

# UCSF

## UC San Francisco Previously Published Works

### Title

Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment

### Permalink

<https://escholarship.org/uc/item/5278j8gj>

### Authors

Bao, S  
Chang, EF  
Teng, C-L  
[et al.](#)

### Publication Date

2013-09-01

### DOI

10.1016/j.neuroscience.2013.05.056

Peer reviewed

Published in final edited form as:

*Neuroscience*. 2013 September 17; 0: 30–42. doi:10.1016/j.neuroscience.2013.05.056.

## Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment

Shaowen Bao<sup>1</sup>, Edward F. Chang, Ching-Ling Teng, Marc A. Heiser, and Michael M. Merzenich

Keck Center for Integrative Neuroscience, University of California, San Francisco, CA 94143

### Abstract

Cortical sensory representations can be reorganized by sensory exposure in an epoch of early development. The adaptive role of this type of plasticity for natural sounds in sensory development is, however, unclear. We have reared rats in a naturalistic, complex acoustic environment and examined their auditory representations. We found that cortical neurons became more selective to spectrotemporal features in the experienced sounds. At the neuronal population level, more neurons were involved in representing the whole set of complex sounds, but fewer neurons actually responded to each individual sound, but with greater magnitudes. A comparison of population-temporal responses to the experienced complex sounds revealed that cortical responses to different renderings of the same song motif were more similar, indicating that the cortical neurons became less sensitive to natural acoustic variations associated with stimulus context and sound renderings. By contrast, cortical responses to sounds of different motifs became more distinctive, suggesting that cortical neurons were tuned to the defining features of the experienced sounds. These effects lead to emergent “categorical” representations of the experienced sounds, which presumably facilitate their recognition.

### Keywords

natural sound; spectrotemporal receptive field; neurogram; sparse coding; unsupervised learning

## INTRODUCTION

The cerebral cortex is highly adaptive to its sensory environment both during development and in adulthood. During an epoch of early life known as the “critical period”, cortical sound representations can be profoundly shaped by the acoustic environment (Zhang et al., 2001, Insanally et al., 2009). Previous studies, using tone pips or noise bursts, demonstrated that the specific effects of the early environment on auditory neurons depend on the

© 2013 IBRO. Published by Elsevier Ltd. All rights reserved.

Correspondence to: Shaowen Bao, Ph.D. Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720-3190, sbao@berkeley.edu, Phone: (510) 643-3549, Fax: (510) 643-4966.

<sup>1</sup>Current address: Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720-3190;

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

spectrotemporal properties of the sensory stimuli (Zhang et al., 2002; Insanally et al., 2010; Chang and Merzenich, 2003, de Villers-Sidani et al., 2008; Kim and Bao, 2009; Nakahara et al., 2004). Notably, auditory cortex over-represents experienced sounds. In stark contrast to exposure to tone pips and noise bursts in experimental conditions, humans and animals in normal acoustic environment experience natural and complex sounds that are highly structured during sensory development. Such enriched experience is necessary for the proper development of language/vocalization and other cognitive skills. In this report, we investigate how a rich and structurally complex auditory environment reorganizes acoustic representations by auditory cortical neurons, and its potential impact for categorical sound perception.

Auditory stimuli in the natural environment are rich, complex and highly structured. Animal vocalizations are an important class of natural sounds. Many animal species show stronger neural responses and better perceptual sensitivity to vocal sounds of their own species (Margoliash, 1983, Doupe and Konishi, 1991, Esser et al., 1997). For instance, neurons in the primary auditory cortex (AI) of marmoset monkeys responded more strongly to their conspecific calls than their time-reversed version (Wang and Kadia, 2001). In cats, both the normal and the reversed marmoset calls activated AI neurons equally weakly (Wang and Kadia, 2001). However, in most plasticity studies, experimental animals were given extensive experience of a single, often simple stimulus, such as a single-frequency tone (Zhang et al., 2001, Han et al., 2007). Those experimental stimuli were impoverished in terms of variability, complexity and structures, which are presumably important for normal sensory development. Thus, it is unclear how those results are applicable in understanding normal sensory development. For instance, it has been shown that animals that had experienced single-frequency tone pips exhibit enlarged cortical representations of the experienced sounds (Zhang et al., 2001), which presumably leads to refined cortical sensory processing (Recanzone et al., 1993, Godde et al., 2000, Bao et al., 2004, Polley et al., 2004a, Tegenthoff et al., 2005). However, recent studies failed to correlate enlarged cortical representations with refined perceptual performance (Talwar and Gerstein, 2001, Brown et al., 2004, Han et al., 2007, Reed et al., 2011), questioning the implication and generality of findings in animals with impoverished sensory experience.

Animal vocalizations exhibit complex statistical structure (Holy and Guo, 2005; Takahashi et al.). For example, rodent pup and adult calls are repeated in bouts with repetition rates in a range near 6 Hz (Liu et al., 2003; Holy and Guo, 2005; Kim and Bao, 2009). Calls of the same type are more likely to occur in the same sequences than do calls of different types (Holy and Guo, 2005). This higher-order statistical structure may provide information for classifying the calls into distinct, categorically perceived groups of sounds (Ehret and Haack, 1981). Indeed, a recent study showed that auditory cortex was shaped to represent sounds in the same sequences more similarly, and represent sounds that never occurred in the same sequences more distinctively, forming a representational boundary (Köver et al., 2013). In addition, perceptual sensitivity was also elevated across the representational boundary, establishing a perceptual boundary (Köver et al., 2013). Such representational and perceptual changes could contribute to categorical perception of the experienced sounds.

Categorical perception of conspecific animal vocalizations has been reported in a wide range of non-human animals (May et al., 1989, Nelson and Marler, 1989, Wyttenbach et al., 1996, Baugh et al., 2008), including rodents (Ehret and Haack, 1981). Although the existence of language-specific differences in speech sound perception (Fox et al., 1995, Iverson and Kuhl, 1996) has long pointed to a role for early experience in shaping human speech perception, the contribution of early experience to categorical perception in non-human animals is still unclear (Kuhl and Miller, 1978, Mesgarani et al., 2008, Prather et al., 2009). In the present study, we examine effects of early experience of a set of animal vocalizations recorded in a natural environment—referred to here as “jungle sounds” —on primary cortical acoustic representations. Our results reveal emergent categorical representation of the experienced heterospecific vocalizations, suggesting a role of early experience in categorical perception of natural complex sounds.

## MATERIAL AND METHODS

### Animals and experimental design

All procedures were approved under University of California San Francisco Animal Care Facility protocols. Four groups of rat pups were used in this study, including three sound exposure groups and one naïve control group. The naïve control group was housed in regular animal room. The three sound exposure groups were exposed to natural animal vocalizations, noises with separately modulated power spectrum and envelop, or pitch-shifted human speech sounds. After sound exposure, responses to natural animal vocalizations were recorded from AI neurons to determine how experience shapes cortical representation of complex sound. The noise group was used to determine effects of simple sound features such as separate spectral and temporal modulations. Speech group was used to determine if effects of complex sounds were specific to the stimuli. Naïve control animals were also mapped at the same age.

### Sound stimuli

A set of natural animal vocal sounds (referred to hereafter as “jungle sounds”, 1-h long, continuously recorded from a natural forest at a sampling rate of 44.2 kHz) was used as the experimental stimuli. In this sound record, there were at least 40 repeated vocalization or communication sounds generated by different mammalian, avian or insect species.

We filtered the jungle sound stimuli with 109 Gaussian filters with center frequency separation and bandwidth (SD) of the filters equal 0.05 octave. Analytical signal decomposition was performed on the signal in each frequency band to obtain its instantaneous amplitude. To construct the spectrogram, the instantaneous amplitude was plotted on a logarithmic scale in frequency and time. Two-dimensional Fourier transforms of the stimulus spectrograms were computed to obtain spectrotemporal modulation spectra of the stimuli. Noises with overall frequency spectra and temporal envelope modulation similar to those of the jungle sounds were used as control stimuli. To make these noise stimuli, white noise was first filtered with the 109 Gaussian filters. The amplitude of the signal in each frequency band was adjusted to match the mean amplitude of the jungle sound stimuli in that frequency band. Signals were then added together and amplitude-modulated to match

the overall amplitude of the jungle sound stimuli. Speech stimuli used for acoustic rearing were 5 min of sentences of Western American English naturally spoken by different male and female speakers, which was obtained from Linguistic Data Consortium at University of Pennsylvania. To match the rat's hearing, spectra of the speech stimuli were shifted 1.46 octaves higher.

### Acoustic exposure

Litters of rat pups were placed (with their mothers) in a sound-shielded chamber with a 12/12-h light/dark cycle from postnatal day 13 (P13) to P30. The jungle sounds, noises and speech stimuli were continuously played. The sound pressure level was rarely above 65 dB and never higher than 80 dB. After sound exposure, the rats were placed back in the animal room.

### Physiological recording

Under anesthesia (ketamine and xylazine), frequency-intensity receptive fields were recorded from multiunits sampled across the right temporal cortex to outline AI (Bao et al., 2001). Multiunits were then recorded from layer 4/5 in AI. Auditory stimuli were delivered through a calibrated earphone to the right ear. Frequency/intensity response areas were reconstructed in detail by presenting 60 pure-tone frequencies (0.5–30 kHz, 25-ms duration, 5-ms ramps) at each of 8 sound intensities. Responses to jungle sound testing stimuli were recorded from these units for 20–30 repetitions.

### Simple response properties

Response properties and receptive field properties were manually determined using custom-made programs. Briefly, frequency-intensity receptive fields were smoothed with a  $3 \times 3$  median-filter (a Matlab function). The tuning curve was defined as the outline of the response area. CF was the frequency that activated the unit at lowest intensity. The lowest intensity of tones that activate the unit is the threshold. Tuning bandwidths were measured at 10 dB about the threshold intensity (Bao et al., 2003).

### STRF analysis

STRFs were derived from spike trains activated by 8 seconds of jungle sound stimuli, using a software package provided by Dr. Frédéric E. Theunissen (STRFPAK-0.1, Theunissen Lab and Gallant Lab, UC Berkeley). Details of the method has been described elsewhere (Theunissen et al., 2000, Theunissen et al., 2001). The noise tolerance was first set at 13 levels (0.5 to 0.0001 with 1 octave spacing). The goodness of fit of the derived STRFs was quantified with cross-correlation between the estimated and actual firing rates of the unit to a second set (8 s) of jungle sounds from the same song ensemble. The tolerance level that resulted in the highest cross-correlation was chosen and the corresponding STRF was selected for that unit. The excitatory and inhibitory areas of the STRF were defined as the areas with values that were 4 SD above or below zero. Temporal duration and frequency bandwidth were calculated accordingly. Singular value decomposition of the STRF was performed to obtain its singular values  $\lambda_i$ , with  $i=1, 2 \dots n$ , and  $\lambda_1 > \lambda_2 > \dots > \lambda_n$ . Spectrotemporal separability of the STRF was defined as  $(\lambda_1)^2 / \sum (\lambda_i)^2$ .

## Population response analysis

Spike trains activated by testing jungle sound stimuli were represented as mean firing rates in 5-ms bins. Spike trains from all units in each group were temporally aligned to form a neurogram (population-temporal responses). A significant response was scored if the mean firing rate of a unit in a bin was 2 standard deviations above the mean baseline firing rate of that unit. The percentage of cortical units responding to stimuli was calculated for each bin and averaged across all bins. The magnitudes of responses (e.g., number of spikes in the bins) were normalized to the standard deviation of the baseline firing rate and averaged over time for a unit only for the bins in which a response occurred, and then averaged across all units. To determine whether sound representations were mediated mainly by a small fraction of the neuronal population, firing rate variances were examined (Willmore et al., 2000, Weliky et al., 2003). The firing rate variances of all units in each group were normalized to their maximum value of the group, rank-ordered and presented in a *scree* plot. To facilitate comparison between groups of different number of units, the unit numbers were normalized to a maximum of 1. A greater mean of the normalized variances would indicate that they were more evenly distributed across the population.

A population-temporal response to a song motif was defined as a section of the neurogram that was 50-ms long (10 bins). Our analysis indicated that a 50-ms window was optimal for sound discrimination by neurogram responses under our experimental conditions. Differences between population-temporal responses to two sounds were estimated using Bray-Curtis dissimilarity (Quinn and Keough, 2002), which was defined as

$$d_{xy} = \sum |X_i - Y_i| / \sum (X_i + Y_i)$$

where  $X_i$  and  $Y_i$  were values of  $i^{\text{th}}$  pixels of the population-temporal responses to sounds X and Y. To illustrate how responses to different sounds were segregated, multidimensional scaling (Rolls and Tovee, 1995, Ohl et al., 2001) was performed to map response dissimilarities to distances in lower-dimensional space. The first two dimensions typically accounted for 70–80% of the variance and were used to graphically represent similarity relations between the responses. The stress of the dimensionality reduction was 13%–19%. A two-step statistical analysis was performed to determine if the two-dimensional distributions of the representations of two song motifs were significantly segregated. First, we determined whether the distribution pattern of the combined populations of representations of sounds was significantly different from two-dimensional random patterns. Distances between pairs of sound representations in the two-dimensional plots were calculated, and a test statistic was computed as the variance of the distances divided by the square of the mean distance. The distributions of the test statistic were obtained from 5000 two-dimensional uniform or Gaussian random patterns (with the same spreads in the two principal dimensions as the actual pattern) using Monte-Carlo methods. The significance levels with which the actual pattern differed from random patterns (uniform or Gaussian) were then estimated. In the second step, we determined whether the two populations of representations of sounds were significantly different in the two dimensions using

multivariate analysis of variance test. The two sets of representations were considered significantly segregated if the differences were significant ( $p < 0.05$ ) for tests in both steps.

Unless stated otherwise, statistical significance was assessed using Kruskal-Wallis tests and post-hoc Bonferroni's multiple-comparisons. Data are presented as mean  $\pm$  SEM.

## RESULTS

In this study, we exposed rat pups to a set of natural animal vocal sounds, referred to here as “jungle sounds”, twenty-four hours a day from P13 to P30, covering a large portion of the critical period for complex sound selectivity (Insanally et al., 2009). The jungle sounds CD loop was an hour of spectrotemporally complex sounds, in which there were at least 40 distinctive repeating motifs of bird songs, mammalian vocalizations and insect sounds. A group of control rats were exposed to synthetic modulated noise. The overall temporal modulation and the spectral composition of the noise were designed to be similar to those of the jungle sounds (Figure 1), but with no frequency-modulated features. This modulated noise served as a control for effects induced by simple features of separate spectral and temporal modulations to reveal the influences of complex features of joint spectrotemporal modulations in the jungle sounds. Following sound exposure, we recorded multiunit AI responses to 16-s long composite jungle sound stimuli with 20 distinctive complex sounds (Figure 2).

### Characterization of the jungle sounds

Many of the sounds were repeated renderings of the same song motif, as evidenced by similar acoustic structures and regular repeating intervals. It is important to understand how well these sound motifs can be recognized based on their acoustic features. Visually examined, these sounds appear to be from five categories of song motifs, with sounds in each category showing substantial variability. We calculated correlation coefficients (CCs) for every pair of sounds to quantify acoustic similarity between these sounds. CCs are less sensitive to changes in overall sound levels (Figure 2B). CCs between sounds of the same song motifs (within-motif) and CCs between sounds of different motifs (between-motif) were highly overlapping due to at least three factors: 1) variations between repeats of the same song motifs; 2) variations in environmental sounds; 3) relatively higher power distribution in high than in low frequency ranges. The first two factors led to lower within-motif CCs, and the third factor resulted in higher between-motif CCs.

We visualized the similarity relations (i.e., CCs) between the sounds, using a multidimensional scaling method, which reduces the dimensionality of the data and allows for two-dimensional illustration of the similarity relations. All sounds were not significantly segregated because of the acoustic variability associated with different renderings of the song motifs (Figure, 2C,  $p > 0.05$ , see Methods for statistical tests), indicating that the sounds cannot be categorized based on the complete spectrograms. These results are consistent with the highly overlapping within-motif and between-motif CCs (Figure 2B).

We further subjected the spectrograms of the sounds to temporal band-pass (20–80 Hz, with 5 dB reduction at 20 and 80 Hz) filtering to enhance distinctive features of the sounds'

spectrotemporal “edges”. After the temporal filtering, the sounds of the same motifs showed highly similar characteristic features, and sounds of different motifs revealed more distinct acoustic structures (Figure 2D). The overlap between the distribution of within-motif CCs and that of between-motif CCs was much less than in Figure 2B. Similarity relations between the filtered spectrographic features of the sounds showed significantly segregated sounds by motif identities (Figure 2F,  $p < 0.05$ ). These results indicate that the testing sounds can be objectively separated into five song categories based on their distinctive dynamic features, providing a basis for later analysis of stimulus categorization. In addition, these results also suggest that, to recognize the song motifs, neurons must be selective for the defining spectrotemporal edges of the complex sounds, but not the random acoustic variations.

### Enhanced temporal response precision to complex sounds

Responses of a total of 250 multiunits (6 animals per group; number of multiunits: 90 for Jungle group; 80 for Noise group; 80 for Naïve Control group) were recorded. Sound exposure had significantly altered tuning bandwidth measured at 10 dB above threshold ( $p < 0.05$ ), with the bandwidth being significantly narrower for the Jungle group ( $0.73 \pm 0.01$  octaves) than for the Naïve group ( $0.86 \pm 0.04$  octaves) and the Noise group ( $0.94 \pm 0.04$ ). Tuning bandwidth were slight broader for the Noise than for the Naïve group, but the difference was not significant ( $p = 0.087$ ). No differences were found in threshold among the three groups (Naïve,  $31.3 \pm 1.5$  dB; Jungle,  $25.3 \pm 2.8$  dB; Noise,  $29.9 \pm 2.0$  dB;  $p = 0.40$ ). CF distributions were not between the Naïve and the Jungle or Noise groups (Kolmogorov-Smirnov Test with correction for multiple comparison,  $p > 0.1$ ).

Response firing rates to the testing complex sounds were calculated in 5-ms windows for each unit and neurograms were constructed with firing rates of all units recorded. While many units in the Naïve and Noise groups responded to the complex sounds with sustained responses, most of the units in the Jungle group showed responses time-locked to distinctive features of the sounds (Figure 3). Similar phasic responses to conspecific vocalization have been reported before (Suta et al., 2008). The reduced sustained responses may be related to the fast temporal modulations of the jungle sounds that the animals were exposed to during their critical period (see later discussion and Figures 2 and 7). The enhanced temporal precision of the cortical responses may reflect improved complex feature selectivity of units of the complex sound-experienced animals (as seen later in Figure 6). These coherent and dynamic responses may be a basis for cortical processing of rapidly modulated complex sound (Nagarajan et al., 1999, Schnupp et al., 2006).

### Emergent categorical representation of experienced complex sounds

To examine how the cortex represents fine differences between sound motifs, we quantified the dissimilarity between neurogram responses to the 20 complex sounds belonging to 5 song motifs, which is the ratio of neurogram response difference over sum of the responses (see methods for details). A dissimilarity of one indicates that none of the units responded to the two sounds with the same delay—a unit may either respond to only one of the sounds, or responds to the two sound with different delays. A dissimilarity of zero means that all of the units showed identical mean post-stimulus time histograms for the two sounds. The

dissimilarity between responses to different song motifs was significantly higher for Jungle than for Noise or Naïve Control animals ( $p < 0.05$ ), whereas the dissimilarity between responses to different renderings of the same song motifs was significantly lower after jungle sound experience ( $p < 0.01$ , Figure 4b).

To illustrate how representations of these sounds were organized, we plotted their similarity relations in Figure 4c, in which distance between representations is proportional to their representational dissimilarity (Rolls and Tovee, 1995, Ohl et al., 2001). In the Jungle animal but not in the two control groups, representations of all sounds were sharply segregated (with  $p < 0.05$ , Material and Methods) by motif identities. These results indicate that after jungle sound experience, cortical neurons became *more* sensitive to features that were important for distinguishing different song motifs, while they became *less* sensitive to acoustic variations associated with each repetition of the same song. These differences contribute together to the more distinct “categorical” representations of different song motifs (please see Discussion).

To determine if this emergent representational segregation was specific to the sounds to which rat pups were exposed in the critical period, we exposed an additional group of animals to naturally spoken American English (shifted upward in sound frequency to rat hearing range; see Material and Methods), then examined their cortical responses to jungle sounds—to which they were not exposed. A total of 82 multiunits from 6 animals were examined. Initial analysis of their receptive fields indicated that their CF distribution was not altered by speech sound exposure (Speech vs. Naïve, Kolmogorov-Smirnov Test with correction for multiple comparison,  $p > 0.1$ ). The threshold was similar between Speech and Naïve groups (Speech,  $29.01 \pm 2.1$  dB; Naïve,  $31.3 \pm 1.5$  dB,  $p > 0.5$ ). The tuning bandwidth measured at 10 dB above threshold was narrower for the Speech than Naïve group (Speech,  $0.66 \pm 0.05$  octaves; Naïve,  $0.86 \pm 0.04$  octaves,  $p < 0.05$ ).

Further analysis of their response to jungle sounds indicated that the levels of dissimilarity between neurogram responses to different jungle song motifs were comparable between the speech-exposed and the jungle sound-exposed groups ( $p > 0.5$ ). However, responses to different renderings of the same jungle song motifs were very different for the speech group—indeed, greater than for any of the other two groups (i.e., had higher dissimilarity,  $p < 0.005$ , Figure 5b), even including the Naïve Control group. Similarity relation plots indicated that representations of all the testing jungle sounds were not significantly segregated in speech-exposed animals (Figure 5c,  $p > 0.1$ ). Therefore, representational segregation appears to be specific to the sounds to which animals were exposed in early development.

### Neuronal feature selectivity underlies population-based changes

To reveal what changes in feature selectivity may underlie the segregated complex sound representations, we derived cortical spectrotemporal receptive fields (STRF) from responses to the testing jungle sounds using a generalized reverse-correlation technique (Theunissen et al., 2000, Theunissen et al., 2001). The STRF-predicted and actual firing rates to a different set of stimuli from the same song ensemble were reasonably correlated (Theunissen et al., 2000) (correlation coefficient: Jungle,  $0.49 \pm 0.02$ ; Noise,  $0.34 \pm 0.03$ ; Naïve Control,  $0.43 \pm$

0.02;  $p < 0.001$  for Jungle vs. Noise groups). Response selectivity to spectrotemporally complex features was assessed using a STRF separability measure, which estimates how well the STRF can be approximated by a multiplicative product of separate spectral and temporal receptive field profiles (see Material and Methods). Frequency-modulated complex features are spectrotemporally inseparable. Although units selective for frequency-modulated (FM) sweeps were rare in Noise and Naïve Control animals (Linden et al., 2003), they were frequently observed in the jungle sound-experienced animals, and typically had a separability measure below 0.3 (Figure 6, a1–a2). Overall, STRFs of jungle sound-experienced animals had significantly lower spectrotemporal separability than did STRFs of Noise and Naïve Control animals (Figure 6d). The overall sizes of the excitatory areas of the STRFs were not different among the three groups (in octave\*ms: Jungle,  $28.1 \pm 1.8$ ; Noise,  $28.2 \pm 2.3$ ; Naïve Control,  $31.0 \pm 2.1$ ;  $p > 0.5$ ). By contrast, the size of the inhibitory areas was significantly greater for Jungle than for Noise and Naïve Control animals (Figure 6e,  $p < 0.05$ ). The duration of the excitatory areas was significantly shorter for Jungle than for Noise and Naïve Control animals (Figure 6f,  $p < 0.005$ ). These results indicate that more units in Jungle animals selected for rapidly frequency-modulated features of the jungle sounds. The enlarged inhibitory areas of STRFs of Jungle animals also improved complex sound response selectivity.

We also derived STRFs for units in speech-sound-reared animals, using the same jungle sound stimuli that were used to derive STRFs in Jungle animals. The units in speech-sound-reared animals tended to select for prolonged, frequency-modulated features (Figure 7A–B). The temporal modulation spectra of their STRFs showed less power above 22 Hz than those of units in jungle sound-reared animals (Figure 7C). Interestingly, temporal modulation spectra of the rearing stimuli indicate that the jungle sounds has greater power than speech sound for modulation above 22 Hz (Figure 7D).

### Complex sound exposure shapes sparse and distinct cortical sound representations

Previous studies have shown that repeated experience of a simple stimulus generally increases its cortical representational size. To determine how natural sensory experience affects cortical representations, we quantified the size of the populations of cortical units responding to the testing jungle sounds. For each sampled unit, spikes were binned with 5-ms windows, and the firing rate was calculated for each bin. A cortical unit was considered responsive to the complex sound stimuli during a 5-ms window if its firing rate in the corresponding bin was two standard deviations above the mean baseline firing rate of the unit. In jungle sound-experienced animals, on average,  $6.75 \pm 0.02\%$  of the sampled units responded to the complex testing stimuli in any given time bin, which was 25% and 32% lower than in Noise and Naïve Control animals ( $8.97 \pm 0.01\%$  and  $9.96 \pm 0.02\%$ , respectively,  $p < 0.0001$ ; see Figure 8a). However, the magnitude of responses was significantly greater for Jungle than for Noise and Naïve control animals (Figure 8b; response magnitude was calculated only for the time bins with firing rate  $>$  mean plus 2 SDs).

The reduced complex sound representations in jungle sound-experienced animals may reflect one of two different possibilities: 1) fewer units represented all complex sounds; or 2)

most units were involved in coding, with each specific sound represented by fewer units. Assuming that neurons code complex sounds by changing their firing rates, the variance of firing rate may be considered as an index of involvement of the neurons in sensory coding (Willmore et al., 2000, Weliky et al., 2003). The firing rate variances of all sampled units were more evenly distributed in Jungle and Noise animals than in Naïve Control animals, indicating that more units vigorously participated in acoustic representations after experience of jungle sound or noise (Figure 8c). Taken together, these results indicate that jungle sound experience increased the total number of neurons representing complex sounds, but reduced the numbers of neurons representing each individual sound, leading to their more distinct sparse representations (Olshausen and Field, 1996).

## DISCUSSION

Our results demonstrate that naturalistic complex sound experience improves cortical selectivity for spectrotemporally complex features of those specific sounds, and “specializes” acoustic processing in the cortex for their specific, selective representations (Rauschecker et al., 1995, Wang et al., 1995, Nelken et al., 1999, Fritz et al., 2003). They also reveal important differences between the effects of naturalistic versus simple experiences on cortical sensory representations. Repeated experience of the same stimulus generally results in expanded representations of the experienced stimulus, and in this non-competitive circumstance, can actually result in a degraded neuronal selectivity for specific stimulus features (Sengpiel et al., 1999, Zhang et al., 2001, Han et al., 2007). In striking contrast, mere environmental exposure to naturalistic stimuli in the critical period leads to their smaller but more distinct cortical representations.

Although sensory representations are more malleable in the critical period, they can also be reorganized in adulthood. For example rats trained with human speech sounds can learn the statistical structures of human speech, and recognize different languages and different phonemes (Toro et al., 2003, Toro and Trobalon, 2005, Engineer et al., 2008). Such adult learning process likely also involves sparse and refined representations of spectrotemporally complex features (Reed et al., 2011, Gdalyahu et al., 2012). By contrast, repeated exposure to unstructured tone pips during critical period or in adulthood can disrupt cortical sound representations and perception (Norena et al., 2006, Han et al., 2007).

Categorical sensory representations are typically demonstrated with categorical neural responses to a series of morphed stimuli from one prototype stimulus to another. We did not examine cortical responses to morphed jungle sounds. In this study, categorical representation refers to the highly segregated cortical representation resulting from early experience (Figure 4, C1), which is indicative of better categorization of the jungle sound motifs by cortical neurons.

The naturalistic sounds used in the present study are not conspecific for the rat, but rather a common ambient natural sound scene (including bird vocalizations). In a natural sound environment, it may be important to distinguish between different types of sounds produced by different sources with highly distinctive representations. Mechanistically, a heavy schedule of complex acoustic inputs engages strong competitive plasticity that results in the

separate representations of categorically distinguishable sounds. These competitive processes are not activated by repetitions of an invariant sound (Sengpiel et al., 1999, Zhang et al., 2001, Han et al., 2007).

The present study also revealed that cortical neurons become more selective for complex features of the sensory environment during early development. More units responded to the complex features in the rearing sounds. This is consistent with earlier results showing expanded representations of experienced sounds (Insanally et al., 2009). However, auditory cortical neurons in animals reared with complex sound in this study were more selective of the sounds and the sound features, resulting in smaller number of neurons representing any particular sound. When these neurons did respond, they responded more vigorously. More importantly, neuronal representation was distributed more evenly among the neurons. These are hallmarks of sparse sensory representations (Olshausen and Field, 1996, Willmore et al., 2000). These results indicate that early experience of the natural acoustic environment leads to sparse sound representations (Klein et al., 2003, Hromadka et al., 2008, Carlson et al., 2012). Sparse somatosensory cortical representations have been reported in animals housed in an enriched environment (Polley et al., 2004b), or following associative learning (Gdalyahu et al., 2012).

The input-specific representational segregation observed in the present studies argues for an auditory system that is specialized in recognizing experienced sounds even in the presence of certain distortions. Animals in their normal environment are likely exposed to large amount of con-specific vocal sounds. These sounds are represented differently than other complex sounds (Wang and Kadia, 2001), and their perception is more refined (Ehret and Haack, 1981, May et al., 1989, Nelson and Marler, 1989, Wytenbach et al., 1996, Baugh et al., 2008). The conspecific representation and perception of animal vocalizations may be intrinsic, or, as suggested by this and other study, they may be shaped by early experience (Kuhl and Miller, 1978, Mesgarani et al., 2008, Prather et al., 2009, Köver et al., 2013). Thus, the history of the animal's sensory experiences must be considered in order to understand the preferential representations of conspecific vocalizations.

Results of the present study suggest that reliable sensory percept may be mediated by a form of *distributed* categorical representation in AI (Kim and Bao, 2008, Köver and Bao, 2010). Such a notion is consistent with previous findings that while cortical responses to novel stimuli may be sensitive to contextual disturbances and small variations, representations of heavily experienced natural stimuli are relatively stable (Dragoi et al., 2001, Bar-Yosef et al., 2002). Experience-dependent specialization also occurs in human speech perception. In the first year of life, human phonemic perception evolves from a general to a language-specific form through experience-dependent processes (Kuhl et al., 1992, Kuhl, 2000). The type of cortical plasticity observed in the present study provides a plausible model for the development of categorical phonemic perception.

## Acknowledgments

We thank Dr. Konrad Körding for providing the jungle sound stimuli. This research was supported by National Institute of Health Grants (NS-10414, NS-34835, DC-009259), The Coleman Fund, The Sandler Fund, The Mental Insight Foundation.

## References

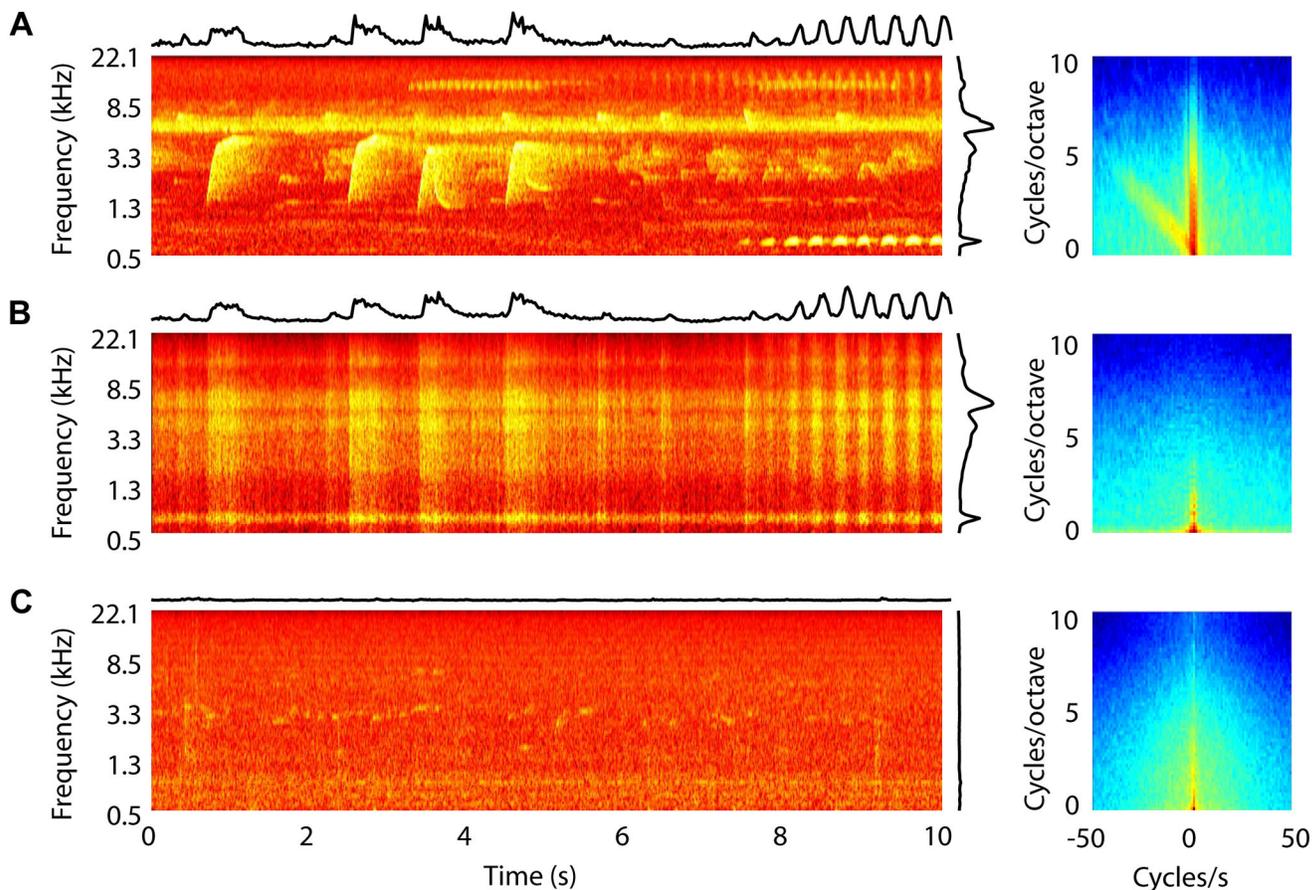
- Bao S, Chan VT, Merzenich MM. Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*. 2001; 412:79–83. [PubMed: 11452310]
- Bao S, Chang EF, Davis JD, Gobeske KT, Merzenich MM. Progressive degradation and subsequent refinement of acoustic representations in the adult auditory cortex. *J Neurosci*. 2003; 23:10765–10775. [PubMed: 14645468]
- Bao S, Chang EF, Woods J, Merzenich MM. Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat Neurosci*. 2004; 7:974–981. [PubMed: 15286790]
- Bar-Yosef O, Rotman Y, Nelken I. Responses of neurons in cat primary auditory cortex to bird chirps: effects of temporal and spectral context. *J Neurosci*. 2002; 22:8619–8632. [PubMed: 12351736]
- Baugh AT, Akre KL, Ryan MJ. Categorical perception of a natural, multivariate signal: mating call recognition in tungara frogs. *Proc Natl Acad Sci U S A*. 2008; 105:8985–8988. [PubMed: 18577592]
- Brown M, Irvine DR, Park VN. Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex. *Cereb Cortex*. 2004; 14:952–965. [PubMed: 15115736]
- Carlson NL, Ming VL, Deweese MR. Sparse codes for speech predict spectrotemporal receptive fields in the inferior colliculus. *PLoS Comput Biol*. 2012; 8:e1002594. [PubMed: 22807665]
- Doupe AJ, Konishi M. Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc Natl Acad Sci U S A*. 1991; 88:11339–11343. [PubMed: 1763048]
- Dragoi V, Turcu CM, Sur M. Stability of cortical responses and the statistics of natural scenes. *Neuron*. 2001; 32:1181–1192. [PubMed: 11754846]
- Ehret G, Haack B. Categorical perception of mouse pup ultrasound by lactating females. *Naturwissenschaften*. 1981; 68:208–209. [PubMed: 7278997]
- Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP. Cortical activity patterns predict speech discrimination ability. *Nat Neurosci*. 2008; 11:603–608. [PubMed: 18425123]
- Esser KH, Condon CJ, Suga N, Kanwal JS. Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proc Natl Acad Sci U S A*. 1997; 94:14019–14024. [PubMed: 9391145]
- Fox RA, Flege JE, Munro MJ. The perception of English and Spanish vowels by native English and Spanish listeners: a multidimensional scaling analysis. *J Acoust Soc Am*. 1995; 97:2540–2551. [PubMed: 7714272]
- Fritz J, Shamma S, Elhilali M, Klein D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci*. 2003; 6:1216–1223. [PubMed: 14583754]
- Gdalyahu A, Tring E, Polack PO, Gruver R, Golshani P, Fanselow MS, Silva AJ, Trachtenberg JT. Associative fear learning enhances sparse network coding in primary sensory cortex. *Neuron*. 2012; 75:121–132. [PubMed: 22794266]
- Godde B, Stauffenberg B, Spengler F, Dinse HR. Tactile coactivation-induced changes in spatial discrimination performance. *J Neurosci*. 2000; 20:1597–1604. [PubMed: 10662849]
- Han YK, Kover H, Insanally MN, Semerdjian JH, Bao S. Early experience impairs perceptual discrimination. *Nat Neurosci*. 2007; 10:1191–1197. [PubMed: 17660815]
- Hromadka T, Deweese MR, Zador AM. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol*. 2008; 6:e16. [PubMed: 18232737]
- Insanally MN, Kover H, Kim H, Bao S. Feature-dependent sensitive periods in the development of complex sound representation. *J Neurosci*. 2009; 29:5456–5462. [PubMed: 19403813]
- Iverson P, Kuhl PK. Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *J Acoust Soc Am*. 1996; 99:1130–1140. [PubMed: 8609297]
- Kim H, Bao S. Distributed representation of perceptual categories in the auditory cortex. *J Comput Neurosci*. 2008; 24:277–290. [PubMed: 17917802]
- Klein DJ, Konig P, Kording KP. Sparse spectrotemporal coding of sounds. *EURASIP J Appl Signal Process*. 2003; 2003:659–667.

- Köver H, Bao S. Cortical plasticity as a mechanism for storing Bayesian priors in sensory perception. *PLoS One*. 2010; 5:e10497. [PubMed: 20463967]
- Köver H, Gill K, Tseng YL, Bao S. Perceptual and Neuronal Boundary Learned from Higher-Order Stimulus Probabilities. *J Neurosci*. 2013; 33:3699–3705. [PubMed: 23426696]
- Kuhl PK. A new view of language acquisition. *Proc Natl Acad Sci U S A*. 2000; 97:11850–11857. [PubMed: 11050219]
- Kuhl PK, Miller JD. Speech perception by the chinchilla: identification function for synthetic VOT stimuli. *J Acoust Soc Am*. 1978; 63:905–917. [PubMed: 670558]
- Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*. 1992; 255:606–608. [PubMed: 1736364]
- Linden JF, Liu RC, Sahani M, Schreiner CE, Merzenich MM. Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *J Neurophysiol*. 2003; 90:2660–2675. [PubMed: 12815016]
- Margoliash D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J Neurosci*. 1983; 3:1039–1057. [PubMed: 6842281]
- May B, Moody DB, Stebbins WC. Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *J Acoust Soc Am*. 1989; 85:837–847. [PubMed: 2925998]
- Mesgarani N, David SV, Fritz JB, Shamma SA. Phoneme representation and classification in primary auditory cortex. *J Acoust Soc Am*. 2008; 123:899–909. [PubMed: 18247893]
- Nagarajan S, Mahncke H, Salz T, Tallal P, Roberts T, Merzenich MM. Cortical auditory signal processing in poor readers. *Proc Natl Acad Sci U S A*. 1999; 96:6483–6488. [PubMed: 10339614]
- Nelken I, Rotman Y, Bar Yosef O. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*. 1999; 397:154–157. [PubMed: 9923676]
- Nelson DA, Marler P. Categorical perception of a natural stimulus continuum: birdsong. *Science*. 1989; 244:976–978. [PubMed: 2727689]
- Norena AJ, Gourevitch B, Aizawa N, Eggermont JJ. Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. *Nat Neurosci*. 2006; 9:932–939. [PubMed: 16783369]
- Ohl FW, Scheich H, Freeman WJ. Change in pattern of ongoing cortical activity with auditory category learning. *Nature*. 2001; 412:733–736. [PubMed: 11507640]
- Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*. 1996; 381:607–609. [PubMed: 8637596]
- Polley DB, Heiser MA, Blake DT, Schreiner CE, Merzenich MM. Associative learning shapes the neural code for stimulus magnitude in primary auditory cortex. *Proc Natl Acad Sci U S A*. 2004a; 101:16351–16356. [PubMed: 15534214]
- Polley DB, Kvasnak E, Frostig RD. Naturalistic experience transforms sensory maps in the adult cortex of caged animals. *Nature*. 2004b; 429:67–71. [PubMed: 15129281]
- Prather JF, Nowicki S, Anderson RC, Peters S, Mooney R. Neural correlates of categorical perception in learned vocal communication. *Nat Neurosci*. 2009; 12:221–228. [PubMed: 19136972]
- Quinn, GP.; Keough, MJ. *Experimental Design and Data Analysis for Biologist*. Cambridge: Cambridge University Press; 2002.
- Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*. 1995; 268:111–114. [PubMed: 7701330]
- Recanzone GH, Schreiner CE, Merzenich MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci*. 1993; 13:87–103. [PubMed: 8423485]
- Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, Kilgard MP. Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron*. 2011; 70:121–131. [PubMed: 21482361]
- Rolls ET, Tovee MJ. Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J Neurophysiol*. 1995; 73:713–726. [PubMed: 7760130]
- Schnupp JW, Hall TM, Kokelaar RF, Ahmed B. Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci*. 2006; 26:4785–4795. [PubMed: 16672651]

- Sengpiel F, Stawinski P, Bonhoeffer T. Influence of experience on orientation maps in cat visual cortex. *Nat Neurosci.* 1999; 2:727–732. [PubMed: 10412062]
- Suta D, Popelar J, Syka J. Coding of communication calls in the subcortical and cortical structures of the auditory system. *Physiol Res.* 2008; 57(Suppl 3):S149–159. [PubMed: 18481905]
- Talwar SK, Gerstein GL. Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. *J Neurophysiol.* 2001; 86:1555–1572. [PubMed: 11600620]
- Tegenthoff M, Ragert P, Pleger B, Schwenkreis P, Forster AF, Nicolas V, Dinse HR. Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 Hz rTMS. *PLoS Biol.* 2005; 3:e362. [PubMed: 16218766]
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network.* 2001; 12:289–316. [PubMed: 11563531]
- Theunissen FE, Sen K, Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci.* 2000; 20:2315–2331. [PubMed: 10704507]
- Toro JM, Trobalon JB. Statistical computations over a speech stream in a rodent. *Percept Psychophys.* 2005; 67:867–875. [PubMed: 16334058]
- Toro JM, Trobalon JB, Sebastian-Galles N. The use of prosodic cues in language discrimination tasks by rats. *Anim Cogn.* 2003; 6:131–136. [PubMed: 12728358]
- Wang X, Kadia SC. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol.* 2001; 86:2616–2620. [PubMed: 11698547]
- Wang X, Merzenich MM, Beitel R, Schreiner CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol.* 1995; 74:2685–2706. [PubMed: 8747224]
- Weliky M, Fiser J, Hunt RH, Wagner DN. Coding of natural scenes in primary visual cortex. *Neuron.* 2003; 37:703–718. [PubMed: 12597866]
- Willmore B, Watters PA, Tolhurst DJ. A comparison of natural-image-based models of simple-cell coding. *Perception.* 2000; 29:1017–1040. [PubMed: 11144817]
- Wytenbach RA, May ML, Hoy RR. Categorical perception of sound frequency by crickets. *Science.* 1996; 273:1542–1544. [PubMed: 8703214]
- Zhang LI, Bao S, Merzenich MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat Neurosci.* 2001; 4:1123–1130. [PubMed: 11687817]

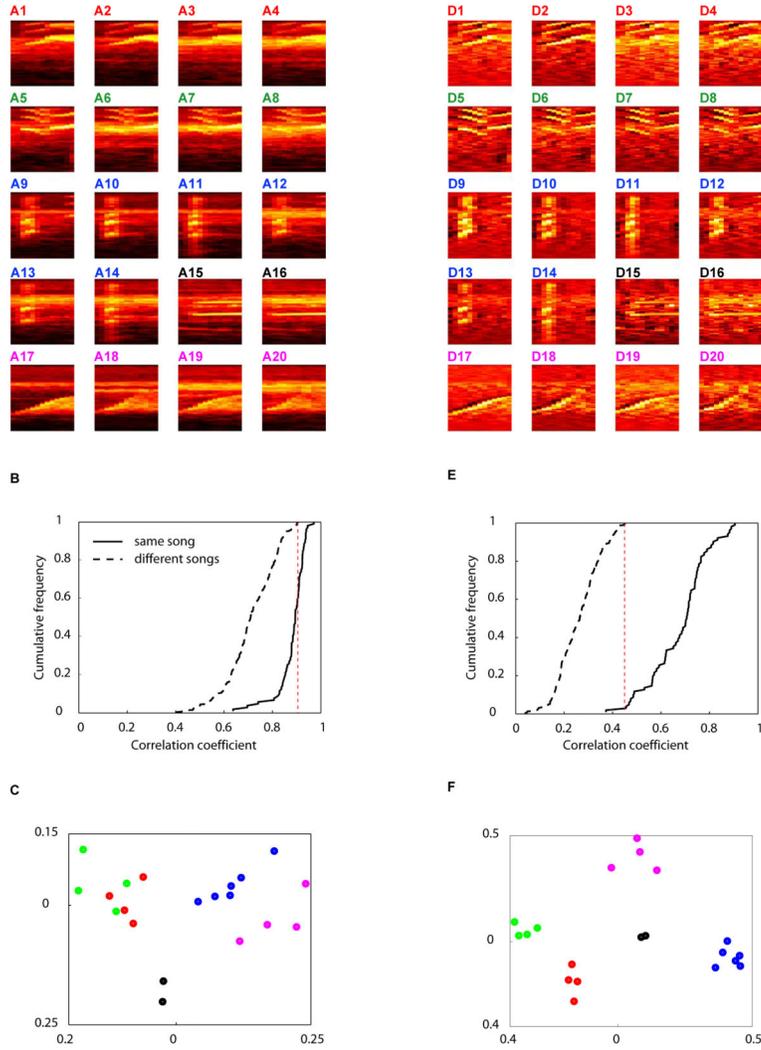
### Highlights

- Early exposure to naturalistic complex sounds shapes refined STRFs.
- Early experience shapes sparse representation of natural sounds.
- Exposure to natural sounds results in categorical representations of the sounds.
- The refinement in sound representations is specific to the experience.



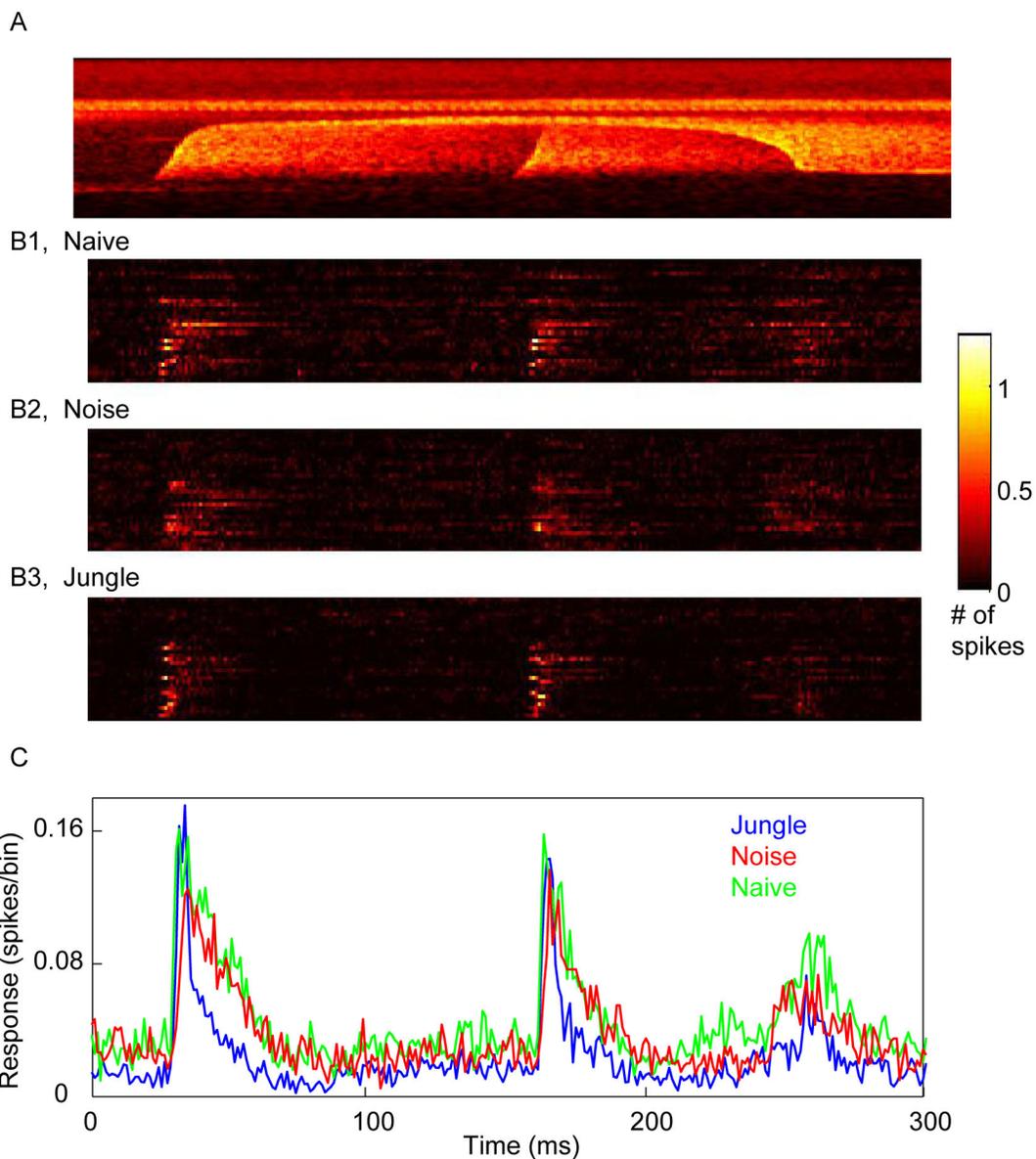
**Figure 1.**

Spectrograms (left) and spectrotemporal modulation spectra (right) of the sounds in three different acoustic rearing environments. (A) jungle sounds; (B) modulated noises; and (C) animal room sounds. The jungle sounds and the modulated noises had similar overall spectral and temporal modulations, which were plotted on the top and the right side of corresponding spectrograms. The spectral and temporal modulations are completely separable for the modulated noise. Consequently, in the modulation spectra, significantly more power was distributed along the two axes (i.e. with temporal or spectral modulations  $< 1$  Hz) for the modulated noise than for the jungle sounds ( $p < 0.0001$ ). The animals room sounds were dominated by relatively constant, low-level fan noise, and had less power for spectral and temporal modulations  $> 1$  Hz ( $p < 0.0001$ , comparing to jungle sounds).

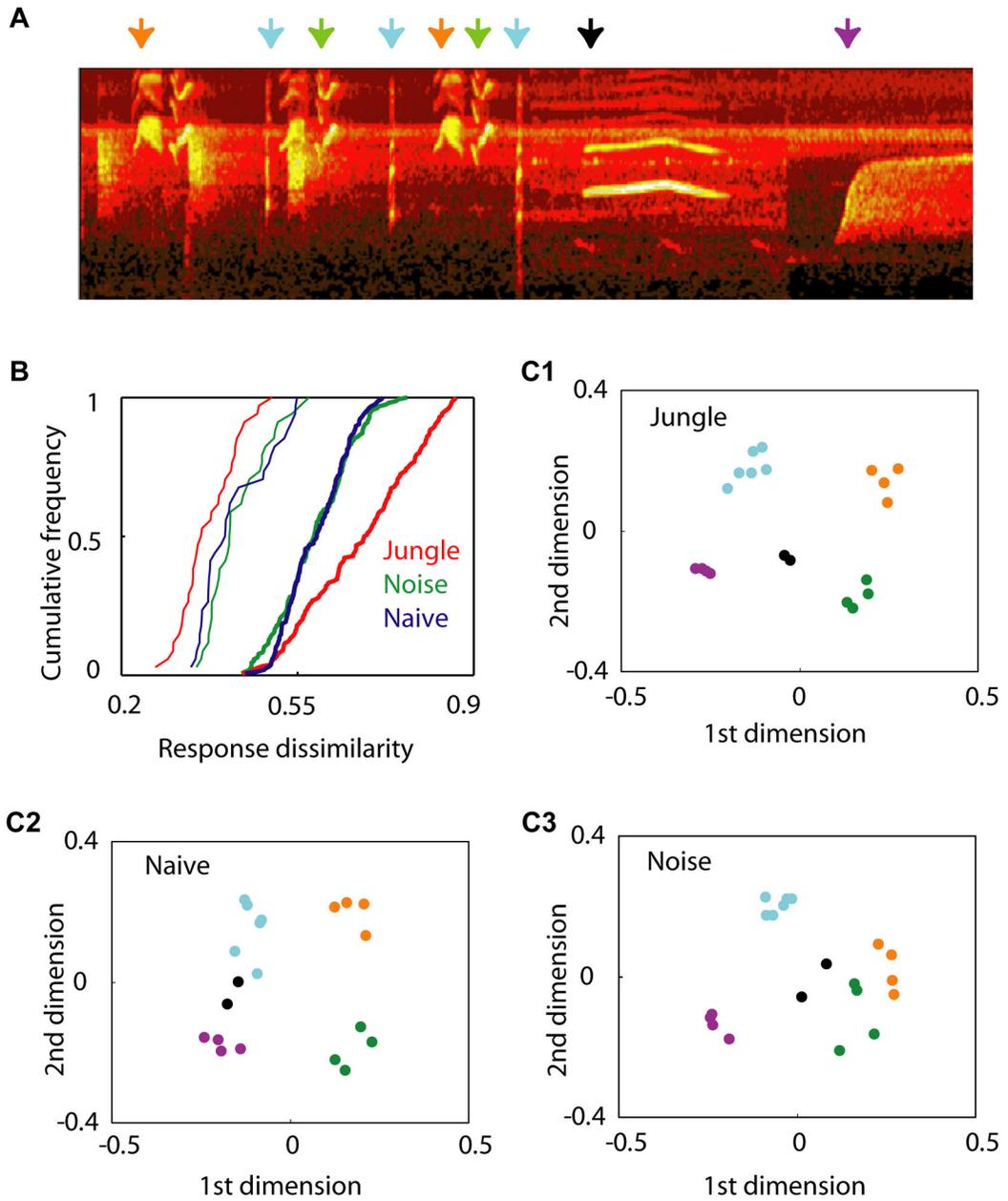


**Figure 2.** Properties of the testing jungle sounds. (A1–A20) Spectrograms of the twenty jungle sounds whose cortical responses were analyzed in Figure 1. Different song motifs were coded with different colors. Spectrogram frequency range, 0.5–22.1 kHz; time scale, 60 ms. (B) Cumulative distributions of pair-wise correlation coefficients (CCs) between the twenty sound spectrograms as measures of their similarity. CCs between sounds of the same song motifs (within-motif) and CCs between sounds of different motifs (between-motif) are plotted separately. The two populations of CCs were highly overlapping. (C) Similarity relations between the sounds obtained with multidimensional scaling methods. Each dot represents one of the sounds. The distances between the dots are proportional to the dissimilarity between the corresponding sounds (defined as  $1 - CC$ ). The color-coding for the sounds is the same as shown in (A). (D1–D20) Spectrograms of the sounds after temporal band-pass filtering to emphasize spectrotemporal “edges”. (E) Cumulative distributions of pair-wise correlation coefficients (CCs) between the twenty sound spectrograms in (D1–D20). (F) Similarity relations between the filtered spectrographic

features of the sounds. All sounds were significantly segregated by motif identities ( $p < 0.05$ ).



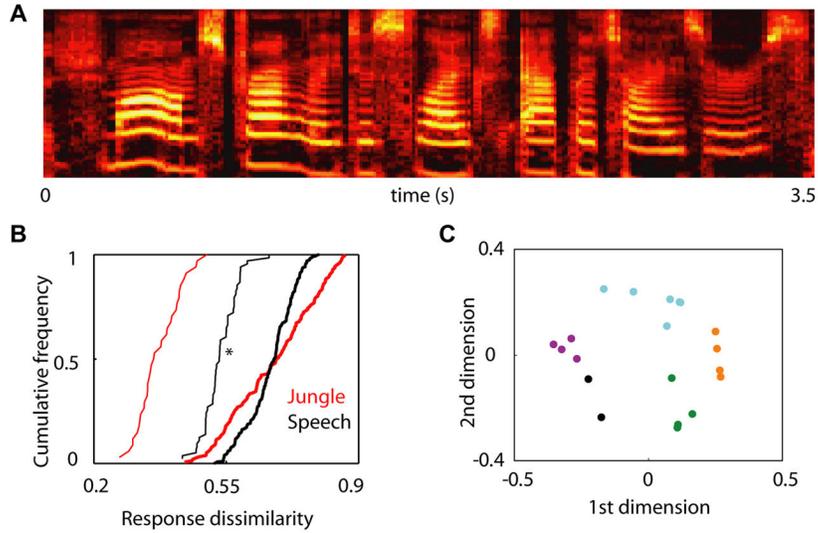
**Figure 3.** Coherent responses shaped by early experience. A. The spectrogram of the testing sounds. B. Neurogram responses. For figure clarity, only 30 multiunits are shown here for each group. C. Mean response magnitude. Although the magnitudes of dynamic responses were not different between the three groups, units in the Jungle group showed weaker sustained responses ( $p < 0.05$ ).



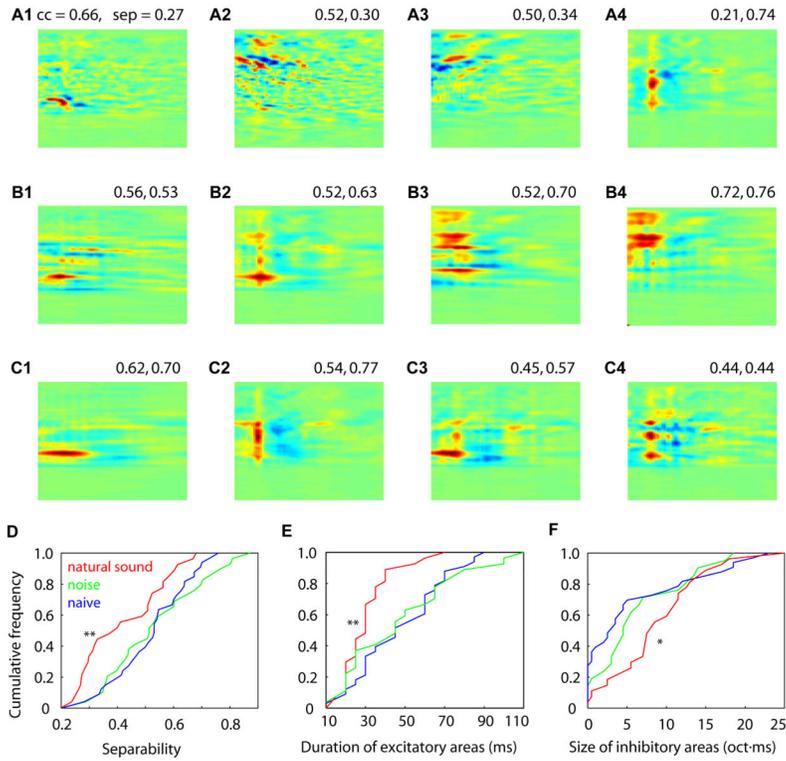
**Figure 4.**

Segregated representations of jungle song motifs instructed by early experience. (A) Spectrogram of a section of the testing jungle sound stimuli (duration, 4.5 s). Arrows with different colors indicate different song motifs. (B) Cumulative distributions of response dissimilarities. Dissimilarities between responses to different song motifs were plotted with thick lines, and those to same song motifs were plotted with thin lines. (\*,  $p < 0.05$ ; \*\*,  $p < 0.005$ ; comparing with the other two groups). (C) Similarity relations between responses to jungle sounds. Each dot represents a cortical population-temporal response to a jungle sound. The distances between dots are proportional to the dissimilarity between corresponding responses. The color-coding for song motifs is the same as shown in (A).

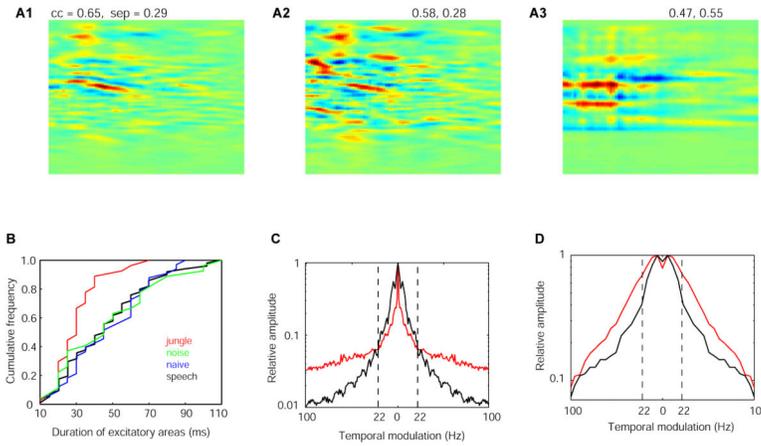
Note that cortical responses to different song motifs were sharply segregated in jungle sounds-experienced animals ( $p < 0.05$ , Material and Methods for statistical tests).



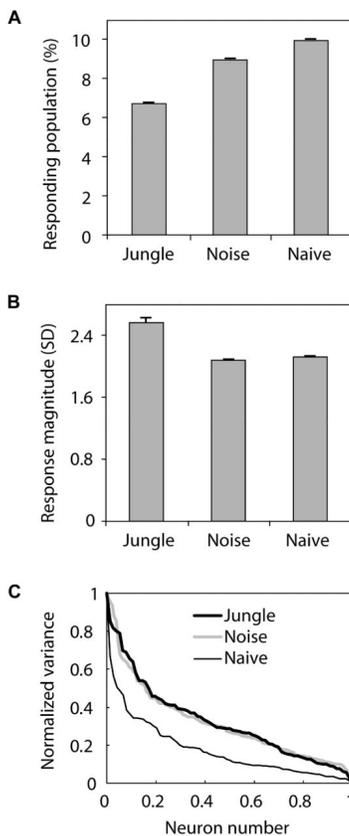
**Figure 5.** Influences of speech sound experience on jungle sound representation. **(A)** The spectrogram of representative speech sounds used for acoustic rearing. The frequencies of the speech sounds were shifted 1.45 octaves higher to rat hearing range. **(B)** Cumulative distributions of response similarities. Similarities between responses to different song motifs were plotted with thick lines, and those to same song motifs were plotted with thin lines. \* indicates  $p < 0.005$ . **(C)** Similarity relations between responses to the jungle sounds. Responses for all song motifs were not significantly segregated ( $p > 0.05$ ).



**Figure 6.** Spectrotemporal receptive fields. (A–C) Representative STRFs derived for units of Jungle (A), Noise (B) and Naïve Control (C) animals. Cross-correlation (cc) between STRF-predicted and actual firing rates and spectrotemporal separability (sep) of the STRF were given. Scales: horizontal, –50 to 300 ms; vertical, 5.45 octaves from 0.5 to 22.1 kHz. (D–F) Cumulative distributions of spectrotemporal separabilities (D), durations of the excitatory response areas (E) and sizes of the inhibitory areas of the STRFs (F). (\*,  $p < 0.05$ ; \*\*,  $p < 0.005$ ; comparing with the other two groups).



**Figure 7.** Properties of speech and jungle sound stimuli and STRFs recorded from speech- and jungle sound-experienced animals. **(A)** Representative STRFs derived from responses to jungle sounds recorded in the speech sound-experienced animals. Cross-correlation (cc) between STRF-predicted and actual firing rates and spectrotemporal separability (sep) of the STRF were given. Scales: horizontal,  $-50$  to  $300$  ms; vertical,  $5.45$  octaves from  $0.5$  to  $22.1$  kHz. Note that units tended to select for prolonged, frequency-modulated features. These STRFs had separability and inhibitory areas similar to those derived from jungle sound-experienced animals (separability: speech,  $0.43 \pm 0.02$ ; jungle,  $0.42 \pm 0.03$ ;  $p > 0.5$ ; size of inhibitory areas: speech,  $14.4 \pm 1.7$ ; jungle  $18.2 \pm 2.6$ ;  $p > 0.05$ ) **(B)** Duration of the excitatory areas of the STRFs. **(C)** Mean temporal modulation spectra of all STRFs of the jungle sound-reared animals or the speech-reared animals. Note that jungle sounds had more power at temporal modulation rates higher than  $22$  Hz, and that STRFs of the jungle sound-reared animals had greater power in that modulation frequency range. **(D)** Temporal modulation spectra of the jungle and the speech sound rearing stimuli. They were obtained through column summation of two-dimensional Fourier transforms of the stimulus spectrograms.



**Figure 8.**

Population response properties. **(A)** Mean percentage of the sampled units that were responsive to the testing jungle sound stimuli in any given 5-ms bin ( $p < 0.0001$ , comparing Jungle with Noise or Naïve Control animals). A unit was considered responsive in a time bin if its firing rate in that bin was two SDs greater than the mean baseline firing rate. **(B)** Response magnitudes normalized to the SD of the firing rates of each unit. Only the bins in which responses occurred were included for that unit. **(C)** Distributions of firing rate variances as an index of dispersed representations. Firing rate variances of all units in each group were normalized to their maximum value, rank-ordered and plotted. The unit numbers were normalized to a range from 0 to 1. The mean normalized variance was greater for Jungle and Noise animals than for Naïve Control animals ( $p < 0.0005$ ), indicating more equally distributed variances of firing rates among units in the Jungle and Noise groups, and hence more dispersed cortical acoustic representations in these groups of animals.