,  $p = 0.674$ ).

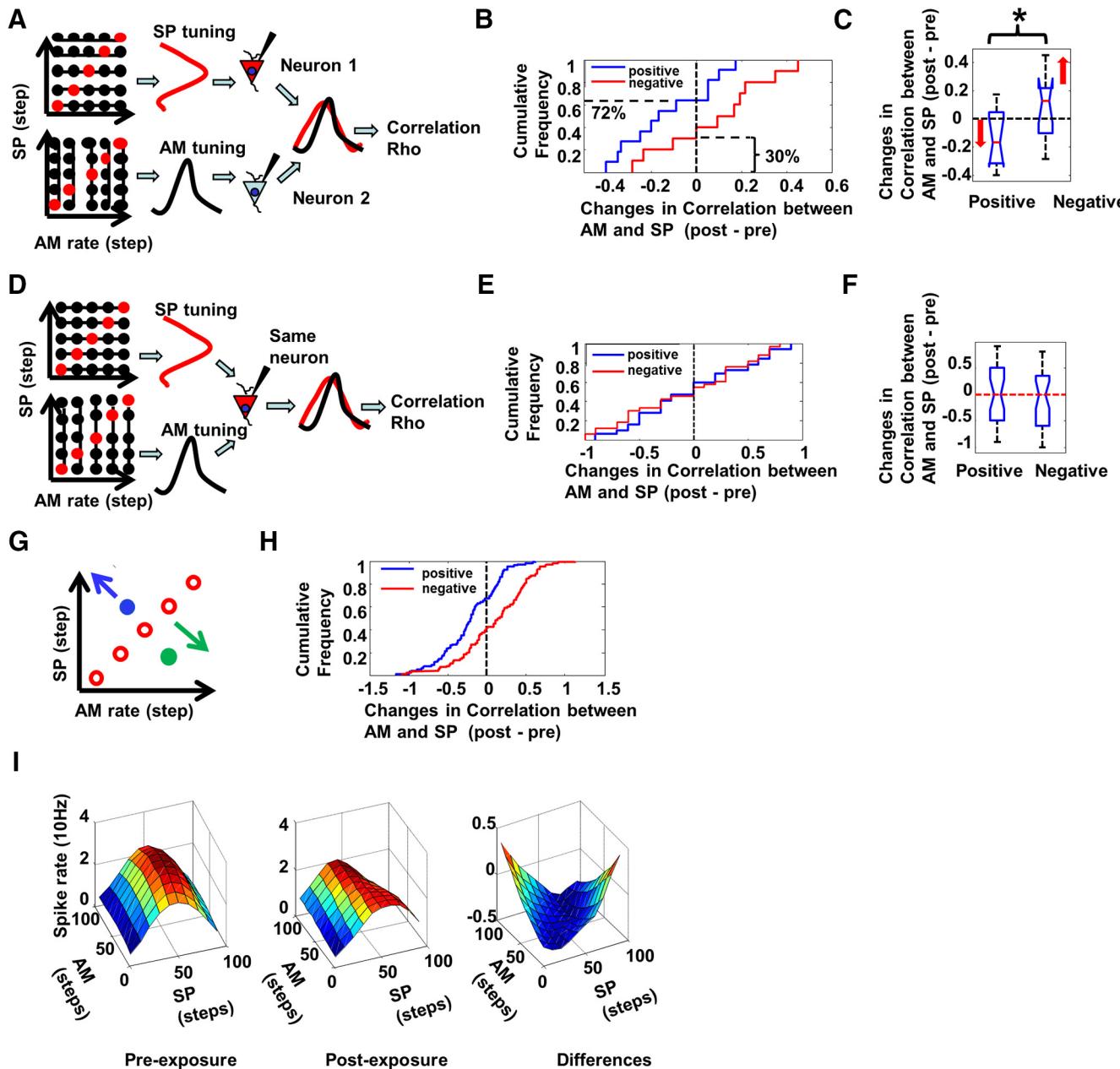
The neural response changes we described following stimulus exposure in Experiment 1 strongly support Barlow's hypothesis that neural tuning to correlated properties becomes decorrelated in a population of A1 neurons after sufficient exposure to the

covariance, even in the passively listening animal. A closely related phenomenon is that the tuning functions of the recorded neuron shifted away from the exposed stimuli so that, at the population level, responses to exposed stimuli decreased, whereas responses to stimuli far away from exposed stimuli increased (Fig. 2). To understand whether the two phenomena are related, we simulated the effect of tuning shifts on correlation between neurons in a computational model of a population of neurons ( $N = 200$ ) that are arbitrarily and uniformly tuned to a 2-dimensional of stimulus matrix (100 AM steps  $\times$  100 SP steps) similar to the test-stimuli matrix in our experiments. Details of a specific implementation are shown in Figure 5*G*. By assuming that the effect of exposure is to move each neuron's tuning function away from the parameters of the exposure stimulus (Fig. 5*G*, red dots) along the orthogonal direction (green and blue arrows), one can recreate the effects of both decorrelations (Fig. 5*H*) and tuning function shifts (Fig. 5*I*). Therefore, it is very likely that the tuning shift along the orthogonal dimension contributes to the decorrelation between neurons. As a conclusion, both our experimental results and the results of our computational model support the hypothesis that tuning to correlated properties becomes decorrelated between neurons after exposure to the covariant features.

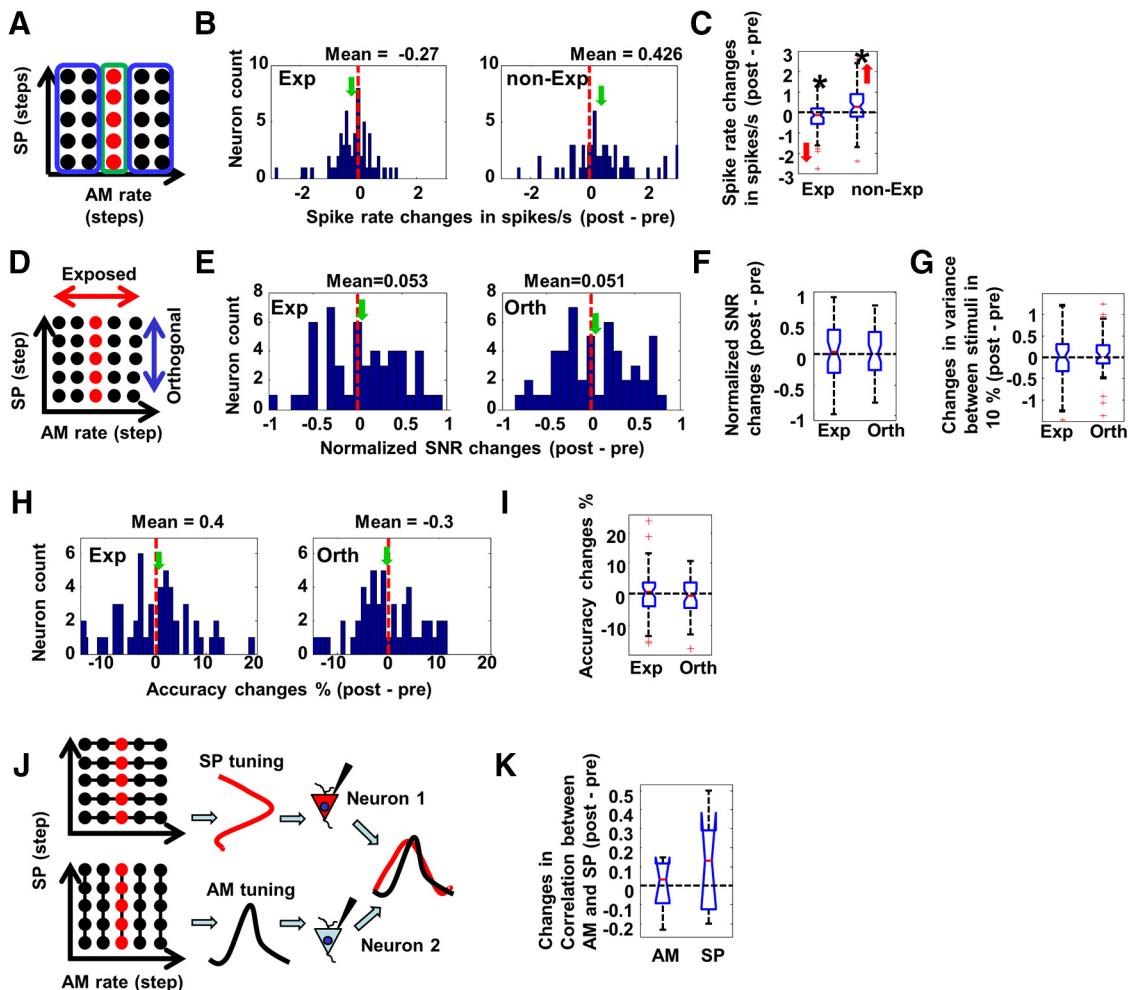
#### Experiment 2: the effects of exposure to stimuli on a single feature dimension

In this control experiment, we sought to determine whether exposure to a covariance of two properties was necessary to cause the adaptation patterns described in Experiment 1, and whether the effects observed in Experiment 1 might have also arisen in the absence of any attribute covariance, but instead simply as the result of adaptation to variance along a single acoustic dimension. To answer this question, we recorded from single-units in A1 ( $n = 56$ ) in two ferrets as they were exposed to stimuli varying along one dimension only, either AM or SP (Fig. 1*E*). We compared the neural responses from Experiment 2 with those from Experiment 1, focusing on a careful comparison of the four measurements (adaptation, SNR, decoder performance, and correlation in tuning functions between neurons) obtained in the two experiments. In the absence of stimulus covariance in the exposure stimuli presented in Experiment 2, the EC hypothesis suggested that the SNR changes along the exposure versus the orthogonal dimension, observed in Experiment 1, would not occur.

At first, we found that the adaptation and tuning effects resembled those of Experiment 1 in that responses to the exposure stimuli (Fig. 6*A*, red dots) were indeed significantly reduced (Wilcoxon test:  $z = -2.3, p = 0.022$ ) compared with the increased response to the non-exposed stimuli (Fig. 6*A*, blue dots; Wilcoxon test:  $z = -2.5, p = 0.015$ ), as illustrated by the histograms and boxplots in Figure 6, *B* and *C*. Thus, we conclude that exposure to sounds varying along one dimension did cause an adaptation effect similar to that observed in Experiment 1 (Fig. 2*B, C*). However, unlike Experiment 1, comparing the SNR changes along the exposure versus the orthogonal dimensions (Fig. 6*D*) revealed no significant changes in coding quality, as summarized by the histograms of Figure 6, *E* and *F* (Wilcoxon test:  $z = -0.1, p = 0.940$  for exposure responses; Wilcoxon test:  $z = -0.25, p = 0.800$  for the orthogonal dimension). There were also no significant changes in the variance between stimuli in either dimension (Fig. 6*G*; Wilcoxon test:  $z = 0.27, p = 0.786$  for the exposed dimension, and Wilcoxon test:  $z = 0.96, p = 0.966$  for the orthogonal dimension). Therefore, it is evident that the



**Figure 5.** *A*, Schematic illustration of the calculation of correlation between different neurons in Experiment 1. AM tuning and SP tuning were calculated from the responses to the 2-dimensional stimuli space. On the top, each line connected with dots indicated one level of SP to be averaged (across different AM levels) to create the tuning curve for SP (red bell shape curve). In this curve, the vertical axis indicates different SP levels. The horizontal axis indicates the amplitude of averaged responses. On the bottom, each line connected with dots indicated one level of AM to be averaged (across different SP levels) to create the tuning curve for AM (black bell shape curve). In this curve, the horizontal axis indicates different AM levels. The vertical axis indicates the amplitude of averaged responses. The two tuning curves were calculated from two different simultaneously recorded neurons: Neuron 1 (red) and Neuron 2 (blue). Then, correlation between the red and the blue tuning curve was computed. Finally, the difference in correlation coefficient obtained before and after exposure was derived. *B*, Cumulative frequency distribution (left) and the boxplot (right) of tuning correlation changes cross neurons in Experiment 1. The y-axis of the cumulative frequency plot shows the percentage of the total sample lower than (or equal to) an associate value on the x-axis. The vertical black dashed line indicates zero. The blue trace shows correlation changes after exposure to positively correlated attributes. The horizontal black dashed line on the left of the blue trace indicates that correlation changes were negative in 72% (8 of 11) recordings after exposure to positively correlated attributes. The red trace shows correlation changes after exposure to negatively correlated attributes. The horizontal black dashed line on the right indicates that only 30% (3 of 10) of red trace was on the negative side. *C*, Boxplot of tuning correlation changes across neurons in Experiment 1. The bracket and the asterisk indicate that the two distributions were significantly different from each other. The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. The red arrow on the side indicates the direction of significant changes.  $*p = 0.018$ . *D*, Schematic illustration of the tuning correlation calculation for the same neuron in Experiment 1. On the top, each line connected with dots indicated one level of SP to be averaged (across different AM levels) to create the tuning curve for SP (red bell shape curve). In this curve, the vertical axis indicates different SP levels. The horizontal axis indicates the amplitude of averaged responses. On the bottom, each line connected with dots indicated one level of AM to be averaged (across different SP levels) to create the tuning curve for AM (black bell shape curve). In this curve, the horizontal axis indicates different AM levels. The vertical axis indicates the amplitude of averaged responses. The two tuning curves were calculated from the same neuron. Then, correlation between the red and the blue tuning curve was computed. At last, difference in correlation coefficient obtained before and after exposure was taken. *E*, Cumulative frequency distribution (left) and the boxplot (right) of tuning correlation changes for the same neuron in Experiment 1: there was no significant change in tuning correlation either for positively correlated exposure (blue trace) or for negative correlated exposure (red trace). *F*, Boxplot of tuning correlation changes for the same neurons in Experiment 1. *G*, Illustration of hypothesis that exposure to correlated attributes pushes neuron's tuning function away from the exposure stimuli along the orthogonal direction. The blue and green dots illustrate the best tuning position of two hypothetical neurons in the stimulus matrix. The arrows indicate the direction of tuning changes. *H*, Results of modeling. Cumulative frequency distribution of tuning correlation changes in 200 simulated neurons. Consistent with results, correlation significantly decreased (Figure legend continues.)



**Figure 6.** *A*, Exemplars of stimuli used in tuning analysis in Experiment 2. The green rectangle indicates exposure stimuli, from which the spike rates were averaged and compared between tests before and after exposure. Two blue rectangles indicate non-exposure stimuli, from which the spike rates were averaged for comparison. *B*, Histograms of spike rate changes after exposure in Experiment 2. Spike rate significantly decreased for the exposure stimuli (left,  $p = 0.022$ ) and significantly increased for non-exposure stimuli (right,  $p = 0.015$ ), which resembled results in Experiment 1. The red dashed line indicates zero. The green arrow indicates the mean. *C*, Boxplot of spike rate changes after exposure in Experiment 2. Spike rate significantly decreased for the exposure stimuli (left) and significantly increased for non-exposure stimuli (right). The star on the top indicates that the distribution was significantly different from zero. The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. Red crosses indicate possible outliers. The red arrow on the side indicates the direction of significant changes. *D*, Schematic illustration of the stimuli tested in Experiment 2. The  $5 \times 5$  matrix of dots represents the 33 stimuli tested. Red dots indicated exposure stimuli that were only changed in a single dimension (AM or SP, balanced across recordings). Red arrows indicated the exposed dimension and blue arrows indicated the orthogonal dimension. *E*, Histograms of normalized SNR changes in Experiment 2. Both the distribution for stimuli along the exposure dimension (left) and the distribution for stimuli along the orthogonal dimension (right) were symmetrically around zero, without any significant bias. The means of the distributions are indicated both on the top of the plot and by the green arrow. *F*, The boxplot of normalized SNR changes in Experiment 2. The median of SNRs for both dimensions overlapped with zero. Thus, there was no change in SNRs in either dimension in Experiment 2. *G*, Boxplot of changes in variance (in spike rates) between stimuli in Experiment 2. *H*, Histogram of accuracy changes of the decoder trained with spike data in Experiment 2. The decoder was trained to discriminate stimuli along the exposed dimension or those along the orthogonal dimension. Contrary to the results of Experiment 1, there was no significant change in accuracy for either dimension in Experiment 2. *I*, The boxplot of accuracy changes in Experiment 2. *J*, Schematic illustration of the calculation of correlation of tuning from different neurons in Experiment 2. *K*, The boxplot of tuning correlation changes in Experiment 2: there was no significant change in tuning correlation for either of the two types of exposure.

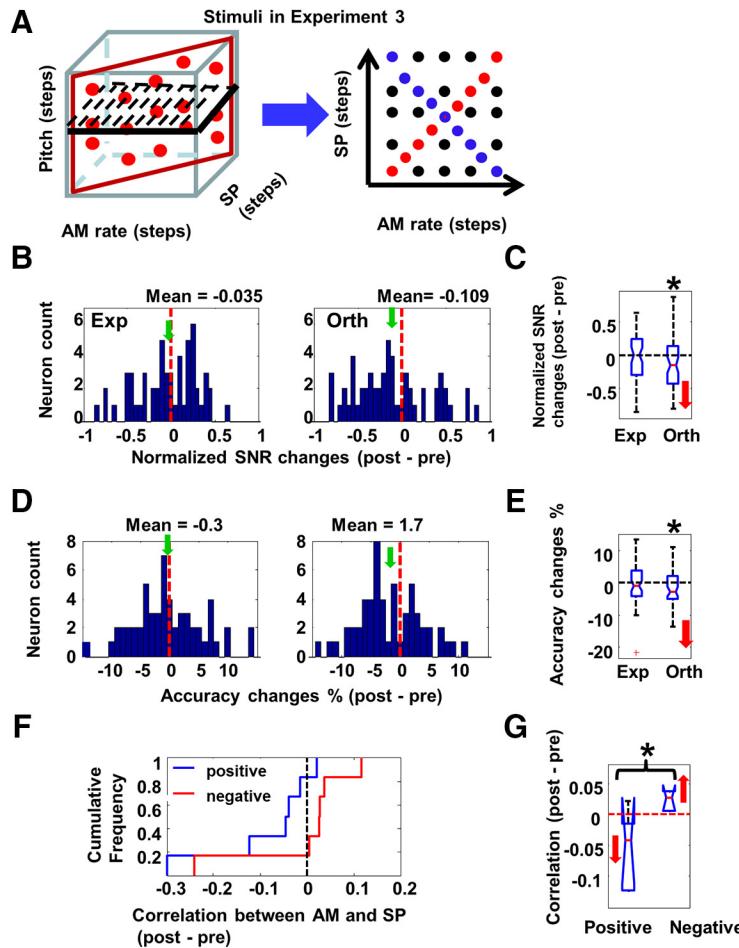
SNR effects observed in Experiment 1 were dependent on exposure to stimuli with correlated attributes.

To verify that the SNR results in Experiment 2 were consistent with discriminability of stimuli, we also measured the effects of single-dimension exposures on the performance of the decoder trained with neural data. As in Experiment 1, a decoder was

trained to discriminate stimuli along the exposed dimension or stimuli along the orthogonal dimension. In this condition, there was no significant change in decoder performance, either on the exposure dimension (Wilcoxon test:  $z = -0.15, p = 0.881$ ) or on the orthogonal dimension (Wilcoxon test:  $z = -0.3, p = 0.764$ ), as shown in Figure 6, *H* and *I*. Comparing the results of Experiments 1 and 2, we conclude that exposure to correlated stimuli is necessary to elicit changes in discriminability. Finally, we also analyzed changes in the correlation between the tuning curves in the neural responses in Experiment 2 (Fig. 6*J*). Data from AM and SP exposures were separately analyzed and compared, and no significantly different correlational changes were observed between the two types of exposure (Fig. 6*K*; Mann–Whitney *U* test,  $p = 0.268$ ).

←

(Figure legend continued.) after exposure to positively correlated attributes (red trace) and significantly increased after exposure to negatively correlated attributes (blue trace), resembling results shown in *B*. *I*, Results of modeling. Surface plot of the 2-dimensional tuning map obtained from 200 simulated neurons: before exposure (left) and after exposure (middle). The differences between the two maps were taken (post – pre) and are shown on the right. The overall pattern of change was consistent with our results in Figure 2*D*.



**Figure 7.** **A**, Testing stimuli space in Experiment 3. The red parallelogram indicates exposure sounds, which had perfect positive or negative correlation between AM and SP, and varied  $f_0$ s. The black parallelogram with dashed lines indicates the testing sounds, in which only one pitch was randomly chosen for testing in each recording. The structure of the testing space was the same as in Experiment 1, indicated on the right by the blue arrow. **B**, Histograms of SNR changes in Experiment 3. SNRs for exposure dimension were symmetrically distributed around zero (left, red dashed line indicates zero). Consistent with Experiment 1, SNRs for the orthogonal dimension were significantly reduced (right), resembling what was found in Experiment 1. The green arrow indicates the mean. **C**, Boxplot of normalized SNR changes in Experiment 3. The star on the top indicates that the distribution was significantly different from zero (\* $p = 0.020$ ). The red line in the middle and the notch indicated the median and its 95% confidence interval. The bottom and top edges of the box indicate the 25th and 75th percentiles. The whiskers indicate the range. Red crosses indicate possible outliers. The red arrow on the side indicates the direction of significant changes. **D**, Histogram of accuracy changes for the decoder trained with spike data in Experiment 3. The decoder was trained to discriminate stimuli along the exposed dimension or those along the orthogonal dimension. Consistent with Experiment 1, accuracy significantly decreased for orthogonal stimuli (right), but not for exposed stimuli (left). **E**, Boxplot of accuracy changes in Experiment 3. \* $p = 0.016$ . **F**, Cumulative frequency distribution of tuning correlation changes across neurons in Experiment 3: consistent with Experiment 1, the blue trace showed that correlation decreased (5 of 6) after exposure to positively correlated attributes. The red trace showed that correlation increased (5 of 6) after exposure to negatively correlated attributes. The contrast between the two distributions was statistically significant. **G**, Boxplot of changes in tuning correlation changes across neurons in Experiment 3. \* $p = 0.032$ .

In summary, although exposure to sounds varying along one dimension in Experiment 2 did lead to adaptation and tuning shifts as observed in Experiment 1, it did not lead to any significant changes in SNR, decoder discriminability performance, or neuronal correlation, as was elicited by exposure to stimuli with two covarying acoustic features.

**Experiment 3: coding efficiency with exposure to stimuli of correlated features embedded in a higher-dimensional space**  
 Here we extended Experiment 1 to explore whether the effects induced by exposure to stimuli of two correlated features would persist if the stimuli had an additional, but uncorrelated, third acoustic feature (e.g., the fundamental frequency of the tone-

complex,  $f_0$ ) that varied randomly. Our design of Experiment 3 parallels human psychophysical studies by Stilp and Kluender (2011, 2012) who also studied the effects of an uncorrelated third acoustic attribute. Our neuronal results in Experiment 3 are consistent with their psychophysical results, thus demonstrating efficient coding of redundant acoustic dimensions even in the presence of unrelated variability in a third feature dimension.

In Experiment 3, we recorded neuronal responses from A1 single units ( $n = 56$ ) in two ferrets that were exposed to stimuli located at the two diagonals of the testing space (Fig. 7A), axes along which we previously found the strongest adaptive effects. The results obtained strongly resembled those of Experiment 1: (1) SNRs decreased for the orthogonal dimension (Fig. 7B,C; Wilcoxon test:  $z = -2.3$ ,  $p = 0.02$ ), while remaining intact along the exposure dimension (Wilcoxon test:  $z = -0.4$ ,  $p = 0.677$ ). (2) Consistent with Experiment 1, the performance of the decoder decreased along the orthogonal dimension after stimulus exposure (Fig. 7D,E; Wilcoxon test:  $z = -2.40$ ,  $p = 0.016$ ), while remaining unchanged for the exposure dimension (Wilcoxon test:  $z = -0.49$ ,  $p = 0.625$ ). Therefore both SNRs and the performance of the decoder were consistently reduced along the orthogonal dimension as in Experiment 1.

Finally, in Experiment 3, correlation of tuning functions to AM and SP between neurons decreased in 5/6 recordings following exposure to positive covariance and increased (in 5/6 recordings) after exposure to negative covariance (Fig. 7F,G; Mann–Whitney  $U$  test,  $p = 0.032$ ). Combining data from Experiments 1 and 3, correlations decreased in 13/17 recordings after exposure to positive covariance (binomial test:  $p = 0.038$ ), whereas 12/16 recordings showed increased correlation after exposure to negative covariance (binomial test:  $p = 0.025$ ). We may conclude, in agreement with the results

of earlier psychophysical studies (Stilp and Kluender, 2011, 2012) that exposure to sounds with an additional randomly varying property like  $f_0$  did not affect the efficient coding of the correlated features, consistent with the results of Experiment 1. Thus, Experiment 3 demonstrates that the auditory system can capture covariant features even in the presence of a third, randomly variant feature dimension.

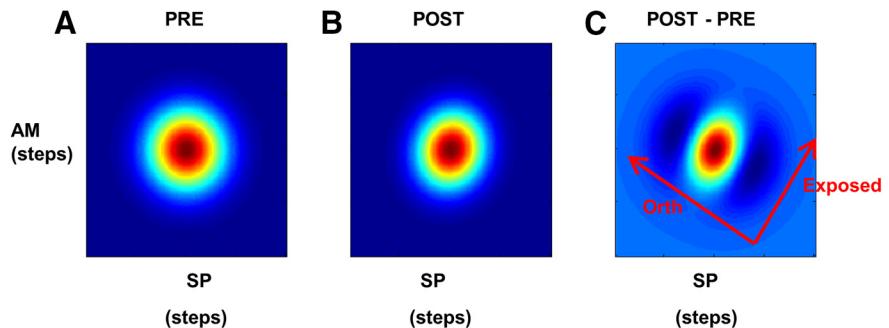
## Discussion

Experiments described here sought to explore the neural correlates of efficient coding, proposed as a key principle of coding in sensory systems. In vision, this principle is exemplified by the McCollough effect in which the visual system, following passive

exposure to correlated properties of visual stimuli (color and orientation), combines them as a single property (McCollough, 1965). The physical acoustics of natural sounds often reveal correlations between acoustic attributes (Lutfi et al., 2011). In auditory perception, the adaptive coding principle has also been supported by the results of a set of psychoacoustic experiments that were the inspiration for this study (Stilp et al., 2010; Stilp and Klunder, 2011, 2012, 2016).

The adaptive effects we observed in the neural recordings in the primary auditory cortex of the ferret were consistent with, and mapped readily to the results of the psychoacoustic experiments in human subjects showing that discrimination along the orthogonal dimension decreased due to reduced coding redundancy, whereas discrimination along the exposed dimension remained intact (Stilp et al., 2010). Furthermore, the changes allowed us to explore the underlying mechanisms that gave rise to this coding efficiency and test whether this neural adaptation was resilient to the addition of other uncorrelated feature dimensions. Our findings indicate that two aspects of the adapted responses contribute to the significant increase in coding efficiency after exposure to stimuli with the AM and SP correlated properties: (1) in single neurons, repeated exposure to stimuli of correlated attributes reduced SNR of responses to stimuli along the orthogonal dimension. This change was consistent with the decreased discrimination performance made by the decoder trained on neural data along the orthogonal dimension. (2) Tuning to the attributes of the AM and SP stimuli in different neurons became less correlated following exposure to stimuli with covariance, enhancing the efficiency of the population coding of stimulus identity. Consequently, as observed in Experiment 2, exposure effects did not materialize when stimulus parameters varied along only one dimension (AM or SP), although, adaptation (decrease) of spike rates still occurred to the exposure stimuli as before. Finally, we found in Experiment 3 that neurons were capable of extracting and enhancing the coding efficiency of two correlated properties even when they were embedded in a higher dimensional space that included independent variation in a third dimension.

Reduced SNR along the orthogonal dimension reflects an adaptive coding mechanism in single neurons in response to dynamically changed feature statistics. This mechanism can be modeled by changes in the receptive field (RF) of single neurons which are schematized as a 2-dimensional Gaussian probability density functions (PDF) as shown in Figure 8. The variance of the PDFs is estimated by the mean of the variances between stimuli along the exposure and orthogonal dimensions, which were calculated from the A1 neurons studied in Experiment 1. Because there were no significant changes in overall response amplitude, changes in PDF widths before and after exposure illustrate how such change may contribute to adaptive coding of a new combined feature that emerged during passive exposure. As shown in Figure 8, the estimated RFs after exposure became tilted along the exposed dimension (Fig. 8, middle). To examine this change further, we took the difference between the two PDFs (Fig. 8, right). The RF change reflects increased responsiveness that takes the shape of an ellipse along the exposure dimension, with reduced responsiveness along the orthogonal dimension. Such a change



**Figure 8.** **A**, The RF of neurons before exposure, modeled with 2-dimensional Gaussian probability density function. Variance between stimuli on the exposed dimension and the orthogonal dimension in Experiment 1 were used in this model. **B**, The RF of neurons after exposure, modeled with variance between stimuli from Experiment 1. **C**, Changes of RFs, by taking the difference between the probability density function before and after exposure. The coordinate of the exposed and orthogonal dimension is indicated by red arrows.

would optimize a neuron's capacity to capture the emergent feature along the exposure dimension at the cost of coding capacity along the orthogonal dimension. This is analogous to the principle component analysis, in which the coordinate of variance measurement is altered to capture the most important feature. This hypothesis is consistent with previous behavioral results and models based on the psychophysical data (Stilp et al., 2010; Stilp and Klunder, 2012, 2016). As schematized in Figure 5G, when the RF shapes change in single units, the tuning center of each neuron in the tested population also moves away from the exposed diagonal in the testing feature space, which in turn leads to an overall decrease in response amplitude to the exposed stimuli (Fig. 2D).

Apparently, enhancing a neuron's capacity to efficiently capture an emergent combined features along the covariant dimension comes at the cost of reducing coding capacity along the orthogonal dimension. Therefore, if the brain develops sparse coding cortical ensembles that are highly sensitive to covaried properties, this would come at the expense of coarser resolution for independent sensory features. An alternative compromise might be that there is a highly adaptive network of cortical neurons that reshape their responses to encode covaried stimulus attributes, and also different cell populations in other cortical networks that maintain sharp tuning for unidimensional sensory features. Consequently, separate networks would provide both a lower-resolution code for covaried sensory attributes and a high-resolution encoding for independent features.

Our results provide the first evidence at the neuronal level that sensory neurons in the primary auditory cortex may adapt to combine different sensory cues. However, it is also evident that not all neurons exhibited the change to capture covariance. Therefore, although we have emphasized efficient coding and covariant coding, it is possible and consistent with our observations that a subset of neurons may retain sensitivity to single cues, while another subset of neurons adapts to capture covariance. In this way, with multiple forms of adaptation, sensory systems could flexibly adopt different perceptual strategies in different situations. Further studies with high-density simultaneous recordings may reveal the presence of multiple, different encoding neuronal networks.

Although the decorrelation between neurons that we observed is consistent with the EC hypothesis (Barlow and Földiák, 1989), a recent study that develops a Bayesian theory of EC that goes beyond the traditional reliance of EC on the information theoretic approach, has questioned whether decorrelation con-

tributes to neural coding in all contexts (Park and Pillow, 2017). For example, by re-examining classic data from blowfly retinal neurons (Laughlin, 1981), it is shown that decorrelation may only be beneficial at high SNRs (Park and Pillow, 2017). Nevertheless, as we argued in the introduction, the changes we have observed in this study of the primary auditory cortex following short-term exposure to correlated features may be highly relevant to the sensory processing of natural stimuli because the covariance of multiple sensory attributes is key to enhancing the reliability of transmitting and perceiving the information in complex stimuli.

Previous studies have shown that sensory systems can combine different sensory cues to improve perception. For example, in vision, shape information from texture, and depth information from disparity cues, can be combined to form a fused percept, leading to improved discrimination performance compared with performance based on single cues (Hillis et al., 2002). Face recognition and detection algorithms use combinations of face parts to accomplish classification (Ullman et al., 2002) and a recent, compelling study of the neural basis for face recognition in the macaque monkey showed that the best model for decoding assumes that face patch neurons are linearly combining different features (Chang and Tsao, 2017). In haptic perception, force and position cues can be integrated for shape perception (Robles-De-La-Torre and Hayward, 2001; Drewing and Ernst, 2006). Furthermore, sensory cue integration may even occur across different sensory modalities (McGurk and MacDonald, 1976; Ernst et al., 2000; Shams et al., 2000).

Recent psychoacoustic studies also show that capturing correlation between different acoustic attributes may improve sound perception (Stilp et al., 2010; Stilp and Kluender, 2011, 2012, 2016). However, in auditory signal processing, the benefits of relying on correlated cues over single cues are even more prominent in vocal communication in animals and in speech (Kluender et al., 2019). A recent study has emphasized the value of sound categorization of marmoset vocalizations using combinations of acoustic features (Liu et al., 2019). Human speech is an even more complex signal characterized by many correlated attributes arising from the physical acoustic constraints of the vocal apparatus. Single acoustic attributes can be modified by the dynamics of articulating sequences of sounds (leading to co-articulation correlated context effects). In contrast, correlation between different attributes, such as power spectrum and manner of articulation (Llanos et al., 2017) and other covariant relationships (Sussman et al., 1998) may provide more reliable and robust information in the perception and acquisition of human speech. In such cases, tuning to the covariance, even at the cost of possible loss of resolution for any single attribute, would still be beneficial for processing speech.

We should note that the adaptive effects we have studied here are induced by short periods of passive exposure. Our previous work (Fritz et al., 2003, 2007) would suggest that the effects may become enhanced and magnified by behavioral training and/or task engagement in an auditory task that requires recognition of stimuli with covarying attributes. Thus, the ability to capture regularities through repetitive passive exposure may not only benefit the organism in sensory coding efficiency, but also function as a simple form of learning. There is considerable evidence that sensory systems can learn different types of regularities through such passive exposure: transition probability in sound sequences (Saffran et al., 1996, 1999; Hauser et al., 2001; Newport and Aslin, 2004; Newport et al., 2004; Abe and Watanabe 2011; Lu and Vicario, 2014) and other complex patterns (Agus et al., 2010;

McDermott et al., 2011; Barascud et al., 2016; Lu et al., 2018; Stilp et al., 2018). Our current study fits into this mold, and demonstrates the underlying neural transformations that make it possible. We note that our observations in this study are also interpretable in the context of implicit learning and habituation (Lu et al., 2018) in which exposed stimuli are contrasted with novel (unexposed or “orthogonal”) stimuli. However, thoroughly exploring this interpretation of our results will require a new study in which exposures and analyses are adapted to the parameters typical of statistical learning paradigms (Lu et al., 2018).

Our current work therefore demonstrates a form of implicit learning at the neuronal level in A1 that binds features from two correlated attributes, even when the pattern was embedded in high dimensional variance. This study therefore adds to the extensive previous research exploring brain areas and neural mechanisms underlying such learning, ranging from the oddball tone detection (Ulanovsky et al., 2003, 2004; Yaron et al., 2012; Nieto-Diego and Malmierca, 2016; Parras et al., 2017), to the encoding of transition probabilities in sound sequences (Lu and Vicario, 2014), to the mismatch negativity due to the detection of violations in acoustic sequences (Paavilainen et al., 2007; Barascud et al., 2014). Finally, implicit learning is widespread and not limited to the auditory system. For example, human subjects can implicitly learn visual transition probabilities in visual scene sequences (Turk-Browne et al., 2009), whereas mice can learn spatiotemporal sequences through exposure at the neuronal level in V1 (Gavornik and Bear, 2014). It is therefore highly likely that a larger network, including higher auditory and other sensory areas, contribute to these encoding and learning phenomena, and the future challenge is to unravel the interlocking roles of these different brains.

## References

- Abe K, Watanabe D (2011) Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat Neurosci* 14:1067–1074.
- Agus TR, Thorpe SJ, Pressnitzer D (2010) Rapid formation of robust auditory memories: insights from noise. *Neuron* 66:610–618.
- Attnave F (1954) Some informational aspects of visual perception. *Psychol Rev* 61:183–193.
- Barascud N, Griffiths TD, McAlpine D, Chait M (2014) “Change deafness” arising from inter-feature masking within a single auditory object. *J Cogn Neurosci* 26:514–528.
- Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M (2016) Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc Natl Acad Sci U S A* 113:E616–E625.
- Barlow H (1961) “Possible principles underlying the transformation of sensory messages” in sensory communication. Cambridge, MA: MIT.
- Barlow HB, Földiák P (1989) Adaptation and decorrelation in the cortex. In: *The computing neurone* (Durbin R, Miall C, Mitchison G, eds), pp 54–72. Wokingham, England: Addison-Wesley.
- Chang L, Tsao DY (2017) The code for facial identity in the primate brain. *Cell* 169:1013–1028.e14.
- Coen-Cagli R, Kohn A, Schwartz O (2015) Flexible gating of contextual influences in natural vision. *Nat Neurosci* 18:1648–1655.
- Dragoi V, Sharma J, Sur M (2000) Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron* 28:287–298.
- Drewing K, Ernst MO (2006) Integration of force and position cues for shape perception through active touch. *Brain Res* 1078:92–100.
- Ernst MO, Banks MS, Bülthoff HH (2000) Touch can change visual slant perception. *Nat Neurosci* 3:69–73.
- Francis NA, Elgueta D, Englitz B, Fritz JB, Shamma SA (2018) Laminar profile of task-related plasticity in ferret primary auditory cortex. *Sci Rep* 8:16375.
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223.

Fritz JB, Elhilali M, Shamma SA (2007) Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J Neurophysiol* 98:2337–2346.

Gavornik JP, Bear MF (2014) Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nat Neurosci* 17:732–737.

Geman S, Hwang C (1982) Nonparametric maximum likelihood estimation by the method of sieves. *Ann Stat* 10:401–414.

Hauser MD, Newport EL, Aslin RN (2001) Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78:B53–B64.

Hillis JM, Ernst MO, Banks MS, Landy MS (2002) Combining sensory information: mandatory fusion within, but not between, senses. *Science* 298:1627–1630.

Kluender KR, Lotto AJ (1999) Virtues and perils of empiricist approaches to speech perception. *J Acoust Soc Am* 105:503–511.

Kluender KR, Kiefe M (2006) Speech perception within a biologically-realistic information theoretic framework. In: *Handbook of psycholinguistics* (Gernsbacher MA, Traxler M, eds), pp 153–199. London: Elsevier.

Kluender KR, Alexander JM (2007) Perception of speech sounds. In: *The senses: a comprehensive reference*, Vol 3 (Dallos P, Oertel D, eds), pp 829–860. San Diego: Academic.

Kluender KR, Stilp CE, Kiefe M (2013) Perception of vowel sounds within a biologically-realistic model of efficient coding. In: *Vowel inherent spectral change* (Morrison G, Assmann PF, eds), pp 117–151. Berlin: Springer.

Kluender KR, Stilp CE, Lucas FL (2019) Long-standing problems in speech perception dissolve within an information-theoretic perspective. *Atten Percept Psychophys* 81:861–883.

Laughlin S (1981) A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch C* 36:910–912.

Lewicki MS (2002) Efficient coding of natural sounds. *Nat Neurosci* 5:356–363.

Liu ST, Montes-Lourido P, Wang X, Sadagopan S (2019) Optimal features for auditory categorization. *Nat Commun* 10:1302–1314.

Llanos F, Alexander JM, Stilp CE, Kluender KR (2017) Power spectral entropy as an information-theoretic correlate of manner of articulation in American English. *J Acoust Soc Am* 141:EL127.

Lu K, Vicario DS (2014) Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. *Proc Natl Acad Sci U S A* 111:14553–14558.

Lu K, Xu Y, Yin P, Oxenham AJ, Fritz JB, Shamma SA (2017) Temporal coherence structure rapidly shapes neuronal interactions. *Nat Commun* 8:13900.

Lu K, Liu W, Zan P, David SV, Fritz JB, Shamma SA (2018) Implicit memory for complex sounds in higher auditory cortex of the ferret. *J Neurosci* 38:9955–9966.

Lutfi RA, Liu CJ, Stoelinga CN (2011) Auditory discrimination of force of impact. *J Acoust Soc Am* 129:2104–2111.

McCollough C (1965) Color adaptation of edge-detectors in the human visual system. *Science* 149:1115–1116.

McDermott JH, Wroblewski D, Oxenham AJ (2011) Recovering sound sources from embedded repetition. *Proc Natl Acad Sci U S A* 108: 1188–1193.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. *Nature* 264:746–748.

Newport EL, Aslin RN (2004) Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cogn Psychol* 48:127–162.

Newport EL, Hauser MD, Spaepen G, Aslin RN (2004) Learning at a distance: II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn Psychol* 49:85–117.

Nieto-Diego J, Malmierca MS (2016) Topographic distribution of stimulus-specific adaptation across auditory cortical fields in the anesthetized rat. *PLoS Biol* 14:e1002397.

Olshausen BA, Field DJ (1997) Sparse coding with an overcomplete data set: a strategy employed by V1. *Vis Res* 37:3311–3325.

Paavilainen P, Arajarvi P, Takegata R (2007) Preattentive detection of nonsalient contingencies between auditory features. *Neuroreport* 18: 159–163.

Park IM, Pillow JW (2017) Bayesian efficient coding. *bioRxiv*. Advance online publication. Retrieved August 25, 2017. doi: 10.1101/178418.

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

Privitera GJ (2017) *Essential statistics for the behavioral sciences*. Los Angeles: Sage.

Robles-De-La-Torre G, Hayward V (2001) Force can overcome object geometry in the perception of shape through active touch. *Nature* 412:445–448.

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.

Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70:27–52.

Shams L, Kamitani Y, Shimojo S (2000) Illusions: what you see is what you hear. *Nature* 408:788.

Smith EC, Lewicki MS (2006) Efficient auditory coding. *Nature* 439: 978–982.

Stilp CE, Kluender KR (2011) Non-isomorphism in efficient coding of complex sound properties. *J Acoust Soc Am* 130:EL352.

Stilp CE, Kluender KR (2012) Efficient coding and statistically optimal weighting of covariance among acoustic attributes in novel sounds. *PLoS One* 7:e30845.

Stilp CE, Kluender KR (2016) Stimulus statistics change sounds from near-indiscriminable to hyperdiscriminable. *PLoS One* 11: e0161001.

Stilp CE, Rogers TT, Kluender KR (2010) Rapid efficient coding of correlated complex acoustic properties. *Proc Natl Acad Sci U S A* 107: 21914–21919.

Stilp CE, Kiefe M, Kluender KR (2018) Discovering acoustic structure of novel sounds with varying predictability. *J Acoust Soc Am* 143:2460–2473.

Sussman HM, Fruchter D, Hilbert J, Sirosh J (1998) Linear correlates in the speech signal: the orderly output constraint. *Behav Brain Sci* 21:241–259; discussion 260–299.

Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK (2009) Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J Cogn Neurosci* 21:1934–1945.

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.

Ullman S, Vidal-Naquet M, Sali E (2002) Visual features of intermediate complexity and their use in classification. *Nat Neurosci* 5:682–687.

Vinje WE, Gallant JL (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287:1273–1276.

Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76:603–615.