Published in final edited form as:

Behav Brain Res. 2014 January 1; 258: 166–178.

Speech training alters tone frequency tuning in rat primary auditory cortex

Crystal T. Engineer, Claudia A. Perez, Ryan S. Carraway, Kevin Q. Chang, Jarod L. Roland, and Michael P. Kilgard

School of Behavioral and Brain Sciences, The University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080, United States

Abstract

Previous studies in both humans and animals have documented improved performance following discrimination training. This enhanced performance is often associated with cortical response changes. In this study, we tested the hypothesis that long-term speech training on multiple tasks can improve primary auditory cortex (A1) responses compared to rats trained on a single speech discrimination task or experimentally naïve rats. Specifically, we compared the percent of A1 responding to trained sounds, the responses to both trained and untrained sounds, receptive field properties of A1 neurons, and the neural discrimination of pairs of speech sounds in speech trained and naïve rats. Speech training led to accurate discrimination of consonant and vowel sounds, but did not enhance A1 response strength or the neural discrimination of these sounds. Speech training altered tone responses in rats trained on six speech discrimination tasks but not in rats trained on a single speech discrimination task. Extensive speech training resulted in broader frequency tuning, shorter onset latencies, a decreased driven response to tones, and caused a shift in the frequency map to favor tones in the range where speech sounds are the loudest. Both the number of trained tasks and the number of days of training strongly predict the percent of A1 responding to a low frequency tone. Rats trained on a single speech discrimination task performed less accurately than rats trained on multiple tasks and did not exhibit A1 response changes. Our results indicate that extensive speech training can reorganize the A1 frequency map, which may have downstream consequences on speech sound processing.

Keywords

Speech; Primary auditory cortex; Plasticity; Training

1. Introduction

Numerous studies have demonstrated that animals can accurately discriminate between human speech sounds [1-4]. We have previously shown that rats can discriminate both consonant sounds, such as /d/ vs. /t/ or /sh/ vs. /ch/ [1] and vowel sounds, such as /æ/ vs. /ε/ ('dad' vs. 'dead') [5]. Although animal models are clearly not appropriate to study higher order aspects of language, auditory processing of basic speech sounds in humans and other

© 2013 Elsevier B.V. All rights reserved.

Corresponding author: Crystal Engineer novitski@utdallas.edu The University of Texas at Dallas 800 West Campbell Road, GR41 Richardson, TX 75080 Phone: 972-883-2376 Fax: 972-883-2491.

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.

mammals is similar [6]. Both rodents and humans can discriminate speech sounds down to a signal to noise ratio of approximately -12 dB, can rapidly generalize to novel exemplars of phonetic categories, and exhibit similar patterns of impairment when temporal or spectral information is removed from speech sounds [3, 7-9].

The behavioral accuracy of the rats is strongly predicted by the spatiotemporal response pattern evoked by each sound in primary auditory cortex (A1). Sounds that evoke very distinct neural patterns of activity (/d/ and /s/) are easy for rats to discriminate, while sounds that evoke very similar neural patterns of activity (/r/ and /l/) are difficult for rats to discriminate. The neural responses used to predict behavioral accuracy were collected from experimentally naïve rats. It is not known whether A1 responses are modified by speech training in rats.

Previous studies have documented that tone frequency discrimination training improves behavioral performance, and this improved performance is associated with significant reorganization of A1 [10-12]. Other studies have shown that discrimination training of vocalization sounds enhances the responses to the trained vocalization sounds [13, 14]. In this study, we compared auditory cortex responses in speech trained rats with responses in passively exposed and experimentally naïve rats. Speech training could result in no changes in A1 responses or could result in auditory cortex map plasticity or speech sound specific plasticity. We hypothesized based on previous studies that speech training would either cause A1 map plasticity or significantly enhance A1 responses to the trained speech sounds.

Performance on novel speech discrimination tasks by humans and animals improves gradually over weeks or months of training [3, 15]. For example, humans require several weeks of training to learn difficult non-native speech contrasts and generate measureable cortical plasticity [15-19]. The specific neural mechanisms responsible for this form of learning are not clear. In this study, we evaluated whether short term speech training can alter response properties of neurons in primary auditory cortex or whether extensive speech training is required.

2. Materials and methods

We trained 29 rats to discriminate speech sounds, after which A1 responses to tones and speech sounds were recorded from 1,110 A1 sites (Table 1). Five groups of rats were trained; (1) the One Task Voicing group was trained to discriminate sounds by voicing (n=5 rats), (2) the One Task Gender group was trained to discriminate female sounds from male sounds (n=5 rats), (3) the Three Tasks Vowels group was trained to discriminate vowels (n=8 rats), (4) the Six Tasks Consonants group was trained to discriminate consonants and discriminate female sounds from male sounds (n=5 rats), and (5) the Six Tasks Consonants and Vowels group was trained to discriminate consonants, vowels, and by voicing (n=6 rats). A sixth Passive Exposure group of rats was passively exposed to the same stimuli used for the voicing and gender discrimination tasks (n=5 rats, 236 A1 sites). A1 responses in these trained and passively exposed rats were compared to the responses recorded in experimentally naïve rats (n=11 rats, 392 A1 sites). Speech sounds, training, and anesthetized recording procedures are identical to the methods used in our previous studies [1, 5]. The last day behavioral performance of rats in the Three Tasks and Six Tasks groups has previously been reported [1, 5]. The University of Texas at Dallas Institutional Animal Care and Use Committee approved all protocols and recording procedures.

2.1 Speech stimuli

Speech stimuli were produced in a CVC (consonant-vowel-consonant) context and were recorded in a soundproof booth. Twenty English consonants were produced in a ‘_æd’

context (as in 'dad'). Additionally, the four vowels /æ/, /ε/, /Λ/, /i/, /u/ were produced in both a 'd_d' and 's_d' context (/æ/ 'dad', /ε/ 'dead', /Λ/ 'dud', /i/ 'deed', /u/ 'dood'). The words 'dad' and 'tad' were also recorded spoken by multiple speakers, 3 adult males and 3 adult females. As in our previous studies, the speech sounds were shifted up by one octave using the STRAIGHT vocoder in order to better match the rat hearing range and presented so that the intensity of the loudest 100 ms of the vowel was 60 dB SPL [1, 5, 21, 22].

2.2 Training timeline

Rats were trained using an operant go/no-go procedure in double-walled booths for two 1-hour sessions a day, 5 days a week. Each booth contained a speaker (Optimus Bullet Horn Tweeter), house light, and cage (8" length × 8" width × 8" height) which included a lever and pellet dish. A pellet dispenser was mounted outside of the booth to minimize the sound of the pellet dispensing. Rats received a 45 mg sugar pellet reward for pressing the lever within 3 seconds in response to the presentation of a target sound. If the rats pressed the lever at any other time (for example, in response to a non-target sound), they received a time out where the house light was extinguished for approximately 6 seconds.

To shape behavior, rats were initially trained to press the lever to receive a sugar pellet reward. After reaching the criteria of obtaining 100 pellets for 2 sessions, rats were moved to the detection stage. In this stage, rats were trained to press the lever in response to the target sound ('dad') to receive a sugar pellet. This stage lasted until the rat was able to obtain a *d'* performance value of 1.5 or greater for 10 sessions. Following the detection stage, rats began the discrimination stage of training. Each discrimination task lasted for 10 days.

2.3 Training groups

Of the five groups of trained rats, two groups of rats were short-term experiments where rats trained for 10 days on a single discrimination task. The One Task Voicing group of rats was trained on a voicing discrimination task with a target of 'dad' and a non-target of 'tad'. Both 'dad' and 'tad' were spoken by multiple male and female speakers. The One Task Gender group of rats was trained on a gender discrimination task using the exact same stimuli as the One Task Voicing rats, but with targets of female 'dad' and 'tad' sounds and non-targets of male 'dad' and 'tad' sounds. Rats presumably use frequency information to discriminate between adult male and adult female speakers. While the pitch of a speech sound can often be used to identify the gender of the speaker, other cues such as the first and second formants can be used to identify speaker gender [23, 24]. An additional group of rats was passively exposed to the same stimuli used for the One Task Voicing and One Task Gender groups of rats, and each Passive Exposure rat was exposed to the same sounds for the same 10 day period.

The remaining 3 groups of trained rats were long-term experiments where rats trained on multiple discrimination tasks over 11-27 weeks. The Three Tasks Vowels group of rats was trained on three vowel discrimination tasks: (1) 'dad' (target) vs. 'dead', 'dud', 'deed', and 'dood' (non-targets); (2) 'sad' (target) vs. 'said', 'sud', 'seed', and 'sood' (non-targets); and (3) a combination of the two previous tasks, where the target was the vowel /æ/ in either the /d/ or /s/ context, and the non-target sounds were the vowels /ε/, /Λ/, /i/, and /u/ in either the /d/ or /s/ context. Half of the rats (n = 4) were trained on the /d/ context first, followed by the /s/ context, while the other rats (n = 4) were trained on the /s/ context first, followed by the /d/ context [5]. The Six Tasks Consonants group of rats was trained on six discrimination tasks: (1) pitch discrimination; (2) gender discrimination of the word 'dad' spoken by multiple male and female speakers; (3) gender discrimination of the words 'dad' and 'tad' spoken by multiple male and female speakers; (4) /m/ vs. /n/, (5) 'shad' vs. 'fad', 'sad', and 'had'; and (6) 'shad' vs. 'chad' and 'jad' [1]. The Six Tasks Consonants and Vowels group

of rats was trained on six discrimination tasks: (1) /d/ vs. /t/; (2) /d/ vs. /t/ when sounds were temporally compressed; (3) /d/ vs. /t/ when sounds were spoken by multiple male and female speakers; (4) /r/ vs. /l/; (5) 'dad' vs. 'dead', 'dud', 'deed', and 'dood'; and (6) /d/ vs. /b/ and /g/ [1]. Four of the Six Tasks Consonants and Vowels rats were trained on an additional /d/ vs. /s/ discrimination task. After training on each task for 10 days, rats underwent a training review period where they were tested on each of their previous tasks for 2 days per task.

2.4 Anesthetized recordings

Following the last day of training, rats were anesthetized and multi-unit and local field potential responses were recorded from the right primary auditory cortex (as in [1]). Rats were anesthetized with pentobarbital (50 mg kg⁻¹), and received dilute pentobarbital (8 mg ml⁻¹) as needed. Responses were recorded 600 µm below the cortical surface using four Parylene-coated tungsten microelectrodes (1–2 MΩ, FHC) simultaneously.

In order to determine the characteristic frequency at each recording site, 25 ms tones were presented at 81 frequencies (1 to 32 kHz) at 16 intensities (0 to 75 dB). A total of 80 speech stimuli were presented at each site for 20 repeats each. Each of the trained speech sounds was presented, and other novel speech sounds were presented at each site in order to determine if response changes were specific to trained stimuli. Stimulus generation and data acquisition were performed with Tucker-Davis hardware (RP2.1 and RX5) and software (SigGen and Brainware). We recorded from a total of 2,662 auditory cortex sites: 627 sites in experimentally naïve rats (392 A1 sites), 316 sites in Passive Exposure rats (236 A1 sites), 595 sites in One Task trained rats (431 A1 sites), 526 sites in Three Task trained rats (252 A1 sites), and 598 sites in Six Task trained rats (427 A1 sites).

2.5 Data analysis

Behavioral performance was quantified using the measure percent correct, which is the average of the correct lever presses to target sounds and correct rejections of non-target sounds. A 2-way ANOVA was used to compare first day performance to last day performance on each of the speech discrimination tasks.

We quantified the receptive field properties for each group, including threshold, bandwidth, latency, and response strength to tones. These properties were quantified by a blind observer using customized MATLAB software. Threshold was defined as the lowest intensity that evoked a response at the characteristic frequency for each site. Bandwidth was measured in octaves as the frequency range that evokes a response 10, 20, 30, and 40 dB above the threshold. The onset, peak, and end of peak latencies were also determined for each site. The latency of each site and characteristic frequency topography were used to classify sites as primary auditory cortex.

The percent of primary auditory cortex responding to each tone at each intensity was calculated for each group and compared to the percent of cortex responding in experimentally naïve rats, as in previous studies [25–28]. Each point on the map was assumed to have the response properties of the nearest recording site using Voronoi tessellations. The response strength at each recording site to specific tone frequency intensity combinations was quantified using the average of tones within 1 frequency or intensity step (i.e. 1/16 of an octave or 5 dB respectively). The percent of A1 sites responding to each tone at each intensity was calculated by adding the areas of sites that responded to that tone/intensity combination divided by the A1 total area. The sampled frequency ranges in each of the experimental groups are not different from the sampled frequency range in experimentally naïve rats (1.5±0.1 – 29.3±0.5 kHz, $F(4,38) = 1.72$, $p = 0.17$). In each of the experimental groups, tones were considered to be significantly

increased ($p < 0.05$) compared to experimentally naïve rats if at least one neighboring tone was also significantly increased. Pearson's correlation coefficient was used to quantify the relationship between the percent of A1 responding to 1.8 or 4.3 kHz tones at 60 dB and the number of trained tasks or days of training.

The response strength to speech sounds was quantified using the number of evoked spikes within 40 ms of consonant onset, within 300 ms of vowel onset, and during a 700 ms window beginning at consonant onset and including the entire sound. A two-sample *F*-test for equal variances was used to examine the response variance to speech sounds. Post stimulus time histograms (PSTHs) were generated for each group for an example trained target sound ('dad'), an example trained non-target sound ('tad'), and an example novel sound ('pad'). The percent of A1 sites responding to speech sounds was calculated by dividing the number of sites that fire significantly more during the first 100 ms following speech sound presentation compared to 100 ms recorded before speech sound presentation by the total number of A1 sites. Neural discrimination between pairs of speech sounds was calculated using a nearest neighbor classifier [1, 5]. The spatiotemporal patterns evoked by each sound at each recording site using the 40 ms onset response with 1 ms bin precision were compared using the Euclidean distance measure. For example, in naïve rats, the spatiotemporal response pattern to each sound consisted of forty 1 ms bins at each of 392 A1 recording sites. For each pair of sounds, templates were created from 19 of the 20 repeats recorded at each site. The remaining single trial response was compared to each template, and the classifier assigned the single trial to the most similar template (smallest Euclidean distance). The difference in classifier percent correct between experimentally naïve rats and trained rats was quantified using the selectivity index $(PC_{\text{Trained}} - PC_{\text{Naive}})/(PC_{\text{Trained}} + PC_{\text{Naive}})$.

3. Results

3.1 Speech sound discrimination performance

The One Task groups of rats trained on either a voicing ($n = 5$ rats) or gender ($n = 5$ rats) discrimination task for two weeks. Both groups of rats heard an identical set of speech sounds: 'dad' and 'tad' spoken by 3 male and 3 female speakers ($n = 12$ sounds). The One Task Voicing rats were trained to lever press in response to 'dad' spoken by both males and females, while the One Task Gender rats were trained to lever press in response to females speaking both 'dad' and 'tad'. Rats were first trained to discriminate a female 'dad' from a male 'tad', and moved to the next task after reaching a performance criteria of a d' greater than 1.5 for 2 training sessions. One Task Voicing rats were not able to successfully discriminate the sounds better than chance performance on the first day of training (58.1 ± 3.7 percent correct, $p = 0.07$), but improved after two weeks of training and were able to discriminate the sounds better than chance performance (67.1 ± 4.9 percent correct, $p = 0.02$, Figure 1a). One Task Gender rats successfully discriminate sounds by gender on the first day of training (60.2 ± 2.6 percent correct, $p = 0.01$), and continued to accurately discriminate sounds better than chance performance on the last day of training (73.1 ± 6.2 percent correct, $p = 0.03$, Figure 1b).

Three groups of rats were trained on multiple speech tasks. Three Tasks Vowels rats ($n = 8$ rats) were trained to discriminate the vowel /æ/ in CVC context ('dad') from the vowels /ε/, /Λ/, /i/, and /u/ ('dead', 'dud', 'deed', and 'dood'). Half of the rats were first trained on vowels with a 'd' initial consonant for two weeks ('dad') followed by vowels with an 's' initial consonant for two weeks ('sad'), while the other four rats trained on vowels with an 's' initial consonant first. Performance was significantly better on the vowel tasks after 2 weeks of training compared to first day vowel discrimination performance (First day 59.6 ± 2.2 percent correct, last day 77.6 ± 2.6 percent correct, $p = 0.003$, Figure 1c,d). Six Tasks

Consonants rats ($n = 5$ rats) and Six Tasks Consonants and Vowels rats ($n = 6$ rats) each trained on 6 speech discrimination tasks for at least 2 weeks per task. After two weeks of training on each task, Six Tasks Consonants rats (First day 60.8 ± 1.3 percent correct, last day 71.7 ± 1.4 percent correct, $F(1,49) = 33.21$, $p < 0.0001$, Figure 1e) and Six Tasks Consonants and Vowels rats (First day 67.4 ± 1.3 percent correct, last day 74.2 ± 1.6 percent correct, $F(1,51) = 11.18$, $p = 0.002$, Figure 1f) were able to accurately discriminate most speech tasks, and overall performance was significantly better on the last day of training compared to the first day of training. Collectively, these results indicate that every speech trained group exhibited significant improvements in speech sound discrimination.

3.2 Speech sound responses

Following speech discrimination training, the rats were anesthetized and responses were recorded from 1,110 multi-unit primary auditory cortex sites. These responses were compared to 236 A1 sites recorded from five rats passively exposed to speech sounds, and 392 A1 sites recorded from eleven experimentally naïve rats (Table 1). The number of A1 sites recorded in each of the experimental groups are not different from each other (ANOVA $F(4,39) = 1.81$, $p = 0.15$).

Humans trained to discriminate speech sounds exhibit an increase in the N1-P2 peak-to-peak amplitude in response to the trained sounds compared to the responses evoked before training [32]. This effect has been documented after training on a single speech discrimination task; based on this, we predicted that all trained groups of rats would exhibit enhanced N1-P2 amplitudes in response to the trained sounds. It is possible that simple passive exposure to the sounds would alter A1 responses to speech sounds. However, the N1 amplitude did not significantly differ from experimentally naïve rats in Passive Exposure rats (-0.42 ± 0.04 mV in Passive Exposure rats vs. -0.43 ± 0.05 mV in experimentally naïve rats, $p = 0.88$, Figure 2a). The N1 amplitude also did not significantly differ from experimentally naïve rats in One Task trained rats (-0.35 ± 0.04 mV, $p = 0.2$, Figure 2b) or Six Tasks trained rats (-0.31 ± 0.03 mV, $p = 0.06$, Figure 2d). Instead, we found that the Three Tasks trained group of rats had a significantly reduced N1 amplitude evoked by the sound ‘dad’ compared to experimentally naïve rats (-0.31 ± 0.03 mV in Three Tasks trained rats vs. -0.43 ± 0.05 mV in experimentally naïve rats, $p = 0.04$, Figure 2c). This effect was not limited to trained sounds: the Three Tasks trained group of rats also exhibited a reduced N1 amplitude to a set of novel sounds (‘pad’, ‘kad’, ‘zad’, ‘wad’, ‘yad’, ‘vad’, ‘dayed’, ‘sid’, and ‘teed’) compared to naïve controls (-0.12 ± 0.01 mV in Three Tasks trained rats vs. -0.17 ± 0.02 mV in experimentally naïve rats, $p = 0.04$), while the N1 amplitude in response to novel sounds was unaffected in the other trained groups ($p > 0.05$). While our finding that moderate speech training decreases the N1 amplitude was not predicted, a previous study also found decreased activation following categorization training [33].

Many previous studies have reported plasticity specific to the trained stimuli. Our previous study demonstrated that the similarity of neural activity patterns in naïve rats using the first 40 ms of the A1 onset response predicts consonant discrimination [1]. After months of training on multiple speech tasks, we predicted that (1) the response strength to the trained sounds would increase, and (2) the neural activity patterns between two sounds that the rats could reliably discriminate would become more distinct. We first quantified the driven response to a set of 20 consonant sounds (‘b’, ‘d’, ‘g’, ‘p’, ‘t’, ‘k’, ‘s’, ‘z’, ‘w’, ‘y’, ‘r’, ‘l’, ‘f’, ‘v’, ‘m’, ‘n’, ‘h’, ‘j’, ‘sh’, and ‘ch’). Each of the trained groups of rats had heard a subset of these sounds, but most were novel sounds. The number of driven spikes evoked in the first 40 ms onset response to these consonant sounds did not significantly increase in any of the trained groups of rats compared to experimentally naïve rats ($p > 0.05$, Figure 3a). The response was also quantified for a set of 9 vowel sounds (‘dad’, ‘deed’, ‘dayed’, ‘dud’,

‘doed’, ‘dood’, ‘dead’, ‘dawed’, ‘did’) using the first 300 ms of the response to the vowel. The vowel sounds were also a combination of trained and novel sounds. The number of driven spikes evoked in response to these vowel sounds also did not significantly increase in any of the trained groups of rats compared to experimentally naïve rats ($p > 0.05$, Figure 3b).

For each group of rats, we calculated the driven response to target sounds that were rewarded, non-target sounds that were not rewarded, and novel sounds (Figure 4). For the voicing discrimination task, the number of driven spikes fired in response to target ‘dad’ sounds and non-target ‘tad’ sounds did not significantly increase in the rats that had trained on this task (the One Task Voicing group of rats and the Six Tasks Consonants and Vowels group of rats) compared to experimentally naïve rats ($p > 0.05$, Figure 5a,b). For the gender discrimination task, the number of driven spikes fired in response to target female sounds and non-target male sounds also did not significantly increase in the trained rats (the One Task Gender group of rats and the Six Tasks Consonants group of rats) compared to experimentally naïve rats ($p > 0.05$, Figure 5c,d). For both trained tasks, the driven response was not significantly increased using either of the duration ranges tested: the 40 ms onset response, or the 700 ms response to the entire sound ($p > 0.05$). Similarly, for the vowel discrimination task, there was no significant increase in the number of driven spikes in response to target and non-target vowels in the trained rats (the Three Tasks Vowels group of rats and the Six Tasks Consonants and Vowels group of rats) compared to experimentally naïve rats using either a 300 or 700 ms response window ($p > 0.05$). The driven response was also calculated for novel speech sounds that were not presented during training to any of the groups of rats. The response strength to novel sounds in any of the trained groups was not significantly different from experimentally naïve rats ($p > 0.05$, Figure 5e). In response to voicing, gender, and novel sounds, the response variance in each of the trained groups was not significantly different from experimentally naïve rats ($p > 0.05$).

We quantified the percent of A1 sites responding to speech sounds to determine how speech training affected the cortical response to both trained and novel speech sounds. The percent of A1 sites responding to a set of six trained speech sounds (‘dad’, ‘tad’, ‘sad’, ‘rad’, ‘shad’, and ‘dead’) was not significantly increased in any of the trained groups of rats compared to experimentally naïve rats ($81 \pm 5\%$ naïve; $79 \pm 3\%$ one task; $86 \pm 2\%$ three tasks; $77 \pm 4\%$ six tasks, $p > 0.05$). The percent of A1 sites responding to novel sounds in any of the trained groups was also not significantly different from experimentally naïve rats ($p > 0.05$).

The neural discrimination of a pair of speech sounds was quantified using a nearest-neighbor classifier. The classifier compares the single trial response to a sound with the average response evoked by each sound in the pair and selects the most similar. The activity pattern for each sound was made up of the 40 ms onset response at each A1 recording site using 1 ms precision. We predicted that following speech training on a task, the increase in behavioral performance would be accompanied by an increase in neural performance. However, classifier performance on each of the trained speech pairs was not significantly improved following speech training compared to classifier performance in experimentally naïve rats. The classifier percent correct was not enhanced in voicing trained rats compared to naïve rats ($p > 0.05$, Figure 6a), gender trained rats compared to naïve rats ($p > 0.05$, Figure 6b), or vowel trained rats compared to naïve rats ($p > 0.05$, Figure 6c). Additionally, the classifier percent correct to a set of novel sounds was not enhanced in any trained groups compared to naïve rats ($p > 0.05$, Figure 6d). Classifier performance on each trained pair in each trained group was plotted against the classifier performance on each trained pair in naïve rats ($R = 0.92$, $p < 0.001$, Figure 7). The difference in classifier performance was quantified using the selectivity index $(PC_{\text{Trained}} - PC_{\text{Naive}})/(PC_{\text{Trained}} + PC_{\text{Naive}})$. The mean selectivity index distribution of -0.0033 ± 0.0028 was not significantly different from the main diagonal ($p = 0.24$, Figure 7), which confirms that speech training does not improve

neural performance. Speech sound discrimination training does not increase the response strength to the trained sounds or make the neural activity patterns more discriminable, suggesting that the neural activity patterns evoked in experimentally naïve rats are sufficient for speech sound discrimination.

3.3 Receptive field plasticity

Previous studies have documented receptive field changes such as sharper tuning and longer latencies following long-term frequency discrimination training [12]. Six Tasks Consonants and Vowels rats and Six Tasks Consonants rats had broader bandwidths, shorter onset latency, decreased driven rate, and larger A1 area compared to naïve controls. Training on both consonants and vowels increased bandwidths by 33% and training on consonant tasks increased bandwidths by 20% (Bandwidth 20 dB above threshold: 2.4 ± 0.3 octaves in Six Tasks Consonants and Vowels rats, 2.2 ± 0.2 octaves in Six Tasks Consonants rats, and 1.8 ± 0.1 octaves in experimentally naïve rats, $p < 0.02$, Table 2). Neurons responded to tones more than 4 ms faster in rats trained on six tasks compared to naïve control rats (Onset latency: 11.9 ± 0.6 ms in Six Tasks Consonants and Vowels rats, 12.2 ± 0.1 ms in Six Tasks Consonants rats, and 16.5 ± 0.7 ms in experimentally naïve rats, $p = 0.0004$, Table 2). These rats exhibited a decrease in the driven rate to tones compared to experimentally naïve control rats (2.0 ± 0.3 spikes in Six Tasks Consonants and Vowels rats, 2.1 ± 0.2 spikes in Six Tasks Consonants rats, 3.0 ± 0.3 spikes in experimentally naïve rats, $p < 0.05$, Table 2). Six Tasks Consonants rats also had a 5.6 dB higher thresholds compared to naïve control rats (22.3 ± 1.5 dB vs. 16.7 ± 1.5 dB, $p = 0.03$, Table 2). Three Tasks Vowels rats did not have any of the bandwidth, latency, driven rate, or threshold differences seen in the other multiple task groups ($p > 0.05$, Table 2). Rats trained on One Task and rats passively exposed to speech sounds did not have any receptive field changes compared to naïve rats. Bandwidth, onset latency, driven rate in response to tones, and threshold were all not significantly different in any of the One Task or exposure groups of rats compared to experimentally naïve control rats ($p > 0.05$, Table 2).

The total area of A1 was larger in Six Tasks trained rats (ANOVA $F(4,39) = 3.9$, $p = 0.009$). Training on Six Tasks increased A1 area by 59% compared to experimentally naïve control rats (3.1 ± 0.3 mm² in Six Tasks rats vs. 1.9 ± 0.2 mm² in experimentally naïve rats, Figure 8). Rats trained on One Task or Three Tasks and rats passively exposed to speech sounds did not exhibit a significant increase in A1 area ($p > 0.05$, post hoc Bonferroni test, Figure 8).

The acoustics of the speech sounds that each group of rats heard were low frequency biased, with peaks in the power spectrum intensity around 1.8 and 4.3 kHz (Figure 9). We quantified the percent of cortex responding to these low frequency tones to determine how speech training affected the cortical response to low frequency tones. The percent of cortex responding to a 1.8 kHz tone at 60 dB was not significantly increased in Passive Exposure rats compared to experimentally naïve rats ($42 \pm 7\%$ vs. $39 \pm 4\%$, $p = 0.61$, Figure 10a,b). The One Task Voicing and One Task Gender groups of rats had an increase in the percent of cortex responding to a 1.8 kHz tone compared to experimentally naïve rats ($49 \pm 4\%$ vs. $39 \pm 4\%$, $p = 0.05$, Figure 10a,c), but not compared to Passive Exposure rats ($49 \pm 4\%$ vs. $42 \pm 7\%$, $p = 0.34$, Figure 10b,c). The percent of cortex responding to a low frequency tone was also significantly increased in the Three Tasks Vowels rats compared to experimentally naïve rats ($50 \pm 4\%$ vs. $39 \pm 4\%$, $p = 0.04$, Figure 10a,d), but not compared to Passive Exposure rats ($50 \pm 4\%$ vs. $42 \pm 7\%$, $p = 0.31$, Figure 10b,d). The Six Tasks Consonants rats and Six Tasks Consonants and Vowels rats also exhibited an increase in the percent of cortex responding to a 1.8 kHz tone compared to both experimentally naïve rats ($63 \pm 5\%$ vs. $39 \pm 4\%$, $p = 0.0003$, Figure 10a,e) and Passive Exposure rats ($63 \pm 5\%$ vs. $42 \pm 7\%$, $p = 0.02$, Figure 10b,e). Six Tasks trained rats also exhibited an increase in the percent of cortex

responding to a 4.3 kHz tone compared to experimentally naïve rats ($64 \pm 5\%$ vs. $49 \pm 3\%$, $p = 0.02$, Figure 10a,e).

In addition to the percent of cortex responding to a low frequency tone, we also quantified the A1 area responding to a low frequency tone. The A1 area responding to a 1.8 kHz tone at 60 dB was significantly increased in Passive Exposure rats compared to experimentally naïve rats ($1.2 \pm 0.2 \text{ mm}^2$ vs. $0.8 \pm 0.1 \text{ mm}^2$, $p = 0.04$, Figure 11a,b). The One Task speech trained rats had an increase in A1 area compared to experimentally naïve rats ($1.3 \pm 0.1 \text{ mm}^2$ vs. $0.8 \pm 0.1 \text{ mm}^2$, $p = 0.0007$, Figure 11a,c), but not Passive Exposure rats ($1.3 \pm 0.1 \text{ mm}^2$ vs. $1.2 \pm 0.2 \text{ mm}^2$, $p = 0.49$, Figure 11b,c). The A1 area responding to a low frequency tone was also significantly increased in the Three Tasks trained rats compared to experimentally naïve rats ($1.3 \pm 0.2 \text{ mm}^2$ vs. $0.8 \pm 0.1 \text{ mm}^2$, $p = 0.01$, Figure 11a,d), but not compared to Passive Exposure rats ($1.3 \pm 0.2 \text{ mm}^2$ vs. $1.2 \pm 0.2 \text{ mm}^2$, $p = 0.49$, $p = 0.75$, Figure 11b,d). The Six Tasks trained rats also exhibited an increase in the A1 area responding to a 1.8 kHz tone ($2.0 \pm 0.3 \text{ mm}^2$ vs. $0.8 \pm 0.1 \text{ mm}^2$, $p = 0.0001$, Figure 11a,e) and a 4.3 kHz tone ($2.0 \pm 0.3 \text{ mm}^2$ vs. $0.9 \pm 0.1 \text{ mm}^2$, $p = 0.0005$, Figure 11a,e) compared to experimentally naïve rats, but not Passive Exposure rats ($2.0 \pm 0.3 \text{ mm}^2$ vs. $1.2 \pm 0.2 \text{ mm}^2$ response to a 1.8 kHz tone, $p = 0.06$, Figure 11b,e).

The percent of cortex responding to low frequency tones was highly correlated with both the number of trained tasks ($R^2 = 0.91$, $p = 0.048$, Figure 12a) and the number of days of training ($R^2 = 0.94$, $p = 0.03$, Figure 12b). Rats that trained for a larger number of days or tasks had a larger percent of cortex responding to a 1.8 kHz tone, while rats trained for a shorter number of days or tasks had a smaller percent of cortex responding to a 1.8 kHz tone. For example, rats trained on one task had a smaller percent of cortex responding to a low frequency tone compared to rats trained on six tasks ($49 \pm 4\%$ vs. $63 \pm 5\%$, $p = 0.03$, Figure 12a). The correlation between map plasticity and training history trended towards significance if responses to a 4.3 kHz tone were used (number of trained tasks $R^2 = 0.76$, $p = 0.13$; number of days of training $R^2 = 0.71$, $p = 0.16$). These results indicate that several months of speech training on multiple tasks results in greater receptive field plasticity and a larger number of neurons that respond to low frequency sounds compared to speech training for two weeks on a single task.

The significantly expanded percent of cortex responding to low frequency tones was not limited to a single tone frequency or tone intensity. Rats trained on One Task had a significantly expanded percent of cortex responding to tones with frequencies between 1 – 1.8 kHz and intensities between 40 – 75 dB ($p < 0.05$, Figure 10c). Rats trained on Three Tasks had a significantly expanded percent of cortex responding to tones with frequencies between 1 – 2 kHz and intensities between 50 – 75 dB ($p < 0.05$, Figure 10d). Rats trained on Six Tasks had a significantly expanded percent of cortex responding to tones with frequencies between 1.1 – 5.7 kHz and intensities between 50 – 75 dB ($p < 0.05$, Figure 10e). These results indicate that cortical plasticity (expansion) in A1 occurs for frequencies that are most prominent in the training speech sounds (i.e. 1.8 and 4.3 kHz), especially as a function of increasing the number of training tasks and/or number of days of training.

While Six Tasks speech trained rats exhibit an increase in both total A1 area (Figure 8) and the amount of A1 area that responds to low frequency tones (Figure 11), they also exhibit a significant decrease in the driven response strength to tones (Table 2). To further examine this, we plotted the evoked response to tones of any frequency and intensity combination (Figure 13). As expected, Six Tasks trained rats exhibited an increase in the strength of response to a low frequency 1.8 kHz tone at 60 dB compared to experimentally naïve rats (2.1 ± 0.1 spikes vs. 1.4 ± 0.1 spikes, $p < 0.0001$, Figure 13a,e) and a decrease in the strength of response to a high frequency 25 kHz tone at 60 dB compared to experimentally

naïve rats (0.7 ± 0.1 spikes vs. 1.1 ± 0.1 spikes, $p = 0.003$, Figure 13a,e). This reduction in the response to a high frequency tone was significantly correlated with amount of training ($R^2 = 0.92$, $p = 0.04$). Collectively, these results demonstrate that training on multiple speech tasks alters tone frequency tuning in rat primary auditory cortex.

4. Discussion

4.1 Summary of results

Training on a tone frequency discrimination task improves performance and often results in A1 map plasticity [10-12]. In this study, we extend those findings by training rats on speech sound discrimination tasks and documenting A1 responses following training. Previous studies examining auditory cortex responses to vocalization sounds after training have shown that responses to the trained sounds are enhanced [13, 14]. Our study tested the hypothesis that speech sound discrimination training would either cause A1 map plasticity or would enhance A1 responses to the trained speech sounds. We found that extensive speech training alters A1 responses, but not in a manner that improves neural discrimination of the trained sounds. The neural discrimination of trained speech sound pairs was not enhanced in any of the trained groups, which suggests that the spatiotemporal activity patterns evoked in A1 of experimentally naïve rats are sufficient for speech discrimination [1]. Alternatively, critical discriminative plasticity for speech sounds might have occurred in other auditory cortical fields. Rats that were trained on six speech discrimination tasks had a larger percentage of A1 that responded to low frequency sounds, broader bandwidths, shorter onset latencies, and a decreased driven response to tones compared to experimentally naïve rats. Rats trained on a single speech discrimination task and passively exposed rats did not exhibit any changes in frequency selectivity or response strength, but did have a larger percentage of A1 that responded to low frequency sounds compared to experimentally naïve rats. Both the number of trained tasks and the number of days of training strongly predict the percent of A1 responding to a low frequency tone.

The use of speech sounds in the present experiment provides a basis for the study of speech processing in animal models. However, it is not yet known whether the findings are specifically applicable to speech sounds per se or to similar complex stimuli that have the same spectral components, with peaks near 1.8 and 4.3 kHz (Figure 9). Future studies with such stimuli can resolve this issue.

4.2 Trained sound plasticity in other auditory fields

Starlings trained to discriminate songs have an increased response strength in the caudomedial ventral hyperstriatum to familiar songs compared to unfamiliar songs [13], but a decreased response strength in the caudomedial nidopallium to familiar songs compared to unfamiliar songs [34]. It is possible that speech training resulted in stimulus specific plasticity in auditory areas other than A1. We have recently shown that behavioral consonant discrimination is correlated with neural discrimination of consonants using activity from anterior auditory field, ventral auditory field, or posterior auditory field of untrained rats [22]. While neural activity in each field could be used to predict consonant discrimination ability, each field represents the speech sounds somewhat differently, with less redundant firing to speech sounds in ventral and posterior auditory fields compared to A1 and anterior auditory field.

While our study recorded responses in primary auditory cortex, it is thought that there is a gradual transformation of information along the auditory pathway, as seen in the visual and somatosensory systems [35-37]. Previous studies have documented trained stimulus specific plasticity in non-primary auditory areas [13, 38] and non-primary visual areas [39, 40]. A

recent study documented the complexity of training-induced plasticity and found that response differences are dependent on both the training time course and the auditory field [38]. Our result showing weaker N1 amplitude but no difference in A1 spike response rate suggests that other nearby auditory fields may exhibit speech sound plasticity [41]. It is possible that the observed changes in the response properties of A1 neurons alter speech responses in other fields and could generate more distinct neural representations of the trained sounds. Future studies in A1 and higher regions are needed to further evaluate our hypothesis that extensive speech training on multiple tasks alters auditory cortex responses more than short term speech training on a single task. Additional studies will be needed to determine whether learning of more complex sounds requires plasticity in non-primary auditory cortex.

4.3 Trained sound plasticity time course

Many studies have documented stimulus specific plasticity following auditory training [11-13, 42]. For example, monkeys that were trained for months to discriminate tone frequency have more A1 neurons responding to the trained frequency compared to untrained monkeys [12]. Our finding that learning occurs without trained sound specific A1 plasticity is similar to previous findings that stimulus specific plasticity in A1 is not needed to maintain enhanced performance [14, 25, 43-45]. Under some training conditions, map reorganization is followed by a return to normal topography without a behavioral decrement [25, 46, 47]. A previous study proposed an Expansion – Renormalization model of map plasticity where an initial map expansion to the frequency of the trained tone occurs during learning, which increases the number of neurons responding to the trained frequency [25]. The most efficient circuits to perform the task are selected, and the map is later renormalized after learning is complete [48]. Since each of the rats in our study had learned a new speech discrimination task within two weeks of our neural recordings, they were likely in the expansion stage of learning [25, 38]. It appears that the Six Tasks group had the largest map plasticity because they had multiple rounds of training. Earlier reports that frequency map plasticity is transient did not vary the task conditions to drive multiple rounds of learning. One earlier study showed that it is possible to maintain A1 plasticity for long periods by continually changing task parameters [12]. It is possible that persistent speech-specific changes occur in brain regions other than A1. It is possible that training specific plasticity occurred in A1 at an earlier time point in our study, but was not present at the time point of our recordings. A1 plasticity only occurs when subjects are appropriately motivated [44, 49, 50]. While changes do occur in A1 after speech training in this study, there is no trained sound specific plasticity, which is consistent with previous work in humans showing generalization to untrained stimuli following speech training [51].

Additional studies will be needed covarying the length of training and the number of trained tasks in order to dissociate the duration of training from the number of trained tasks. With the current experimental design, it is not possible to determine whether the number of tasks or the duration of training was responsible for the observed response changes. Further experiments will be necessary training on a single task for many weeks as well as training on multiple tasks in a short amount of time.

4.4 Other factors affecting trained sound plasticity

It is possible that anesthesia masked training-induced changes in A1 speech responses. Auditory cortex responses to repetitive broadband click stimuli are suppressed in anesthetized recordings compared to awake recordings [52]. Responses to the first click in the train were not significantly different between awake and anesthetized recording states, but responses to subsequent clicks occurred later and were weaker in anesthetized recordings compared to awake recordings. For this study, our analysis used only the initial

onset response to the speech sounds, which is unlikely to be significantly affected. Although anesthesia alters the response properties to repetitive stimuli, anesthesia does not mask the training induced changes in A1 responses to tones [12]. Additionally, our previous study documented that anesthetized A1 speech responses can be used to predict behavioral discrimination just as accurately as A1 speech responses recorded from awake passively listening animals [1]. Additional studies in awake speech trained animals are needed to determine if anesthesia masks trained sound plasticity in A1.

It is also possible that we were unable to see a trained sound specific effect due to recording multi-units instead of single units. Our finding that multiple task trained rats exhibited decreased N1 amplitude is consistent with a recent study showing that enhanced sparse coding results in overall decreased activity, while the individual neurons that respond to the trained stimulus exhibited enhanced response strength [53]. If speech training enhances sparse coding, we would not be able to see this trained sound enhancement effect using our multi-unit recordings.

4.5 Neural plasticity after speech discrimination training

Speech sounds are broadband sounds, and neurons located all across the A1 frequency map can contribute to speech sound discrimination [1]. It was possible that speech training could enhance neural speech discrimination without altering A1 topography. Since speech sounds have more spectral power at low frequencies, it was possible that the population of neurons that respond strongly to speech sounds would be expanded (i.e. low frequency map expansion). It was also possible that speech training could result in both enhanced speech discrimination and A1 map reorganization, or speech training could result in neither. Our results show that the population of A1 neurons that respond to pure tones of 1.8 or 4.3 kHz, which were the peaks of spectral power for speech sounds, was expanded. This finding rejects the hypothesis that speech discrimination training would improve the A1 response to the target sounds and reduce the response to the non-target sounds. Speech discrimination learning may 1) require plasticity in a brain region other than A1, 2) generate changes in A1 that are only observable during behavior, or 3) result in changes in a small subpopulation of A1 neurons. All of these possibilities suggest that the earlier models of cortical plasticity may need to be revised and new explanations of the neural basis of auditory learning should be considered and tested.

References

1. Engineer CT, Perez CA, Chen YTH, 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]
2. Kluender KR, Diehl RL, Killeen PR. Japanese quail can learn phonetic categories. *Science*. 1987; 237:1195–7. [PubMed: 3629235]
3. Kuhl PK, Miller JD. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science*. 1975; 190:69–72. [PubMed: 1166301]
4. Reed P, Howell P, Sackin S, Pizzimenti L, Rosen S. Speech perception in rats: use of duration and rise time cues in labeling of affricate/fricative sounds. *Journal of the experimental analysis of behavior*. 2003; 80:205–15. [PubMed: 14674729]
5. Perez CA, Engineer CT, Jakkamsetti V, Carraway RS, Perry MS, Kilgard MP. Different Timescales for the Neural Coding of Consonant and Vowel Sounds. *Cereb Cortex*. 2013; 23:670–683. [PubMed: 22426334]
6. Steinschneider, M. Anonymous Neural Correlates of Auditory Cognition. Springer; 2013. Phonemic Representations and Categories.; p. 151-191.

7. Ranasinghe KG, Vrana WA, Matney CJ, Kilgard MP. Neural mechanisms supporting robust discrimination of spectrally and temporally degraded speech. *J Assoc Res Otolaryngol.* 2012; 13:527–42. [PubMed: 22549175]
8. Shetake JA, Wolf JT, Cheung RJ, Engineer CT, Ram SK, Kilgard MP. Cortical activity patterns predict robust speech discrimination ability in noise. *The European journal of neuroscience.* 2011; 34:1823–38. [PubMed: 22098331]
9. Engineer, CT.; Kilgard, MP.; Assmann, PF.; Atzori, M.; Thompson, LT. *Speech Sound Coding and Training-Induced Plasticity in Primary Auditory Cortex.* Citeseer; 2008.
10. Edeline JM, Pham P, Weinberger NM. Rapid development of learning-induced receptive field plasticity in the auditory cortex. *Behavioral neuroscience.* 1993; 107:539–51. [PubMed: 8397859]
11. Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. *J Neurosci.* 2006; 26:4970–82. [PubMed: 16672673]
12. 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]
13. Gentner TQ, Margoliash D. Neuronal populations and single cells representing learned auditory objects. *Nature.* 2003; 424:669–74. [PubMed: 12904792]
14. Schnupp JW, Hall TM, Kokelaar RF, Ahmed B. Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci.* 2006; 26:4785–95. [PubMed: 16672651]
15. Callan DE, Tajima K, Callan AM, Kubo R, Masaki S, Akahane-Yamada R. Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *Neuroimage.* 2003; 19:113–124. [PubMed: 12781731]
16. Bradlow AR, Akahane-Yamada R, Pisoni DB, Tohkura Y. Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Percept Psychophys.* 1999; 61:977–985. [PubMed: 10499009]
17. Lively SE, Pisoni DB, Yamada RA, Tohkura Y, Yamada T. Training Japanese listeners to identify English/r/and/l/. III. Long-term retention of new phonetic categories. *J Acoust Soc Am.* 1994; 96:2076. [PubMed: 7963022]
18. MacKain KS, Best CT, Strange W. Categorical perception of English/r/and/l/by Japanese bilinguals. *Applied Psycholinguistics.* 1981; 2:369–390.
19. Strange W, Dittmann S. Effects of discrimination training on the perception of/r/and/l/by Japanese adults learning English. *Percept Psychophys.* 1984; 36:131–145. [PubMed: 6514522]
20. Kawahara H. Speech representation and transformation using adaptive interpolation of weighted spectrum: vocoder revisited. 1997; 2:1303–1306.
21. Kawahara H. Speech representation and transformation using adaptive interpolation of weighted spectrum: Vocoder revisited. *Proceedings of the ICASSP.* 1997; 2:1303–1306.
22. Centanni TM, Engineer CT, Kilgard MP. Cortical speech-evoked response patterns in multiple auditory fields are correlated with behavioral discrimination ability. *J Neurophysiol.* 2013
23. Peterson GE, Barney HL. Control Methods Used in a Study of the Vowels. *J Acoust Soc Am.* 1952; 24:175–184.
24. Wright, R. A review of perceptual cues and cue robustness.. In: Hayes, B.; Kirchner, R.; Steriade, D., editors. *Phonetically based phonology.* Cambridge University Press; 2004.
25. 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; 70:121–31. [PubMed: 21482361]
26. Kilgard MP, Merzenich MM. Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hear Res.* 1999; 134:16–28. [PubMed: 10452372]
27. Kilgard MP, Merzenich MM. Cortical map reorganization enabled by nucleus basalis activity. *Science.* 1998; 279:1714–1718. [PubMed: 9497289]
28. 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]

29. Porter BA, Rosenthal TR, Ranasinghe KG, Kilgard MP. Discrimination of brief speech sounds is impaired in rats with auditory cortex lesions. *Behavioural brain research*. 2011; 219:68–74. [PubMed: 21167211]
30. Ranasinghe KG, Vrana WA, Matney CJ, Kilgard MP. Neural Mechanisms Supporting Robust Discrimination of Spectrally and Temporally Degraded Speech. *JARO-Journal of the Association for Research in Otolaryngology*. 2012:1–16.
31. Shetake JA, Wolf JT, Cheung RJ, Engineer CT, Ram SK, Kilgard MP. Cortical activity patterns predict robust speech discrimination ability in noise. *Eur J Neurosci*. 2011
32. Tremblay K, Kraus N, McGee T, Ponton C, Otis B. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and hearing*. 2001; 22:79–90. [PubMed: 11324846]
33. Guenther FH, Nieto-Castanon A, Ghosh SS, Tourville JA. Representation of sound categories in auditory cortical maps. *Journal of Speech, Language and Hearing Research*. 2004; 47:46.
34. Thompson JV, Gentner TQ. Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. *J Neurophysiol*. 2010; 103:1785–1797. [PubMed: 20107117]
35. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci*. 2003; 23:5235–46. [PubMed: 12832548]
36. Romo R, Salinas E. Flutter discrimination: neural codes, perception, memory and decision making. *Nature reviews*. 2003; 4:203–18.
37. Russ BE, Ackelson AL, Baker AE, Cohen YE. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *Journal of neurophysiology*. 2008; 99:87–95. [PubMed: 18003874]
38. Takahashi H, Yokota R, Funamizu A, Kose H, Kanzaki R. Learning-stage-dependent, field-specific, map plasticity in the rat auditory cortex during appetitive operant conditioning. *Neuroscience*. 2011; 199:243–258. [PubMed: 21985937]
39. Moran J, Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science*. 1985; 229:782–784. [PubMed: 4023713]
40. Spitzer H, Desimone R, Moran J. Increased attention enhances both behavioral and neuronal performance. *Science*. 1988; 240:338–340. [PubMed: 3353728]
41. Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*. 1987; 24:375–425. [PubMed: 3615753]
42. Bakin JS, Weinberger NM. Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. *Brain Res*. 1990; 536:271–286. [PubMed: 2085753]
43. 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; 14:952–65. [PubMed: 15115736]
44. Engineer ND, Engineer CT, Reed AC, Pandya PK, Jakkamsetti V, Moucha R, Kilgard MP. Inverted-U function relating cortical plasticity and task difficulty. *Neuroscience*. 2012; 205:81–90. [PubMed: 22249158]
45. Wong SW, Schreiner CE. Representation of CV-sounds in cat primary auditory cortex: intensity dependence. *Speech Communication*. 2003; 41:93–106.
46. Molina-Luna K, Hertler B, Buitrago MM, Luft AR. Motor learning transiently changes cortical somatotopy. *Neuroimage*. 2008; 40:1748–1754. [PubMed: 18329289]
47. Takahashi H, Funamizu A, Mitsumori Y, Kose H, Kanzaki R. Progressive plasticity of auditory cortex during appetitive operant conditioning. *BioSystems*. 2010; 101:37–41. [PubMed: 20398726]
48. Kilgard MP. Harnessing plasticity to understand learning and treat disease. *Trends Neurosci*. 2012
49. Bieszczad KM, Weinberger NM. Representational gain in cortical area underlies increase of memory strength. *Proceedings of the National Academy of Sciences*. 2010; 107:3793–3798.
50. David SV, Fritz JB, Shamma SA. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proceedings of the National Academy of Sciences*. 2012; 109:2144–2149.

51. Tremblay K, Kraus N, Carrell TD, McGee T. Central auditory system plasticity: generalization to novel stimuli following listening training. *The Journal of the Acoustical Society of America*. 1997; 102:3762–73. [PubMed: 9407668]
52. Rennaker R, Carey H, Anderson S, Sloan A, Kilgard M. Anesthesia suppresses nonsynchronous responses to repetitive broadband stimuli. *Neuroscience*. 2007; 145:357–369. [PubMed: 17207583]
53. Gdalyahu A, Tring E, Polack P, 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]
54. Stevens, KN. *Acoustic Phonetics*. The MIT press; 2000.

- Extensive speech training can reorganize the A1 frequency map.
- The amount of speech training was well correlated with the extent of map expansion.
- Training did not improve neural discrimination or responses to speech sounds.

Figure 1.

Time course of speech sound training for the (a) One Task Voicing group (n=5 rats), (b) One Task Gender group (n=5 rats), (c & d) Three Tasks Vowels groups (n=8 rats), (e) Six Tasks Consonants group (n=5 rats), and (f) Six Tasks Consonants and Vowels group (n=6 rats). The first ten days of training on each task are plotted. Symbols indicate the mean percent correct on each day of training. Error bars indicate s.e.m. across rats.

Figure 2.

Auditory cortex local field potential recordings in response to the word 'dad'. (a) The N1 amplitude in Passive Exposure rats (gray line) does not significantly differ from experimentally naïve rats (black line). Gray shading behind each line indicates s.e.m. across rats. (b) The N1 amplitude in One Task trained rats (gray line) does not significantly differ from experimentally naïve rats. (c) The N1 amplitude in Three Task trained rats (gray line) is significantly reduced compared to experimentally naïve rats (black line). (d) The N1 amplitude in Six Task trained rats (gray line) does not significantly differ from experimentally naïve rats.

Figure 3.

There was no difference in the number of evoked spikes in response to consonant and vowel sounds following speech training. **(a)** Driven response to a set of 20 consonant sounds in experimentally naïve rats, Passive Exposure rats, and rats trained on one, three, or six speech discrimination tasks. The number of driven spikes was calculated for the 40 ms onset response. Error bars indicate s.e.m. across rats. **(b)** Driven response to a set of 9 vowel sounds. The number of driven spikes was calculated for the 300 ms onset response.

Figure 4.

Post stimulus time histograms (PSTHs) in response to speech sounds in each of the groups.

(a) PSTHs of the driven response to the target speech sound ‘dad’, the non-target speech sound ‘tad’, and a novel speech sound ‘pad’ in the experimentally naïve group of rats. The sound ‘dad’ evokes a single peak of activity in response to the consonant onset, the sound ‘tad’ evokes a response to both the onset of the consonant and the onset of the vowel, and the sound ‘pad’ evokes a response to both the onset of the consonant and the onset of the vowel. Gray shading behind each PSTH indicates s.e.m. across rats. (b) PSTHs to speech sounds in the Passive Exposure group of rats. (c) PSTHs to speech sounds in the One Task

trained group of rats. **(d)** PSTHs to speech sounds in the Three Tasks trained group of rats.
(e) PSTHs to speech sounds in the Six Tasks trained group of rats.

Figure 5.

No difference in the number of driven spikes evoked by target, non-target, and novel sounds after speech training. (a) For the voicing discrimination task, target 'dad' sounds were spoken by 3 male and 3 female speakers. Both the Six Tasks rats and the One Task rats were trained on this task for 10 days, while the Passive Exposure rats (0 tasks right square) were passively exposed to these same sounds for 10 days. The number of driven spikes was calculated for the 40 ms onset response. Error bars indicate s.e.m. across rats. Square symbols indicate groups of rats that were not trained on this task, while circle symbols indicate groups of rats that were trained on the voicing discrimination task. (b) For the voicing discrimination task, non-target 'tad' sounds were also spoken by the same 3 male

and 3 female speakers. **(c)** For the gender discrimination task, target female sounds were the words ‘dad’ and ‘tad’ spoken by 3 female speakers. **(d)** For the gender discrimination task, non-target male sounds were the words ‘dad’ and ‘tad’ spoken by 3 male speakers. **(e)** The nine novel sounds were never presented during training to any of the groups of rats, and included the words ‘pad’, ‘kad’, ‘zad’, ‘wad’, ‘yad’, ‘vad’, ‘dayed’, ‘sid’, and ‘teed’.

Figure 6.

There is no improvement in classifier percent correct on multiple speech tasks following speech training. (a) Average classifier performance on six 'dad' vs. 'tad' pairs for each group of rats. Square symbols indicate groups of rats that were not trained on this task, while circle symbols indicate groups of rats that were trained on the voicing discrimination task. Both the Six Tasks rats (6 tasks right circle) and the One Task rats (1 task left circle) were trained on this task for 10 days, while the Passive Exposure rats (0 tasks right square) were passively exposed to these same sounds for 10 days. The classifier was provided with the 40 ms onset response using 1 ms bins. Error bars indicate s.e.m. across rats. (b) Average classifier performance on six female vs. male pairs for each group of rats. Both the Six Tasks rats (6 tasks left circle) and the One Task rats (1 task right circle) were trained on this

task for 10 days, while the Passive Exposure rats (0 tasks right square) were passively exposed to these same sounds for 10 days. The classifier was provided with the 40 ms onset response using 1 ms bins. **(c)** Average classifier performance on eight vowel pairs for each group of rats. Both the Six Tasks rats (6 tasks right circle) and the Three Task rats (3 tasks circle) were trained on this task. The classifier was provided with the first 300 ms response using a single 300 ms bin. **(d)** Average classifier performance on six novel pairs for each group of rats. The nine novel sounds were never presented during training to any of the groups of rats, and included the words 'pad', 'kad', 'zad', 'wad', 'yad', 'vad', 'dayed', 'sid', and 'teed'. The classifier was provided with the 40 ms onset response using 1 ms bins.

Figure 7.

Speech training does not improve neural discrimination of trained pairs. The classifier percent correct for each of the trained pairs is plotted in experimentally naïve control rats (*x* axis) compared to speech trained rats (*y* axis). Neural discrimination performance in experimentally naïve rats on each trained pair was highly correlated with neural discrimination performance in speech trained rats ($R = 0.92$, $p < 0.001$). The mean selectivity index confirmed that speech trained did not improve neural discrimination compared to experimentally naïve rats (-0.0033 ± 0.0028 , $p = 0.24$).

Figure 8.

A1 area (mm²) is larger in Six Tasks speech trained rats compared to experimentally naïve rats. Training on Six Tasks increased A1 area by 59% compared to experimentally naïve control rats ($p < 0.05$). Rats trained on One Task or Three Tasks and rats passively exposed to speech sounds did not exhibit a significant increase in A1 area ($p > 0.05$). Error bars indicate s.e.m. across rats.

Figure 9.

Speech sound spectrograms and power spectra. **(a)** Spectrograms of a set of five speech sounds. Time is represented on the x axis (–50 to 500 ms) and frequency is represented on the y axis (0 to 35 kHz). As in our previous studies, speech sounds were shifted up by one octave in order to better match the rat hearing range **(b)** Power spectra for the target sound ‘dad’ spoken by a female speaker, and the average power spectrum for all 33 trained sounds used in the study. Frequency is represented on the x axis (0.5 to 32 kHz) and relative power is represented on the y axis (–35 to 0 dB). The black line indicates the mean power spectrum, and the gray lines indicate the s.e.m. across sounds. The power spectrum peak of 1.8 kHz is one octave higher than the human power spectrum peak of 0.9 kHz [54].

Figure 10.

The percent of cortex responding to low frequency tones increases in speech trained rats. (a) The percentage of A1 neurons that respond to a tone of any frequency and intensity combination in experimentally naïve rats. Black contour lines indicate 20, 40, and 60 percent responses. The filled white circles indicate the 1.8 and 4.3 kHz tones used to compare the change in percent of cortex responding to the tone across groups. (b) The percentage of A1 neurons that respond to a tone of any frequency and intensity combination in Passive Exposure rats. White contour lines surround the regions of tones that are significantly increased ($p < 0.05$ with at least one neighboring significant tone) compared to experimentally naïve rats. (c) The percentage of A1 neurons that respond to a tone of any frequency and intensity combination in One Task trained rats. (d) The percentage of A1 neurons that respond to a tone of any frequency and intensity combination in Three Tasks trained rats. (e) The percentage of A1 neurons that respond to a tone of any frequency and intensity combination in Six Tasks trained rats.

Figure 11.

The A1 area responding to low frequency tones increases in speech trained rats. (a) The A1 area (mm²) that responds to a tone of any frequency and intensity combination in experimentally naïve rats. The filled white circles indicate the 1.8 or 4.3 kHz tones used to compare the change in A1 area responding to the tone across groups. (b) The A1 area that responds to a tone of any frequency and intensity combination in Passive Exposure rats. White contour lines surround the regions of tones that are significantly increased ($p < 0.05$ with at least one neighboring significant tone) compared to experimentally naïve rats. (c) The A1 area that responds to a tone of any frequency and intensity combination in One Task trained rats. (d) The A1 area that responds to a tone of any frequency and intensity combination in Three Tasks trained rats. (e) The A1 area that responds to a tone of any frequency and intensity combination in Six Tasks trained rats.

Figure 12.

Both the number of trained tasks and the number of days of training strongly predict the percent of A1 responding to a low frequency tone. **(a)** The percent of A1 responding to a low frequency tone increases as the number of trained tasks increases. Zero trained tasks includes both experimentally naïve rats and Passive Exposure rats, one trained task includes One Task Voicing and One Task Gender rats, three trained tasks includes Three Tasks Vowels rats, and six trained tasks includes Six Tasks Consonants rats and Six Tasks Consonants and Vowels rats. Error bars indicate s.e.m. across rats. **(b)** The percent of A1 responding to a low frequency tone increases as the number of days of training increases.

Figure 13.

The average number of spikes per tone in speech trained rats increases for low frequency tones and decreases for high frequency tones. (a) The number of spikes evoked in response to a tone of any frequency and intensity combination in experimentally naïve rats. The filled white circles indicate the 1.8 and 4.3 kHz tones used to compare the number of spikes per tone across groups. (b) The number of spikes that respond to a tone of any frequency and intensity combination in Passive Exposure rats. White contour lines surround the regions of tones that are significantly increased ($p < 0.05$ with at least two neighboring significant tones) compared to experimentally naïve rats, while black contour lines surround the regions of tones that are significantly decreased compared to experimentally naïve rats. (c) The number of spikes that respond to a tone of any frequency and intensity combination in One Task trained rats. (d) The number of spikes that respond to a tone of any frequency and intensity combination in Three Tasks trained rats. (e) The number of spikes that respond to a tone of any frequency and intensity combination in Six Tasks trained rats.

Table 1

Experimental groups.

Experimental Groups	Subgroups	Trained tasks	Total Days of Training	# of rats	# of A1 sites
Naïve control	-	-	-	11	392
Passive Exposure	-	-	27	5	235
One Task	Voicing	voicing multiple speaker	29	5	209
	Gender	gender with 'd' and 't' sounds	29	5	222
Three Tasks	Vowels	dad vs. dead, dud, deed, dood sad vs. said, sud, seed, sood dad or sad vs. distractors	57	8	252
Six Tasks	Consonants	pitch gender with 'c' sounds gender with 'd' and 't' sounds mad vs. nad shad vs. fad, sad, had shad vs. chad, jad	88	5	269
	Consonants & Vowels	dad vs. tad voicing compression voicing multiple speaker rad vs. lad dad vs. dead, dud, deed, dood dad vs. bad, gad	136	6	158
Totals:				45	1738

Table 2

Receptive field plasticity in the Six Task trained rats.

Group	Threshold (dB)	BW10 (octaves)	BW20 (octaves)	BW30 (octaves)	BW40 (octaves)	Onset Latency (ms)	Driven Rate (spikes/tone)
Naïve Control	16.7	1.3	1.8	2.1	2.4	16.5	3.0
Passive Exposure	14.3	1.3	1.8	2.1	2.4	14.9	3.2
One Task Gender	15.7	1.4	1.9	2.3	2.5	15.8	3.3
One Task Voicing	18.1	1.2	1.6	1.9	2.2	16.1	3.3
Three Tasks Vowel	19.6	1.4	1.8	2.1	2.4	17.1	3.4
Six Tasks Consonants	22.3*	1.5	2.2*	2.6*	2.8*	12.2 [#]	2.1 [#]
Six Tasks Consonant Vowel	19.1	1.8*	2.4*	2.8*	3.1*	11.9 [#]	2.0 [#]

* Numbers marked with a star are significantly increased compared to experimentally naïve rats ($p < 0.05$)

[#] while numbers marked with a pound sign are significantly decreased compared to experimentally naïve rats ($p < 0.05$).