UCSF

Title

Auditory Cortical Plasticity Dependent on Environmental Noise Statistics

Permalink

Journal

ISSN

Authors

Homma, Natsumi Y
Hullett, Patrick W
Atencio, Craig A
et al.

Publication Date

DOI

Peer reviewed

Auditory Cortical Plasticity Dependent on Environmental Noise Statistics

Natsumi Y. Homma[1,2], **Patrick W. Hullett**[1,2], **Craig A. Atencio**[1,2], **Christoph E. Schreiner**[1,2,3,*]

[1]Coleman Memorial Laboratory, Department of Otolaryngology - Head and Neck Surgery, University of California, San Francisco, San Francisco, CA 94143, USA

[2]Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, CA 94143, USA

[3]Lead Contact

SUMMARY

During critical periods, neural circuits develop to form receptive fields that adapt to the sensory environment and enable optimal performance of relevant tasks. We hypothesized that early exposure to background noise can improve signal-in-noise processing, and the resulting receptive field plasticity in the primary auditory cortex can reveal functional principles guiding that important task. We raised rat pups in different spectro-temporal noise statistics during their auditory critical period. As adults, they showed enhanced behavioral performance in detecting vocalizations in noise. Concomitantly, encoding of vocalizations in noise in the primary auditory cortex improves with noise-rearing. Significantly, spectro-temporal modulation plasticity shifts cortical preferences away from the exposed noise statistics, thus reducing noise interference with the foreground sound representation. Auditory cortical plasticity shapes receptive field preferences to optimally extract foreground information in noisy environments during noise-rearing. Early noise exposure induces cortical circuits to implement efficient coding in the joint spectral and temporal modulation domain.

In Brief

After rearing rats in moderately loud spectro-temporally modulated background noise, Homma et al. investigated signal-in-noise processing in the primary auditory cortex. Noise-rearing improved vocalization-in-noise performance in both behavioral testing and neural decoding. Cortical plasticity shifted neuronal spectro-temporal modulation preferences away from the exposed noise statistics.

Graphical Abstract

INTRODUCTION

Investigating how the brain develops in specific auditory environments can aid in understanding how hearing abilities are shaped. Developmental "critical periods" are instrumental in forming auditory processing capabilities, including speech processing and the acquisition of language (Friedmann and Rusou, 2015; Kuhl, 2010). During maturation, neural circuits adjust to environmental demands generating receptive fields (RFs) that optimally process relevant sounds (Chang et al., 2005; Froemke and Jones, 2011; Takesian et al., 2018).

Rearing animals in a modified sound environment can influence RF properties, such as frequency tuning or temporal preference (e.g., Chang and Merzenich, 2003; Insanally et al., 2009; Oliver et al., 2011; Pysanenko et al., 2018). How stimulus statistics during development affect neural coding and the resulting consequences for perceptual performance, however, is still unresolved but can be assessed by manipulating the acoustic environment during the critical period.

An essential ability of listeners is the perception of sounds in the face of competing background noise, which is achieved by perceptual segregation of signal from noise. Signal-in-noise (SiN) processing mechanisms have been studied extensively and point to a germane role of the auditory cortex (Bar-Yosef and Nelken, 2007; Malone et al., 2017; Moore et al.,

2013; Rabinowitz et al., 2013; Schneider and Woolley, 2013). However, we do not know the specific neural mechanisms that are involved in solving the SiN problem and how they are shaped by the acoustic environment during development. Important sound features include spectral and temporal modulations (Chi et al., 1999; Elliott and Theunissen, 2009), which are encoded in cortical neurons and contribute to intelligibility and the discrimination of natural sounds (Atencio and Schreiner, 2010; Chi et al., 1999; Hullett et al., 2016).

Our objective was to generate animals that excel in performing SiN analyses and to evaluate the cortical functional properties specialized for that task. Therefore, we explored the influence of exposure to complex sounds with specified spectro-temporal modulation content during the rat auditory critical period (de Villers-Sidani and Merzenich, 2011). After development, we behaviorally tested rats on a vocalization-in-noise task followed by extracellular recordings in the primary auditory cortex (A1). We found that noise-rearing improved both the adult behavioral SiN performance and the neuronal SiN decoding capacity. Significantly, the RF preferences of neurons shifted away from the spectro-temporal modulation range of the exposure sound. The results reveal parallel perceptual and physiological changes following noise exposure and uncover essential aspects of SiN processing.

RESULTS

First, we assessed the spectral and temporal modulation content in representative communication sounds of our model animals to design several background noises. The frequencies of various Sprague-Dawley rat vocalizations used in this study ranged from 20 to 30 kHz (Figure 1A). The modulation power spectrum (MPS) from 18 stereotypical rat vocalizations (see Method Details) had the highest power at low spectral modulation frequencies (SMFs) (0 to 0.3 cyc/oct) and low temporal modulation frequencies (TMFs) (0 to 10 Hz) (Figure 1B) similar to other animal vocalizations and human speech (Chi et al., 1999; Elliott and Theunissen, 2009; Singh and Theunissen, 2003).

The experimental noises were broadband dynamic moving ripples (DMRs) with specified spectral and temporal modulation combinations. DMRs offer several advantages, including short-time properties common in vocalizations and speech, rigorous control of the stimulus statistics, and suitability for estimating spectro-temporal RFs (STRFs) (Atencio and Schreiner, 2010; Escabi and Schreiner, 2002). We generated four biased DMRs (Figures 1C and 1E) corresponding to four regions of the rat vocalization MPS (Figures 1B and 1D). Noise1 contained TMFs from 0 to 5 Hz and SMFs from 0 to 0.3 cyc/oct (hereafter referred to as "LL"), covering the most powerful modulations in the vocalization MPS. Three additional DMRs covered non-overlapping regions of the MPS (Noise2: 5–20 Hz, 0–0.3 cyc/oct, "HL"; Noise3: 5–20 Hz, 0.3–2 cyc/oct, "HH"; Noise4: 0–5 Hz, 0.3–2 cyc/oct, "LH"; Figure 1D). For RF evaluation, we used a wider, unbiased DMR (Noise5, "full") with SMFs of 0–4 cyc/oct and TMFs of 0–40 Hz (Atencio et al., 2008; Escabi and Schreiner, 2002).

Improvement in Vocalization-in-DMR Detection for Noise-Reared Animals

To investigate the influence of environmental noise statistics during the auditory critical period (P6–45) (de Villers-Sidani and Merzenich, 2011) on sound processing development,

we raised rat pups in different biased DMR noises followed by training in a vocalization-in-DMR detection task. Subsequently, we measured their cortical neural activity either at the end of noise-rearing or at the end of behavioral testing (Figure 2A).

Animals were trained on a rat vocalization-in-DMR detection task using a go/no-go paradigm (n = 17; Figure 2B). One-third of the animals were raised in a typical colony environment without noise exposure and trained for the behavioral task (hereafter referred to as "Trained" group or "T," n = 5). The rest of the animals were raised in one of two biased DMRs prior to their behavioral training, subdivided into the "Noise2-reared-trained" ("N2T," n = 6) and the "Noise4-reared-trained" ("N4T," n = 6) groups.

Behavioral training commenced at P45 for these animals (n = 17). To ensure comparable task performance regardless of background noise sound or exposure history, the animals were first trained to respond to vocalization targets without DMR background until the standard deviation (SD) of their false alarm (FA) rate became <10%, which indicates consistent behavior over sessions (see Method Details). This simple vocalization detection task without background noise revealed no difference in the SD of FA rate (Figure 2C) or sensitivity index (d') (Figure 2D) of the last sessions (ANOVA: groups, $F < 4.65$, p > 0.09 for both SD and d' values).

Subsequently, animals were tested for their ability to detect vocalizations in the presence of the four biased and one unbiased DMR noises (Noise1–5) at 10 different signal-to-noise ratios (SNRs). Individual psychometric functions and d' values were plotted as a function of the SNR (Figures 3A–3E; curve fitting based on the cumulative Gaussian distribution).

Influence of Noise Type on Detection Performance—Before considering noise-rearing effects, we investigated the influence of the background noise statistics on the signal detection performance. Detection performance decreased with decreasing SNR for all noise types (Figures 3A–3E). Testing in Noise3(HH), Noise4(LH), and Noise5(full), however, yielded psychometric functions with higher thresholds (Figures 3C–3F) and lower sensitivities (Figure 3G) than for Noise1(LL) and Noise2(HL) (Figures 3A and 3B) (ANOVA, noise effect, $\chi^2 > 93.79$, p < 0.0001). Performance threshold was defined as falling below $d' = 2.4$ (gray lines in Figures 3A–3E), a value corresponding to hit rates of 87%–94% with 10%–20% FA rates.

These findings demonstrate that the background modulation statistics influence the ability to detect target sounds in noise. The better behavioral performance for Noise1(LL) and Noise2(HL) compared to Noise3(HH), Noise4(LH) and Noise5(full) suggests that the lower spectral modulation rates in Noisel(LL) and Noise2(HL) create more favorable detection conditions due to the more widely spaced energy peaks.

Influence of Noise-Rearing on Detection Performance—Our main question was to determine whether and how noise-rearing affects vocalization detection performance. Indeed, the noise-reared animals (Figures 3A–3E, blue and green symbols/lines) showed better detection performance than the unexposed animals (red symbols/line). This is reflected in a shift to higher d' values for all five noise conditions relative to the unexposed

Trained group. This shift corresponded to 3–16-dB improvements in the threshold SNR for noise-reared animals (group effect, $\chi^2 = 20.08$, p < 0.0001; T versus N2T, p = 0.02; T versus N4T, p = 0.002; N2T versus N4T, p = 0.87) (Figure 3F). Testing with Noise5(full) showed the smallest noise-rearing effect. Rearing in Noise2(HL) or Noise4(LH) resulted in similar detection improvement, independent of the testing noises (Figure 3F). No significant performance differences were observed at the no-noise conditions (Figures 3A–3E, "Inf"). The area under the curves (AUCs) of the psychometric functions supported better sensitivities for noise-reared animals (group effect, $\chi^2 = 13.86$, p = 0.001; T versus N2T, p = 0.03, T versus N4T, p = 0.01, N2T versus N4T, p = 1.00) (Figure 3G).

These findings demonstrate that (1) noise-rearing can improve adult SiN performance by decreasing the vocalization-in-DMR detection threshold values by ~10 dB toward lower SNRs, and (2) the behavioral benefits of noise exposure were not narrowly focused on the statistics of the noise used for exposure.

For detection thresholds, we found some significant interactions between the noise types used for testing and those present during noise-rearing (group x testing noise, $\chi^2 = 17.32$, p = 0.03). The N4T group showed a ~13-dB-lower threshold than Trained animals when tested in Noise4(LH) (Figures 3D and 3F). The improvement was ~5 dB larger than for N2T animals (mean thresholds in Noise4: T, –4 dB SNR; N2T, –12 dB SNR; N4T, –17 dB SNR), suggesting that behavioral improvement may be strongest when measured in the noise matching the raising experience. There was no difference between N2T and N4T animals, however, when testing with Noise2(HL) (mean thresholds in Noise2: –20 dB SNR for both N2T and N4T) (Figures 3B and 3F). Further, noise-rearing improved performance for some noises not heard during development. N2T animals performed better in Noise4 than Trained animals. Testing in Noise3 (containing high TMFs and high SMFs) yielded the highest detection thresholds for the Trained group, and both N2T and N4T groups showed the strongest improvements of 14–16 dB (mean thresholds in Noise3: T, 3 dB SNR; N2T, –11 dB SNR; N4T, –13 dB SNR; Figures 3C and 3F). This indicates a particularly robust noise-rearing effect when testing environments with high SMFs (i.e., Noise3(HH) and Noise4(LH)).

Overall, these interactions suggest that (1) the highest benefit of exposure plasticity occurred in backgrounds similar to the exposure environment, and (2) the largest threshold improvements were achieved for testing noises that induced poorer performance (i.e., noises containing high SMFs).

Behavioral Response Timing—The trends noted in detection accuracy were also observed in the behavioral response time (see Method Details). Despite high individual variability in response time distribution (Figures S1A and S1B), the response time was longer for low SNR conditions (Figures S1C–S1G). Differences between the tested noise types reached significance between low and high spectral modulations (*Bootstrap*: Noise1 versus Noise3, Noise1 versus Noise5, Noise2 versus Noise3, Noise2 versus Noise5, p < 0.01). The Trained group took longer to respond in low SNR conditions compared to N2T and N4T groups (p < 0.03). This trend was enhanced for biased DMRs with high SMFs

(Noise3(HH) and Noise4(LH)), showing a 0.5- to 1-s delay for the −20-dB SNR condition in the Trained group (Figures S1E and S1F).

The behavioral response time findings corroborate the notions that (1) detecting signals in noises containing low SMFs is an easier task, and (2) rearing in modulated noises can improve the behavioral detection performance for a range of noise statistics.

Neuronal Vocalization Decoding Improves in Noise-Reared Animals

We next determined whether noise-rearing affected cortical responses to foreground sounds in different noise background conditions. Following behavioral testing, anesthetized animals underwent extracellular recordings in A1. In addition, we obtained A1 recordings from a control group with no noise-rearing and no behavioral training ("Control" or "C," n = 8) and two noise-reared but untrained groups ("Noise2-reared" group, "N2," n = 4; "Noise4-reared" group, "N4," n = 3).

After identifying the high frequency region in A1 (20–30 kHz, corresponding to the trained and tested rat vocalizations), we presented three vocalizations embedded in the four biased DMR noises (Figure 1C) for six different SNR conditions (see Method Details). In each 6-s trial, the first 3 s consisted of a DMR noise token, and the following 3 s contained the same noise token plus one of the three vocalizations (Figure 1A) at 60 dB sound pressure level (SPL). Across our recordings, we identified 2239 multi-units (MUs) that responded to vocalizations in the six groups (T, 254 MUs; N2T, 652; N4T, 428; C, 170; N2, 587; N4, 148).

A1 neurons responded to vocalizations without background noise with specific temporal patterns (e.g., Figure 4A). Using a nearest-neighbor linear decoder, we estimated which of the three vocalizations elicited the response on each trial (Foffani and Moxon, 2004; Malone et al., 2017; see Method Details). For the example no-noise condition, the confusion matrix reflected high accuracy for the predicted vocalizations that correctly matched the actual ones in each trial (Figure 4B).

Influence of Noise Type on Cortical Decoding Performance—Added background noise substantially reduced the neuronal decoding accuracy (e.g., Figures 4C–4F). We computed group medians of decoding accuracy for each combination of SNR, testing-noise type, rearing condition, and training condition, using all 2239 MUs (Figures 4H–4K). Signal decoding accuracy depended on the SNR as well as on the noise types used for testing. Use of Noise1(LL) as background noise showed the highest vocalizations decoding accuracy and the lowest SNR values at the "threshold" (assessed at 50% correct) (Figure 4L) (ANOVA, testing-noise effect, χ^2 = 3477.40, p < 0.0001). The decoding accuracy in Noise2(HL), Noise3(HH), and Noise4(LH) for very low SNRs was ~30%–40%, close to chance performance (33.33%) for all groups. The difference in threshold between Noise1(LL) and Noise3(HH) and Noise4(LH) corresponded closely to the testing-noise effect in the behavioral performance (Figure 3). Unlike the behavioral performance, decoding performance in Noise2(HL) was more similar to that in Noise3(HH) and Noise4(LH) than to Noise1(LL) (Figure 4L), suggesting a more detrimental interference of higher-temporal modulation backgrounds with the vocalization encoding of A1 neurons.

Overall, the testing-noise statistics influenced the decoding performance of A1 responses, as predicted from the behavioral performance. The relatively low decoding performance for testing in Noise2(HL) indicates that behavioral performance cannot be fully predicted based on considering only a single cortical field.

Influence of Noise-Rearing on Cortical Decoding Performance—Next, we determined decoding accuracy for the "no-noise" ("Inf" dB SNR) condition to evaluate the contributions of noise-rearing (Figure 4G). Both N2T and N4T groups slightly improved their decoding accuracy for the no-noise condition by 5%–7.5%, compared to the unexposed but trained animals (median, T, 81.67%; N2T, 89.17%; N4T, 86.67%; C, 74.17%; N2, 90.83%; N4, 79.58%; Kruskal-Wallis, $\chi^2 = 204.45$, p < 0.0001). This is consistent with the increased sensitivity revealed by behavioral testing after noise-rearing (Figures 3A–3F). Similarly, for the untrained animals, rearing in Noise2(HL) showed an increase in decoding accuracy over the unexposed group (C). However, this did not hold for rearing in Noise4(LH) (Figure 4G, green and olive boxes).

Further, when tested in noise, the vocalization-in-DMR decoding functions shifted upward for both the N2T and N4T groups (solid lines in Figures 4H–4K), resulting in lower SNR thresholds than for the Trained group (ANOVA, group effect, group x testing-noise, $\chi^2 > 174.58$, p < 0.0001) (Figures 4L and S3A). Significant decoding threshold improvements with noise-rearing were 7–13 dB in N2T and N4T animals, which are comparable to the behavioral improvements (Figures 3F, 4M, and 4N).

The experimental groups with no behavioral training also showed some decoding improvement (dashed lines in Figures 4H–4K) but with a more prominent dependence on the type of exposure noise. Whereas N2 animals showed a 7–11-dB decoding threshold improvement, N4 animals did not reach significant differences compared to the Control group, with the exception of testing in Noise3(HH) (Figures 4L and S3A).

The decoding accuracy between the trained and untrained groups (C versus T; N2 versus N2T) showed no significant differences (Figure S3B). Thus, vocalization detection training had no evident influence on vocalization-in-DMR decoding for these conditions. N4T animals, however, did show a benefit of behavioral training compared to the untrained N4 group. This interaction between noise-rearing and training indicates that at least for some noise statistics, active behavioral engagement, either during or after the critical period, can enhance SiN processing by improving cortical signal representations.

Overall, these findings indicate that SiN performance of A1 neurons can be improved by noise-rearing and can depend on the noise statistics experienced by the animals. Adding behavioral training can further enhance noise tolerance depending on the exposure statistics. The similarity of improvements for cortical decoding and behavioral performance with noise-rearing suggest high correlations between cortical response properties and behavioral capacity (Figures 4M and 4N).

Plasticity of Modulation Transfer Functions with Noise-Rearing

The improvements in behavioral and neuronal SiN processing following noise-rearing suggested that cortical RFs were adjusted during development to optimally contribute to this task. We therefore investigated the structure of STRFs and, particularly, the modulation or "ripple" transfer functions (RTFs) to reveal potential modulation preference changes due to exposure and training. We studied 842 single units (SUs) in A1 across the six experimental groups (T, 59 SUs; N2T, 246; N4T, 90; C, 46; N2, 359; N4, 42). As the trained and tested rat vocalization frequencies ranged from 20 to 30 kHz, we targeted the corresponding characteristic frequency (CF) region (median CF, 25.5 kHz; 25th to 75th percentile range, 20.8–27.8 kHz). The CF distribution in each group was similar and uniform with only slight variations for the N2T group (Kolmogorov-Smirnov test, T versus N2T, p = 0.01; N2T versus N4T, p = 0.002; N2T versus C, p = 0.01; N2T versus N2, p < 0.0001) (Figure 5A).

STRFs were obtained through the spike-triggered average of the responses to the unbiased DMR (Noise5), and RTFs were computed using the 2D Fourier transform of each STRF (Atencio and Schreiner, 2010, 2016; Escabi and Schreiner, 2002) (Figure 5B). From the RTFs, we computed the best TMFs (bTMFs) and best SMFs (bSMFs) (see Method Details).

The median bTMFs of nearly all trained or exposed groups (T, 9.21 Hz; N2T, 7.85 Hz; N4T, 9.03 Hz; N2, 8.53 Hz; N4, 8.83 Hz) showed only a modest increase relative to the untrained, unexposed group (C, 7.19 Hz) (Kruskal-Wallis, group, $\chi^2 = 27.60$, p < 0.0001, post hoc test shown in Figure 5C). The bTMFs of the Trained group showed the largest increase (by ~2.5 Hz) over the Control group.

For bSMFs, noise-rearing and/or training resulted in moderate shifts either up or down, depending on rearing conditions. The N4T group yielded the largest effect with a significantly lower bSMFs compared to most of the other groups (median: T, 0.24 cyc/s; N2T, 0.35 cyc/s; N4T, 0.18 cyc/s; C, 0.24 cyc/s; N2, 0.42 cyc/s; N4, 0.32 cyc/s; group, $\chi^2 = 78.39$, p < 0.0001, post hoc test shown in Figure 5D). By contrast, the two groups with Noise2(HL)-rearing (N2T and N2) tended to show higher bSMFs.

Maximum TMFs (maxTMFs) and SMFs (maxSMFs) and bandwidths of temporal and spectral modulation transfer functions (tMTFs and sMTFs) (see Method Details) were also affected by noise-rearing (Figure S5). tMTF bandwidths broadened most prominently in the Trained group (Figure S5C). maxSMFs and sMTF bandwidths showed opposite changes for Noise2(HL)- and Noise4(LH)-rearing (Figures S5B and S5D). None of these modulation properties were dependent on CFs (for all: Spearman correlation coefficient, R < |0.04|, p > 0.26).

Although comparing modulation properties across groups indicated significant effects of noise-rearing and behavioral training, collapsing RTFs along the temporal or spectral axis loses information regarding the joint temporal and spectral modulation preferences that characterize the different exposure noises. Therefore, we constructed the population RTF (pRTF) for each experimental group and computed the paired differences of the joint modulation preferences (Figures 6, 7, and S7). For statistical analysis, we divided the spectral and temporal modulation ranges into nine distinct sectors of the pRTF, taking into

account the regions corresponding to the two rearing noises (bold dashed lines in Figures 6D–6F, 7D–7F, S7B, and S7C; see Method Details).

The pRTF for the Trained group (Figure 6A) showed generally low temporal and low spectral modulations, similar to the distribution of modulation power in rat vocalizations (Figure 1B). Noise-rearing induced specific pRTF changes (Figures 6B–6F). N2T animals reduced the modulation power at SMFs of the exposure noise and increased it above the range of the exposure noise compared to the Trained group (Figure 6B). Temporal modulation power also increased when associated with higher SMFs not contained in Noise2(HL) (yellow asterisk on red region in Figure 6D; Bootstrap, $p = 0.009$; see Method Details), but was reduced when associated with SMFs contained in Noise2(HL) (asterisk in blue region; $p = 0.003$). The pRTF region corresponding to the exposure noise (bold dashed outline in Figure 6D) appeared to be slightly reduced, although without reaching statistical significance. Similarly, for Noise4(LH)-rearing, the difference between N4T and T shows an increase in spectral modulation power outside the exposure range combined with an increase in lower TMFs for lower SMFs (green asterisk, $p = 0.045$) (Figure 6E). This also supports the notion that noise-rearing suppressed the representation of SMFs in the exposure range and induced a gain for the surrounding modulation combinations. These divergent, noise-specific effects for N2T and N4T groups are reflected in the pRTF difference between the two conditions ($p < 0.03$) (Figure 6F).

For the groups without behavioral training, Noise2(HL)-rearing similarly altered the pRTF with a reduction in the spectral modulation power contained in Noise2(HL) and an increase of joint spectro-temporal modulations outside the exposure range (Figures 7A, 7B, and 7D). This is consistent with the behaviorally trained group. The N4 group indicated several gain changes, including losses at higher SMFs and gains at lower SMFs (Figures 7A and 7C). Joint gains in lower SMFs and higher TMFs, again, were outside the range of the exposure stimulus ($p < 0.05$) (Figure 7E). pRTF differences between the N2 and N4 groups (Figure 7F) show a similar pattern to the groups with behavioral training (Figure 6F), particularly regarding the increase in higher spectral modulation power for Noise2(HL)-rearing and an increase in lower spectral modulation power for Noise4(LH)-rearing.

When examining the effects of the two exposure stimuli (Figures 6F and 7F), it appears that developmental exposure pushes the neuronal modulation preference of neurons away from the exposure range and increases the number of neurons with modulation preferences outside that range. A consequence of noise-rearing would be a reduced number of cortical neurons tuned to modulations of the exposure noise and an enhanced SNR in the neurons tuned to modulations falling outside of the exposure noise.

Finally, we evaluated how behavioral training affected pRTFs (Figure S7). Trained animals without prior noise exposure showed an increase of higher TMFs (20–40 Hz) and lower SMFs (0–0.3 cyc/oct) accompanied by a decrease of lower TMFs (0–20 Hz) and higher SMFs (>2 cyc/oct) ($p < 0.005$) compared to unexposed, untrained animals (T versus C) (Figure S7A). Behavioral training after Noise2(HL)-rearing showed gains at lower SMFs and suppression at higher SMFs ($p < 0.01$) relative to the untrained N2 group (N2T versus N2) (Figure S7B). Behavioral training after Noise4(LH)-rearing showed gains at lower and

higher SMFs and a loss of power at the middle SMFs (0.3–2 cyc/oct), especially for lower TMFs (p < 0.05) (N4T versus N4) (Figure S7C). Thus, the effect of behavioral training was dependent on the prior, exposure-mediated modulation power distribution. Training-induced gains could be interpreted as aiding vocalization representation (~5 Hz and ~0.1 cyc/oct; see Figure 1B).

Collectively, our physiological results show that noise-rearing influenced the development of auditory STRFs by reducing responsiveness to the background statistics presented during the critical period. By contrast, behavioral training shifted RFs toward enhancing responsiveness to a trained foreground sound. Both changes—the reduction of neuronal responsiveness to the noise and the increase in responsiveness to the foreground — contribute to increasing the SNR in the neuronal responses and likely account for the animals' improved behavioral performance following noise exposure.

DISCUSSION

The aim of this study was to identify emerging principles of cortical SiN processing by creating animals that excel at such a task, thereby revealing advantageous RF structures to detecting foreground signals in the presence of background noise. This was accomplished by rearing animals in the presence of different modulated noises and testing their behavioral and neuronal vocalization-in-DMR processing ability. We found that experiencing moderately loud modulated noise during the auditory critical period improved the behavioral SiN detection ability of adult rats. Similarly, A1 neurons improved their decoding performance of vocalizations embedded in noise. This was accompanied by a shift of the joint spectro-temporal modulation preferences toward values different from the dominant modulations presented during noise-rearing. As a consequence, fewer neurons were responsive to the experienced noise, thus improving the neuronal SNR. These results highlight (1) the functional utility of the joint spectro-temporal modulation properties beyond that of foreground encoding, (2) a beneficial plasticity to environmental noise statistics, and (3) a well-matched similarity of the induced perceptual and neural noise tolerance.

Approach

To create specialized animals, we reared rats in synthesized naturalistic noise throughout the course of development and into adulthood. In this natural training paradigm, rats were forced to perceive signals of interest in the face of noise during a time when the brain is most plastic and can adapt to the acoustic environment (de Villers-Sidani et al., 2007, 2008; Insanally et al., 2009; Nakahara et al., 2004; Zhang et al., 2002; Zhou and Merzenich, 2008). To avoid potentially pathologic sound pressure levels while still creating a challenging listening condition, we exposed animals to slowly changing, biased noises at a moderate level of 60 dB SPL (Cappaert et al., 2000; Ward et al., 1976). Alternatively, one can create specialized animals by directly training them to perform an auditory task in the face of noise (Shetake et al., 2011; Whitton et al., 2014). This more direct approach could result in training-induced changes more specific to the particular foreground signal rather than general adaptive changes to the noise. In our present study, plasticity was not solely based on

passive exposure, but driven by active behavioral engagement for the animals to be able to improve their sound-dependent behavior in such a noisy environment. Furthermore, although we focused on the critical period to maximize the noise-rearing effect, cortical plasticity to our modulated noise is likely to occur in a similar manner in adult animals (Noreña et al., 2006; Pienkowski and Eggermont, 2009; Pienkowski et al., 2011).

Noise-Rearing Effects

We hypothesized that noise-rearing improves SiN processing specific to the exposed spectro-temporal modulation statistics. Behavioral SiN performance indeed improved thresholds by up to 16 dB in noise-reared animals. This beneficial effect of noise-rearing differs from previous reports that showed some behavioral deficits or no clear improvement following critical period exposure to various sounds, such as narrow or broadband signals or speech (Han et al., 2007; Ranasinghe et al., 2012; Sun et al., 2011; Zhang et al., 2008). This discrepancy may be due to some differences in the exposure sounds. We raised animals in moderately loud DMRs without sharp onsets and offset transients. This avoided the highly synchronized activity shown to disrupt cortical development (Zhang et al., 2002) while allowing exposed pups to remain responsive to vocalizations by their mother and siblings. That is, the observed behavioral improvements may not solely be due to passive exposure effects but may also be influenced by active learning processes invoked by the pups and young adult animals in their attempt to detect vocalizations (Zheng, 2012). Overall, performance improvements generalized to all noise types with modest trends indicate better sensitivity for noises with statistics similar to the rearing environment (Figure 3).

Similar to the behavioral results, decoding accuracy for A1 neurons generally improved for noise-reared animals, with better thresholds for all tested noises and small variations with the exposure statistics. Previous studies in A1 also had shown that noise invariance can depend on the noise types (Mesgarani et al., 2014; Narayan et al., 2007; Ni et al., 2017). The relatively large scatter in the decoding data may be related to the fact that not all cortical neurons process SiN in a linear manner, as some neurons show nonmonotonic behavior (Malone et al., 2017). Overall, the average enhancement of cortical decoding paralleled the behavioral changes in direction and magnitude (Narayan et al., 2007; Schneider and Woolley, 2013).

Exposure-Induced RF Plasticity

Our results showed that rearing with biased DMRs boosted the cortical modulation power for ranges different from the presented modulation statistics. While the marginal distribution of the RTFs was only moderately affected, the joint distribution of spectro-temporal modulations showed extensive changes. Specifically, the noise-rearing did not result in a gain of the representation of the presented noise parameters, as would have been predicted from plasticity studies of other sound features (Bao et al., 2013; de Villers-Sidani et al., 2007; Insanally et al., 2009; Zhang et al., 2002) and enriched auditory environments (Grécové et al., 2009; Pysanenko et al., 2018). While previous studies showed an overrepresentation of the exposed sounds, we observed an underrepresentation of the exposed noise parameters. There was also no sign of maturational delays or deteriorations, as has been seen for noise-rearing with stimuli dominated by sharp onsets or lacking defined

structure (Chang and Merzenich, 2003; de Villers-Sidani et al., 2008; Zhou and Merzenich, 2008).

We designed our noise-rearing stimuli to be continuous (i.e., without clear onsets or offsets) and with a modulation content found in many vocalizations. This also differed from other natural sounds that usually have wider ranges of modulation statistics (Chi et al., 1999; McDermott and Simoncelli, 2011; Singh and Theunissen, 2003). The selected temporal and spectral modulations were chosen to produce opposing effects on the pRTFs. Indeed, the boost of modulation power away from the presented sound statistics can be interpreted as repellent with an increase of modulation power outside the exposure range. This is similar to frequency tuning shifts induced by band-limited noise in adult cats (Noreña et al., 2006; Pienkowski and Eggermont, 2009; Pienkowski et al., 2011) with suppression in the exposed frequency range and enhancement at its edges. Combined, this indicates that the relative representation of less informative—"background"—sounds is sacrificed in favor of more informative—"foreground"—signal aspects, which could be seen as a manifestation of efficient coding principles (Barlow, 1961) in the SiN problem. Temporal and spectral modulation filters play a key role in achieving noise tolerance. Use of modulation filter banks, a series of 2D filters arranged by SMFs and TMFs, can improve automatic speech recognition (Ganapathy et al., 2009; Mesgarani et al., 2006; Schädler et al., 2012). Our observed RF changes are consistent with models of STRF preferences that promote noise invariance (Moore et al., 2013). The results provide explicit physiological evidence for neuronal adaptation to maximize differences in the representation of foreground and background sounds by reshaping RTFs to reduce responses to background noise and enhance responses to foreground sounds, resulting in a more favorable neural SNR.

Effect of Background Noise Statistics

Both behavioral performance of SiN detection and electrophysiological decoding accuracy for vocalizations in DMRs were better when testing with noises containing low SMFs. Modulated noise enables "dip-listening" or "glimpsing" in which the momentary SNR condition for a foreground signal is more favorable by either falling into a wide spectral valley or a longer temporal modulation minimum (Cooke, 2006; Howard-Jones and Rosen, 1993; Stone and Canavan, 2016). Behavioral thresholds were ~10 dB lower for noises with low SMFs (Noise1&2), regardless of TMFs. A similar effect was observed in neuronal decoding accuracy for Noise1 (with low SMFs and low TMFs). Interestingly, Noise2 (containing low SMFs but high TMFs) did not show better thresholds than the noises with high SMFs (Noise3&4). While this does suggest that low TMFs do provide an advantage over higher TMFs, the spectral modulation content seems to influence detection performance (but see Souffi et al., 2019). A potential cause of this discrepancy between behavioral and neural performance may be related to cortical area specificity, as spectral and temporal properties substantially vary between different core and belt areas (Eggermont, 1998; Schönwiesner and Zatorre, 2009; Schreiner and Urbas, 1988). Background noise influences on signal encoding, however, are present throughout the auditory system, and more noise-tolerant responses emerge toward the thalamus and cortex (Las et al., 2005; Rabinowitz et al., 2013; Schneider and Woolley, 2013). Further studies are necessary to better understand the link between neuronal and perceptual performance in SiN processing.

Training Effects

Perceptual learning can alter cortical stimulus representations (e.g., Ohl and Scheich, 1996; Polley et al., 2006; Schnupp et al., 2006), but changes are not always detected (e.g., Brown et al., 2004; Reed et al., 2011; Talwar and Gerstein, 2001). Here, the RFs of animals that underwent behavioral testing showed a moderate increase of modulation power characteristic of the trained vocalizations compared to untrained animals. Thus, behavioral training interacted with some of the changes induced by noise-rearing. Nonetheless, RF structures and neural responses in A1 changed to improve behavioral performance during task engagement (Atiani et al., 2009; Fritz et al., 2003; Lee and Middlebrooks, 2011). To capture training effects more fully, attentional control of signal decoding may need to be considered.

Consequences for Human Performance

We showed that noise-rearing altered RTFs and achieved better SiN processing for a range of noise statistics. When moving from one background context to another, SiN performance can be maintained either by fast adjustments of neurons to sound statistics (e.g., Dean et al., 2005; Kvale and Schreiner, 2004; Malone et al., 2015) or by selecting a subpopulation of neurons that provide the most favorable neural SNR (i.e., neurons not tuned to the modulation properties of the background). Training adult humans in a speech-in-noise task is effective and is likely to influence the modulation representation, at least temporarily (Song et al., 2012; Wayne et al., 2016; Whitton et al., 2014, 2017). Elderly people and hearing-impaired listeners commonly experience difficulty in segregating speech from background noise, often accompanied by reduced sensitivities to modulated sounds (Füllgrabe et al., 2015; Mehraei et al., 2014; Ruggles et al., 2011). In the present study, exposure to biased noise during maturation shaped the modulation representation and enhanced SiN processing for that noise and similar noises. These changes, however, may reduce performance for environmental noises that substantially differ from those presented during maturation. Furthermore, strongly dominant and biased noise distributions during maturation (e.g., incubator noise, traffic sounds, or air conditioning noises) may create long-term conditions not equally favorable for performance in other ambient noise conditions. Both human infants and adults are capable of active and passive statistical learning (Kang et al., 2014; Saffran et al., 1996, 1999; Whitton et al., 2017). Thus, a combination of perceptual training and exposure to noise with relevant statistics may improve adult modulation sensitivities and enhance SiN processing abilities at all ages.

STAR★METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Christoph Schreiner (chris@phy.ucsf.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of California San Francisco and were carried out in accordance with NIH guidelines.

Thirty-two female Sprague-Dawley rats (wild-type) sourced from Charles River were used in this study. Nineteen of them (19/32) were raised in a sound-shielded test chamber under a biased DMR presentation, and twelve of the noise-reared animals (12/19) were tested for a rat vocalization-in-DMR detection task. The rest of thirteen (13/32) were raised in a typical rat housing, and five of them (5/13) were underwent for the behavioral training. A 12 h light/12 h dark cycle was established for all of the groups, and they were co-housed in groups of 2-3. At the end of noise-rearing or behavior testing, an extracellular recording was performed on all of them (age: P79-152, weight: 242-330 g) (Figure 2A). For the Control group with no noise-rearing and no behavioral training (8/32), recording was performed matching their ages to other experimental groups.

METHOD DETAILS

All sound stimulus generation, data analysis and statistical analysis were performed with MATLAB (Mathworks). RStudio (https://www.rstudio.com/) was also used for specific statistical tests.

Biased DMR noise synthesis—We prepared an unbiased DMR comprised of 50 sinusoidal carrier frequencies per octave, each in random phase, in the range of 0.5 to 40 kHz. The carrier structure was modulated by a spectro-temporal envelope with 0 to 4 cyc/oct spectral modulation frequencies (SMFs), 0 to 40 Hz temporal modulation frequencies (TMFs), and a maximum 40 dB spectro-temporal modulation depth (Atencio et al., 2008; Escabi and Schreiner, 2002; See et al., 2018). For the biased DMRs, the minimum and maximum SMFs were restricted to 0 to 0.3 cyc/oct for low SMFs and 0.3 to 2 cyc/oct for high SMFs. Similarly, the minimum and maximum TMFs were 0 to 5 Hz for low TMFs and 5 to 20 Hz for high TMFs. We refer to the biased DMR with low SMFs and low TMFs as "Noise1(LL)," and the noise with low SMFs and high TMFs as "Noise2(HL)." "Noise3(HH)" has high SMFs and high TMFs; "Noise4(LH)" has high SMFs and low TMFs. For convenience, unbiased DMR with full range of SMFs and TMFs is referred to as "Noise5(full)."

Modulation power spectrum estimation—We visualized a time-frequency representation of rat vocalizations using short time Fourier transform with a 10 ms Hanning window. Because we wanted the modulation power spectrums (MPSs) of rat vocalizations and DMR noises to be comparable, rat vocalization spectrograms were obtained using a set of 128 filters equally spaced on an ERB_N scale (Glasberg and Moore, 1990) spanning 0.5 to 40 kHz, with 8 ms half-cosine frequency responses (McDermott et al., 2011); the MPS was obtained from a time-frequency decomposition on time and frequency (on octave scale) after applying a 2D Fourier transform with 125 ms time window. The rat vocalization MPS was computed as an average of 18 stereotypical vocalization segments after normalizing the cumulative area of each MPS to be exactly 1 to weight them equally. Similarly, MPSs of

DMR noises were obtained as an average of 332 files of 4 s-noise segment applying 2D Fourier transform on the DMR spectrograms. Adjacent filters and time windows overlapped by 50%.

DMR noise-rearing—Litters of rat pups and their mother were placed in a sound-shielded test chamber at P6. A DMR with different stimulus statistics was presented to rat litters 24 h/d at a sound level of ~60 dB SPL from a loudspeaker mounted at 15 cm above the cage using a sound amplifier (DV-610AV, Pioneer) and a DVD player (Servo120a, Samson) at a 96 kHz sampling rate. No abnormal behavior of either the mother or pups was detected during noise-rearing. A biased DMR (either Noise2(HL) or Noise4(LH)) was used for noise-rearing based on rat vocalization MPS to contrast the developmental effects of noise structures.

Rat vocalization-in-DMR detection behavior paradigm—The animals were trained by positive reinforcement with food reward (dustless precision pellet, Bio-Serv) using a go/no-go paradigm. Blocks of 5 days of training were interspersed with 2 days off. All the behavioral training was conducted in an operant training chamber (26 cm x 32 cm x 20 cm, length x width x height) within a sound attenuated chamber controlling a photobeam detector, food dispenser, sound presentation and house light by MED-PC IV (Med Associates, Inc.) (Figure 2B). Animals' behavior was monitored via a USB camera (1.3 Megapixel, ELP).

The target stimuli were two of the rat vocalizations (upper two of Figure 1A). The vocalization was presented at 60 dB SPL, and the DMR noise were varied in 10 levels (0, 30, 45, 50, 55, 60, 65, 70, 75 or 80 dB SPL) to produce Inf, +30, +15, +10, +5, 0, −5, −10, −15, −20 dB SNRs. The SNR values were computed by contrasting the signal level and overall noise level. If only the power of vocalization frequency bands in the noise is considered, noise level at the band is estimated as:

$$10 \times \log_{10}(29 \, / \, 316 \times (10^{\wedge}(ONL \, / \, 10)))dB,$$

where ONL is overall noise level, since the total number of carrier frequencies in the DMR is 316 and 29 of them fall in the vocalization frequency band (20-30 kHz), and the hardest condition is assumed to be −10 dBSNR. Four biased and one normal DMR noises were used as a background noise. Auditory stimuli were presented using an audio interface (QUAD-CAPTURE, Roland) at a 96 kHz sampling rate with a power amplifier (Servo200, Samson). Sound calibration was performed with an 1/4-inch pressure field microphone (Type4939, Bruel and Kjar) prior to behavioral training and re-calibrated regularly over testing periods.

The animals were first trained to initiate a nose poke at the nose-poke hole when rat vocalization stimulus was presented in silence and to obtain a reward at the food dispenser. In the testing, the SNR levels were varied from −20 to Inf dB SNR. The waiting time preceding the vocalization stimuli was randomized from 3 to 12 s to reduce the predictability of the timing of the go signal. After 12 warm-up trials with no DMR noise, we presented blocks of 6 trials, where the identity of DMR noise was randomized on each trial, and the SNR levels were randomized for each block.

After misses (no initiation at the nose-poke hole after the vocalizations were presented) or false alarms (initiation at the nose-poke hole while no vocalization was presented), house light was turned off as feedback and to signal the lack of reward. Following a miss, a 3 s time-out was given, whereas a 30 s time-out was used after false alarms to reinforce the performance of the animals on trials that required a long waiting time. The probability of trials that comprised vocalization (go trials) was 60%, and the probability of trials that comprised no vocalization (no-go trials) was 40% based on previous work (Schnupp et al., 2006).

Analysis of rat vocalization-in-DMR detection behavior—Behavioral analysis for go/no-go paradigm was previously described in detail (Homma et al., 2016, 2017). Behavioral performance was assessed by hit rate (number of correct responses on go trials / total number of go trials) and false alarm (FA) rate (number of incorrect responses on no-go trials / total number of no-go trials). Animals were advanced from training to testing when their hit and FA rates became higher than 80% and lower than 20%, respectively, for four consecutive days with < 10% standard deviation (SD). Individual sessions with FA rates > 0.30, indicative of a lack of attention, were excluded from the analysis (< 14% of collected data for each animal). FA rate for all SNR conditions was 0.13 ± 0.04 (MEAN ± SEM) (T, 0.19 ± 0.02; N2T, 0.11 ± 0.02; N4T, 0.07 ± 0.006). When animals find the task difficult, they tend to increase FA rate to maintain hit rate. Thus, the larger FA rate for group T reflected the lower sensitivity for that group.

Using signal detection theory (Wickens, 2002), a sensitivity index (d') was calculated for each DMR noise and SNR condition from the z-transformed hit rate and FA rate. Thresholds were defined as the first point to cross a criterion of $d' = 2.4$, which corresponds to a hit rate of 0.87 to 0.94 with a 0.10 to 0.20 FA rate. When all the points were above the criteria, the threshold value was set to the lowest SNR condition value, −20 dB SNR. Although a lower d' value is typically used for threshold criteria, we chose $d' = 2.4$ in order to obtain threshold values from each noise type at equivalent criteria. This was necessary because the sensitivities for Noise1(LL) and Noise2(HL) were higher than the other noise types, and the comparisons of noise types in this present study were conducted well above classical just noticeable differences. Patients with profoundly impaired hearing show difficulty of hearing speech in noisy background with increased threshold of 15-20 dB SNR, and use of hearing aids improved it by ~10 dB SNR (Killion and Niquette, 2000; Taylor, 2003). Thus, 9 ± 4 dB (MEAN ± SD) improvements in our present study indicate an effect size of clinically significant impact. In addition to threshold values, we estimated the area under the curve (AUC) using a trapezoidal rule for numerical integration over the d' values on a log scale from the minimum to maximum degree of SNR conditions used. Mean psychometric functions were derived by fitting the mean d' values using unconstrained nonlinear optimization to a cumulative Gaussian distribution:

$$D'(x \mid \alpha, \mu, \sigma) = \alpha \left(\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt \right),$$

where x is SNR condition (in dB SNR), α is constant to rescale the function in the y direction, μ is the mean, and σ the standard deviation.

Response time was measured as the time from target vocalization onset until animals made a nose poke. Because there was a large individual variability in their response times and the distributions, we calculated a median response delay for each animal for the combination of DMR testing-noises and SNR conditions by individually subtracting a unique median response time at Inf dB SNR (no noise condition).

Electrophysiological recording surgical procedure—The detailed procedures were described in previous studies (Polley et al., 2007; See et al., 2018). Anesthesia was induced with a single intraperitoneal injection of ketamine hydrochloride (Ketathesia, 100 mg/kg, HenrySchein) and xylazine hydrochloride (AnaSed, 3.33 mg/kg body weight, Akorn), and supplemented as needed with a mixture of ketamine (10-50 mg/kg) and xylazine (0-20 mg/kg). Atropine sulfate (AtroJectSA, 0.54 mg/kg, s.c., HenrySchein) and dexamethasone sodium phosphate (Dexium-SP, 4 mg/kg, s.c., Bimeda) were administered to minimize pulmonary secretions and prevent cerebral edema, respectively. Perioperative analgesia was provided with meloxicam (Eloxiject, 2 mg/kg, s.c., HenrySchein). Lidocaine hydrochloride (Lidoject, 2%, s.c., HenrySchein) was used as a local anesthetic prior to secessions. Atropine and lactated ringer solution were administered every 6 h to reduce the viscosity of bronchial secretions and to prevent dehydration. Depth of anesthesia and respiratory rate were monitored and maintained throughout the experiment. Temperature was monitored using a rectal probe, and maintained at 37°C using a homeothermic blanket system (55-7020, Harvard Apparatus). The eyes were protected with lubricant ophthalmic ointment (Aritificial tears, HenrySchein).

A tracheotomy was performed to access the airway and maintain breathing during recording, and the rat was placed in a head clamp manufactured in-house that allowed unimpeded access to the external auditory meati. One side of the temporal muscle was removed, and the skull was revealed. A craniotomy window 3 mm x 5 mm was made over the auditory cortex. Once the dura was removed, the cortex was covered with silicone oil. Recordings with tungsten electrode (WE3003A5, 0.5-0.75 mΩ, MicroProbes) using pure tones were conducted to locate the primary auditory cortex (A1) based on the tonotopic organization of rat auditory fields and their properties (Polley et al., 2007). Typically 20-30 sites were investigated to identify the frequency reversal border of A1 and the ventral auditory field, and then the targeted high frequency areas (20-30 kHz) suited for the rat vocalization stimulus. Once A1 was identified, we inserted a multi-contact silicon probe array (Neuronexus or Cambridge NeuroTech), with 32 or 64 recording sites vertically arranged and spaced at 20 or 50 mm intervals in the dorso-ventral direction and 50 mm in the medio-lateral direction. The probe was slowly advanced perpendicularly to the pial surface until all the electrode contacts were in the cortex (~0.9 to 1.35 mm).

Electrophysiological recording stimulus presentation—Stimulus generation and data acquisition were controlled with a PC using RHD2000 Interface software (Intan Technologies). Auditory stimuli were presented using an audio interface (QUAD-CAPTURE, Roland) at a 96 kHz sampling rate with a power amplifier (SLA1, Applied

Research and Technology). All auditory stimuli were delivered contralaterally from the recording site using an electrostatic speaker (EC1, Tucker-Davis Technologies), with an inverse filter generated by closed field calibrations using an 1/2-inch pressure field microphone (Type4192, Brüel and Kjær).

In mapping experiments with tungsten electrodes, we presented pure tones (50 ms, with 5 ms cosine ramps) ranging from 0.5 to 32 kHz in 0.13 octave steps in frequency and from 0 and 70 dB SPL in 5 dB increments in level. Each frequency-level combination was presented pseudorandomly once at a rate 0.5 Hz. We did not observe any abnormal tonotopic organization. After the high frequency region of A1 was identified, the recording setup was switched to multi-electrode arrays, and an unbiased DMR (Noise5) was presented at 70 dB SPL. A sequence of rat vocalizations was also presented in the presence of biased DMR (Noise1–4) (Figure 1C), in which one of the four different noise types was randomly presented every 6 s at a different level to compose six different SNR conditions (−20, −10, 0, 10, 20, Inf dB SNR). In other words, one of three vocalizations (Figure 1A) was presented at 60 dB SPL after 3 s of the noise onset in a 6 s segment, and the noise level was randomized between -Inf (no noise), 40, 50, 60, 70, and 80 dB SPL. Each vocalization was repeated 10 times for each noise and SNR condition.

Electrophysiological recording pre data processing—Neural traces were band-pass filtered between 500 and 6000 Hz and were recorded to disk at 20 kHz sampling rate with an Intan RHD2132 Amplifier system (Intan Technologies). Spike trains were sorted offline using MountainSort, a fully-automated spike sorter (Chung et al., 2017). To obtain single units, we used the same metric threshold as Chung et al. (noise overlap < 0.03, isolation > 0.95, firing rate > 0.1 Hz, SNR > 1.5), and only analyzed units with a percentage of interspike intervals indicative of refractory period violations (1.5 ms) less than 2%. For multi-units, spikes of the sorted units from the same electrode channel were compiled. The metric thresholds used for this analysis were noise overlap < 0.35, isolation > 0.8, firing rate >0.1 Hz, SNR >1.2.

Analysis of neural responses to vocalization-in-DMR stimuli—All the data presented for vocalization-in-DMR stimuli are multi-units whose spike counts following vocalization onset significantly differed from the spike counts in windows of the same duration just preceding stimulus onset (Wilcoxon signed rank test, p < 0.0001). We used a nearest-neighbor linear decoder (Foffani and Moxon, 2004) to estimate which vocalizations elicited the response on each trial. The details of this technique have been described in previous studies (e.g., Malone et al., 2017). Briefly, responses to each stimulus were binned to form a bin-dimensional vector representing the response across time and averaged across trials to form a response template for each stimulus. Individual trials were similarly binned, and the Euclidean distance was computed between the trial and each stimulus template vectors. After identifying the stimulus associated with the nearest template, each decoded trial was indexed by a confusion matrix whose columns represent the actual stimulus and whose rows represent the stimulus estimated. Each cell in the matrix indicates the number of counts (out of 10, for 10 trials) that a spike train elicited by the actual stimulus (row) was assigned to each estimated stimulus (column). Correctly identified trials fall along the

diagonal, and the sum of the diagonal entries divided by the total number of trials is referred as decoding accuracy, which indicates the percentage of trials whose responses were correctly assigned to the stimulus that elicited them.

For vocalization-in-DMR decoding, we analyzed the vocalization presentation window (3 to 4.795 s) in a 6 s segment. We report the percentage of correctly classified trials for a 10 ms binning resolution. We assigned significance values for decoding accuracy by comparing the actual confusion matrices obtained against simulated confusion matrices generated by Monte Carlo methods that perturbed the stimulus-response relationship by circular shifting the spike trains. Decoding accuracy for the actual confusion matrices was compared against the distribution of simulated decoding accuracies. *P values* were assigned by counting the fraction of simulated accuracies that exceeded the actual accuracy, divided by the total number of simulated values (n = 100). We computed median decoding accuracies for each combination of group and SNR condition from units with significant vocalization decoding performance (p < 0.01).

Analysis of auditory receptive fields and modulation properties—STRFs were obtained based on the spike-triggered averaging of the responses to a DMR containing a wide range of modulations (Noise5) that covered the modulation preferences of A1 neurons, and RTFs were computed using 2D Fourier transform of each STRF (Atencio and Schreiner, 2010, 2016; Escabi and Schreiner, 2002). We used the RTFs to estimate temporal and spectral modulation transfer functions (tMTFs and sMTFs) by summing the RTF along the spectral modulation axis for the tMTF and along the temporal modulation axis for the sMTF, respectively. bTMF and bSMF were computed as peak values of tMTFs and sMTFs, respectively. In addition to bTMF and bSMF, the maxTMF and maxSMF were computed as a maximum frequency above 70% of the power in each MTF; tMTF and sMTF bandwidths were obtained as a cumulative distance above 70% of the power.

pRTFs were estimated by averaging individual RTFs after normalizing each of them to sum to 100% to avoid biasing the pRTF because of systematic differences in firing strength (Escabi and Schreiner, 2002; Hullett et al., 2016). Significant differences between pRTFs were determined for the nine areas by dividing each pRTFs into a combination of three spectral (0 to 0.3, 0.3 to 2, 2 to 4 cyc/oct, horizontal dashed lines in Figures 6, 7, and S7) and temporal (0 to 5,5 to 20,20 to 40 Hz, perpendicular dashed lines in Figures 6, 7, and S7) modulation frequencies based on unbiased DMR noise parameters. First, the sum of normalized power was obtained for each RTF sector, and the mean values were computed. We then compared the actual difference of mean values to simulated ones for the pairs of areas and groups. The simulated differences were generated by resampling group labels. *P values* were assigned by counting the fraction of simulated differences that exceeded the actual difference, divided by the total number of simulated values (n = 1000). In order to control unbalanced sampling number, we repeated the bootstrap test with subsets of data using the smallest number of units for the compared groups. This procedure corrected unbalanced CF distribution (*Kolmogorov-Smirnov* test with bootstrapping, p > 0.05). Although the statistical power slightly decreased, the major outcome of pRTF differences was consistent.

QUANTIFICATION AND STATISTICAL ANALYSIS

All the data was first tested for a standard normal distribution. If the null hypothesis was not rejected, data was analyzed using a one-way ANOVA as a parametric method to examine the differences in groups, and means and SDs were reported. Otherwise, data were analyzed using non-parametric methods on ranks, and medians with percentiles were computed. For comparing different groups, the Kruskal-Wallis test, a non-parametric one-way ANOVA on ranks, was used, followed by Dunn's test for post hoc comparisons. In order to investigate the effects of different treatments for groups and testing-noise structures on the vocalization-in-DMR detection/discrimination, a linear mixed effects model was fitted to the data of threshold/AUC values with background noise and group as fixed effects using RStudio. Individual rat or unit number was used as a random effect for behavioral or electrophysiological data, respectively. Two model fits with and without interaction:

$$y \sim \text{testing-noise} \times \text{group} + (1 \mid r)$$

and

$$y \sim \text{testing-noise} + \text{group} + (1 \mid r)$$

where y is threshold or AUC values and r is rat or unit number, were compared using a chi-square test. If no significant interaction was found, it was dropped off and the data were fitted to the simpler model. Normal QQ-plots confirmed the goodness of the model fit by showing that the residuals were normally distributed; 2.5%–97.5% confidence intervals were obtained by bootstrapping. For response delays in the vocalization-in-DMR detection task, bootstrap multiple comparisons were performed in order to evaluate the differences of paired testing-noises or groups. Bonferroni correction was applied on all the multiple comparisons. A significance level (alpha) of 0.05 was used as a conventional criterion for null hypothesis rejection. All statistical tests and outcomes are provided in Table S1 with *P values*.

DATA AND CODE AVAILABILITY

The dataset and code supporting the current study have not been deposited in a public repository because of on-going other projects using this dataset but are available from the Lead Contact on request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

Atencio CA, and Schreiner CE (2010). Laminar diversity of dynamic sound processing in cat primary auditory cortex. J. Neurophysiol 103, 192–205. [PubMed: 19864440]

Atencio CA, and Schreiner CE (2016). Functional congruity in local auditory cortical microcircuits. Neuroscience 316, 402–419. [PubMed: 26768399]

Atencio CA, Sharpee TO, and Schreiner CE (2008). Cooperative nonlinearities in auditory cortical neurons. Neuron 58, 956–966. [PubMed: 18579084]

Atiani S, Elhilali M, David SV, Fritz JB, and Shamma SA (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron 61, 467–480. [PubMed: 19217382]

Bar-Yosef O, and Nelken I (2007). The effects of background noise on the neural responses to natural sounds in cat primary auditory cortex. Front. Comput. Neurosci 1, 3. [PubMed: 18946525]

Bao S, Chang EF, Teng C-L, Heiser MA, and Merzenich MM (2013). Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment. Neuroscience 248, 30–42. [PubMed: 23747304]

Barlow HB (1961). Possible Principles Underlying the Transformations of Sensory Messages In Sensory Communication (The MIT Press), pp. 217–234.

Brown M, Irvine DRF, and Park VN (2004). Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex. Cereb. Cortex 14, 952–965. [PubMed: 15115736]

Cappaert NLMM, Klis SFLL, Muijser H, Kulig BM, and Smoorenburg GF (2000). Noise-induced hearing loss in rats. Noise Health 3, 23–32. [PubMed: 12689440]

Chang EF, and Merzenich MM (2003). Environmental noise retards auditory cortical development. Science 300, 498–502. [PubMed: 12702879]

Chang EF, Bao S, Imaizumi K, Schreiner CE, and Merzenich MM (2005). Development of spectral and temporal response selectivity in the auditory cortex. Proc. Natl. Acad. Sci. USA 102, 16460–16465. [PubMed: 16263924]

Chi T, Gao Y, Guyton MC, Ru P, and Shamma S (1999). Spectro-temporal modulation transfer functions and speech intelligibility. J. Acoust. Soc. Am 106,2719–2732. [PubMed: 10573888]

Chung JE, Magland JF, Barnett AH, Tolosa VM, Tooker AC, Lee KY, Shah KG, Felix SH, Frank LM, and Greengard LF (2017). A Fully Automated Approach to Spike Sorting. Neuron 05, 1381–1394.e6.

Cooke M (2006).Aglimpsing model ofspeech perception in noise. J. Acoust. Soc. Am 110, 1562–1573.

de Villers-Sidani E, and Merzenich MM (2011). Lifelong plasticity in the rat auditory cortex: basic mechanisms and role of sensory experience. Prog. Brain Res. 191, 119–131. [PubMed: 21741548]

de Villers-Sidani E, Chang EF, Bao S, and Merzenich MM (2007). Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat. J. Neurosci 27, 180–189. [PubMed: 17202485]

de Villers-Sidani E, Simpson KL, Lu Y-F, Lin RCS, and Merzenich MM (2008). Manipulating critical period closure across different sectors of the primary auditory cortex. Nat. Neurosci 11, 957–965. [PubMed: 18604205]

Dean I, Harper NS, and McAlpine D (2005). Neural population coding of sound level adapts to stimulus statistics. Nat. Neurosci 8, 1684–1689. [PubMed: 16286934]

Eggermont JJ (1998). Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. J. Neurophysiol 80, 2743–2764. [PubMed: 9819278]

Elliott TM, and Theunissen FE (2009). The modulation transfer function for speech intelligibility. PLoS Comput. Biol. 5, e1000302. [PubMed: 19266016]

Escabi MA, and Schreiner CE (2002). Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. J. Neurosci 22, 4114–4131. [PubMed: 12019330]

Foffani G, and Moxon KA (2004). PSTH-based classification of sensory stimuli using ensembles of single neurons. J. Neurosci. Methods 135, 107–120. [PubMed: 15020095]

Friedmann N, and Rusou D (2015). Critical period for first language: the crucial role of language input during the first year of life. Curr. Opin. Neurobiol 35, 27–34. [PubMed: 26111432]

Fritz J, Shamma S, Elhilali M, and Klein D (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci 6, 1216–1223. [PubMed: 14583754]

Froemke RC, and Jones BJ (2011). Development of auditory cortical synaptic receptive fields. Neurosci. Biobehav. Rev 35, 2105–2113. [PubMed: 21329722]

Füllgrabe C, Moore BCJ, and Stone MA (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. Front. Aging Neurosci. 6, 347. [PubMed: 25628563]

Ganapathy S, Thomas S, and Hermansky H (2009). Modulation frequency features for phoneme recognition in noisy speech. J. Acoust. Soc. Am 125, EL8–EL12. [PubMed: 19173383]

Glasberg BR, and Moore BCJ (1990). Derivation of auditory filter shapes from notched-noise data. Hear. Res 47, 103–138. [PubMed: 2228789]

Grécová J, Bures Z, Popelár J, Suta D, and Syka J (2009). Brief exposure of juvenile rats to noise impairs the development of the response properties of inferior colliculus neurons. Eur. J. Neurosci 29, 1921–1930. [PubMed: 19473244]

Han YK, Köver H, Insanally MN, Semerdjian JH, and Bao S (2007). Early experience impairs perceptual discrimination. Nat. Neurosci 10, 1191–1197. [PubMed: 17660815]

Homma NY, Bajo VM, Happel MFK, Nodal FR, and King AJ (2016). Mistuning detection performance of ferrets in a go/no-go task. J. Acoust. Soc. Am 139, EL246–EL251. [PubMed: 27369180]

Homma NY, Happel MFK, Nodal FR, Ohl FW, King AJ, and Bajo VM (2017). A Role for Auditory Corticothalamic Feedback in the Perception of Complex Sounds. J. Neurosci 37, 6149–6161. [PubMed: 28559384]

Howard-Jones PA, and Rosen S (1993). Uncomodulated glimpsing in "checkerboard" noise. J. Acoust. Soc. Am 93, 2915–2922. [PubMed: 8315155]

Hullett PW, Hamilton LS, Mesgarani N, Schreiner CE, and Chang EF (2016). Human Superior Temporal Gyrus Organization of Spectrotemporal Modulation Tuning Derived from Speech Stimuli. J. Neurosci 36, 2014–2026. [PubMed: 26865624]

Insanally MN, Köver H, Kim H, and Bao S (2009). Feature-dependent sensitive periods in the development of complex sound representation. J. Neurosci 29, 5456–5462. [PubMed: 19403813]

Kang R, Sarro EC, and Sanes DH (2014). Auditory training during development mitigates a hearing loss-induced perceptual deficit. Front. Syst. Neurosci 8, 49. [PubMed: 24772071]

Killion MC, and Niquette PA (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? Hear. J 53, 46–48.

Kuhl PK (2010). Brain mechanisms in early language acquisition. Neuron 67, 713–727. [PubMed: 20826304]

Kvale MN, and Schreiner CE (2004). Short-Term Adaptation of Auditory Receptive Fields to Dynamic Stimuli. J. Neurophysiol 91, 604–612. [PubMed: 14762146]

Las L, Stern EA, and Nelken I (2005). Representation of tone in fluctuating maskers in the ascending auditory system. J. Neurosci 25, 1503–1513. [PubMed: 15703404]

Lee CC, and Middlebrooks JC (2011). Auditory cortex spatial sensitivity sharpens during task performance. Nat. Neurosci 14, 108–114. [PubMed: 21151120]

Malone BJ, Beitel RE, Vollmer M, Heiser MA, and Schreiner CE (2015). Modulation-frequency-specific adaptation in awake auditory cortex. J. Neurosci 35, 5904–5916. [PubMed: 25878263]

Malone BJ, Heiser MA, Beitel RE, and Schreiner CE (2017). Background noise exerts diverse effects on the cortical encoding of foreground sounds. J. Neurophysiol 118, 1034–1054. [PubMed: 28490644]

McDermott JH, and Simoncelli EP (2011). Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. Neuron 71, 926–940. [PubMed: 21903084]

McDermott JH, Wrobleski D, and Oxenham AJ (2011). Recovering sound sources from embedded repetition. Proc. Natl. Acad. Sci. USA 108, 1188–1193. [PubMed: 21199948]

Mehraei G, Gallun FJ, Leek MR, and Bernstein JGW (2014). Spectro-temporal modulation sensitivity for hearing-impaired listeners: dependence on carrier center frequency and the relationship to speech intelligibility. J. Acoust. Soc. Am 136, 301–316. [PubMed: 24993215]

Mesgarani N, Slaney M, and Shamma SA (2006). Discrimination of speech from nonspeech based on multiscale spectro-temporal Modulations. IEEE Trans. Audio Speech Lang. Process. 14, 920–930.

Mesgarani N, David SV, Fritz JB, and Shamma SA (2014). Mechanisms of noise robust representation of speech in primary auditory cortex. Proc. Natl. Acad. Sci. USA 111, 6792–6797. [PubMed: 24753585]

Moore RC, Lee T, and Theunissen FE (2013). Noise-invariant neurons in the avian auditory cortex: hearing the song in noise. PLoS Comput. Biol. 9, e1002942. [PubMed: 23505354]

Nakahara H, Zhang LI, and Merzenich MM (2004). Specialization of primary auditory cortex processing by sound exposure in the "critical period". Proc. Natl. Acad. Sci. USA 101, 7170–7174. [PubMed: 15118079]

Narayan R, Best V, Ozmeral E, McClaine E, Dent M, Shinn-Cunningham B, and Sen K (2007). Cortical interference effects in the cocktail party problem. Nat. Neurosci 10, 1601–1607. [PubMed: 17994016]

Ni R, Bender DA, Shanechi AM, Gamble JR, and Barbour DL (2017). Contextual effects of noise on vocalization encoding in primary auditory cortex. J. Neurophysiol 117, 713–727. [PubMed: 27881720]

Noreña AJ, Gourévitch B, Aizawa N, and Eggermont JJ (2006). Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex. Nat. Neurosci 9, 932–939. [PubMed: 16783369]

Ohl FW, and Scheich H (1996). Differential frequency conditioning enhances spectral contrast sensitivity of units in auditory cortex (field Al) of the alert Mongolian gerbil. Eur. J. Neurosci 8, 1001–1017. [PubMed: 8743748]

Oliver DL, Izquierdo MA, and Malmierca MS (2011). Persistent effects of early augmented acoustic environment on the auditory brainstem. Neuroscience 184, 75–87. [PubMed: 21496479]

Pienkowski M, and Eggermont JJ (2009). Long-term, partially-reversible reorganization of frequency tuning in mature cat primary auditory cortex can be induced by passive exposure to moderate-level sounds. Hear. Res 257, 24–40. [PubMed: 19647789]

Pienkowski M, Munguia R, and Eggermont JJ (2011). Passive exposure of adult cats to bandlimited tone pip ensembles or noise leads to long-term response suppression in auditory cortex. Hear. Res 277, 117–126. [PubMed: 21316436]

Polley DB, Steinberg EE, and Merzenich MM (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. J. Neurosci 26, 4970–4982. [PubMed: 16672673]

Polley DB, Read HL, Storace DA, and Merzenich MM (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. J. Neurophysiol 97, 3621–3638. [PubMed: 17376842]

Pysanenko K, Bureš Z, Lindovský J, and Syka J (2018). The Effect of Complex Acoustic Environment during Early Development on the Responses of Auditory Cortex Neurons in Rats. Neuroscience 371, 221–228. [PubMed: 29229554]

Rabinowitz NC, Willmore BDB, King AJ, and Schnupp JWH (2013). Constructing noise-invariant representations of sound in theauditory pathway. PLoS Biol. 11, e1001710. [PubMed: 24265596]

Ranasinghe KG, Carraway RS, Borland MS, Moreno NA, Hanacik EA, Miller RS, and Kilgard MP (2012). Speech discrimination after early exposure to pulsed-noise or speech. Hear. Res 289, 1–12. [PubMed: 22575207]

Reed A, Riley J, Carraway R, Carrasco A, Perez C, Jakkamsetti V, and Kilgard MP (2011). Cortical map plasticity improves learning but is not necessary for improved performance. Neuron 70, 121–131. [PubMed: 21482361]

Ruggles D, Bharadwaj H, and Shinn-Cunningham BG (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. Proc. Natl. Acad. Sci. USA 108, 15516–15521. [PubMed: 21844339]

Saffran JR, Aslin RN, and Newport EL (1996). Statistical learning by 8-month-old infants. Science 274, 1926–1928. [PubMed: 8943209]

Saffran JR, Johnson EK, Aslin RN, and Newport EL (1999). Statistical learning of tone sequences by human infantsand adults. Cognition 70, 27–52. [PubMed: 10193055]

Schädler M, Meyer BT, and Kollmeier B (2012). Spectro-temporal modulation subspace-spanning filter bank features for robust automatic speech recognition. J. Acoust. Soc. Am 131, 4134–4151. [PubMed: 22559385]

Schneider DM, and Woolley SMN (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. Neuron 79, 141–152. [PubMed: 23849201]

Schnupp JWH, Hall TM, Kokelaar RF, and Ahmed B (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. J. Neurosci 26, 4785–4795. [PubMed: 16672651]

Schönwiesner M, and Zatorre RJ (2009). Spectro-temporal modulation transfer function of single voxels in the human auditory cortex measured with high-resolution fMRI. Proc. Natl. Acad. Sci. USA 106, 14611–14616. [PubMed: 19667199]

Schreiner CE, and Urbas JV (1988). Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hear. Res 32, 49–63. [PubMed: 3350774]

See JZ, Atencio CA, Sohal VS, and Schreiner CE (2018). Coordinated neuronal ensembles in primary auditory cortical columns. eLife 7, 1–33.

Shetake JA, Wolf JT, Cheung RJ, Engineer CT, Ram SK, and Kilgard MP (2011). Cortical activity patterns predict robust speech discrimination ability in noise. Eur. J. Neurosci 34, 1823–1838. [PubMed: 22098331]

Singh NC, and Theunissen FE (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. J. Acoust. Soc. Am 114, 3394–3411. [PubMed: 14714819]

Song JH, Skoe E, Banai K, and Kraus N (2012). Training to improve hearing speech in noise: biological mechanisms. Cereb. Cortex 22, 1180–1190. [PubMed: 21799207]

Souffi S, Lorenzi C, Huetz C, and Edeline J-M (2019). Cortical and subcortical neurons discriminate sounds in noise on the sole basis of acoustic amplitude modulations. bioRxiv. 10.1101/528653.

Stone MA, and Canavan S (2016). The near non-existence of "pure" energetic masking release for speech: Extension to spectro-temporal modulation and glimpsing. J. Acoust. Soc. Am 140, 832–842. [PubMed: 27586715]

Sun W, Tang L, and Allman BL (2011). Environmental noise affects auditory temporal processing development and NMDA-2B receptor expression in auditory cortex. Behav. Brain Res. 218, 15–20. [PubMed: 21094188]

Takesian AE, Bogart LJ, Lichtman JW, and Hensch TK (2018). Inhibitory circuit gating of auditory critical-period plasticity. Nat. Neurosci 21, 218–227. [PubMed: 29358666]

Talwar SK, and Gerstein GL (2001). Reorganization in awake rat auditory cortex by local microstimulation and its effect on frequency-discrimination behavior. J. Neurophysiol 86, 1555–1572. [PubMed: 11600620]

Taylor B (2003). Speech-in-noise tests: How and why to include them in your basic test battery. Hear. J 56, 40.

Ward ED, Cushing EM, and Burns EM (1976). Effective quiet and moderate TTS: Implications for noise exposure statndars. J. Acoust. Soc. Am 59, 160–165. [PubMed: 1249315]

Wayne RV, Hamilton C, Jones Huyck J, and Johnsrude IS (2016). Working Memory Training and Speech in Noise Comprehension in Older Adults. Front. Aging Neurosci. 8, 49. [PubMed: 27047370]

Whitton JP, Hancock KE, and Polley DB (2014). Immersive audiomotor game play enhances neural and perceptual salience of weak signals in noise. Proc. Natl. Acad. Sci. USA 111, E2606–E2615. [PubMed: 24927596]

Whitton JP, Hancock KE, Shannon JM, and Polley DB (2017). Audiomotor Perceptual Training Enhances Speech Intelligibility in Background Noise. Curr. Biol 27, 3237–3247.e6. [PubMed: 29056453]

Wickens T (2002). Elementary Signal Detection Theory (Oxford University Press).

Zhang LI, Bao S, and Merzenich MM (2002). Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. Proc. Natl. Acad. Sci. USA 99, 2309–2314. [PubMed: 11842227]

Zhang J, Chen L, Gao F, Pu Q, and Sun X (2008). Noise exposure at young age impairs the auditory object exploration behavior of rats in adulthood. Physiol. Behav 95, 229–234. [PubMed: 18598709]

Zheng W (2012). Auditory map reorganization and pitch discrimination in adult rats chronically exposed to low-level ambient noise. Front. Syst. Neurosci 6, 65. [PubMed: 22973201]

Zhou X, and Merzenich MM (2008). Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. Proc. Natl. Acad. Sci. USA 105, 4423–4428. [PubMed: 18332439]
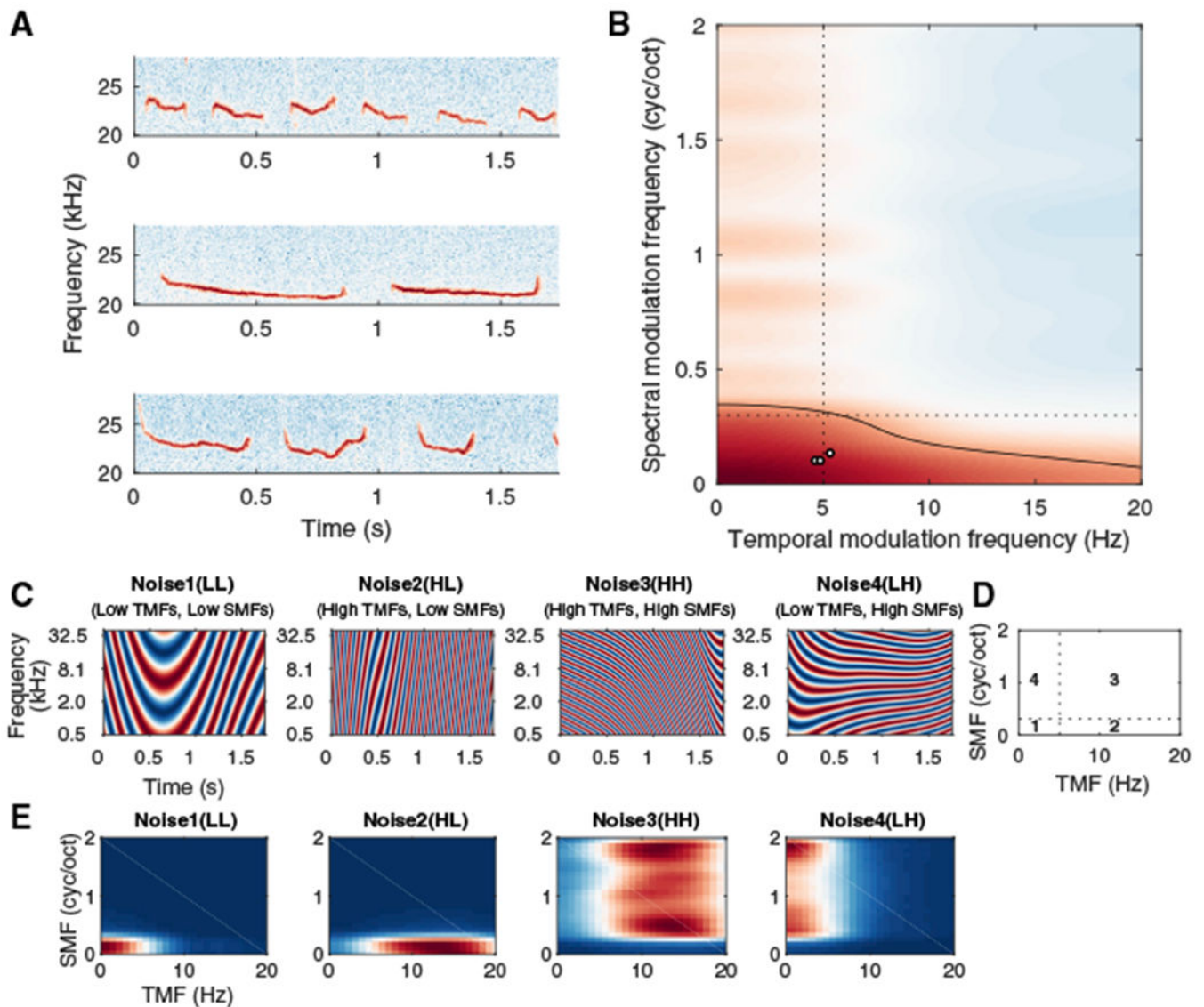
Highlights

- Rats were reared with spectro-temporally modulated noise

- Behavioral vocalization-in-noise detection performance improves with noise-rearing

- Neural encoding of vocalizations-in-noise improves with noise-rearing

- Cortical modulation preferences shift away from the exposed noise statistics

Figure 1. Sound Stimuli of Rat Vocalizations and Biased DMR Noises
(A) Spectrograms of three rat vocalizations used in this study.
(B) Rat vocalization modulation power spectrum (MPS, mean of 18 different vocalizations).
Black contour: 70% of power strength. Black circles: temporal and spectral modulation
frequencies (TMFs and SMFs) of the three vocalizations in (A). Dashed lines separate the
value ranges of the four biased DMR noises (see D).
(C) Example spectrograms of short segments for the biased DMRs.
(D) Modulation ranges for four biased DMR noises. Noise1, TMFs: 0–5 Hz, SMFs: 0–0.3
cyc/oct; Noise2, TMFs: 5–20 Hz, SMFs: 0–0.3 cyc/oct; Noise3, TMFs: 5–20 Hz, SMFs:
0.3–2 cyc/oct; Noise4, TMFs: 0–5 Hz, SMFs: 0.3–2 cyc/oct.
(E) MPSs of the four biased noises. Color: power on a log scale.
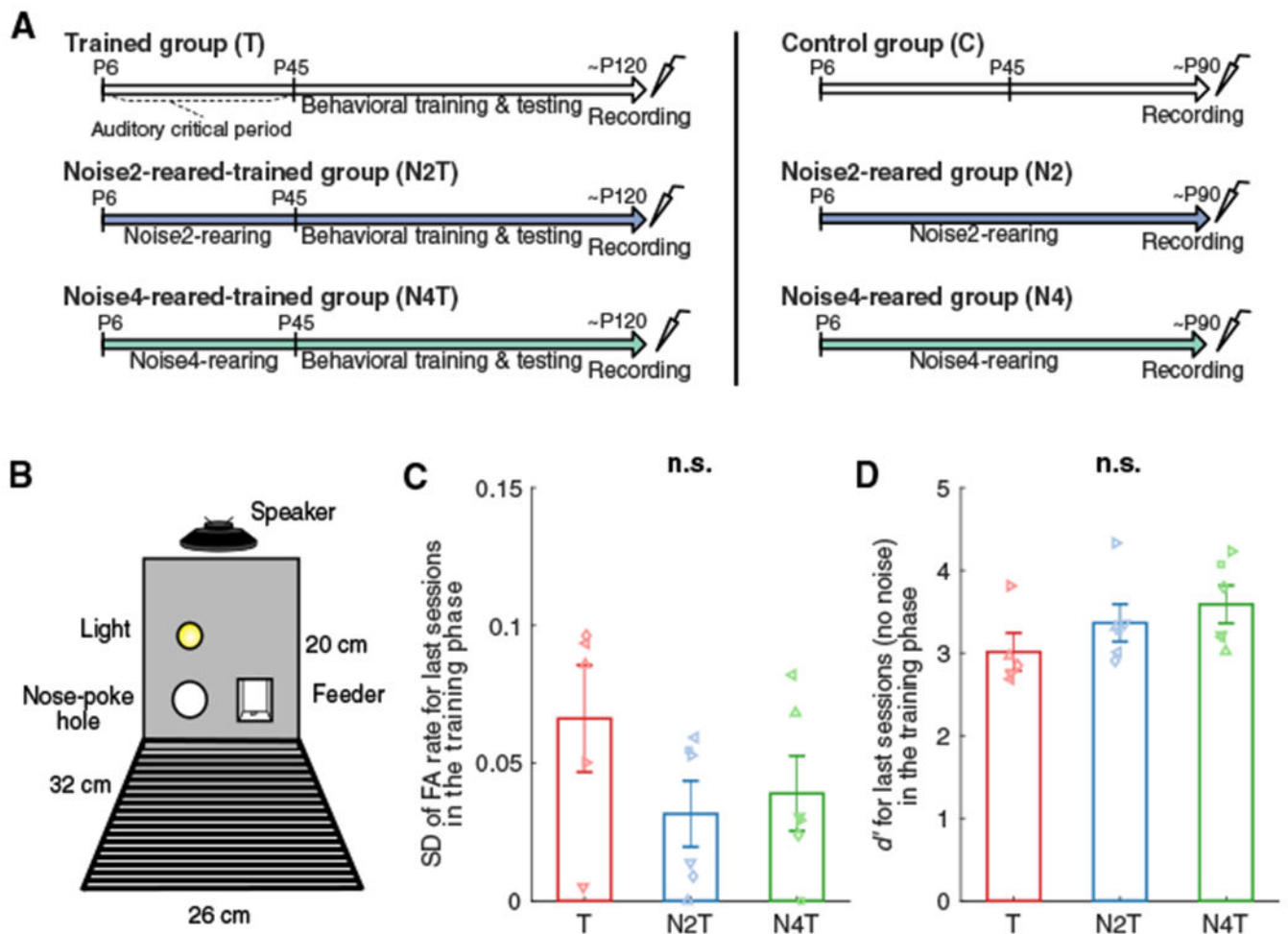
Figure 2. Experimental Design

(A) Trained and Control groups were raised in a typical rat colony environment while the groups with noise-rearing (N2T, N4T, N2, N4) were raised in the presence of biased DMR noise (~60 dB SPL) from P6. Behavioral training was started at P45 (T, N2T, N4T). Cortical activity was recorded at ~P120 for trained groups and ~P90 for untrained groups.

(B) The behavioral testing chamber contained one nose poke hole (left) and one feeder (right). A loudspeaker was mounted at a height of 20 cm, and a house light was placed 5 cm above the nose poke hole.

(C and D) Standard deviation (SD) of false alarm (FA) rate (C) or sensitivity index (d') (D) was computed for the last sessions in the training phase (no noise). Bars represent mean values (± SEM) for T (red; n = 5), N2T (blue; n = 6), and N4T (green; n = 6) groups. No group difference was found in SD or d' (n.s.).
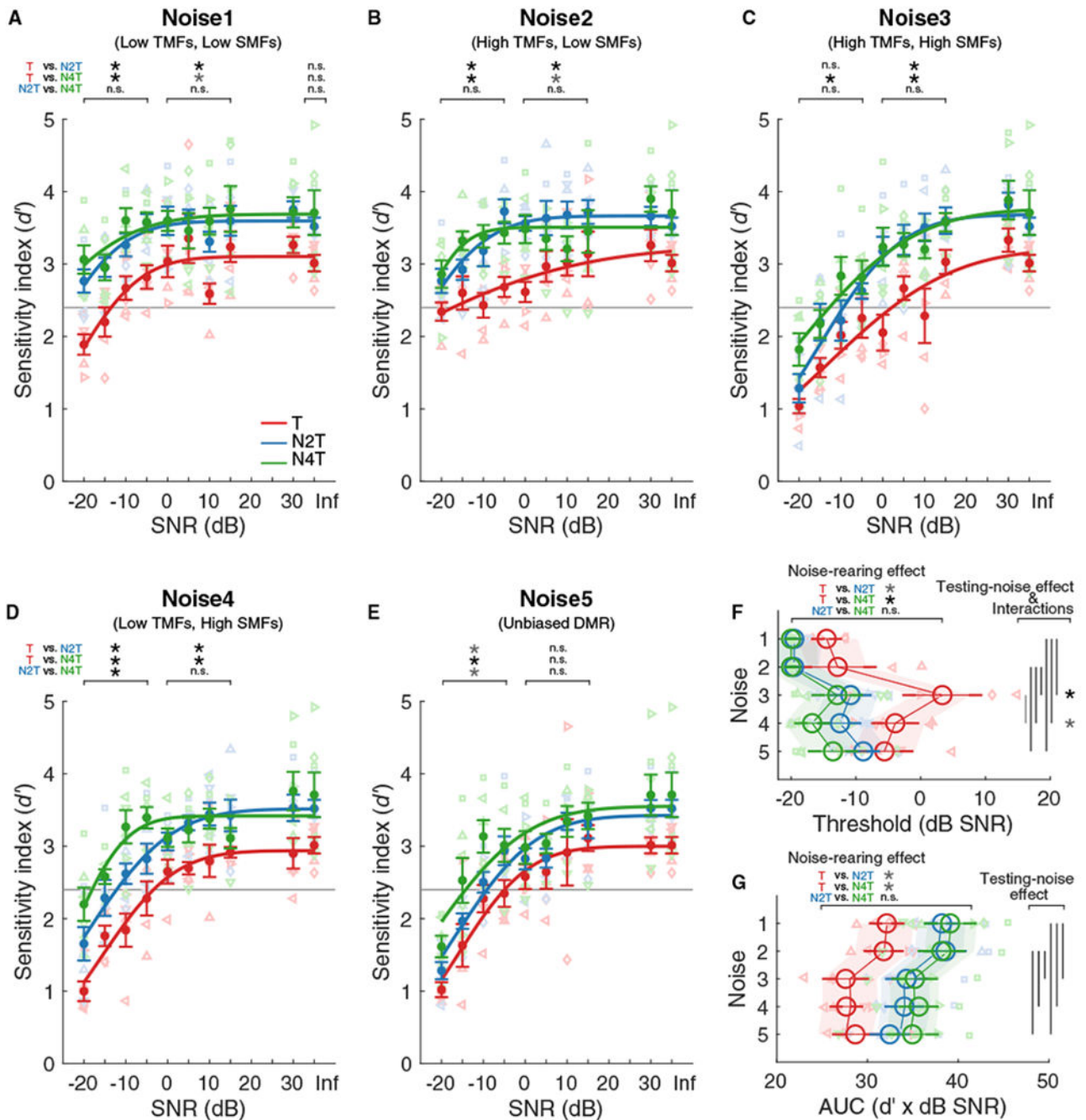
Figure 3. Noise-Rearing Improved Vocalization-in-DMR Detection Performance

(A–E) Psychometric functions for vocalization detection in Noise1 (A), Noise2 (B), Noise3 (C), Noise4 (D), or Noise5 (E). Individual d' values derived for the T (red; n = 5), N2T (blue; n = 6), and N4T (green; n = 6) groups are plotted against signal-to-noise ratio (SNR). A cumulative Gaussian distribution was used to fit the functions for the group means (lines). Dots represent mean (±SEM) for each group. The horizontal gray lines indicate the threshold criterion of d' = 2.4. The asterisks indicate paired comparisons for the range of −20 to −5, 0 to 15, or Inf dB SNR.

(F and G) Large open circles represent mean (±SEM) of threshold (F) or AUC (G) values for each noise type and group, which were obtained by fitting to the functions for individual animals. The shaded areas represent the 2.5%–97.5% confidence intervals of the linear mixed effects model. Vertical lines: testing-noise effects ($p < 0.05$). Asterisks: differences for T versus N2T and T versus N4T for each noise type; black, $p < 0.01$, gray, $p < 0.05$; n.s., not significant.

See Table S1 for details of statistical analysis. See also Figures S1 and S2.
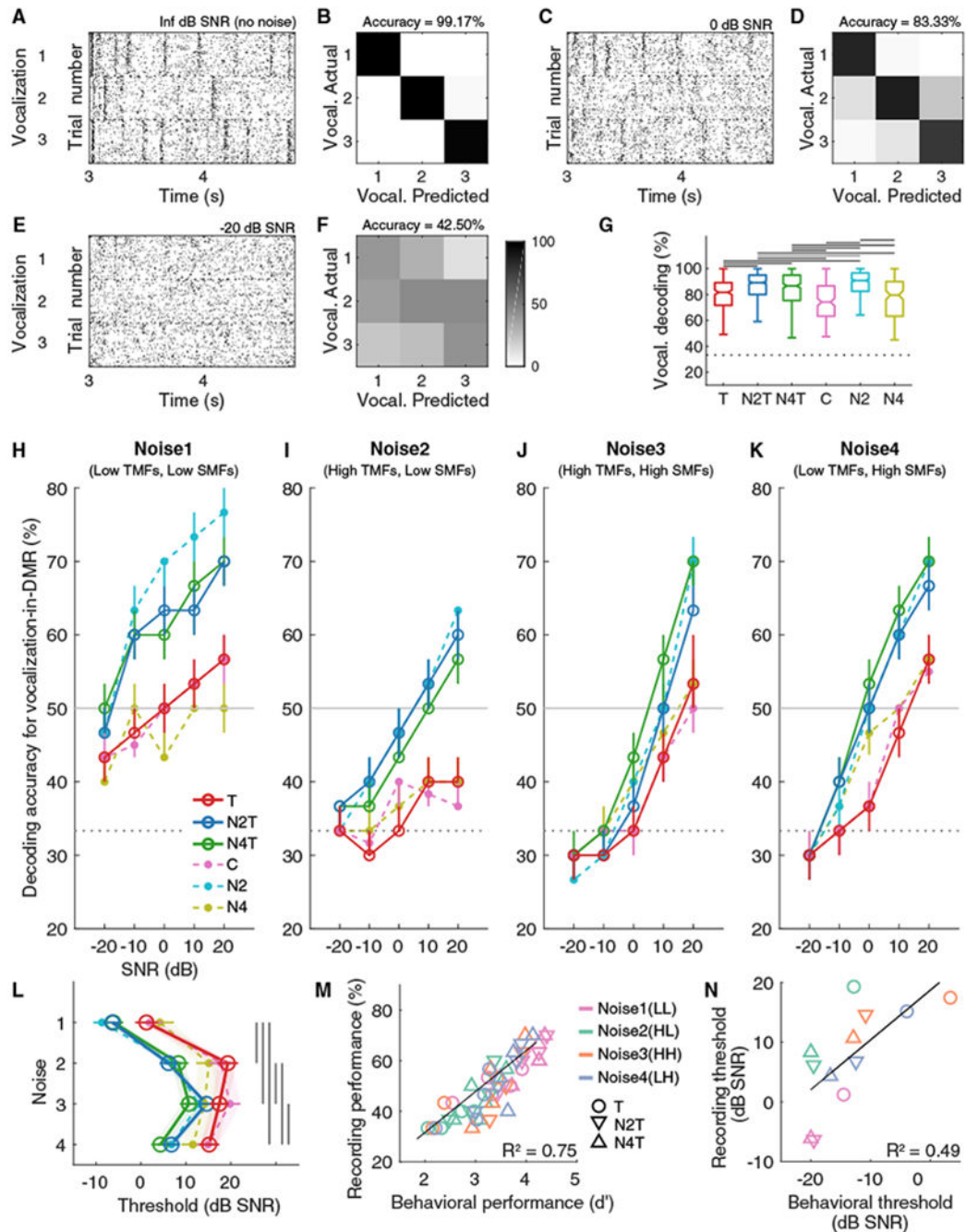
Figure 4. Noise-Rearing Improved Cortical Vocalization-in-DMR Decoding

(A, C, and E) Examples of responses to three vocalizations for Inf dB SNR (no noise) (A), 0 dB SNR (C), and −20 dB SNR (E). Each dot indicates a neural spike.

(B, D, and F) Examples of confusion matrices for vocalization-in-DMR decoding for Inf dB SNR(B), 0 dB SNR(D), and −20 dB SNR(F) corresponding to(A), (C), and (E).

(G) Boxplots of median values plus 25th–75th percentiles of vocalization decoding accuracy (Inf dB SNR) for T (red; 254 MUs), N2T (blue; 652), N4T (green; 428), C (pink; 170), N2 (cyan; 587), and N4 (olive; 148) groups. Dashed horizontal line indicates the chance level of

33.33%. Solid horizontal lines indicate significant post hoc multiple comparisons (black, p < 0.01; gray, p < 0.05).

(H–K) Vocalization decoding accuracy in the presence of Noise1 (H), Noise2 (I), Noise3 (J), or Noise4 (K) is plotted for the six groups. Circles(open, T, N2T or N4T; filled, C, N2 or N4) represent median values (with 45th–55th percentile ranges) across animals. The horizontal gray lines indicate the threshold criterion of decoding accuracy = 50%.

(L) Decoding thresholds were obtained as the first point to cross the criterion, and the mean (± SEM) value is plotted for each combination of groups and testing-noise types with open or filled circle. The shaded areas represent the 2.5%–97.5% confidence intervals of the linear mixed effects model. The vertical lines indicate significant post hoc multiple comparisons for noise type (p < 0.0001). See Figure S3 and Table S1 for details of statistical analysis. See also Figure S4.

(M) Behavioral sensitivity for vocalization-in-DMR detection and physiological decoding accuracy for each SNR condition of vocalization-in-DMR task were correlated (Pearson correlation coefficient, R = 0.78, p < 0.0001). Symbols represent mean values for each noise (Noise1, pink; Noise2, green; Noise3, orange; Noise4, blue).

(N) Behavioral and decoding thresholds were highly correlated (R = 0.70, p = 0.01).
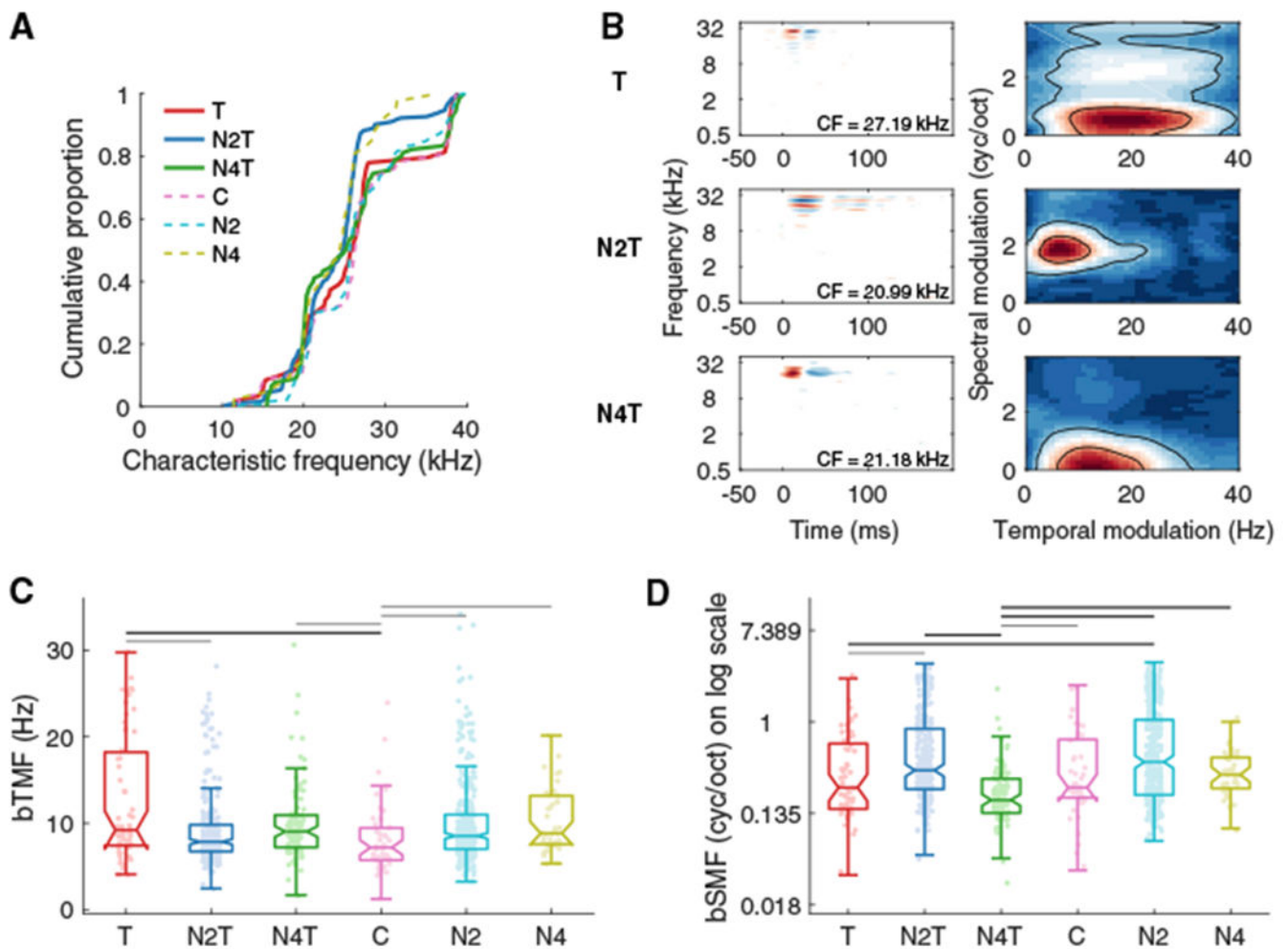
Figure 5. Noise-Rearing and Behavior Training Shifted Modulation Parameter Values
(A) Cumulative distributions of characteristic frequency (CF) are plotted for the T (red; 59 SUs), N2T (blue; 246), N4T (green; 90), C (pink; 46), N2 (cyan; 359), and N4 (olive; 42) groups.
(B) Examples of STRFs and RTFs. RTFs were estimated using 2D Fourier transform on each STRF.
(C and D) Boxplots (plus 25th to 75th percentiles) for bTMFs (C) and bSMFs (D) obtained as the peak values after summing each RTF along spectral modulation or temporal modulation axes. Noise-rearing and behavioral training shifted both temporal and spectral modulations to higher frequencies, although bSMFs for the N4T group tended to be smaller than other groups. Horizontal lines, black, $p < 0.01$, gray, $p < 0.05$.
See also Figures S5 and S6.
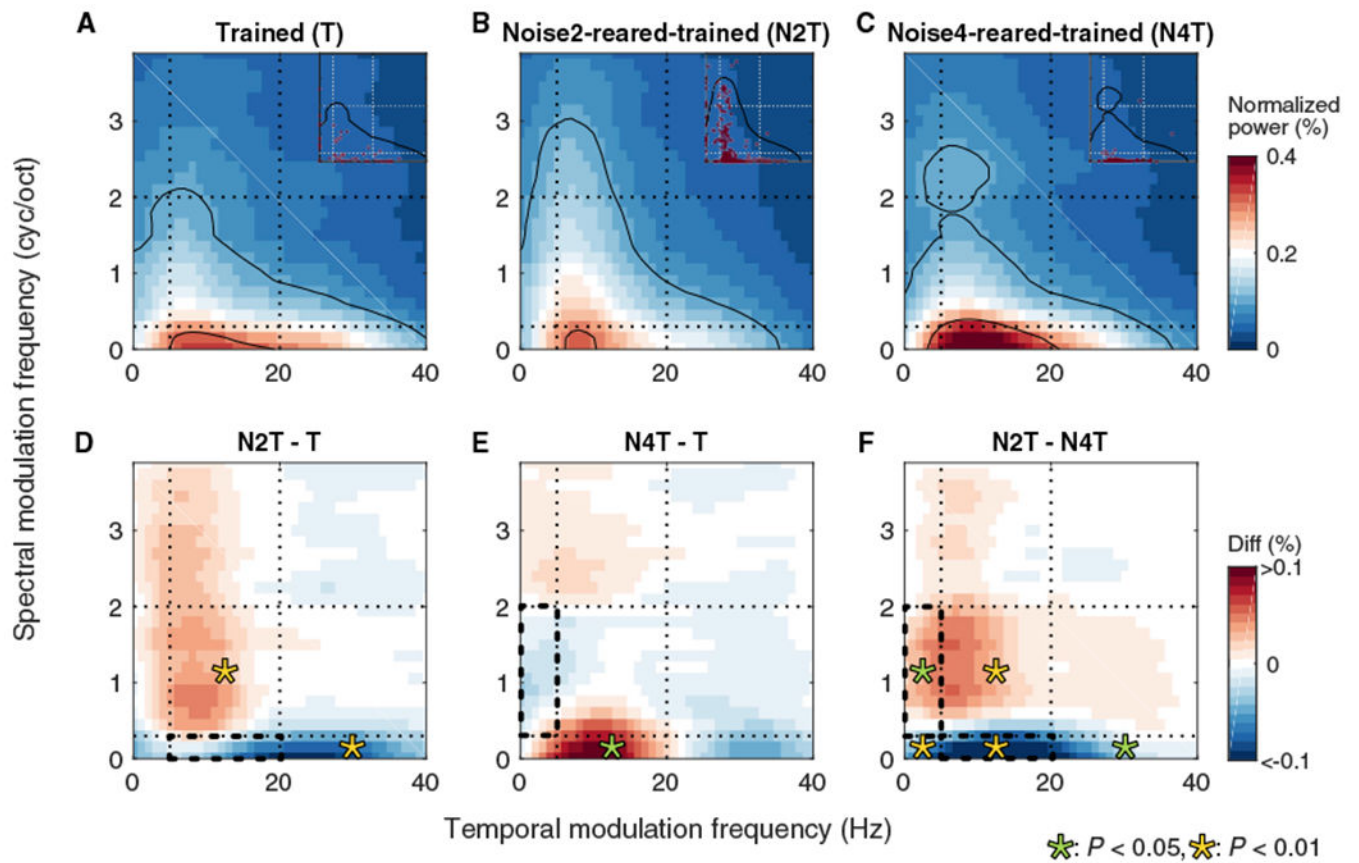
Figure 6. Noise-Rearing Combined with Behavioral Training Modified RTF Preferences

(A–C) Population RTFs (pRTFs) were obtained for T (59 SUs) (A), N2T (246 SUs) (B), and N4T (90 SUs) (C) groups. Color bar indicates normalized power strength in percentage. Black contours indicate 25% and 75% of the power. Dotted lines indicate the boundaries of biased DMR noises. Insets indicate bTMF and bSMF of each neuron.

(D–F) Differences of pRTFs between N2T and T(D), between N4T and T(E), and between N2Tand N4T(F) groups. Bold dashed boxes indicate modulation ranges of the exposure noises. Color bar indicates the values of differences in percentage. Asterisks: significant regional differences, green, p < 0.05, yellow, p < 0.01.
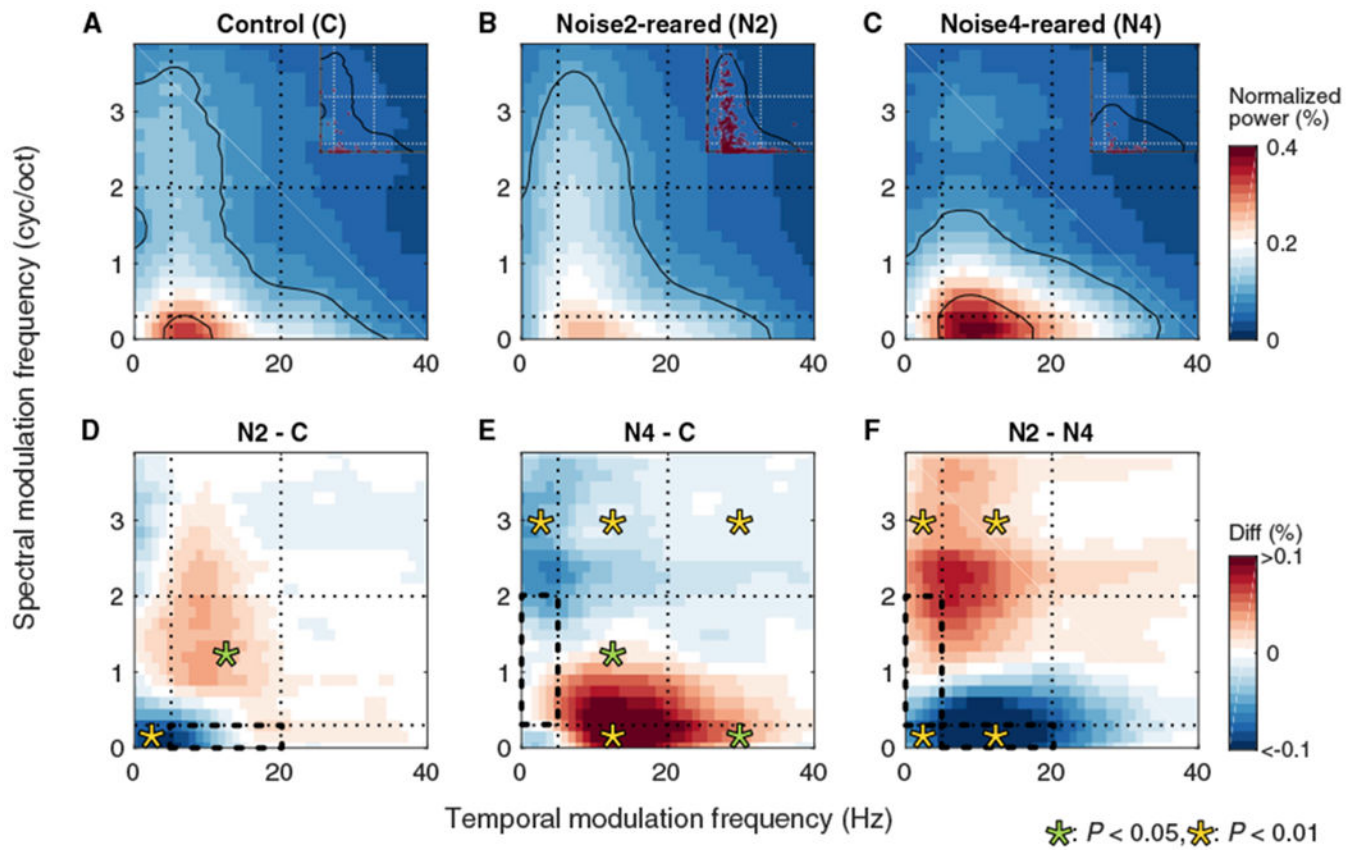
Figure 7. Noise-Rearing without Behavioral Training Modified RTF Preferences
(A–C) pRTFs were obtained for C (46 SUs) (A), N2 (359 SUs) (B), and N4 (42 SUs) (C)
groups.

(D–F) Differences of pRTFs were obtained between N2 and C (D), between N4 and C (E),
and between N2 and N4 (F) groups. The presentation scheme is the same as Figure 6.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Sprague-Dawley rat (wild type)	Charles River	001CD
Software and Algorithms		
MATLAB (versions, 2015b, 2017b)	MathWorks	https://www.mathworks.com/
MED-PC IV	Med Associates, Inc.	https://www.med-associates.com/
MountainSort	Chung et al., 2017	https://doi.org/10.1016/j.neuron.2017.08.030
RHD2000 Interface software	Intan Technologies	http://intantech.com/
RStudio	RStudio Inc.	https://rstudio.com/
Other		
Electrostatic speaker	Tucker-Davis Technologies	EC1
QUAD-CAPTURE	Roland	https://www.roland.com/us/products/quad-capture/
RHD2132 Amplifier system	Intan Technologies	http://intantech.com/